J Optim Theory Appl (2010) 146: 267–303
DOI 10.1007/s10957-010-9667-4

Hybrid Approximate Proximal Method with Auxiliary Variational Inequality for Vector Optimization

L.C. Ceng · B.S. Mordukhovich · J.C. Yao

Published online: 18 February 2010

Abstract This paper studies a general vector optimization problem of finding weakly efficient points for mappings from Hilbert spaces to arbitrary Banach spaces, where the latter are partially ordered by some closed, convex, and pointed cones with nonempty interiors. To find solutions of this vector optimization problem, we introduce an auxiliary variational inequality problem for a monotone and Lipschitz continuous mapping. The approximate proximal method in vector optimization is extended to develop a hybrid approximate proximal method for the general vector optimization problem under consideration by combining an extragradient method to find a solution of the variational inequality problem and an approximate proximal point method for

Communicated by F. Giannessi.

L.C. Ceng research was partially supported by the Leading Academic Discipline Project of Shanghai Normal University (DZL707), Innovation Program of Shanghai Municipal Education Commission Grant (09ZZ133), National Science Foundation of China (10771141), PhD Program Foundation of Ministry of Education of China (20070270004), Science and Technology Commission of Shanghai Municipality Grant (075105118), and Shanghai Leading Academic Discipline Project (S30405).
B.S. Mordukhovich research was partially supported by the USA National Science Foundation under Grant DMS-0603896.
J.C. Yao research was partially supported by Grant NSC 98-2923-E-110-003-MY3.

L.C. Ceng
Department of Mathematics, Shanghai Normal University, Shanghai 200234, China
e-mail: zenglc@hotmail.com

L.C. Ceng
Scientific Computing Key Laboratory of Shanghai Universities, Shanghai, China

B.S. Mordukhovich
Department of Mathematics, Wayne State University, Detroit, MI 48202, USA
e-mail: boris@math.wayne.edu

J.C. Yao (✉)
Department of Applied Mathematics, National Sun Yat-sen University, 804 Kaohsiung, Taiwan
e-mail: yaojc@math.nsysu.edu.tw

finding a root of a maximal monotone operator. In this hybrid approximate proximal method, the subproblems consist of finding approximate solutions to the variational inequality problem for monotone and Lipschitz continuous mapping, and then finding weakly efficient points for a suitable regularization of the original mapping. We present both absolute and relative versions of our hybrid algorithm in which the subproblems are solved only approximately. The weak convergence of the generated sequence to a weak efficient point is established under quite mild assumptions. In addition, we develop some extensions of our hybrid algorithms for vector optimization by using Bregman-type functions.

Keywords Vector optimization · Proximal points · Hybrid inexact algorithms · Auxiliary variational inequalities · Banach spacies

1 Introduction and Overview

In this introductory section we first discuss recent extensions of the proximal point method to solve problems of vector optimization, and then overview some developments on extragradient methods of solving variational inequalities. The main contributions of this paper involve developing efficient *hybrid* algorithms to solve vector optimization problems in infinite dimensions that combine and unify some constructions for finding roots of maximal monotone operators used in proximal point methods with those employed in iterative extragradient schemes of finding approximate solutions to auxiliary variational inequalities associated with the vector optimization problems under consideration.

We start with an overview of recently developed iterative algorithms to solve vector optimization and related problems. In their paper [1], Bonnel, Iusem and Svaiter introduced and studied some extensions to vector-valued optimization of several iterative methods for scalar-valued functions. In those extensions, they defined iterates in the vector-valued case by considering the order $\preceq_C$ on a Banach space Y, mimicking whenever it is possible a role of the usual order on the real line $\mathcal{R}$ in the corresponding algorithms for scalar-valued optimization. Meantime, they admitted the possibility that mappings $F : X \to Y$ take the value ∞_C (this is made precise in Sect. 2), where X is a Hilbert space, and where C is a closed, convex, and pointed cone in Y with $\operatorname{int} C \neq \emptyset$; $\operatorname{int} C$ denotes as usual the interior of the set C. Such extensions can be traced back to the fashion of extensions, which always exist in the finite-dimensional setting; compare in $\mathcal{R}^m$, e.g., the steepest descent method for multiobjective optimization [2], the same method for general finite-dimensional vector optimization [3], and the projected gradient method for convexly constrained vector optimization [4].

Let us recall some basic versions of the *proximal point method* in scalar-valued convex optimization, which are strongly related to finding *zeros of maximal monotone operators*. Given a Hilbert space X and a set-valued operator $T : X \to 2^X$, the classical proximal point method, in its so-called *exact version*, is an iterative procedure for finding a zero of T, i.e., a point $z \in X$ such that $0 \in T(z)$. The method generates a sequence $\{x_n\} \subset X$, starting with 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, the next iterate x_{n+1}

is a unique vector in X such that $0 \in T_n(x_{n+1})$, where $T_n : X \to 2^X$ is defined as $T_n(x) := T(x) + \alpha_n(x - x_n)$. In other words, whenever T is a maximal monotone operator, the proximal point method means that, starting with any vector $x_0 \in X$, iteratively updates x_{n+1} conforming to the following recursion:

$$x_{n+1} + c_n T(x_{n+1}) \ni x_n, \tag{1}$$

where $\{c_n\} \subset [c, \infty)$, $c > 0$, is a sequence of scalars. However, as pointed out in [5], the ideal form of this method is often impractical, since in many cases solving problem (1) exactly is either impossible or as difficult as solving the original problem $0 \in T(x)$. On the other hand, there seems to be little justification of the effort required to solve the problem accurately when the iterate is far from the solution point.

In [6], Rockafellar developed an *inexact* variant of the method:

$$x_{n+1} + c_n T(x_{n+1}) \ni x_n + \theta_{n+1}, \tag{2}$$

where θ_{n+1} is regarded as an error sequence. This method is called the *inexact proximal point algorithm*. It is proved in [6], in the finite-dimensional setting, that if $\theta_n \to 0$ sufficiently fast so that $\sum_{n=1}^{\infty} \|\theta_n\| < \infty$, then $x_n \to z \in \mathcal{R}^m$ with $0 \in T(z)$. Because of its relaxed requirement, the inexact proximal point algorithm is more practical than the exact one. Thus it has been studied widely, and various forms of the method have been developed; see, e.g., [4, 7–12]. In most of these papers, conditions ensuring that the error term being summable are essential requirements for the convergence of the method. In [6] and some subsequent papers (e.g., [13]), a criterion for this is as follows:

$$\|\theta_{n+1}\| \le \sigma_n \|x_{n+1} - x_n\|, \quad \text{with } \sum_{n=0}^{\infty} \sigma_n < \infty. \tag{3}$$

In the further development, Eckstein [5] extended the inexact proximal point method by Rockafellar to the corresponding algorithm based on *Bregman functions*; see more references in [5]. It is proved therein that the sequence $\{x_n\}$ generated by the latter algorithm converges to a root of T under the conditions

$$\sum_{n=1}^{\infty} \|\theta_n\| < \infty \quad \text{and} \quad \sum_{n=1}^{\infty} \langle \theta_n, x_n \rangle \quad \text{exists and is finite} \tag{4}$$

(see (18) and (19) in [5]). Conditions (4) involve assumptions on the whole generated sequence $\{x_n\}$ and the error term sequence $\{\theta_n\}$, but nevertheless they can be checked and enforced in practice more easily than those existed earlier. Subsequently Solodov and Svaiter [15–17] proposed more accurate criteria for proximal point algorithms. Their criteria are different from (3) requiring that

$$\sup_{n \ge 0} \sigma_n < 1. \tag{5}$$

However, in [15–17] this comes at the cost of involving an additional projection or "extragradient" step in the algorithm, and it turns out that the applicable portion of [14] is efficient mainly for problems of convex minimization.

Note also that He [9] proposed a different inexact criterion for the study of monotone variational inequalities, which involves relationships between the error term and the residual function. In particular, the restriction $\sum_{n=0}^{\infty}\sigma_n < \infty$ in (3) is replaced in [9] by

$$\sum_{n=0}^{\infty}\sigma_n^2 < \infty. \tag{6}$$

The main attention in what follows is paid to the following problem of *constrained vector optimization*. Let Ω be a nonempty, closed, and convex subset of a Hilbert space X, and let $F:\Omega\to Y\cup\{\infty_C\}$. Utilizing the ordering cone C, we have a partial order $\preceq_C$ on Y given by $y\preceq_C y'$ if and only if $y'-y\in C$ and the associate relation $\prec_C$ given by $y\prec_C y'$ if and only if $y'-y\in \operatorname{int} C$. It is said as usual that $\bar{x}\in\Omega$ is a *weakly efficient minimizer* of F with respect to $\preceq_C$ if there exists no $x\in\Omega$ satisfying $F(x)\prec_C F(\bar{x})$.

Considering a closed and convex set $K\subset X$ and an extended-real-valued convex function $f:X\to\mathcal{R}\cup\{\infty\}$, recall that the *indicator function* of K is defined by

$$I_K(x) := \begin{cases} 0, & \text{if } x\in K,\\ \infty, & \text{otherwise}\end{cases} \tag{7}$$

and the (approximate) ε-*subdifferential* of f at x is defined, for any $\varepsilon\ge 0$, by

$$\partial_\varepsilon f(x) := \{u\in X|\ f(y)-f(x)-\langle u, x-y\rangle \ge -\varepsilon,\ \text{for all } y\in X\}. \tag{8}$$

When $\varepsilon=0$, the set $\partial_\varepsilon f(x)$ reduces to the classical subdifferential of convex analysis $\partial f(x)$.

Employing constructions (7) and (8), the authors of [1] developed extensions of both the exact proximal method (1) and its inexact counterpart to the vector-valued optimization problem formulated above. Basically, in the exact case the nth subproblem consists of finding weakly efficient minimizers of $F_n: X\to Y$ with

$$F_n(x) := F(x) + \alpha_n\|x-x_n\|^2 e_n \tag{9}$$

restricted to the set $\Omega_n := \{x\in X|\ F(x)\preceq_C F(x_n)\}$, where e_n is an exogenously selected vector belonging to $\operatorname{int} C$ and such that $\|e_n\|=1$. On the other hand, for their inexact version they considered the topological dual space Y^* of Y, the *positive polar cone*

$$C^+ := \{z\in Y^*|\ \langle y, z\rangle \ge 0,\ \text{for all } y\in C\}$$

with $\langle\cdot,\cdot\rangle: Y\times Y^*\to\mathcal{R}$ standing for the usual duality pairing, and the indicator function I_{Ω_n} of the set Ω_n. To describe the latter version, consider an exogenous sequence $\{\hbar_n\}\subset C^+$ with $\|\hbar_n\|=1$ for all $n\ge 0$ and define, at iteration n, a function $f_n: X\to\mathcal{R}\cup\{\infty\}$ by

$$f_n(x) := \langle F(x), \hbar_n\rangle + I_{\Omega_n}(x). \tag{10}$$

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

$$0 \in \partial_{\varepsilon_n} f_n(x) + \alpha_n \langle e_n, \hbar_n \rangle (x - x_n), \tag{11}$$

$$\varepsilon_n \le \sigma \frac{\alpha_n}{2} \langle e_n, \hbar_n \rangle \|x_n - x\|^2, \tag{12}$$

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

It is proved in [1] that any sequence of iterates generated by either the exact or inexact version converge 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 $a_0 = x_0$ such that $F(a_{n+1}) \preceq_C F(a_n)$ there is $a \in X$ such that $F(a) \preceq_C F(a_n)$ for all $n \ge 0$.

Note that a vectorial proximal method is also discussed in Sect. 4.2 of [18]. It is a generalization of algorithms for specific classes of problems in vector optimization: a particular control approximation problem in [19] and certain location problems in [20]. In fact, the authors of [18] dealt with a problem more general than the vector optimization problem formulated above; namely, with some vector *equilibrium* problem (VEP). It can be seen that solutions to the scalarized equilibrium problem for a real bifunction f defined on $M \times M$ (i.e., points $\bar{x} \in M$ such that $f(\bar{x}, x) \ge 0$ for all $x \in M$) are solutions to VEP provided that M is a closed and convex subset of $X = \mathcal{R}^m$. We refrain here from making explicit the iterative formula of the scalarized method proposed in [18] for solving VEP, since in the case of vector optimization of our interest it ends up as the standard scalar proximal point method applied to the scalarized function $\langle F(x), \hbar_n \rangle$ with $\hbar_n \in C^+$ and $\|\hbar_n\| = 1$. The convergence analysis in [18] is restricted to the finite-dimensional case. The fact that the method in [18] is essentially a scalar proximal method is a crucial difference from the algorithms of [1], which essentially exploit some characteristic features of vector optimization.

Motivated by [1], Ceng and Yao [22] have recently introduced and studied new versions of the proximal point method in vector optimization called the *absolute approximate proximal method* and the *relative approximate proximal method*. To describe these methods, we take a bounded sequence of positive numbers $\{\alpha_n\}$. In the absolute case the nth subproblem consists of finding first weakly efficient minimizer $\tilde{x}_n$ for the problem

$$\widetilde{F}_n(x) := F(x) + \alpha_n \|x - x_n - \theta_n\|^2 e_n,$$
$$\text{s.t.} \quad \Omega_n := \{x \in X | F(x) \preceq_C F(x_n)\}$$

and then computing the $(n+1)$th iterate by

$$x_{n+1} = \beta_n x_n + (1 - \beta_n)\tilde{x}_n, \tag{13}$$

where β_n is a relaxation parameter in [0, 1], θ_n is an error term in X satisfying

$$\|\theta_n\| \le \sigma_n \|\tilde{x}_n - x_n\|, \quad \text{with} \sum_{n=0}^{\infty} \sigma_n^2 < \infty, \tag{14}$$

and e_n is an exogenously selected vector belonging to $\operatorname{int} C$ and such that $\|e_n\| = 1$. The relative version of [22] deals with the positive polar cone $C^+ \subset Y^*$ of the topologically dual space Y^* of Y and the indicator function I_{Ω_n} of the set Ω_n defined above. Taking an exogenous sequence $\{\hbar_n\} \subset C^+$ with $\|\hbar_n\| = 1$ for all $n \ge 0$, the function

$$f_n(x) := \langle F(x), \hbar_n \rangle + I_{\Omega_n}(x) \tag{15}$$

is defined at iteration n. Then the next iterate x_{n+1} is selected as a vector $x \in X$ such that there exists $\varepsilon_n \in \mathcal{R}_+$ satisfying the conditions

$$0 \in \partial_{\varepsilon_n} f_n(x) + \alpha_n \langle e_n, \hbar_n \rangle (x - x_n - \theta_n), \tag{16}$$

$$\varepsilon_n \le \sigma \frac{\alpha_n}{2} \langle e_n, \hbar_n \rangle \|x_n + \theta_n - x\|^2, \tag{17}$$

where $\{\theta_n\} \subset X$ is an error sequence given in (14), and where $\sigma \in [0, 1)$ is again a measure of the relative error. It is proved in [22] that any sequence generated by either absolute or relative approximate proximal method converges in the weak topology of X to a weakly efficient minimizer of F under the two assumptions from [1] presented above.

It is worth reminding that the exact version of the proximal method in [1] is indeed a particular case of the aforementioned absolute approximate proximal method corresponding to the choice of $\theta_n = 0$ and $\beta_n = 0$ for all n. Furthermore, the inexact version of the proximal method in [1] is actually a particular case of the relative approximate proximal method corresponding to the choice of $\theta_n = 0$ and $\beta_n = 0$ for all n. Observe also that the absolute version of the proximal algorithm is a particular case of the relative one corresponding to the choice of $\sigma = 0$, or equivalently $\varepsilon_n = 0$ for all n, in the sense that any vector x_{n+1} satisfying relationship (15)–(17) with $\sigma = 0$ is a weakly efficient minimizer of $\widetilde{F}_n$ as defined in (12). Thus a separate analysis of the absolute version might seem superfluous. However, both versions are presented somewhat differently and deserve a special attention from the viewpoint of subsequent implementations; namely, the subproblems of the absolute one are vector-valued optimization problems while in each subproblem of the relative version the focus is on finding zeros of approximate subdifferentials for scalar-valued convex functions.

Next we discuss the other lines of development related to numerical algorithms of solving a special class of *variational inequalities*. Let Ω be a nonempty, closed, and convex subset of a Hilbert space X, and let P_Ω be the *metric projection* from X onto Ω. Given $x \in X$ a sequence $\{x_n\} \subset X$, the symbols $x_n \to x$ and $x_n \rightharpoonup x$ indicate the strong convergence in X and the weak convergence in X, respectively. Recall that a mapping $A\colon \Omega \to X$ is called *monotone* if

$$\langle Ax - Ay, x - y \rangle \ge 0, \quad \text{for all } x, y \in \Omega.$$

Given $k > 0$, such a mapping is called *k-Lipschitz continuous* if

$$\|Ax - Ay\| \le k\|x - y\| \quad \text{for all } x, y \in \Omega.$$

Recall finally that $A: \Omega \to X$ is *α-inverse-strongly monotone* with modulus $\alpha > 0$ (cf. [23]) if

$$\langle Ax - Ay, x - y\rangle \ge \alpha\|Ax - Ay\|^2, \quad \text{for all } x, y \in \Omega.$$

It is easy to see that every α-inverse-strongly monotone mapping under consideration is monotone and k-Lipschitz continuous with some modulus $k = \frac{1}{\alpha}$.

Given Ω and $A: \Omega \to X$ as above, we define the *variational inequality* problem as follows: find a $x \in \Omega$ such that

$$\langle Ax, y - x\rangle \ge 0, \quad \text{for all } y \in \Omega$$

and denote the set of its solutions by $VI(\Omega, A)$.

In 1976, Korpelevich [24] introduced the following *extragradient method* for solving the above variational inequality generated by a closed and convex set $\Omega \subset \mathcal{R}^m$ and a monotone, k-Lipschitz continuous mapping $A: \Omega \to \mathcal{R}^m$:

$$\begin{aligned}
&\text{choose } x_0 \in \Omega \text{ arbitrarily,}\\
&\bar{x}_n = P_\Omega(x_n - \lambda A x_n),\\
&x_{n+1} = P_\Omega(x_n - \lambda A\bar{x}_n), \quad n \ge 0,
\end{aligned}$$

where $\lambda \in (0, 1/k)$. She proved that if $VI(\Omega, A) \ne \emptyset$, then the sequence $\{x_n\}$ generated by her iterative scheme converges to an element of $VI(\Omega, A)$.

Developing Korpelevich's extragradient ideas and combining them with the outer approximation method by Burachik, Lopez and Svaiter [25] for solving variational inequalities, Nadezhkina and Takahashi [23] introduced an iterative process for finding a common element of the *fixed-point* set for nonexpansive self-mappings on Ω and the set of solutions to variational inequalities generated by monotone, k-Lipschitz continuous operators. Subsequently Ceng, Cubiotti and Yao [26] developed another iterative process for solving the aforementioned fixed-point problem for variational inequalities by some combination of extragradient and approximate proximal methods.

The main objective of this paper is to introduce and develop both *absolute* and *relative* versions the *hybrid approximate proximal method* (HAPM) for solving the general vector-valued optimization problem formulated above. Let $\{\alpha_n\}$ be a bounded sequence of positive numbers, and let $\{\gamma_n\} \subset (0, 1)$. In the *absolute version* of HAPM the nth subproblem consists of finding first an approximate solution z_n of the variational inequality problem for a monotone, k-Lipschitz continuous mapping A via

$$\begin{aligned}
y_n &= P_\Omega(x_n - \lambda_n A x_n),\\
z_n &= \gamma_n x_n + (1 - \gamma_n) P_\Omega(x_n - \lambda_n A y_n)
\end{aligned}$$

with some $\{\lambda_n\} \subset (0, 1/k)$. Then we find a weakly efficient minimizer $\tilde{x}_n$ of the mapping $\widetilde{F}_n : \Omega \to Y$ given by

$$\widetilde{F}_n(x) := F(x) + \alpha_n\|x - x_n - \theta_n\|^2 e_n$$

restricted to $\Omega_n := \{x \in \Omega | F(x) \preceq_C F(x_n)\}$ and finally compute the $(n+1)$th iterate by

$$x_{n+1} = \beta_n x_n + (1 - \beta_n)\tilde{x}_n,$$

where β_n is a relaxation parameter in $[0, 1]$, θ_n is an error term in X satisfying

$$\|\theta_n\| \le \sigma_n \|\tilde{x}_n - z_n\|, \quad \text{with} \sum_{n=0}^{\infty} \sigma_n^2 < \infty,$$

and where e_n is an exogenously selected vector belonging to $\operatorname{int} C$ and such that $\|e_n\| = 1$.

In the *relative version* of HAPM we first construct z_n as in the absolute version described above. Picking then an exogenous sequence $\{\hbar_n\} \subset C^+$ with $\|\hbar_n\| = 1$, define the scalarized function $f_n : \Omega \to \mathcal{R} \cup \{\infty\}$ by (15) and take as x_{n+1} any vector $x \in \Omega$ such that there exists $\varepsilon_n \in \mathcal{R}_+$ satisfying the conditions

$$0 \in \partial_{\varepsilon_n} f_n(x) + \alpha_n \langle e_n, \hbar_n \rangle (x - z_n - \theta_n),$$
$$\varepsilon_n \le \sigma \frac{\alpha_n}{2} \langle e_n, \hbar_n \rangle \|z_n + \theta_n - x\|^2,$$

where $\{\theta_n\} \subset \Omega$ is an error sequence yielding some requirements similar to (14), and where $\sigma \in [0, 1)$ is a measure of the relative error.

It is shown in what follows that any sequence generated by either the absolute or relative algorithm of HAPM *converges* in the *weak* topology of X to a *weakly efficient minimizer* of F on Ω and simultaneously to a solution of an *auxiliary variational inequality* under rather mild assumptions imposed on the initial data.

Considering in this paper weakly efficient solutions to the vector optimization problem, we have to require the *nonempty interior* of the ordering cone in Y. This assumption conventional in vector optimization theory is satisfied, in particular, for positive cones in finite dimensions as well as in Banach spaces of continuous functions and of the L^∞ type. However, it is generally restrictive and never holds, e.g., for positive cones in L^p, $1 \le p < \infty$. We draw the reader's attention that the nonempty interior of ordering cones is not required in the approach to vector and set-valued optimization problems developed by Mordukhovich [27] that is mainly based on the *extremal principle*. Furthermore, the recent paper by Bao and Mordukhovich [28] studies certain notions of *relative Pareto minimizers*, which are close in spirit to weak minimizers while do not require the nonempty interior of the ordering cone. In our subsequent publications, we are going to extend and develop the numerical algorithms of the present paper to broad classes of vector optimization problems with possibly empty interiors of ordering cones.

The rest of paper is organized as follows. In Sect. 2 we formulate the underlying vector optimization problem and present some required preliminary material. The absolute version of our HARP algorithm is developed in Sect. 3. The subsequent Sect. 4 is devoted to an appropriate extension of the absolute version of HARP to the case of generated Bregman functions. Finally, Sect. 5 develops the relative version of HARP. The notation used in this paper is basically standard; see, e.g., [1, 23, 28].

2 Problem Formulation and Preliminaries

Let X be a Hilbert space, Y be a Banach space, and $\langle\cdot,\cdot\rangle$ signify the scalar product in X as well as the standard canonical pairing between Y and its topological dual Y^*. For simplicity, any norm is denoted by $\|\cdot\|$. We usually denote by F an extended-valued mapping from X to $Y\cup\{\infty_C\}$. The extended space $\overline{Y}:=Y\cup\{-\infty_C,\infty_C\}$ is introduced in [29], where a neighborhood of ∞_C is defined as a set $\mathcal{N}\subset\overline{Y}$ containing $a+C\cup\{\infty_C\}$ for some $a\in Y$; its opposite $-\mathcal{N}$ is a neighborhood of $-\infty_C$. The binary relations $\preceq_C$ and $\prec_C$, defined in the previous section, are extended to $\overline{Y}$ by

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

Observe that the embedding $Y\subset\overline{Y}$ is continuous and dense.

Mappings F are assumed to be *proper*, i.e., not identically equal to ∞_C. The *effective domain* of F is denoted by $\operatorname{dom}F:=\{x\in X|F(x)\neq\infty_C\}$. By putting $\langle\pm\infty_C,z\rangle:=\pm\infty$ (see [29, 30] for more details), we extend by continuity every $z\in C^+\setminus\{0\}$ to $\overline{Y}$. Given a set $U\subset\overline{Y}$, denote its topological closure in $\overline{Y}$ by $\overline{U}$. Let us further associate with a given set $U\subset\overline{Y}$ the following three collections of minimizers:

- the *infimal set* $C\text{-INF}(U):=\{y\in\overline{U}|\ \nexists z\in U\setminus\{y\}: z\preceq_C y\}$;
- the *weakly infimal set* $C\text{-INF}_w(U):=\{y\in\overline{U}|\ \nexists z\in U: z\prec_C y\}$;
- the *properly infimal set*

$$C\text{-INF}_p(U):=\{y\in\overline{U}|\exists K\subset Y \text{ pointed closed convex cone such that } C\setminus\{0\}\subset\operatorname{int}K,\ y\in K\text{-INF}(U)\}.$$

For the vector optimization problem

$$C\text{-Min } G(x),\quad \text{s.t.}\quad x\in S,$$

where $G:S\to Y\cup\{\infty_C\}$ and $S\subset X$, a point $\bar{x}\in X$ is called:

- *efficient* (or Pareto) if $\bar{x}\in S$ and $G(\bar{x})\in C\text{-INF}(G(S))$,
- *weakly efficient* if $\bar{x}\in S$ and $G(\bar{x})\in C\text{-INF}_w(G(S))$,
- *properly efficient* if $\bar{x}\in S$ and $G(\bar{x})\in C\text{-INF}_p(G(S))$.

Thus the sets of efficient (resp., weakly efficient and properly efficient) solutions, which are denoted by $C\text{-ArgMin}\{G(x)|x\in S\}$ (resp., $C\text{-ArgMin}_w\{G(x)|x\in S\}$ and $C\text{-ArgMin}_p\{G(x)|x\in S\}$), we have the following relations:

$$\begin{aligned} C\text{-ArgMin}\{G(x)|(x)\in S\}&=S\cap G^{-1}(C\text{-INF}(G(S))),\\ C\text{-ArgMin}_w\{G(x)|x\in S\}&=S\cap G^{-1}(C\text{-INF}_w(G(S))),\\ C\text{-ArgMin}_p\{G(x)|x\in S\}&=S\cap G^{-1}(C\text{-INF}_p(G(S))).\end{aligned}$$

It is easy to check that

$$\begin{aligned} C\text{-ArgMin}_p\{G(x)|x\in S\}&\subset C\text{-ArgMin}\{G(x)|x\in S\}\\ &\subset C\text{-ArgMin}_w\{G(x)|x\in S\}.\end{aligned}$$

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

In this paper we pay the main attention to the following *vector optimization problem*

$$\text{(VOP)} \quad C\text{-Min}\{F(x) | x \in \Omega\},$$

where Ω is a nonempty, closed, and convex subset of X. The set of *weakly efficient solutions* of this VOP is denoted by $VO(\Omega, F)$. Throughout this paper we assume that the defined *VOP is C-convex*, i.e., the cost mapping F is C-convex on the convex constraint set Ω.

Recall that a map $G : X \to Y \cup \{\infty_C\}$ is *positively lower semicontinuous* if for each $z \in C^+$ the scalarized extended-real-valued function $x \mapsto \langle G(x), z\rangle$ is lower semicontinuous.

In the sequel we need following scalarization result (cf. [22, 31]), where

$$C_s^+ := \{z \in Y^* | \langle y, z\rangle > 0, \text{ for all } y \in C \setminus \{0\}\}.$$

Proposition 2.1 (Argminimum Sets under Scalarization) *Let $S \subset X$ be a convex set, and let $G : S \to Y \cup \{\infty_C\}$ be a C-convex proper map. Then we have the relationships*

$$C\text{-ArgMin}_w\{G(x) | x \in S\} = \bigcup_{z \in C^+ \setminus \{0\}} \operatorname{argmin}\{\langle G(x), z\rangle | x \in S\},$$

$$C\text{-ArgMin}_p\{G(x) | x \in S\} = \bigcup_{z \in C_s^+} \operatorname{argmin}\{\langle G(x), z\rangle | x \in S\}.$$

It is worth mentioning that, as observed in [4], the set $\operatorname{argmin}\{\langle G(x), z\rangle | x \in S\}$ in Proposition 2.1 may be empty for some $z \in C^+ \setminus \{0\}$.

We also need the following convergence property established in [32].

Proposition 2.2 (Convergence Property) *Let X be a Hilbert space, let $\{\alpha_n\}$ be a sequence of real numbers such that $0 < a \le \alpha_n \le b < 1$ for every $n = 0, 1, 2, \ldots$, and let $\{v_n\}$ and $\{w_n\}$ be sequences in X such that $\limsup_{n\to\infty} \|v_n\| \le c$, $\limsup_{n\to\infty} \|w_n\| \le c$, and $\lim_{n\to\infty} \|\alpha_n v_n + (1 - \alpha_n)w_n\| = c$ for some $c \ge 0$. Then $\lim_{n\to\infty} \|v_n - w_n\| = 0$.*

Recall that if Ω is a nonempty, closed, and convex subset of a Hilbert space X, then for every point $x \in X$ there exists a *unique nearest* point in Ω, denoted by $P_\Omega x$, such that $\|x - P_\Omega x\| \le \|x - y\|$ for all $y \in \Omega$. It is known that P_Ω is a *nonexpansive* mapping from H onto Ω. It is also known that $P_\Omega x \in \Omega$ and

$$\langle x - P_\Omega x, P_\Omega x - y\rangle \ge 0, \quad \forall x \in H, \ y \in \Omega; \tag{18}$$

see [12, 32] for more details. It is easy to see that (18) is equivalent to

$$\|x - y\|^2 \ge \|x - P_\Omega x\|^2 + \|y - P_\Omega x\|^2, \quad \forall x \in H, \ y \in \Omega. \tag{19}$$

Consider further a monotone mapping $A : \Omega \to X$, which generates the variational inequality problem formulated above. In the context of the latter problem the characterization of projection (18) implies that

$$u \in VI(\Omega, A) \quad \Longleftrightarrow \quad u = P_\Omega(u - \lambda Au), \quad \forall \lambda > 0.$$

Recall that a mapping $T : \Omega \to \Omega$ is *pseudocontractive* if for all $x, y \in \Omega$ we have

$$\|Tx - Ty\|^2 \le \|x - y\|^2 + \|(I - T)x - (I - T)y\|^2.$$

Observe that, if $T : \Omega \to \Omega$ is pseudocontractive and k-Lipschitz continuous, then the mapping $A = I - T$ is monotone and $(k + 1)$-Lipschitz continuous; moreover, $\mathrm{Fix}(T) = VI(\Omega, A)$, where $\mathrm{Fix}(T)$ is the *fixed-point set* of T; see, e.g., [23, proof of Theorem 4.5].

Recall also that a set-valued mapping $T : X \to 2^X$ *monotone* if

$$\langle x - y, u - v\rangle \ge 0, \quad \text{for all } x, y \in X \text{ and } u \in T(x),\ v \in T(y).$$

A monotone mapping $T : X \to 2^X$ is *maximal* if its graph $\mathrm{gph}\, T$ is not properly contained in the graph of any other monotone mapping. It is well known that a monotone mapping T is maximal *if and only if* we have the implication

$$\big[(x, u) \in X \times X,\ \langle x - y, u - v\rangle \ge 0,\ \forall (y, v) \in \mathrm{gph}\, T\big] \quad \Longrightarrow \quad u \in Tx.$$

Finally in this section, define a set-valued mapping $T : X \to 2^X$ on a Hilbert space X by

$$Tv := \begin{cases} Av + N_\Omega v, & \text{if } v \in \Omega, \\ \emptyset, & \text{if } v \notin \Omega, \end{cases}$$

where A is a monotone and k-Lipschitz continuous mapping of Ω into X, where $\Omega \subset X$ is a closed and convex set, and where $N_\Omega v$ is the *normal cone* to Ω at $v \in \Omega$ given by

$$N_\Omega v := \{w \in X \mid \langle w, u - v\rangle \le 0,\ \text{for all } u \in \Omega\}.$$

It is well known from Rockafellar [33] that the set-valued mapping T defined above is *maximal monotone* and that $0 \in Tv$ *if and only if* $v \in VI(\Omega, A)$.

3 Absolute Hybrid Approximate Proximal Algorithm

In this section we introduce and develop the aforementioned *absolute version of HAPM* called for simplicity Algorithm 1. Our main goal is to find a weakly efficient solution to the underlying vector optimization problem, i.e., an element of $VO(\Omega, F)$, which we find by solving an auxiliary variational inequality. Algorithm 1 requires some *exogenous sequences*: an error sequence $\{\theta_n\} \subset X$, two relaxation sequences $\{\beta_n\}$ and $\{\gamma_n\}$ in $[0, 1]$, two bounded sequences of positive numbers $\{\alpha_n\}$ and $\{\sigma_n\}$, and a sequence $\{e_n\} \subset \mathrm{int}\, C$ such that $\|e_n\| = 1$ for all n. We always assume that $\Omega \cap \mathrm{dom}\, F \ne \emptyset$. Algorithm 1 generates an iterative sequence $\{x_n\} \subset \Omega$ in the following way:

Algorithm 1

Initialization: Choose $x_0 \in \Omega \cap \operatorname{dom} F$.
Stopping Rule: Given x_n, if $x_n \in C\text{-ArgMin}_w\{F(x)|x \in \Omega\}$ $(= VO(\Omega, F))$, then we let $x_{n+p} := x_n$ for all $p \geq 1$.
Iterative Step: Given x_n, whenever $x_n \notin C\text{-ArgMin}_w\{F(x)|x \in \Omega\}$, we first compute

$$\begin{aligned} y_n &= P_\Omega(x_n - \lambda_n A x_n), \\ z_n &= \gamma_n x_n + (1-\gamma_n) P_\Omega(x_n - \lambda_n A y_n) \end{aligned} \tag{20}$$

for every $n = 0, 1, 2, \ldots$, where $\{\lambda_n\} \subset (0, 1)$ and $\{\gamma_n\} \subset [0, 1]$, and then take as $\tilde{x}_n$ any vector $u \in \Omega$ satisfying the condition

$$u \in C\text{-ArgMin}_w\left\{F(x) + \frac{\alpha_n}{2}\|x - z_n - \theta_n\|^2 e_n | x \in \Omega_n\right\} \tag{21}$$

with $\Omega_n := \{x \in \Omega | F(x) \preceq_C F(x_n)\}$. Finally, the next iterate $x_{n+1} \in \Omega$ is computed by

$$x_{n+1} = \beta_n x_n + (1-\beta_n)\tilde{x}_n. \tag{22}$$

To justify the *well-posedness* and *convergence* of Algorithm 1, we impose the following assumptions on the initial data (F, Ω) of VOP and the starting point x_0 of the algorithm:

(A) The set $(F(x_0) - C) \cap F(\Omega)$ is C-*quasicomplete* for Ω, which means that for each sequences $\{a_n\} \subset \Omega$ with $a_0 = x_0$ and $F(a_{n+1}) \preceq_C F(a_n)$ as $n \geq 0$, we have $F(u) \preceq_C F(a_n)$ for all $u \in VO(\Omega, F) \cap VI(\Omega, A)$ and $n \geq 0$.
(B) The map F is C^+-*uniformly semicontinuous* on Ω, which means that for every sequence $\{x_n\} \subset \Omega$ converging weakly to some $\hat{x} \in \Omega$ and for every sequence $\{\hbar_n\} \subset C^+$ converging weakly to some $\hbar \in C^+$, we have the implication

$$\|x_n - y_n\| \to 0 \quad \Longrightarrow \quad |\langle F(x_n) - F(y_n), \hbar_n\rangle - \langle F(\hat{x}) - F(y_n), \hbar\rangle| \to 0$$

whenever a sequence $\{y_n\} \subset \Omega$ is selected.

Now we are ready to establish the well-posedness and convergence of Algorithm 1 under condition (14) and the imposed assumptions (A) and (B) on the initial data.

Theorem 3.1 (Well-Posedness and Convergence of the Absolute Version of HAPM) *Let $F : X \to Y \cup \{\infty_C\}$ be a proper, C-convex, and positively lower semicontinuous mapping with $\Omega \cap \operatorname{dom} F \neq \emptyset$, and let $A : \Omega \to X$ be a monotone and k-Lipschitz continuous mapping such that $VO(\Omega, F) \cap VI(\Omega, A) \neq \emptyset$. Suppose also that condition* (14), *that assumptions* (A) *and* (B) *are satisfied, and that the exogenous sequence in Algorithm 1 are selected as follows*:

(i) $\{\beta_n\} \subset [\epsilon, 1-\delta]$ *for some* $\epsilon, \delta \in (0, 1)$;
(ii) $\{\lambda_n\} \subset [a, b]$ *for some* $a, b \in (0, 1/k)$;
(iii) $\{\gamma_n\} \subset [0, c]$ *for some* $c \in [0, 1)$.

Then Algorithm 1 is well defined, and the sequence of iterates $\{x_n\}$ weakly converges in X to some element of $VO(\Omega, F) \cap VI(\Omega, A)$ provided that $x_n \notin C\text{-}\mathrm{ArgMin}_w\{F(x)|x \in \Omega\}$ whenever $n \geq 0$.

Prior to proving Theorem 3.1, we present two examples that illustrate the fulfillment of all the assumptions of this theorem including the "joint solvability condition"

$$VO(\Omega, F) \cap VI(\Omega, A) \neq \emptyset.$$

For simplicity we consider the cases of vector cost mappings F defined on $\mathcal{R}$ and $\mathcal{R}^2$, respectively, with values in $\mathcal{R}^2$.

Example 3.2 (Algorithm 1 for One-Dimensional Vector Problems) Let $X = \mathcal{R}$ with inner product $\langle\cdot,\cdot\rangle$, let $Y = \mathcal{R}^2$ with the Euclidean norm, and let

$$C := \mathcal{R}^2_+ = \{(a, b) \big| \, a, b \in [0, \infty)\}.$$

Utilizing C, we have a partial order $\preceq_C$ in Y given by $y \preceq_C y'$ if and only if $y' - y \in C$ with the associate relation $\prec_C$ given by $y \prec_C y'$ if and only if $y' - y \in \operatorname{int} C$. Moreover, $F : X \to Y \cup \{\infty_C\}$, $\Omega \subset X$, and $A : \Omega \to X$ are defined by

$$\begin{aligned} &F(a) := (a, a), \quad \text{for all } a \in X,\\ &\Omega := [0, \infty),\\ &A(a) := a - \sin a, \quad \text{for all } a \in \Omega. \end{aligned}$$

Then we have the following required properties:

(i) F is a proper, C-convex, and positively lower semicontinuous map with $\Omega \cap \operatorname{dom} F \neq \emptyset$;
(ii) $A : \Omega \to X$ is a monotone and 2-Lipschitz continuous map such that the intersection $VO(\Omega, F) \cap VI(\Omega, A)$ is nonempty, since it contains zero.

Example 3.3 (Algorithm 1 for Multidimensional Vector Problems) Let$X = Y = \mathcal{R}^2$ with the inner product $\langle\cdot,\cdot\rangle$ and the norm $\|\cdot\|$ defined by

$$\langle x, y\rangle := ac + bd \quad \text{and} \quad \|x\| := \sqrt{a^2 + b^2}$$

for all $x, y \in \mathcal{R}^2$ with $x = (a, b)$ and $y = (c, d)$. Utilizing $C := \mathcal{R}^2_+ = \{(a, b)|\, a, b \in [0, \infty)\}$, we have the same partial order $\preceq_C$ in Y as in Example 3.2. Furthermore, define $F : X \to Y \cup \{\infty_C\}$, $\Omega \subset X$, and $A : \Omega \to X$ by

$$\begin{aligned} &F(x) := \left(\frac{2}{3}a + \frac{1}{3}b, \frac{1}{3}a + \frac{2}{3}b\right), \quad \text{for all } x = (a, b) \in X,\\ &\Omega := \{(a, a)|a \in [0, \infty)\},\\ &Ax := \left(a - \frac{1}{2}\sin a, a - \frac{1}{2}\sin a\right), \quad \text{for all } x = (a, a) \in \Omega. \end{aligned}$$

Then the following required properties hold:

(i) F is a proper, C-convex, and positively lower semicontinuous map with $\Omega \cap \operatorname{dom} F \neq \emptyset$;
(ii) $A : \Omega \to X$ is a monotone and $\frac{3}{2}$-Lipschitz continuous map for which we have the inclusion $0 \in VO(\Omega, F) \cap VI(\Omega, A)$.

Let us now proceed with a detailed proof of Theorem 3.1.

Proof of Theorem 3.1 We split the proof into several steps.

Step 1: *For every $u \in VO(\Omega, F) \cap VI(\Omega, A)$, we get*

$$\|z_n - u\|^2 \le \|x_n - u\|^2 + (1-\gamma_n)(\lambda_n^2 - 1)\|x_n - y_n\|^2, \quad \forall n \ge 0.$$

Indeed, put $t_n := P_\Omega(x_n - \lambda_n A y_n)$ for every $n = 0, 1, 2, \dots$. It follows from (19), monotonicity of A, and $u \in VI(\Omega, A)$ that

$$\begin{aligned}
\|t_n - u\|^2 &\le \|x_n - \lambda_n A y_n - u\|^2 - \|x_n - \lambda_n A y_n - t_n\|^2 \\
&= \|x_n - u\|^2 - \|x_n - t_n\|^2 + 2\lambda_n \langle A y_n, u - t_n\rangle \\
&= \|x_n - u\|^2 - \|x_n - t_n\|^2 \\
&\quad + 2\lambda_n (\langle A y_n - Au, u - y_n\rangle + \langle Au, u - y_n\rangle + \langle A y_n, y_n - t_n\rangle) \\
&\le \|x_n - u\|^2 - \|x_n - t_n\|^2 + 2\lambda_n \langle A y_n, y_n - t_n\rangle \\
&= \|x_n - u\|^2 - \|x_n - y_n\|^2 - 2\langle x_n - y_n, y_n - t_n\rangle \\
&\quad - \|y_n - t_n\|^2 + 2\lambda_n \langle A y_n, y_n - t_n\rangle \\
&= \|x_n - u\|^2 - \|x_n - y_n\|^2 - \|y_n - t_n\|^2 + 2\langle x_n - \lambda_n A y_n - y_n, t_n - y_n\rangle.
\end{aligned}$$

Further, since $y_n = P_\Omega(x_n - \lambda_n A x_n)$ and A is k-Lipschitz continuous, we have

$$\begin{aligned}
&\langle x_n - \lambda_n A y_n - y_n, t_n - y_n\rangle \\
&\quad = \langle x_n - \lambda_n A x_n - y_n, t_n - y_n\rangle + \langle \lambda_n A x_n - \lambda_n A y_n, t_n - y_n\rangle \\
&\quad \le \langle \lambda_n A x_n - \lambda_n A y_n, t_n - y_n\rangle \\
&\quad \le \lambda_n k \|x_n - y_n\| \|t_n - y_n\|.
\end{aligned}$$

The latter implies the estimates

$$\begin{aligned}
\|t_n - u\|^2 &\le \|x_n - u\|^2 - \|x_n - y_n\|^2 - \|y_n - t_n\|^2 + 2\lambda_n k \|x_n - y_n\| \|t_n - y_n\| \\
&\le \|x_n - u\|^2 - \|x_n - y_n\|^2 - \|y_n - t_n\|^2 + \lambda_n^2 k^2 \|x_n - y_n\|^2 + \|y_n - t_n\|^2 \\
&\le \|x_n - u\|^2 + (\lambda_n^2 k^2 - 1)\|x_n - y_n\|^2 \le \|x_n - u\|^2. \qquad (23)
\end{aligned}$$

Therefore, it follows from (23) and $z_n = \gamma_n x_n + (1-\gamma_n) t_n$ that

$$\begin{aligned}
\|z_n - u\|^2 &= \|\gamma_n x_n + (1-\gamma_n) t_n - u\|^2 \\
&= \|\gamma_n (x_n - u) + (1-\gamma_n)(t_n - u)\|^2
\end{aligned}$$

$$\begin{aligned}
&\le \gamma_n\|x_n-u\|^2+(1-\gamma_n)\|t_n-u\|^2\\
&\le \gamma_n\|x_n-u\|^2+(1-\gamma_n)(\|x_n-u\|^2+(\lambda_n^2k^2-1)\|x_n-y_n\|^2)\\
&= \|x_n-u\|^2+(1-\gamma_n)(\lambda_n^2k^2-1)\|x_n-y_n\|^2\\
&\le \|x_n-u\|^2, \quad \text{for every } n=0,1,2,\ldots. \qquad (24)
\end{aligned}$$

Step 2: *Existence of iterates.* Choosing x_0 as in the initialization step and assuming that the algorithm reaches iteration n, let us show that the next iterate x_{n+1} indeed exists in Algorithm 1. By the stopping rule this is certainly the case if $x_n \in VO(\Omega, F)$. Otherwise, take any $z \in C^+ \setminus \{0\}$ and by $e_n \in \operatorname{int} C$ get from the definition of C^+ that $\langle e_n, z\rangle > 0$. Define $\varphi_n : X \to \mathcal{R} \cup \{\infty\}$ by

$$\varphi_n(x) := \langle F(x), z\rangle + I_{\Omega_n}(x) + \frac{\alpha_n}{2}\langle e_n, z\rangle\|x - z_n - \theta_n\|^2. \qquad (25)$$

Observe that the C-convexity of F implies the convexity of $\langle F(\cdot), z\rangle$ and of Ω_n. Further, the lower semicontinuity of F implies the closedness of Ω_n. Thus $\langle F(\cdot), z\rangle + I_{\Omega_n}$ is convex and lower semicontinuous. Since $\langle e_n, z\rangle > 0$, we have that φ_n is strongly convex. Hence the existence of minimizers for φ_n follows from the standard arguments ensuring the existence of iterates in the scalar-valued proximal method; see, e.g., [20]: the subdifferential of φ_n is maximal monotone and strongly monotone, and so it is onto by *Minty's theorem.* Therefore this subdifferential has some zero, which is a minimizer for φ_n. By Theorem 2.1 such a minimizer satisfies (21) and can be taken as $\tilde{x}_n$. By (22) we thus compute the next iterate x_{n+1}.

Step 3: *Fejér convergence to the set of lower bounds of the initial section*. Observe first that if the stopping rule applies at some iteration, then the sequence remains constant thereafter. Thus it is strongly convergent to the stopping iterate, which is an element of $VO(\Omega, F)$. So we assume from now on that the stopping rule never applies. Therefore, since $\tilde{x}_n$ solves the vector optimization problem in (21), by Proposition 2.1 there exists $\hbar_n \in C^+ \setminus \{0\}$ such that $\tilde{x}_n$ solves also the problem

$$\min\ \eta_n(x), \qquad (26)$$

$$\text{s.t.}\quad x \in \Omega_n, \qquad (27)$$

where $\eta_n : X \to \mathcal{R} \cup \{\infty\}$ is defined by

$$\eta_n(x) := \langle F(x), \hbar_n\rangle + \frac{\alpha_n}{2}\langle e_n, \hbar_n\rangle\|x - z_n - \theta_n\|^2.$$

Since furthermore the solution to (26) and (27) is not altered through multiplication of $\hbar_n$ by positive scalars, we can assume without loss of generality that $\|\hbar_n\| = 1$ for all $n \ge 0$. Note that by definition we have $\Omega_n \subset \operatorname{dom}(\eta_n) = \operatorname{dom} F$, and thus $\emptyset \ne \operatorname{dom}(I_{\Omega_n}) \subset \operatorname{dom}(\eta_n)$. It follows from [34, Theorem 3.23] that $\tilde{x}_n$ satisfies the first-order optimality conditions for problem (26) and (27); i.e., there exists $u_n \in X$ such that

$$u_n \in \partial\eta_n(\tilde{x}_n), \qquad (28)$$

$$0 \le \langle u_n, x - \tilde{x}_n\rangle, \quad \text{for all } x \in \Omega_n. \qquad (29)$$

Define next $\psi_n : X \to \mathcal{R} \cup \{\infty\}$ by

$$\psi_n(x) := \langle F(x), \hbar_n \rangle. \tag{30}$$

In view of (25) and (28) we have

$$u_n = v_n + \alpha_n \langle e_n, \hbar_n \rangle (\tilde{x}_n - z_n - \theta_n), \tag{31}$$

for some

$$v_n \in \partial \psi_n(\tilde{x}_n). \tag{32}$$

Fixing an arbitrary element $u \in VO(\Omega, F) \cap VI(\Omega, A)$, we get by condition (A) that $u \in \Omega_n$ for all $n \geq 0$. Combining (29) with $x = u$ and (31) gives us

$$\begin{aligned} 0 &\leq \langle v_n, u - \tilde{x}_n \rangle + \alpha_n \langle e_n, \hbar_n \rangle \langle \tilde{x}_n - z_n - \theta_n, u - \tilde{x}_n \rangle \\ &\leq \langle F(u) - F(\tilde{x}_n), \hbar_n \rangle + \alpha_n \langle e_n, \hbar_n \rangle \langle \tilde{x}_n - z_n - \theta_n, u - \tilde{x}_n \rangle \\ &\leq \alpha_n \langle e_n, \hbar_n \rangle \langle \tilde{x}_n - z_n - \theta_n, u - \tilde{x}_n \rangle, \end{aligned} \tag{33}$$

by using (30) and (32) in the second inequality and the fact that $\hbar_n \in C^+ \setminus \{0\}$ in the third; it is clear that $F(u) - F(\tilde{x}_n) \preceq_C 0$ and therefore $\langle F(u) - F(\tilde{x}_n), \hbar_n \rangle \leq 0$.

Now define $\nu_n := \alpha_n \langle e_n, \hbar_n \rangle$ and note that $\nu_n > 0$ due to $\alpha_n > 0$, $e_n \in \operatorname{int} C$, and $\hbar_n \in C^+ \setminus \{0\}$. From (33), we obtain

$$\langle z_n - \tilde{x}_n + \theta_n, \tilde{x}_n - u \rangle \geq 0. \tag{34}$$

Moreover, the identity

$$\|x + y\|^2 = \|x\|^2 - \|y\|^2 + 2\langle y, x + y \rangle, \quad \text{for all } x, y \in X$$

allows us to derive from (34) the relationships

$$\begin{aligned} \|\tilde{x}_n - u\|^2 &= \|z_n - u\|^2 - \|\tilde{x}_n - z_n\|^2 + 2\langle \tilde{x}_n - z_n, \tilde{x}_n - u \rangle \\ &= \|z_n - u\|^2 - \|\tilde{x}_n - z_n\|^2 + 2\langle \theta_n, \tilde{x}_n - u \rangle - 2\langle z_n - \tilde{x}_n + \theta_n, \tilde{x}_n - u \rangle \\ &\leq \|z_n - u\|^2 - \|\tilde{x}_n - z_n\|^2 + 2\langle \theta_n, \tilde{x}_n - u \rangle. \end{aligned} \tag{35}$$

Taking further $\sigma_n > 0$, observe that

$$2\langle \theta_n, \tilde{x}_n - u \rangle \leq \frac{1}{2\sigma_n^2} \|\theta_n\|^2 + 2\sigma_n^2 \|\tilde{x}_n - u\|^2. \tag{36}$$

Since $\sigma_n \to 0$ as $n \to \infty$, there exists an integer $N_0 \geq 0$ such that $1 - 2\sigma_n^2 > 0$ for all $n \geq N_0$. Substituting (36) in (34), we get therefore

$$\begin{aligned} \|\tilde{x}_n - u\|^2 &\leq \left(1 + \frac{2\sigma_n^2}{1 - 2\sigma_n^2}\right) \|z_n - u\|^2 - \frac{1}{2(1 - 2\sigma_n^2)} \|\tilde{x}_n - z_n\|^2 \\ &\leq \left(1 + \frac{2\sigma_n^2}{1 - 2\sigma_n^2}\right) \|z_n - u\|^2 - \frac{1}{2} \|\tilde{x}_n - z_n\|^2. \end{aligned} \tag{37}$$

Note that for all $x, y \in X$ and $0 \le \lambda \le 1$ the following identity holds:

$$\|\lambda x + (1-\lambda)y\|^2 = \lambda\|x\|^2 + (1-\lambda)\|y\|^2 - \lambda(1-\lambda)\|x-y\|^2.$$

Thus it follows from (22), (24), and (37) that

$$\begin{aligned}
&\|x_{n+1}-u\|^2\\
&\quad= \|\beta_n(x_n-u) + (1-\beta_n)(\tilde{x}_n-u)\|^2\\
&\quad\le \beta_n\|x_n-u\|^2 + (1-\beta_n)\|\tilde{x}_n-u\|^2\\
&\quad\le \beta_n\|x_n-u\|^2 + (1-\beta_n)\left\{\left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)\|z_n-u\|^2 - \frac{1}{2}\|\tilde{x}_n-z_n\|^2\right\}\\
&\quad\le \beta_n\|x_n-u\|^2 + (1-\beta_n)\left\{\left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)\|x_n-u\|^2 - \frac{1}{2}\|\tilde{x}_n-z_n\|^2\right\}\\
&\quad\le \left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)\|x_n-u\|^2 - \frac{1}{2}(1-\beta_n)\|\tilde{x}_n-z_n\|^2.
\end{aligned}$$

Since $0 \le \beta_n \le 1-\delta$ for some $\delta \in (0,1)$, it follows that $\frac{1}{2}(1-\beta_n) \ge \frac{1}{2}\delta$. Hence, we get

$$\|x_{n+1}-u\|^2 \le \left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)\|x_n-u\|^2 - \frac{\delta}{2}\|\tilde{x}_n-z_n\|^2, \quad \text{for all } n \ge N_0. \tag{38}$$

Finally, from (22) we derive the relationships

$$\begin{aligned}
&x_{n+1}-x_n = (1-\beta_n)(\tilde{x}_n-x_n),\\
&\|\tilde{x}_n-x_n\| = \frac{1}{1-\beta_n}\|x_{n+1}-x_n\| \ge \frac{1}{1-\epsilon}\|x_{n+1}-x_n\|,
\end{aligned} \tag{39}$$

which justify the claimed Fejér-type convergence.

Step 4: *Boundedness of the sequence and proximity of consecutive iterates*. Next we claim that for every $u \in VO(\Omega, F) \cap VI(\Omega, A)$ the sequence $\{\|x_n-u\|^2\}$ is convergent. To proceed, observe from (38) that

$$\|x_{n+1}-u\|^2 \le \left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)\|x_n-u\|^2, \quad \text{for all } n \ge N_0. \tag{40}$$

Since $\sum_{n=0}^{\infty}\sigma_n^2 < \infty$, it follows that

$$K_0 := \sum_{n=N_0}^{\infty}\frac{2\sigma_n^2}{1-2\sigma_n^2} < \infty \quad \text{and} \quad K_1 := \prod_{n=N_0}^{\infty}\left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right) < \infty.$$

Observe further that for all $n \ge N_0$ we have

$$\|x_{n+1}-u\|^2 \le \left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)\|x_n-u\|^2$$

$$
\begin{aligned}
&\le \left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)\left(1+\frac{2\sigma_{n-1}^2}{1-2\sigma_{n-1}^2}\right)\|x_{n-1}-u\|^2 \\
&\vdots \\
&\le \prod_{j=N_0}^{n}\left(1+\frac{2\sigma_j^2}{1-2\sigma_j^2}\right)\|x_{N_0}-u\|^2 \\
&\le \prod_{j=N_0}^{\infty}\left(1+\frac{2\sigma_j^2}{1-2\sigma_j^2}\right)\|x_{N_0}-u\|^2 \\
&= K_1\|x_{N_0}-u\|^2.
\end{aligned}
$$

This shows that $\{x_n\}$ is bounded. Thus it follows from (23) and (24) that both $\{t_n\}$ and $\{z_n\}$ are bounded. Letting $M := \sup_{n\ge 0}\|x_n-u\|$, we get from (40) that

$$
\|x_{n+1}-u\|^2 \le \|x_n-u\|^2 + \frac{2\sigma_n^2}{1-2\sigma_n^2}M^2, \quad \forall n\ge N_0,
$$

which implies that, for all $n, m \ge N_0$, the following inequalities hold:

$$
\begin{aligned}
\|x_{n+m}-u\|^2 &\le \|x_{n+m-1}-u\|^2 + \frac{2\sigma_{n+m-1}^2}{1-2\sigma_{n+m-1}^2}M^2 \\
&\le \|x_{n+m-2}-u\|^2 + \frac{2\sigma_{n+m-2}^2}{1-2\sigma_{n+m-2}^2}M^2 + \frac{2\sigma_{n+m-1}^2}{1-2\sigma_{n+m-1}^2}M^2 \\
&\vdots \\
&\le \|x_n-u\|^2 + \sum_{j=n}^{n+m-1}\frac{2\sigma_j^2}{1-2\sigma_j^2}M^2.
\end{aligned}
$$

Since $\sum_{n=0}^{\infty}\frac{2\sigma_n^2}{1-2\sigma_n^2} < \infty$, we have the estimate

$$
\limsup_{m\to\infty}\|x_m-u\|^2 \le \|x_n-u\|^2 + \sum_{j=n}^{\infty}\frac{2\sigma_j^2}{1-2\sigma_j^2}M^2,
$$

and hence the limit $\lim_{n\to\infty}\|x_n-u\|^2$ exists for every $u\in VO(\Omega,F)\cap VI(\Omega,A)$. In addition, rewriting (38) as

$$
\frac{\delta}{2}\|\tilde{x}_n-z_n\|^2 \le \left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)\|x_n-u\|^2 - \|x_{n+1}-u\|^2 \tag{41}
$$

and observing that the right-hand side of (41) converges to 0 as $n\to\infty$ because the sequence $\{\|x_n-u\|^2\}$ is convergent, we conclude that

$$
\lim_{n\to\infty}\|\tilde{x}_n-z_n\| = 0. \tag{42}
$$

Let now $d := \lim_{n\to\infty} \|x_n - u\|$ and derive from (24) and (37) that

$$\begin{aligned}\limsup_{n\to\infty} \|\tilde{x}_n - u\|^2 &\le \limsup_{n\to\infty}\left\{\left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)\|z_n-u\|^2 - \frac{1}{2}\|\tilde{x}_n - z_n\|^2\right\}\\ &= \limsup_{n\to\infty}\left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)\|z_n-u\|^2 - \frac{1}{2}\lim_{n\to\infty}\|\tilde{x}_n - z_n\|^2\\ &\le \limsup_{n\to\infty}\left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)\|x_n-u\|^2 = d^2.\end{aligned}$$

Observe also the relationships

$$\lim_{n\to\infty} \|\beta_n(x_n-u) + (1-\beta_n)(\tilde{x}_n - u)\| = \lim_{n\to\infty}\|x_{n+1}-u\| = d.$$

Since $\{\beta_n\} \subset [\epsilon, 1-\delta]$ for some $\epsilon, \delta \in (0,1)$, we deduce from Proposition 2.2 that

$$\lim_{n\to\infty} \|x_n - \tilde{x}_n\| = 0 \tag{43}$$

and therefore arrive at

$$\lim_{n\to\infty} \|x_{n+1} - x_n\| = 0. \tag{44}$$

Noting that $\|x_n - z_n\| \le \|x_n - \tilde{x}_n\| + \|\tilde{x}_n - z_n\|$ allows us to conclude from (43) that

$$\lim_{n\to\infty} \|x_n - z_n\| = 0. \tag{45}$$

Furthermore, it follows from (24) that

$$\|z_n - u\|^2 \le \|x_n - u\|^2 + (1-\gamma_n)(\lambda_n^2 k^2 - 1)\|x_n - y_n\|^2$$

for every $u \in VO(\Omega, F) \cap VI(\Omega, A)$. This gives

$$\begin{aligned}\|x_n - y_n\|^2 &\le \frac{1}{(1-\gamma_n)(1-\lambda_n^2k^2)}(\|x_n-u\|^2 - \|z_n-u\|^2)\\ &= \frac{1}{(1-\gamma_n)(1-\lambda_n^2k^2)}(\|x_n-u\| - \|z_n-u\|)(\|x_n-u\| + \|z_n-u\|)\\ &\le \frac{1}{(1-\gamma_n)(1-\lambda_n^2k^2)}(\|x_n-u\| + \|z_n-u\|)\|x_n-z_n\|.\end{aligned} \tag{46}$$

Since $\lim_{n\to\infty}\|x_n - z_n\| = 0$ and the sequences $\{x_n\}$ and $\{z_n\}$ are bounded, we obtain $\lim_{n\to\infty}\|x_n - y_n\| = 0$. Arguing similarly to the proof of (23) yields

$$\begin{aligned}\|t_n - u\|^2 &\le \|x_n-u\|^2 - \|x_n-y_n\|^2 - \|y_n-t_n\|^2 + 2\lambda_n k\|x_n-y_n\|\|t_n-y_n\|\\ &\le \|x_n-u\|^2 - \|x_n-y_n\|^2 - \|y_n-t_n\|^2 + \|x_n-y_n\|^2 + \lambda_n^2k^2\|y_n-t_n\|^2\\ &\le \|x_n-u\|^2 + (\lambda_n^2k^2-1)\|y_n-t_n\|^2,\end{aligned}$$

which implies the relationships

$$
\begin{aligned}
\|z_n - u\|^2 &= \|\gamma_n x_n + (1-\gamma_n)t_n - u\|^2 \\
&= \|\gamma_n(x_n - u) + (1-\gamma_n)(t_n - u)\|^2 \\
&\le \gamma_n\|x_n - u\|^2 + (1-\gamma_n)\|t_n - u\|^2 \\
&\le \gamma_n\|x_n - u\|^2 + (1-\gamma_n)(\|x_n - u\|^2 + (\lambda_n^2k^2 - 1)\|y_n - t_n\|^2) \\
&= \|x_n - u\|^2 + (1-\gamma_n)(\lambda_n^2k^2 - 1)\|y_n - t_n\|^2 \le \|x_n - u\|^2.
\end{aligned}
$$

Rearranging the latter as in (46) gives us

$$
\begin{aligned}
\|t_n - y_n\|^2 &\le \frac{1}{(1-\gamma_n)(1-\lambda_n^2k^2)}(\|x_n - u\|^2 - \|z_n - u\|^2) \\
&= \frac{1}{(1-\gamma_n)(1-\lambda_n^2k^2)}(\|x_n - u\| - \|z_n - u\|)(\|x_n - u\| + \|z_n - u\|) \\
&\le \frac{1}{(1-\gamma_n)(1-\lambda_n^2k^2)}(\|x_n - u\| + \|z_n - u\|)\|x_n - z_n\|.
\end{aligned}
$$

Since $\lim_{n\to\infty}\|x_n - z_n\| = 0$ and the sequences $\{x_n\}$ and $\{z_n\}$ are bounded, we obtain $\lim_{n\to\infty}\|t_n - y_n\| = 0$. It gives, by taking into account the k-Lipschitz continuity of the operator A, that $\lim_{n\to\infty}\|Ay_n - At_n\| = 0$. Noting finally that $\|x_n - t_n\| \le \|x_n - y_n\| + \|y_n - t_n\|$, we arrive at $\lim_{n\to\infty}\|x_n - t_n\| = 0$ and conclude the proof of the claim in Step 4.

Step 5: *Optimality of weak cluster points of* $\{x_n\}$. Observe first that the bounded sequence of iterates $\{x_n\}$ has weak cluster points in the Hilbert space X under consideration. We intend to justify that all of them lie in $VO(\Omega, F) \cap VI(\Omega, A)$.

To proceed, let $\hat{x} \in X$ be a weak cluster point of $\{x_n\}$, and let $\{x_{k_n}\}$ be a subsequence weakly convergent to it. Having in mind to prove that $\hat{x} \in VO(\Omega, F) \cap VI(\Omega, A)$, we show first that $\hat{x} \in VI(\Omega, A)$. It follows from $\lim_{n\to\infty}\|x_n - t_n\| = 0$ and $\lim_{n\to\infty}\|x_n - y_n\| = 0$ that $t_{k_n} \rightharpoonup \hat{x}$ and $y_{k_n} \rightharpoonup \hat{x}$. Consider next the set-valued mapping

$$
Tv := \begin{cases} Av + N_\Omega v, & \text{if } v \in \Omega, \\ \emptyset, & \text{if } v \notin \Omega, \end{cases}
$$

where $N_\Omega v$ is the normal cone to Ω at $v \in \Omega$. As already mentioned, the mapping T is *maximal monotone* and so $0 \in Tv$ *if and only if* $v \in VI(\Omega, A)$; see [15]. Taking a pair $(v, w) \in \operatorname{gph} T$ from the graph of F, we have $w \in v = Av + N_\Omega v$ and hence $w - Av \in N_\Omega v$. This gives $\langle v - t, w - Av\rangle \ge 0$ for all $t \in \Omega$. On the other hand, for $t_n := P_\Omega(x_n - \lambda_n Ay_n)$ and every $v \in \Omega$ we get $\langle x_n - \lambda_n Ay_n - t_n, t_n - v\rangle \ge 0$ and thus

$$
\left\langle v - t_n, \frac{t_n - x_n}{\lambda_n} + Ay_n \right\rangle \ge 0.
$$

It follows from $\langle v - t, w - Av\rangle \ge 0$ that

$$
\langle v - t_{k_n}, w\rangle \ge \langle v - t_{k_n}, Av\rangle
$$

$$\geq \langle v - t_{k_n}, Av\rangle - \left\langle v - t_{k_n}, \frac{t_{k_n} - x_{k_n}}{\lambda_{k_n}} + Ay_{k_n}\right\rangle$$
$$= \langle v - t_{k_n}, Av - At_{k_n}\rangle + \langle v - t_{k_n}, At_{k_n} - Ay_{k_n}\rangle - \left\langle v - t_{k_n}, \frac{t_{k_n} - x_{k_n}}{\lambda_{k_n}}\right\rangle$$
$$\geq \langle v - t_{k_n}, At_{k_n} - Ay_{k_n}\rangle - \left\langle v - t_{k_n}, \frac{t_{k_n} - x_{k_n}}{\lambda_{k_n}}\right\rangle,$$

whenever $t \in \Omega$ and $t_{k_n} \in \Omega$, which implies $\langle v - \hat{x}, w\rangle \geq 0$ as $n \to \infty$. Since T is maximal monotone, we have $\hat{x} \in T^{-1}0$ and hence $\hat{x} \in VI(\Omega, A)$.

Define further $\psi_z : X \to \mathcal{R}$ by $\psi_z(x) := \langle F(x), z\rangle$ and show that

$$\psi_z(\hat{x}) \leq \psi_z(x_n) \tag{47}$$

for all $z \in C^+$ and all $n \geq 0$. Indeed, it follows from the positive lower semicontinuity and C-convexity of F that the scalar function ψ_z is lower semicontinuous and convex, and therefore $\psi_z(\hat{x}) \leq \lim_{n\to\infty} \psi_z(x_{k_n})$. Note that (22) implies that

$$\begin{aligned} F(x_{n+1}) &= F(\beta_n x_n + (1-\beta_n)\tilde{x}_n) \\ &\preceq_C \beta_n F(x_n) + (1-\beta_n)F(\tilde{x}_n) \\ &\preceq_C \beta_n F(x_n) + (1-\beta_n)F(x_n) = F(x_n). \end{aligned}$$

Thus $F(x_{n+1}) \preceq_C F(x_n)$ and $x_{n+1} \in \Omega_n$ for all n. Consequently we have $\psi_z(x_{n+1}) \leq \psi_z(x_n)$ for all n, and hence $\lim_{n\to\infty} \psi_z(x_{k_n}) = \inf\{\psi_z(x_n)\}$. This shows that $\psi_z(\hat{x}) \leq \inf\{\psi_z(x_n)\}$, and so (47) holds. It follows easily from (47) that

$$F(\hat{x}) \preceq_C F(x_n) \quad \text{for all } n \geq 0. \tag{48}$$

Suppose now that $\hat{x}$ is *not* weakly efficient for VOP, i.e., there exists $\bar{x} \in \Omega$ such that $F(\bar{x}) \prec_C F(\hat{x})$. Take $\hbar_n$ as chosen right before (26). Since $\|\hbar_n\| = 1$ for all n, by the classical Bourbaki-Alaoglu theorem there exists a weak* cluster point of $\{\hbar_{k_n}\}$, say $\hbar \in Y^*$, which is a weak* limit of some subnet $\{\hbar_{j_n}\}$ of $\{\hbar_{k_n}\}$. We claim now that the positive polar cone C^+ is weak* closed. Observing that $C^+ = \bigcap_{y\in C}\{z \in Y^* | \langle y, z\rangle \geq 0\}$ and taking into account that the linear forms $z \mapsto \langle y, z\rangle$ are weak* continuous for all $y \in Y$, we represent C^+ as an intersection of weak* closed sets and thus justify the claim. It follows therefore that $\hbar \in C^+$. Note further that ψ_z is convex for each $z \in C^+$. Hence from (22) we get

$$\begin{aligned} \langle F(x_{j_n+1}), \hbar_{j_n}\rangle &= \psi_{j_n}(x_{j_n+1}) = \psi_{j_n}(\beta_{j_n} x_{j_n} + (1-\beta_{j_n})\tilde{x}_{j_n}) \\ &\leq \beta_{j_n}\psi_{j_n}(x_{j_n}) + (1-\beta_{j_n})\psi_{j_n}(\tilde{x}_{j_n}) \\ &= \beta_{j_n}(\psi_{j_n}(x_{j_n}) - \psi_{j_n}(\tilde{x}_{j_n})) + \psi_{j_n}(\tilde{x}_{j_n}) \\ &= \beta_{j_n}\langle F(x_{j_n}) - F(\tilde{x}_{j_n}), \hbar_{j_n}\rangle + \psi_{j_n}(\tilde{x}_{j_n}) \\ &= \beta_{j_n}(\langle F(x_{j_n}) - F(\tilde{x}_{j_n}), \hbar_{j_n}\rangle - \langle F(\hat{x}) - F(\tilde{x}_{j_n}), \hbar\rangle) \\ &\quad - \beta_{j_n}(\psi_{\hbar}(\tilde{x}_{j_n}) - \psi_{\hbar}(\hat{x})) + \psi_{j_n}(\tilde{x}_{j_n}). \end{aligned} \tag{49}$$

Observe also that $\lim_{n\to\infty} \psi_\hbar(x_{j_n}) = \inf\{\psi_\hbar(x_n)\} \geq \psi_\hbar(\hat{x})$ and that $\liminf_{n\to\infty} \psi_\hbar(\tilde{x}_{j_n}) \geq \psi_\hbar(\hat{x})$, since $\|\tilde{x}_{j_n} - x_{j_n}\| \to 0$ and $\psi_\hbar$ is weakly lower semicontinuous. Moreover, we get

$$\begin{aligned}\psi_{j_n}(\bar{x}) - \psi_{j_n}(\tilde{x}_{j_n}) &\geq \langle v_{j_n}, \bar{x} - \tilde{x}_{j_n}\rangle \\ &= \langle u_{j_n}, \bar{x} - \tilde{x}_{j_n}\rangle - \nu_{j_n}\langle \tilde{x}_{j_n} - z_{j_n} - \theta_{j_n}, \bar{x} - \tilde{x}_{j_n}\rangle \\ &\geq -\nu_{j_n}\langle \tilde{x}_{j_n} - z_{j_n} - \theta_{j_n}, \bar{x} - \tilde{x}_{j_n}\rangle \\ &\geq -\nu_{j_n}\|\tilde{x}_{j_n} - z_{j_n} - \theta_{j_n}\|\|\bar{x} - \tilde{x}_{j_n}\| \end{aligned} \tag{50}$$

by using (32) in the first inequality, (31) in the second equality, and (29) in the third inequality together with the fact that $\bar{x} \in \Omega_n$ for all $n \geq 0$ due to $F(\bar{x}) \prec_C F(\hat{x}) \preceq_C F(x_n)$ by (48). Employing consequently (47), we conclude from (49) and (50) that

$$\begin{aligned}\langle F(\bar{x}) - F(\hat{x}), \hbar_{j_n}\rangle &\geq \langle F(\bar{x}) - F(x_{j_n+1}), \hbar_{j_n}\rangle = \psi_{j_n}(\bar{x}) - \psi_{j_n}(x_{j_n+1}) \\ &\geq \psi_{j_n}(\bar{x}) - \psi_{j_n}(\tilde{x}_{j_n}) - \beta_{j_n}(\langle F(x_{j_n}) - F(\tilde{x}_{j_n}), \hbar_{j_n}\rangle \\ &\quad - \langle F(\hat{x}) - F(\tilde{x}_{j_n}), \hbar\rangle) + \beta_{j_n}(\psi_\hbar(\tilde{x}_{j_n}) - \psi_\hbar(\hat{x})) \\ &\geq -\nu_{j_n}\|\tilde{x}_{j_n} - z_{j_n} - \theta_{j_n}\|\|\bar{x} - \tilde{x}_{j_n}\| - \beta_{j_n}(\langle F(x_{j_n}) - F(\tilde{x}_{j_n}), \hbar_{j_n}\rangle \\ &\quad - \langle F(\hat{x}) - F(\tilde{x}_{j_n}), \hbar\rangle) + \beta_{j_n}(\psi_\hbar(\tilde{x}_{j_n}) - \psi_\hbar(\hat{x})). \end{aligned} \tag{51}$$

Note further that $\lim_{n\to\infty} \|\tilde{x}_{j_n} - z_{j_n}\| = 0$ by (42), $\lim_{n\to\infty} \|x_{j_n} - \tilde{x}_{j_n}\| = 0$ by (51), and $\|\theta_{j_n}\| \leq \sigma_{j_n}\|\tilde{x}_{j_n} - z_{j_n}\|$ by (14). Now we take the lower limits in the first and last expressions of (51). Regarding the first term of the rightmost expression in (51), since $\{\alpha_n\}$ is bounded and $\|\hbar_n\| = \|e_n\| = 1$, we have that $\{\nu_n\}$ is bounded as well. Note again that $\{x_n\}$ is bounded, and so $\{\|\bar{x} - \tilde{x}_{j_n}\|\}$ is also bounded. Finally, it is easy to see that $\lim_{n\to\infty} \|\tilde{x}_{j_n} - z_{j_n} - \theta_{j_n}\| = 0$, which allows us to conclude that

$$\lim_{n\to\infty} |\langle F(x_{j_n}) - F(\tilde{x}_{j_n}), \hbar_{j_n}\rangle - \langle F(\hat{x}) - F(\tilde{x}_{j_n}), \hbar\rangle| = 0$$

according to assumption (B). Since the function $\psi_\hbar$ is convex and lower semicontinuous, it is weakly lower semicontinuous. Thus we get $\liminf_{n\to\infty} \psi_\hbar(\tilde{x}_{j_n}) \geq \psi_\hbar(\hat{x})$; i.e., $\liminf_{n\to\infty}(\psi_\hbar(\tilde{x}_{j_n}) - \psi_\hbar(\hat{x})) \geq 0$. The latter yields that for any given $\varepsilon > 0$ there is an integer $N_0 \geq 1$ such that

$$\psi_\hbar(\tilde{x}_{j_n}) - \psi_\hbar(\hat{x}) \geq -\varepsilon, \quad \text{whenever } n \geq N_0.$$

Using this with together $\{\beta_n\} \subset [\epsilon, 1-\delta]$ implies that

$$\beta_{j_n}(\psi_\hbar(\tilde{x}_{j_n}) - \psi_\hbar(\hat{x})) \geq -\varepsilon\beta_{j_n} \geq -\varepsilon(1-\delta), \quad n \geq N_0.$$

Consequently we have

$$\liminf_{n\to\infty} \beta_{j_n}(\psi_\hbar(\tilde{x}_{j_n}) - \psi_\hbar(\hat{x})) \geq -\varepsilon(1-\delta)$$

and, by an arbitrary choice of $\varepsilon > 0$, arrive at

$$\liminf_{n\to\infty} \beta_{j_n}(\psi_\hbar(\tilde{x}_{j_n}) - \psi_\hbar(\hat{x})) \geq 0.$$

Thus the lower limit of the rightmost expression in (51) as $n \to \infty$ is not less than zero. Since $\hbar$ is the weak* limit of $\{\hbar_{j_n}\}$, we get from (51) that

$$\langle F(\bar{x}) - F(\tilde{x}), \hbar \rangle \geq 0. \tag{52}$$

Next we claim that $\hbar \neq 0$. Indeed, taking $e \in \operatorname{int} C$ and using Lemma 2.2 from [29] implies that $\langle e, \hbar_n \rangle \geq d(e, Y \setminus C) > 0$ for all $n \geq 0$. Since $\hbar$ is the weak* limit of $\{\hbar_{j_n}\}$, we get that $\langle e, \hbar \rangle > 0$, which thus yields $\hbar \neq 0$. The latter shows that (52) contradicts the fact that $\hbar$ belongs to C^+ and so the assumption that $F(\bar{x}) \prec_C F(\hat{x})$. Thus this assumption is false, and so $\hat{x}$ is indeed weakly efficient for VOP.

Step 6: *Uniqueness of the weak cluster point of iterates*. This part of the proof, presented for the sake of completeness, is rather similar to the scalar-valued case in [6] using Brézis' uniqueness arguments. Take two limiting points $\hat{x}$ and $\tilde{x}$ of the sequence $\{x_n\}$. As shown above, both limiting points $\hat{x}$ and $\tilde{x}$ belong to $VO(\Omega, F) \cap VI(\Omega, A)$ and both sequences $\{\|\hat{x} - x_n\|\}$ and $\{\|\tilde{x} - x_n\|\}$ converge. This means that there are $\hat{\beta}, \tilde{\beta} \in \mathcal{R}_+$ such that

$$\lim_{n\to\infty} \|x_n - \hat{x}\| = \hat{\beta} \quad \text{and} \quad \lim_{n\to\infty} \|x_n - \tilde{x}\| = \tilde{\beta}. \tag{53}$$

By the Hilbert space identity,

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

we conclude from (53) that

$$\lim_{n\to\infty} \langle x_n - \tilde{x}, \tilde{x} - \hat{x}\rangle = \frac{1}{2}(\hat{\beta}^2 - \tilde{\beta}^2 - \|\tilde{x} - \hat{x}\|^2). \tag{54}$$

The left-hand side of (54) vanishes because $\tilde{x}$ is a weak cluster point of $\{x_n\}$, and thus

$$\hat{\beta}^2 - \tilde{\beta}^2 = \|\tilde{x} - \hat{x}\|^2. \tag{55}$$

Reversing the roles of $\hat{x}$ and $\tilde{x}$, we get $\tilde{\beta}^2 - \hat{\beta}^2 = \|\tilde{x} - \hat{x}\|^2$. Combining the latter with (55) allows us to conclude that $\|\tilde{x} - \hat{x}\| = 0$, i.e., $\tilde{x} = \hat{x}$. This justifies the uniqueness of the weak cluster point of $\{x_n\}$ and thus shows that the sequence $\{x_n\}$ weakly converges to an element of $VO(\Omega, F) \cap VI(\Omega, A)$, which completes the proof of the theorem. □

Remark 3.4 (Stopping Rule for Vectorial Algorithms) Observe that the "stopping rule" in the above Algorithm 1 (resp. Algorithms 2 and 3 in Sects. 4 and 5) requires that $x_{n_p} = x_p$ for all $p \geq 1$ if for given x_n we have $x_n \in C\text{-Min}\{F(x)|x \in \Omega\}$. In general, the requirement $x_{n+1} = x_n$ is sufficient as the usual stopping rule in scalar proximal point method. But, for the above vectorial Algorithm 1 (resp. Algorithms 2 and 3 in Sects. 4 and 5) we specifically indicate and emphasize that "the method generates a sequence $\{x_n\}$", i.e., an *infinite* sequence $\{x_n\}$.

Indeed, the goal of this paper is to solve the VOP: $C\text{-Min}\{F(x)|x \in \Omega\}$. In the proceeding of iterations we meet the two possible cases.

Case I. At each iteration step we have $x_n \notin C\text{-Min}\{F(x)|x \in \Omega\}$. Hence the process of iterations continues infinitely producing an infinite sequence $\{x_n\}$. Under the conditions of Theorem 3.1 (resp. Theorems 4.3 and 5.1 in Sects. 4 and 5), $\{x_n\}$ converges weakly to a solution of the VOP. This achieves our aim.

Case II. There exists some iteration step such that we have $x_n \in C\text{-Min}\{F(x)| x \in \Omega\}$. This actually achieves our aim. However, in order to obtain an infinite sequence $\{x_n\}$, we take $x_{n+p} = x_n$ for all $p \geq 1$. In this case there is no doubt that the sequence $\{x_n\}$ converges weakly to a solution of the VOP under consideration.

4 Extension to Bregman-Function-Based Hybrid Approximate Proximal Algorithms

A number of research during recent years has focused on nonlinear generations of recursion (1) based on *Bregman functions*, which are discussed, e.g., in [5]; see also the references therein. In this section we continue these lines of research concerning an appropriate extension of the absolute version of HAPM developed in Sect. 3.

Let $h : \Omega \to \mathcal{R}$ be a strictly convex function, which is Gâteaux differentiable on Ω. The *Bregman distance* between x and y is defined via the "D-function"

$$D_h(x, y) = h(x) - h(y) - \langle \nabla h(y), x - y \rangle, \tag{56}$$

where $x, y \in \Omega$. It follows from the strict convexity of h that $D_h(x, y) \geq 0$ and that $D_h(x, y) = 0$ if and only if $x = y$. We refer the reader to [17] as to a good source of information on Bregman functions, their properties, and some applications. It is worth reminding here the basis definition taken from [17].

Definition 4.1 (Bregman Functions) Let Ω be a nonempty closed convex subset of X such that the convex open subset $\operatorname{int}\Omega \subset X$ satisfies the property $\Omega = \operatorname{cl}(\operatorname{int}\Omega)$, where the cl denotes as usual the closure of a set. We say that $h : \Omega \to \mathcal{R}$ is a Bregman function if the following hold:

(i) h is strictly convex and continuous on Ω;
(ii) h is continuously differentiable on $\operatorname{int}\Omega$;
(iii) the partial sublevel set

$$L(x, a) = \{y \in \operatorname{int}\Omega | D_h(x, y) \leq \alpha\}$$

is bounded for any $x \in \Omega$ and $\alpha \in \mathcal{R}$;
(iv) if $\{y_n\}$ is a sequence in $\operatorname{int}\Omega$ converging to y, then $\lim_{n\to\infty} D_h(y, y_n) = 0$.

Note that $D_h(x, y) = \frac{1}{2}\|x - y\|^2$ for $h(x) = \frac{1}{2}\|x\|^2$. In what follows we use a class of functions that are represented as

$$h(x) = h_0(x) + \frac{1}{2}\|x\|^2,$$

where h_0 is a Bregman function. It is easy to see that h satisfies the conditions of Definition 4.1, and thus h is also a Bregman function. Furthermore, for all $x, y \in \Omega$

we have

$$\begin{aligned}D_h(x,y)&=h(x)-h(y)-\langle\nabla h(y),x-y\rangle\\&=h_0(x)-h_0(y)-\langle\nabla h_0(y),x-y\rangle+\frac{1}{2}\|x\|^2-\frac{1}{2}\|y\|^2-\langle y,x-y\rangle\\&=D_{h_0}(x,y)+\frac{1}{2}\|x-y\|^2\ge\frac{1}{2}\|x-y\|^2.\end{aligned}\tag{57}$$

Let us next describe a Bregman-function extension of the absolute version of HARM (Algorithm 1) from Sect. 3, which we call for brevity Algorithm 2. It requires some *exogenous sequences*: an error sequence $\{\theta_n\}\subset X$, two bounded sequences of positive real numbers $\{\alpha_n\}$ and $\{\sigma_n\}$, a relaxation sequence $\{\gamma_n\}\subset[0,1]$, and a sequence $\{e_n\}\subset\operatorname{int}C$ such that $\|e_n\|=1$ for all n. Assume that $\Omega\cap\operatorname{dom}F\neq\emptyset$. Algorithm 2 generates a sequence of iterates $\{x_n\}\subset\Omega$ in the following way:

Algorithm 2

Initialization: Choose $x_0\in\Omega\cap\operatorname{dom}F$.
Stopping Rule: Given x_n, if $x_n\in C\text{-}\mathrm{ArgMin}_w\{F(x)|x\in\Omega\}$ $(=VO(\Omega,F))$, then we let $x_{n+p}:=x_n$ for all $p\ge1$.
Iterative Step: Given x_n, if $x_n\notin C\text{-}\mathrm{ArgMin}_w\{F(x)|x\in\Omega\}$, we first compute

$$\begin{cases}y_n=P_\Omega(x_n-\lambda_nAx_n),\\ z_n=\gamma_nx_n+(1-\gamma_n)P_\Omega(x_n-\lambda_nAy_n)\end{cases}\tag{58}$$

for every $n=0,1,2,\ldots$, where $\{\lambda_n\}\subset(0,1)$ and $\{\gamma_n\}\subset[0,1]$, and then take as the next iterate any $x_{n+1}\in\Omega$ satisfying

$$x_{n+1}\in C\text{-}\mathrm{ArgMin}_w\left\{F(x)+\frac{\alpha_n}{2}(2h(x)+\|x-\nabla h(z_n)-\theta_n\|^2-\|x\|^2)e_n|x\in\Omega_n\right\}\tag{59}$$

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

In Algorithm 2, instead of condition (14) imposed in Algorithm 1, for every $u\in VO(\Omega,F)\cap VI(\Omega,A)$ we take the condition

$$\begin{aligned}&\langle\nabla h(z_n)-\nabla h(x_{n+1})+\theta_n,x_{n+1}-u\rangle\ge0\\&\quad\Rightarrow\quad\langle\nabla h(x_n)-\nabla h(x_{n+1})+\theta_n,x_{n+1}-u\rangle\ge0,\\&D_h(u,z_n)-\frac{1}{2}\|u-z_n\|^2\le\|\theta_n\|^2\le2\sigma_n^2D_h(x_{n+1},x_n),\quad\text{with }\sum_{n=0}^{\infty}\sigma_n<\infty\end{aligned}\tag{60}$$

as the approximate criterion corresponding to recursion (59).

Assume the following for the initial data of the VOP under consideration and the initial iterate x_0 in Algorithm 2:

(A) The set $(F(x_0)-C)\cap F(\Omega)$ is C-*quasicomplete* for Ω, which means that for all sequences $\{a_n\}\subset\Omega$ with $a_0=x_0$ such that $F(a_{n+1})\preceq_C F(a_n)$ for all $n\ge 0$, it holds that $F(u)\preceq_C F(a_n)$ for all $u\in VO(\Omega,F)\cap VI(\Omega,A)$ and all $n\ge 0$.

To establish the convergence of iterates in Algorithm 2, we need some properties of Bregman functions that are derived in [13].

Proposition 4.2 (Properties of Bregman Functions) *The following hold for the Bregman functions under consideration*:

(i) *For any $x,y,z\in X$, we have*

$$D_h(y,x)=D_h(z,x)+D_h(y,z)+\langle\nabla h(x)-\nabla h(z),z-y\rangle.$$

(ii) *For any $x,y,z,s\in X$, we have*

$$D_h(s,z)=D_h(s,x)+\langle\nabla h(x)-\nabla h(z),s-y\rangle+D_h(y,z)-D_h(y,x).$$

Now we are ready to prove the convergence of Algorithm 2 under condition (60) and the assumptions imposed in (A).

Theorem 4.3 (Well-Posedness and Convergence of the Bregman-Function Version of HAPM) *Let $F:\Omega\to Y\cup\{\infty_C\}$ be a proper, C-convex, and positively lower semicontinuous mapping with $\Omega\cap\operatorname{dom}F\neq\emptyset$, let $h:\Omega\to\mathcal{R}$ be the Bregman function defined above and such that $\nabla h(\cdot)$ is uniformly continuous from the strong topology of X to the strong topology of X, and let $A:\Omega\to X$ be a monotone and k-Lipschitz continuous mapping such that $VO(\Omega,F)\cap VI(\Omega,A)\neq\emptyset$. Assume the fulfillment of the assumptions imposed in* (A), *the implication in* (60), *and the following conditions on the exogenous sequences*:

(i) $\{\lambda_n\}\subset[a,b]$ *for some* $a,b\in(0,1/k)$;
(ii) $\{\gamma_n\}\subset[c,d]$ *for some* $c,d\in(0,1)$.

Then the sequence of iterates $\{x_n\}$ generated by Algorithm 2 is well defined and has the convergence properties:

(I) $\{x_n\}$ *converges with respect to the weak topology of X to a weakly efficient solution of the vector optimization problems VOP*;
(II) $\{x_n\}$ *converges with respect to the weak topology of X to an element of the set $VO(\Omega,F)\cap VI(\Omega,A)$ provided that $x_n\notin C$-$\operatorname{ArgMin}_w\{F(x)|x\in\Omega\}$ for all $n\ge 0$.*

Proof We split the proof into the following steps.

Step 1: *For every $u\in VO(\Omega,F)\cap VI(\Omega,A)$ we have*

$$\|z_n-u\|^2\le\|x_n-u\|^2+(1-\gamma_n)(\lambda_n^2k^2-1)\|x_n-y_n\|^2,\quad \text{whenever } n\ge 0.$$

The proof of this assertion is similar to Step 1 in Theorem 3.1 and is omitted here.

Step 2: *Existence of iterates*. This assertion can be proved by using the same arguments as in Step 2 of Theorem 3.1 with $\varphi_n : X \to \mathcal{R} \cup \{\infty\}$ defined now by

$$\varphi_n(x) := \langle F(x), z\rangle + I_{\Omega_n}(x) + \frac{\alpha_n}{2}\langle e_n, z\rangle(2h(x) + \|x - \nabla h(z_n) - \theta_n\|^2 - \|x\|^2). \quad (61)$$

Step 3: *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 an element of $VO(\Omega, F)$. We assume from now on that the stopping rule never applies. Since x_{n+1} solves the vector optimization problem in (59), by Proposition 2.1 there is $\hbar_n \in C^+ \setminus \{0\}$ such that x_{n+1} solves the problem:

$$\min \eta_n(x), \quad (62)$$

$$\text{s.t.} \quad x \in \Omega_n, \quad (63)$$

where $\eta_n : X \to \mathcal{R} \cup \{+\infty\}$ is defined by

$$\eta_n(x) := \langle F(x), \hbar_n\rangle + \frac{\alpha_n}{2}\langle e_n, \hbar_n\rangle(2h(x) + \|x - \nabla h(z_n) - \theta_n\|^2 - \|x\|^2).$$

Since the solution of (62) and (63) is not altered through multiplication of $\hbar_n$ by positive scalars, we can assume without loss of generality that $\|\hbar_n\| = 1$ for all $n \geq 0$. Note that by definition we have $\Omega_n \subset \mathrm{dom}(\eta_n) = \mathrm{dom}\, F$, so that $\emptyset \neq \mathrm{dom}(I_{\Omega_n}) \subset \mathrm{dom}(\eta_n)$. According to [20, Theorem 3.23], it follows that x_{n+1} satisfies the first-order optimality conditions for problem (62) and (63), i.e., there exists $u_n \in X$ such that

$$u_n \in \partial\eta_n(x_{n+1}), \quad (64)$$

$$0 \leq \langle u_n, x - x_{n+1}\rangle, \quad \text{for all } x \in \Omega_n. \quad (65)$$

Define now $\psi_n : X \to \mathcal{R} \cup \{\infty\}$ by

$$\psi_n(x) := \langle F(x), \hbar_n\rangle. \quad (66)$$

In view of (61) and (64) we have

$$u_n = v_n + \alpha_n\langle e_n, \hbar_n\rangle(\nabla h(x_{n+1}) - \nabla h(z_n) - \theta_n) \quad (67)$$

for some subgradient

$$v_n \in \partial\psi_n(x_{n+1}). \quad (68)$$

Next fix an arbitrary element $u \in VO(\Omega, F) \cap VI(\Omega, A)$ and get by (A) that $u \in \Omega_n$ for all $n \geq 0$. Combining (65) with $x = u$ and (67), we have

$$\begin{aligned} 0 &\leq \langle v_n, u - x_{n+1}\rangle + \alpha_n\langle e_n, \hbar_n\rangle\langle\nabla h(x_{n+1}) - \nabla h(z_n) - \theta_n, u - x_{n+1}\rangle \\ &\leq \langle F(u) - F(x_{n+1}), \hbar_n\rangle + \alpha_n\langle e_n, \hbar_n\rangle\langle\nabla h(x_{n+1}) - \nabla h(z_n) - \theta_n, u - x_{n+1}\rangle \\ &\leq \alpha_n\langle e_n, \hbar_n\rangle\langle\nabla h(x_{n+1}) - \nabla h(z_n) - \theta_n, u - x_{n+1}\rangle, \end{aligned} \quad (69)$$

by using (66) and (68) in the second inequality and the fact that $\hbar_n \in C^+ \setminus \{0\}$ in the third; it is clear that $F(u) - F(x_{n+1}) \preceq_C 0$ and therefore $\langle F(u) - F(x_{n+1}), \hbar_n\rangle \leq 0$.

Letting $\nu_n := \alpha_n \langle e_n, \hbar_n\rangle$, note that $\nu_n > 0$ by $\alpha_n > 0$, $e_n \in \operatorname{int} C$, and $\hbar_n \in C^+ \setminus \{0\}$. From (69) we obtain the inequality

$$\langle \nabla h(z_n) - \nabla h(x_{n+1}) + \theta_n, x_{n+1} - u\rangle \geq 0, \tag{70}$$

which together with (60) imply that

$$\langle \nabla h(x_n) - \nabla h(x_{n+1}) + \theta_n, x_{n+1} - u\rangle \geq 0. \tag{71}$$

Furthermore, by Proposition 4.2 we derive from (71) that

$$\begin{aligned} D_h(u, x_{n+1}) &= D_h(u, z_n) + \langle \nabla h(z_n) - \nabla h(x_{n+1}), u - x_{n+1}\rangle \\ &\quad + D_h(x_{n+1}, x_{n+1}) - D_h(x_{n+1}, z_n) \\ &= D_h(u, z_n) - D_h(x_{n+1}, z_n) - \langle \nabla h(z_n) - \nabla h(x_{n+1}), x_{n+1} - u\rangle. \end{aligned} \tag{72}$$

Observe that putting $x = x_n,\ y = u,\ z = z_n$ and $s = x_{n+1}$ in Proposition 4.2 gives us

$$\begin{aligned} D_h(x_{n+1}, z_n) &= D_h(x_{n+1}, x_n) + \langle \nabla h(x_n) - \nabla h(z_n), x_{n+1} - u\rangle \\ &\quad + D_h(u, z_n) - D_h(u, x_n). \end{aligned}$$

Substituting the last equality in (72), we get from (71) that

$$\begin{aligned} D_h(u, x_{n+1}) &= D_h(u, z_n) - D_h(x_{n+1}, x_n) - \langle \nabla h(x_n) - \nabla h(z_n), x_{n+1} - u\rangle \\ &\quad - D_h(u, z_n) + D_h(u, x_n) - \langle \nabla h(z_n) - \nabla h(x_{n+1}), x_{n+1} - u\rangle \\ &= D_h(u, x_n) - D_h(x_{n+1}, x_n) + \langle \nabla h(x_{n+1}) - \nabla h(x_n), x_{n+1} - u\rangle \\ &= D_h(u, x_n) - D_h(x_{n+1}, x_n) - \langle \nabla h(x_n) - \nabla h(x_{n+1}) + \theta_n, x_{n+1} - u\rangle \\ &\quad + \langle \theta_n, x_{n+1} - u\rangle \\ &\leq D_h(u, x_n) - D_h(x_{n+1}, x_n) + \langle \theta_n, x_{n+1} - u\rangle. \end{aligned} \tag{73}$$

Taking next an arbitrary sequence of $\sigma_n > 0$ and using (57) and (60), we obtain

$$\begin{aligned} \langle \theta_n, x_{n+1} - u\rangle &\leq \frac{1}{4\sigma_n^2}\|\theta_n\|^2 + \sigma_n^2\|x_{n+1} - u\|^2 \\ &\leq \frac{1}{2} D_h(x_{n+1}, x_n) + 2\sigma_n^2 D_h(u, x_{n+1}). \end{aligned} \tag{74}$$

Since $\sigma_n \to 0$ as $n \to \infty$, there exists an integer $N_0 \geq 0$ such that $1 - 2\sigma_n^2 > 0$ for all $n \geq N_0$. Substituting (74) in (72) and (73) gives us

$$\begin{aligned}D_h(u,x_{n+1}) &\le \left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)D_h(u,x_n)-\frac{1}{2(1-2\sigma_n^2)}D_h(x_{n+1},x_n)\\&\le \left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)D_h(u,x_n)-\frac{1}{2}D_h(x_{n+1},x_n).\end{aligned}\tag{75}$$

On the other hand, it follows from (24), (70), (72), and (60) that

$$\begin{aligned}\frac{1}{2}\|u-x_{n+1}\|^2 &\le D_h(u,x_{n+1})\\&= D_h(u,z_n)-D_h(x_{n+1},z_n)-\langle\nabla h(z_n)-\nabla h(x_{n+1}),x_{n+1}-u\rangle\\&= D_h(u,z_n)-D_h(x_{n+1},z_n)-\langle\nabla h(z_n)-\nabla h(x_{n+1})+\theta_n,x_{n+1}-u\rangle\\&\quad+\langle\theta_n,x_{n+1}-u\rangle\\&\le D_h(u,z_n)-D_h(x_{n+1},z_n)+\langle\theta_n,x_{n+1}-u\rangle\\&\le D_h(u,z_n)+\|\theta_n\|\|x_{n+1}-u\|\\&\le \frac{1}{2}\|u-z_n\|^2+\|\theta_n\|^2+\|\theta_n\|\|x_{n+1}-u\|\\&\le \frac{1}{2}\|u-x_n\|^2+2\sigma_n^2D_h(x_{n+1},x_n)\\&\quad+\sqrt{2}\sigma_nD_h^{1/2}(x_{n+1},x_n)\|x_{n+1}-u\|,\end{aligned}\tag{76}$$

which justify the assertion of Step 3.

Step 4: *Boundedness of the sequence and proximity of consecutive iterates*. Let us prove first that for every $u\in VO(\Omega,F)\cap VI(\Omega,A)$ the sequence $\{D_h(u,x_n)\}$ is convergent. Indeed, we have in terms of (76) that

$$D_h(u,x_{n+1})\le\left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)D_h(u,x_n),\quad \text{whenever } n\ge N_0.\tag{77}$$

Since $\sum_{n=0}^{\infty}\sigma_n^2<\infty$, it follows that

$$K_0:=\sum_{n=N_0}^{\infty}\frac{2\sigma_n^2}{1-2\sigma_n^2}<\infty \quad\text{and}\quad K_1:=\prod_{n=N_0}^{\infty}\left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)<\infty.$$

Observe further that for all $n\ge N_0$ we have the estimates

$$\begin{aligned}D_h(u,x_{n+1})&\le\left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)D_h(u,x_n)\\&\le\left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)\left(1+\frac{2\sigma_{n-1}^2}{1-2\sigma_{n-1}^2}\right)D_h(u,x_{n-1})\\&\ \ \vdots\end{aligned}$$

$$\begin{aligned}
&\le \prod_{j=N_0}^{n}\left(1+\frac{2\sigma_j^2}{1-2\sigma_j^2}\right)D_h(u,x_{N_0})\\
&\le \prod_{j=N_0}^{\infty}\left(1+\frac{2\sigma_j^2}{1-2\sigma_j^2}\right)D_h(u,x_{N_0})\\
&= K_1 D_h(u,x_{N_0}).
\end{aligned}$$

Consequently, $\{D_h(u,x_n)\}$ is bounded and so is $\{x_n\}$ due to (57). Hence it follows from (23) and (24) that both sequences $\{t_n\}$ and $\{z_n\}$ are bounded. Set $\widetilde{M} := \sup_{n\ge 0} D_h(u,x_n)$ and get from (77) the inequality

$$D_h(u,x_{n+1}) \le D_h(u,x_n) + \frac{2\sigma_n^2}{1-2\sigma_n^2}\widetilde{M}, \quad n \ge N_0,$$

which implies that for all $n,m \ge N_0$ the following hold:

$$\begin{aligned}
D_h(u,x_{n+m}) &\le D_h(u,x_{n+m-1}) + \frac{2\sigma_{n+m-1}^2}{1-2\sigma_{n+m-1}^2}\widetilde{M}\\
&\le D_h(u,x_{n+m-2}) + \frac{2\sigma_{n+m-2}^2}{1-2\sigma_{n+m-2}^2}\widetilde{M} + \frac{2\sigma_{n+m-1}^2}{1-2\sigma_{n+m-1}^2}\widetilde{M}\\
&\;\;\vdots\\
&\le D_h(u,x_n) + \sum_{j=n}^{n+m-1}\frac{2\sigma_j^2}{1-2\sigma_j^2}\widetilde{M}.
\end{aligned}$$

Since $\sum_{n=0}^{\infty}\frac{2\sigma_n^2}{1-2\sigma_n^2} < \infty$, we further have

$$\limsup_{m\to\infty} D_h(u,x_m) \le D_h(u,x_n) + \sum_{j=n}^{\infty}\frac{2\sigma_j^2}{1-2\sigma_j^2}\widetilde{M},$$

and hence $\lim_{n\to\infty} D_h(u,x_n)$ exists for every element $u \in VO(\Omega,F)\cap VI(\Omega,A)$. In addition, rewriting (75) as

$$\frac{1}{2}D_h(x_{n+1},x_n) \le \left(1+\frac{2\sigma_n^2}{1-2\sigma_n^2}\right)D_h(u,x_n) - D_h(u,x_{n+1}) \tag{78}$$

and observing that the right-hand side of (78) converges to 0 as $n\to\infty$ due to the convergence of $\{D_h(u,x_n)\}$, we conclude that

$$\lim_{n\to\infty} D_h(x_{n+1},x_n) = 0 \tag{79}$$

and therefore arrive at $\lim_{n\to\infty}\|x_{n+1}-x_n\| = 0$ by (57).

On the other hand, since $\sum_{n=0}^{\infty}\sigma_n < \infty$, both sequences $\{D_h(x_{n+1}, x_n)\}$ and $\{\|x_n - u\|\}$ are bounded, and hence we have

$$\sum_{n=0}^{\infty}(2\sigma_n^2 D_h(x_{n+1}, x_n) + \sqrt{2}\sigma_n D_h^{1/2}(x_{n+1}, x_n)\|x_{n+1} - u\|) < \infty.$$

Thus it follows from (76) that $\lim_{n\to\infty} \frac{1}{2}\|u - x_n\|^2$ exists and so $\lim_{n\to\infty}\|u - x_n\|$ exists as well. We put $\tau := \lim_{n\to\infty}\|x_n - u\|$, and letting $n \to \infty$, obtain from (76) that

$$\lim_{n\to\infty}\frac{1}{2}\|u - z_n\|^2 = \lim_{n\to\infty}\frac{1}{2}\tau^2,$$

which gives the limiting relationship

$$\lim_{n\to\infty}\|\gamma_n(u - x_n) + (1 - \gamma_n)(u - t_n)\| = \lim_{n\to\infty}\|u - z_n\| = \tau.$$

Note that (23) implies that $\limsup_{n\to\infty}\|u - t_n\| \le \tau$. Utilizing Proposition 2.2, we have

$$\lim_{n\to\infty}\|x_n - t_n\| = 0,$$

which together with (58) imply that

$$\lim_{n\to\infty}\|z_n - x_n\| = \lim_{n\to\infty}(1 - \gamma_n)\|t_n - x_n\| = 0.$$

Picking any $u \in VO(\Omega, F) \cap VI(\Omega, A)$ allows us to derive from (24) that

$$\begin{aligned}
\|x_n - y_n\|^2 &\le \frac{1}{(1 - \gamma_n)(1 - \lambda_n^2 k^2)}(\|x_n - u\|^2 - \|z_n - u\|^2) \\
&\le \frac{1}{(1 - \gamma_n)(1 - \lambda_n^2 k^2)}(\|x_n - u\| + \|z_n - u\|)\|x_n - z_n\|.
\end{aligned}$$

Since $\lim_{n\to\infty}\|x_n - z_n\| = 0$ and the sequences $\{x_n\}$ and $\{z_n\}$ are bounded, we get that $\lim_{n\to\infty}\|x_n - y_n\| = 0$. By the same process as in (23) we also have

$$\begin{aligned}
\|t_n - u\|^2 &\le \|x_n - u\|^2 - \|x_n - y_n\|^2 - \|y_n - t_n\|^2 + 2\lambda_n k\|x_n - y_n\|\|t_n - y_n\| \\
&\le \|x_n - u\|^2 + (\lambda_n^2 k^2 - 1)\|y_n - t_n\|^2.
\end{aligned}$$

Then the following relationships hold:

$$\begin{aligned}
\|z_n - u\|^2 &\le \gamma_n\|x_n - u\|^2 + (1 - \gamma_n)\|t_n - u\|^2 \\
&\le \gamma_n\|x_n - u\|^2 + (1 - \gamma_n)(\|x_n - u\|^2 + (\lambda_n^2 k^2 - 1)\|y_n - t_n\|^2) \\
&= \|x_n - u\|^2 + (1 - \gamma_n)(\lambda_n^2 k^2 - 1)\|y_n - t_n\|^2 \\
&\le \|x_n - u\|^2. \qquad (80)
\end{aligned}$$

Rearranging the terms in (80), we arrive at

$$\begin{aligned}\|t_n - y_n\|^2 &\le \frac{1}{(1-\gamma_n)(1-\lambda_n^2 k^2)}(\|x_n - u\|^2 - \|z_n - u\|^2)\\ &\le \frac{1}{(1-\gamma_n)(1-\lambda_n^2 k^2)}(\|x_n - u\| + \|z_n - u\|)\|x_n - z_n\|.\end{aligned}$$

Since $\lim_{n\to\infty}\|x_n - z_n\| = 0$ and the sequences $\{x_n\}$ and $\{z_n\}$ are bounded, we get $\lim_{n\to\infty}\|t_n - y_n\| = 0$. Finally, it follows from the k-Lipschitz continuity of the operator A that $\lim_{n\to\infty}\|Ay_n - At_n\| = 0$, which completes the proof of the assertions in Step 4.

Step 5: *Optimality of weak cluster points of* $\{x_n\}$. Since the sequence $\{x_n\}$ is bounded, it has weak cluster points. Let us prove that all of them lie in $VO(\Omega, F) \cap VI(\Omega, A)$. Pick a weak cluster point $\hat{x}$ of $\{x_n\}$ and let $\{x_{k_n}\}$ be a subsequence weakly convergent to it. We are going to justify that $\hat{x} \in VO(\Omega, F) \cap VI(\Omega, A)$. First, similarly to Step 5 in the proof of Theorem 3.1, we have $\hat{x} \in VI(\Omega, A)$.

Second, we define $\psi_z : X \to \mathcal{R}$ by $\psi_z(x) := \langle F(x), z\rangle$ and claim that

$$\psi_z(\hat{x}) \le \psi_z(x_n) \tag{81}$$

for all $z \in C^+$ and all $n \ge 0$. Indeed, since the cost mapping F is positively lower semicontinuous and C-convex, the function ψ_z is lower semicontinuous and convex, and so we get $\psi_z(\hat{x}) \le \lim_{n\to\infty}\psi_z(x_{k_n})$. Since $F(x_{n+1}) \preceq_C F(x_n)$ for all n, we have $\psi_z(x_{n+1}) \le \psi_z(x_n)$ for all n, and thus $\lim_{n\to\infty}\psi_z(x_{k_n}) = \inf\{\psi_z(x_n)\}$. This shows that $\psi_z(\hat{x}) \le \inf\{\psi_z(x_n)\}$, which gives (81). It easily follows from (81) that

$$F(\hat{x}) \preceq_C F(x_n), \quad \text{for all } n \ge 0. \tag{82}$$

Suppose now that $\hat{x}$ is *not* weakly efficient for VOP, i.e., there exists $\bar{x} \in \Omega$ such that $F(\bar{x}) \prec_C F(\hat{x})$. Then it follows from (82) that $F(\bar{x}) \prec_C F(\hat{x}) \preceq_C F(x_n)$ for all $n \ge 0$. Choose further $\hbar_n$ as before (62). Since $\|\hbar_n\| = 1$ for all n, by the Bourbaki-Alaoglu theorem there is a weak* cluster point of $\{\hbar_{k_n}\}$, say $\hbar$, which is a weak* limit of the subnet $\{\hbar_{j_n}\}$ of $\{\hbar_{k_n}\}$. We claim that the positive polar cone C^+ is weak* closed in Y^*. To proceed, observe that the latter set admits the representation

$$C^+ = \bigcap_{y\in C}\big\{z \in Y^* \big|\ \langle y, z\rangle \ge 0\big\}.$$

Since the linear forms $z \mapsto \langle y, z\rangle$ are weak* continuous for all $y \in Y$, we have that C^+ as an intersection of weak* closed sets that justifies its weak* closedness in Y^*. It follows so that $\hbar \in C^+$. Thus we have

$$\begin{aligned}\langle F(\bar{x}) - F(\hat{x}), \hbar_{j_n}\rangle &\ge \langle F(\bar{x}) - F(x_{j_n+1}), \hbar_{j_n}\rangle\\ &= \psi_{j_n}(\bar{x}) - \psi_{j_n}(x_{j_n+1})\\ &\ge \langle v_{j_n}, \bar{x} - x_{j_n+1}\rangle\end{aligned}$$

$$= \langle u_{j_n}, \bar{x} - x_{j_n+1}\rangle - \nu_{j_n}\langle \nabla h(x_{j_n+1}) - \nabla h(x_{j_n}) - \theta_{j_n}, \bar{x} - x_{j_n+1}\rangle \geq -\nu_{j_n}\langle \nabla h(x_{j_n+1}) - \nabla h(x_{j_n}) - \theta_{j_n}, \bar{x} - x_{j_n+1}\rangle. \quad (83)$$

Note that $\|\theta_n\| \leq \sqrt{2}\sigma_n D_h^{1/2}(x_{n+1}, x_n) \to 0$ by (60). Therefore, utilizing the uniform continuity of ∇h, we deduce that

$$|\langle \nabla h(x_{j_n+1}) - \nabla h(x_{j_n}) - \theta_{j_n}, \bar{x} - x_{j_n+1}\rangle| \leq \|\nabla h(x_{j_n+1}) - \nabla h(x_{j_n}) - \theta_{j_n}\| \|\bar{x} - x_{j_n+1}\| \to 0,$$

due to $\|x_{n+1} - x_n\| \to 0$ and the boundedness of $\{\|\bar{x} - x_n\|\}$.

Let us next pass to the limit in the first and last expressions of (83). Regarding the rightmost one in (83), we get that $\{\nu_n\}$ is bounded, since $\{\alpha_n\}$ is bounded and since $\|\hbar_n\| = \|e_n\| = 1$. Thus the limit of the rightmost expression in (83) as $n \to \infty$ vanishes, and so we easily get from (83) that

$$\langle F(\bar{x}) - F(\hat{x}), \hbar\rangle \geq 0. \quad (84)$$

Now we show that $\hbar \neq 0$. Indeed, take $e \in \operatorname{int} C$ and deduce from [29, Lemma 2.2] that $\langle e, \hbar_n\rangle \geq d(e, Y \setminus C) > 0$ for all $n \geq 0$. Since $\hbar$ is the weak* limit of $\{\hbar_{j_n}\}$, we get that $\langle e, \hbar\rangle > 0$, which implies that $\hbar \neq 0$. The latter clearly yields that (84) contradicts the inclusion $\hbar \in C^+$ and hence also the assumption of $F(\bar{x}) \prec_C F(\hat{x})$. Thus this assumption is false, and so the point $\hat{x}$ is indeed weakly efficient for VOP, which justifies the claim of Step 5.

Step 6: *Uniqueness of the weak cluster point of* $\{x_n\}$. This part of the proof is closely related to the scalar-valued case, as given in [21], and it uses Brézis's uniqueness argument. Taking two cluster points $\hat{x}$ and $\tilde{x}$ of the sequence $\{x_n\}$, we conclude by the same arguments as in Step 5 above that both $\hat{x}$ and $\tilde{x}$ belong to $VO(\Omega, F) \cap VI(\Omega, A)$. It implies, as in Step 6 of Theorem 3.1, that $\tilde{x} = \hat{x}$. Thus the sequence of iterates $\{x_n\}$ weakly converges to an element of $VO(\Omega, F) \cap VI(\Omega, A)$, which completes the proof of the theorem. □

Remark 4.4 (Scalar Counterparts) Let us mention a scalar version of proximal point method developed in [21, Chap. 3] by using Bregman distances. This method is different from our Algorithm 2 even for scalar minimization problems. A more closely related version of Algorithm 2 in the scalar case is developed in [7, Chap. 3].

5 Relative Hybrid Approximate Proximal Algorithm

In the concluding section of the paper we present the *relative version* of our hybrid approximate proximal method, which is called Algorithm 3. It requires several *exogenous sequences*: in addition to those in Algorithm 2 (an error sequence $\{\theta_n\} \subset \Omega$, two bounded sequences of positive real numbers $\{\alpha_n\}$ and $\{\sigma_n\}$, a sequence $\{e_n\} \subset \operatorname{int}(C)$ such that $\|e_n\| = 1$ for all n), we now include a sequence $\{\hbar_n\} \subset C^+$ such that $\|\hbar_n\| = 1$ for all $n \geq 0$. The method generates a sequence of iterates $\{x_n\} \subset \Omega$ in the following way:

Algorithm 3

Initialization: Choose $x_0 \in \Omega \cap \operatorname{dom} F$.
Stopping Rule: Given x_n, if $x_n \in C\text{-}\operatorname{ArgMin}_w\{F(x)|x \in \Omega\}$ $(= VO(\Omega, F))$, then we let $x_{n+p} := x_n$ for all $p \geq 1$.
Iterative Step: Given x_n, if $x_n \notin C\text{-}\operatorname{ArgMin}_w\{F(x)|x \in \Omega\}$, we first compute

$$\begin{aligned} y_n &= P_\Omega(x_n - \lambda_n A x_n), \\ z_n &= \gamma_n x_n + (1-\gamma_n) P_\Omega(x_n - \lambda_n A y_n), \end{aligned} \tag{85}$$

for every $n = 0, 1, \ldots,$ where $\{\lambda_n\} \subset (0,1)$ and $\{\gamma_n\} \subset [0,1]$. Also let $\Omega_n = \{x \in \Omega | F(x) \preceq_C F(x_n)\}$ and define $f_n(x) := \langle F(x), \hbar_n\rangle + I_{\Omega_n}(x)$. Take as the next iterate x_{n+1} any vector $x \in \Omega$ such that there exists $\varepsilon_n \in \mathcal{R}_+$ satisfying

$$0 \in \partial_{\varepsilon_n} f_n(x) + \alpha_n \langle e_n, \hbar_n\rangle (x - z_n - \theta_n), \tag{86}$$

$$\varepsilon_n \leq \sigma \frac{\alpha_n}{2} \langle e_n, \hbar_n\rangle \|x - z_n\|^2. \tag{87}$$

For this algorithm, instead of condition (14), we impose

$$\|\theta_n\| \leq \sigma_n \|x_{n+1} - z_n\|, \quad \text{with } \sum_{n=0}^{\infty} \sigma_n^2 < \infty \tag{88}$$

as the approximate criterion corresponding to recursion (86).

The following theorem establishes the well-posedness and convergence of Algorithm 3.

Theorem 5.1 (Well-Posedness and Convergence of the Relative Version of HAPM) *Let $F : \Omega \to Y \cup \{\infty_C\}$ be a proper, C-convex, and positively lower semicontinuous mapping with $\Omega \cap \operatorname{dom} F \neq \emptyset$, and let $A : \Omega \to X$ be a monotone and k-Lipschitz continuous mapping such that $VO(\Omega, F) \cap VI(\Omega, A) \neq \emptyset$. In addition to condition* (88) *and assumptions in* (A) *formulated in Sect.* 4, *suppose that the exogenous sequences in Algorithm* 3 *satisfy the following requirements*:

(i) $\{\lambda_n\} \subset [a, b]$ *for some* $a, b \in (0, 1/k)$;
(ii) $\{\gamma_n\} \subset [c, d]$ *for some* $c, d \in (0, 1)$.

Then the sequence of iterates $\{x_n\}$ generated by Algorithm 3 *is well defined and has the convergence properties*:

(I) $\{x_n\}$ *converges with respect to the weak topology of X to a weakly efficient solution of the vector optimization problem VOP*;
(II) $\{x_n\}$ *converges with respect to the weak topology of X to an element of the set $VO(\Omega, F) \cap VI(\Omega, A)$ provided that $x_n \notin C\text{-}\operatorname{ArgMin}_w\{F(x)|x \in \Omega\}$ for all $n \geq 0$.*

Proof Similarly to the proofs of Theorems 3.1 and 4.3, we split the proof of this theorem into the following steps.

Step 1: *For every* $u \in VO(\Omega, F) \cap VI(\Omega, A)$, *we have*

$$\|z_n - u\|^2 \leq \|x_n - u\|^2 + (1-\gamma_n)(\lambda_n^2 k^2 - 1)\|x_n - y_n\|^2 \quad \text{whenever } n \geq 0.$$

The proof of this claim is similar to the one in Step 1 of Theorem 3.1.

Step 2: *Existence of the iterates*. By the subdifferential definition given in (8) we have

$$\partial f(x) := \partial_0 f(x) \subset \partial_\varepsilon f(x),$$

for any convex function $f : X \to \mathcal{R} \cup \{\infty\}$, any $x \in X$, and any $\varepsilon \in \mathcal{R}_+$. By the classical Minty theorem, the strongly convex function

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

admits a zero subgradient at some point x_{n+1}. It is obvious that such a point x_{n+1} satisfies the inclusion in (86) with $\varepsilon_n = 0$ satisfying (87).

Step 3: *Fejér convergence to the set of lower bounds of the initial section*. For any $n \geq 0$ denote $\nu_n := \alpha_n\langle e_n, \hbar_n\rangle$ and conclude by (86) that

$$\nu_n(z_n - x_{n+1} + \theta_n) \in \partial_{\varepsilon_n} f_n(x_{n+1}). \tag{89}$$

Pick any $u \in VO(\Omega, F) \cap VI(\Omega, A)$ and get by (89) and the definition of ∂_{ε_n} that

$$-\varepsilon_n + \nu_n\langle z_n - x_{n+1} + \theta_n, u - x_{n+1}\rangle \leq f_n(u) - f_n(x_{n+1}). \tag{90}$$

Since $u \in VO(\Omega, F) \cap VI(\Omega, A) \subset \Omega_n$, we have $u \in \Omega_n$ so that

$$f_n(u) = \langle F(u), \hbar_n\rangle \tag{91}$$

and also, by $F(u) \preceq_C F(x_{n+1})$ and $\hbar_n \in C^+$, that

$$\langle F(u) - F(x_{n+1}), \hbar_n\rangle \leq 0. \tag{92}$$

It follows from the definitions of f_n and I_{Ω_n} that $\partial_{\varepsilon_n} f_n(x) = \emptyset$ for all $x \notin \Omega_n$. Thus $x_{n+1} \in \Omega_n$, which justifies the relationship

$$f_n(x_{n+1}) = \langle F(x_{n+1}), \hbar_n\rangle. \tag{93}$$

Combining (90)–(93), we arrive at the inequalities

$$-\varepsilon_n + \nu_n\langle z_n - x_{n+1} + \theta_n, u - x_{n+1}\rangle \leq \langle F(u) - F(x_{n+1}), \hbar_n\rangle \leq 0. \tag{94}$$

Further, it follows from (94) that

$$\begin{aligned} 0 &\leq \varepsilon_n + \nu_n\langle x_{n+1} - z_n - \theta_n, u - x_{n+1}\rangle \\ &= \varepsilon_n + \nu_n(\|z_n - u\|^2 - \|x_{n+1} - u\|^2 - \|z_n - x_{n+1}\|^2) + \nu_n\langle\theta_n, x_{n+1} - u\rangle. \end{aligned} \tag{95}$$

Now for $\sigma_n > 0$ we get from (88) that

$$\langle \theta_n, x_{n+1} - u\rangle \leq \frac{1}{4\sigma_n^2}\|\theta_n\|^2 + \sigma_n^2\|x_{n+1} - u\|^2 \leq \frac{1}{2}\|x_{n+1} - z_n\|^2 + 2\sigma_n^2\|x_{n+1} - u\|^2. \tag{96}$$

Combining (87), (95), and (96) allows us to deduce that

$$0 \leq \nu_n\left(\|z_n - u\|^2 - (1 - 2\sigma_n^2)\|x_{n+1} - u\|^2 - \frac{1-\sigma}{2}\|z_n - x_{n+1}\|^2\right). \tag{97}$$

Since $\sigma_n \to 0$, there exists $N_0 \geq 0$ such that for all $n \geq N_0$ we have $1 - 2\sigma_n^2 > 0$. Hence it follows from (97) and (24) that

$$\begin{aligned} \|x_{n+1} - u\|^2 &\leq \left(1 + \frac{2\sigma_n^2}{1 - 2\sigma_n^2}\right)\|z_n - u\|^2 - \frac{1-\sigma}{2(1 - 2\sigma_n^2)}\|z_n - x_{n+1}\|^2 \\ &\leq \left(1 + \frac{2\sigma_n^2}{1 - 2\sigma_n^2}\right)\|x_n - u\|^2 - \frac{1-\sigma}{2}\|z_n - x_{n+1}\|^2 \end{aligned} \tag{98}$$

for all $n \geq N_0$, which justifies the assertion of Step 3.

Step 4: *Boundedness of* $\{x_n\}$ *and proximity of consecutive iterates*. Similarly to the proof of Step 4 in Theorem 4.3, we can show that the limit $\lim_{n\to\infty}\|x_n - u\|$ exists and that

$$\lim_{n\to\infty}\|z_n - u\| = \lim_{n\to\infty}\|x_n - u\|.$$

Then Proposition 2.2 ensures that $\lim_{n\to\infty}\|x_n - t_n\| = 0$, and so $\lim_{n\to\infty}\|z_n - x_n\| = 0$ due to (85). The further arguments follow the lines in the proof of Step 4 of Theorem 4.3.

Steps 5–6: *Optimality of the weak cluster points and the existence of the weak limit of iterates.* These steps are similar to the proofs of the corresponding steps of Theorem 4.3, and thus they are omitted here. The proof of the theorem is complete. □

Acknowledgements We would like to thank anonymous referees for their valuable suggestions and remarks that allowed us to improve the original presentation.

References

1. Bonnel, H., Iusem, A.N., Svaiter, B.F.: Proximal methods in vector optimization. SIAM J. Optim. **15**(4), 953–970 (2005)
2. Fliege, J., Svaiter, B.F.: Steepest descent methods for multicriteria optimization. Math. Methods Oper. Res. **51**, 479–494 (2000)
3. Grana Drummond, L.M., Svaiter, B.F.: A steepest descent method for vector optimization. J. Comput. Appl. Math. **175**, 395–414 (2005)
4. Grana Drummond, L.M., Iusem, A.N.: A projected gradient method for vector optimization problems. Comput. Optim. Appl. **28**, 5–30 (2004)
5. Eckstein, J.: Approximate iterations in Bregman-function-based proximal algorithms. Math. Program. **83**, 113–123 (1998)
6. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. **14**, 877–898 (1976)

7. Burachik, R.S., Iusem, A.N., Svaiter, B.F.: Enlargement of monotone operators with applications to variational inequalities. Set-Valued Anal. **5**, 159–180 (1997)
8. Burachik, R.S., Scheimberg, S., Svaiter, B.F.: Robustness of the hybrid extragradient proximal point algorithm. J. Optim. Theory Appl. **111**, 117–136 (2001)
9. He, B.S.: Inexact implicit methods for monotone general variational inequalities. Math. Program. **86**, 199–217 (1999)
10. Jahn, J.: Theory of vector maximization: Various concepts of efficient solutions. In: Gal, T., Stewart, T.J., Hanne, T. (eds.): Multicriteria Decision Making. Advances in MCDM Models, Algorithms, Theory, and Applications, pp. 2-1–2-32. Kluwer, Boston (1999)
11. Solodov, M.V., Svaiter, B.F.: An inexact hybrid extragradient-proximal point algorithm using the enlargement of a maximal monotone operator. Set-Valued Anal. **7**, 323–345 (1999)
12. Burachik, R.S., Iusem, A.N.: Set-Valued Mappings and Enlargements of Monotone Operators. Optimization and Its Applications, vol. 8. Springer, Berlin (2008)
13. Chen, G., Teboulle, M.: A proximal-based decomposition method for convex minimization problems. Math. Program. **64**, 81–100 (1994)
14. Da Silva Silva, P.J., Eckstein, J., Humes, C.: Rescaling and stepsize selection in proximal methods using separable generalized distance. SIAM J. Optim. **12**, 238–261 (2001)
15. Solodov, M.V., Svaiter, B.F.: A hybrid projection-proximal point algorithm. J. Convex Anal. **6**, 59–70 (1999)
16. Solodov, M.V., Svaiter, B.F.: Forcing strong convergence of proximal point iterations in a Hilbert space. Math. Program. **87**, 189–202 (2000)
17. Solodov, M.V., Svaiter, B.F.: An inexact hybrid generalized proximal point algorithm and some new results on the theory of Bregman functions. Math. Oper. Res. **25**(2), 214–230 (2000)
18. Göpfert, A., Riahi, H., Tammer, C., Zălinescu, C.: Variational Methods in Partially Ordered Spaces. Springer, New York (2003)
19. Benker, H., Hamel, A., Tammer, C.: An algorithm for vectorial control approximation problems. In: Fandel, G., Gal, T. (eds.) Multiple Criteria Decision Making, pp. 3–12. Springer, Berlin (1997)
20. Kyono, M., Fukushima, M.: Nonlinear proximal decomposition method for convex programming. J. Optim. Theory Appl. **106**, 357–372 (2000)
21. Butnariu, D., Iusem, A.N.: Totally Convex Functions for Fixed Point Computation and Infinite Dimensional Optimization. Kluwer, Dordrecht (1997)
22. Ceng, L.C., Yao, J.C.: Approximate proximal methods in vector optimization. Eur. J. Oper. Res. **183**(1), 1–19 (2007)
23. Nadezhkina, N., Takahashi, W.: Strong convergence theorem by a hybrid method for nonexpansive mappings and Lipschitz continuous monotone mappings. SIAM J. Optim. **16**, 1230–1241 (2006)
24. Korpelevich, G.M.: The extragradient method for finding saddle points and other problems. Matecon **12**, 747–756 (1976)
25. Burachik, R.S., Lopes, J.O., Svaiter, B.F.: An outer approximation method for the variational inequality problem. SIAM J. Control Optim. **43**, 2071–2088 (2005)
26. Ceng, L.C., Cubiotti, P., Yao, J.C.: An implicit iterative scheme for monotone variational inequalities and fixed point problems. Nonlinear Anal. **69**(8), 2445–2457 (2008)
27. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications. Springer, Berlin (2006)
28. Bao, T.Q., Mordukhovich, B.S.: Relative Pareto minimizers for multiobjective problems: existence and optimality conditions. Math. Program. (2009), published online
29. Bolintineanu, S.: Approximate efficiency and scalar stationarity in unbounded nonsmooth convex vector optimization problems. J. Optim. Theory Appl. **106**, 265–296 (2000)
30. Bolintineanu, S.: Vector variational principles, ε-efficiency and scalar stationarity. J. Convex Anal. **8**, 71–85 (2001)
31. Luc, D.T.: Theory of Vector Optimization. Lecture Notes in Econom. and Math. Syst., vol. 319. Springer, Berlin (1989)
32. Schu, J.: Weak and strong convergence to fixed points of asymptotically nonexpansive mappings. Bull. Austral. Math. Soc. **43**, 153–159 (1991)
33. Rockafellar, R.T.: On the maximality of sums of nonlinear monotone operators. Trans. Am. Math. Soc. **149**, 75–88 (1970)
34. Phelps, R.R.: Convex Functions, Monotone Operators and Differentiability. Lecture Notes in Math., vol. 1364. Springer, Berlin (1988)