1-1-2009

Hybrid Approximate Proximal Method with Auxiliary Variational Inequality for Vector Optimization

L C. Ceng

Shanghai Normal University, China, zenglc@hotmail.com

Boris S. Mordukhovich

Wayne State University, boris@math.wayne.edu

Jen-Chih Yao

National Sun Yat-sen University, Kaohsiung, Taiwan, yaojc@math.nsysu.edu.tw

Recommended Citation

Ceng, L C.; Mordukhovich, Boris S.; and Yao, Jen-Chih, "Hybrid Approximate Proximal Method with Auxiliary Variational Inequality for Vector Optimization" (2009). *Mathematics Research Reports*. Paper 63.
http://digitalcommons.wayne.edu/math_reports/63

This Technical Report is brought to you for free and open access by the Mathematics at DigitalCommons@WayneState. It has been accepted for inclusion in Mathematics Research Reports by an authorized administrator of DigitalCommons@WayneState.

**HYBRID APPROXIMATE PROXIMAL METHOD WITH
AUXILIARY VARIATIONAL INEQUALITY FOR
VECTOR OPTIMIZATION**

L. C. CENG, B. S. MORDUKHOVICH and J. C. YAO

**WAYNE STATE
UNIVERSITY**

Detroit, MI 48202

**Department of Mathematics
Research Report**

**2009 Series
#1**

This research was partly supported by the National Science Foundation

Hybrid Approximate Proximal Method with Auxiliary Variational Inequality for Vector Optimization

L. C. Ceng¹, B. S. Mordukhovich² and J. C. Yao^{3,4}

Communicated by F. Giannessi

¹ Professor, Department of Mathematics, Shanghai Normal University, Shanghai 200234, China. This research was partially supported by the National Science Foundation of China (10771141), Ph.D. Program Foundation of Ministry of Education of China (20070270004), and Science and Technology Commission of Shanghai Municipality grant (075105118). Email: zenglc@hotmail.com

² Distinguished University Professor, Department of Mathematics, Wayne State University, Detroit, MI 48202, USA. This research was partially supported by the USA National Science Foundation under grant DMS-0603896. Email: boris@math.wayne.edu

³ Professor, Department of Applied Mathematics, National Sun Yat-sen University, 804 Kaohsiung, Taiwan. This research was partially supported by the grant NSC 97-2115M-110-001. Email: yaojc@math.nsysu.edu.tw

⁴ Corresponding author.

Abstract. This paper studies the general vector optimization problem of finding weakly efficient points for mappings in a Banach space Y , with respect to the partial order induced by a closed, convex, and pointed cone $C \subset Y$ with nonempty interior. In order to find a solution of this problem, we introduce an auxiliary variational inequality problem for monotone, 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 by the combination of extragradient method for finding a solution to the variational inequality problem and approximate proximal point method for 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, Lipschitz-continuous mapping, and finding weakly efficient points for suitable regularizations of the original mapping. We present both an absolute and a relative version in which the subproblems are solved only approximately. Weak convergence of the generated sequence to a weak efficient point is established under quite mild conditions. In addition, we also discuss an extension to Bregman-function-based hybrid approximate proximal algorithms for finding weakly efficient points for mappings.

Key Words: Vector optimization; Proximal point; Hybrid inexact algorithm; Auxiliary variational inequality; Banach space.

1. Introduction and Discussion

Recently, Bonnel, Iusem and Svaiter [1] introduced and studied the extension to vector-valued optimization of several iterative methods for scalar-valued methods. In those extensions, they defined the iterates in the vector-valued case by considering the order $\preceq_C$ in a real Banach space Y , mimicking, whenever it is possible, a role of the usual order in $\mathcal{R}$ (the set of real numbers) in the corresponding algorithm for scalar-valued optimization. Meantime, they admitted the possibility that $F : X \rightarrow Y$ takes the value ∞_C (this is made precise in Section 2), where X is a Hilbert space and C is a closed, convex, and pointed cone in Y with $\text{int}C \neq \emptyset$, where $\text{int}C$ denotes the interior of the set C . Such extensions can be traced back to the fashion of extension which existed in a finite-dimensional setting. For example, in $\mathcal{R}^n$, see 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 Ω be a nonempty closed convex subset of X and $F : \Omega \rightarrow Y \cup \{\infty_C\}$. 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 its associate relation $\prec_C$, given by $y \prec_C y'$ if and only if $y' - y \in \text{int}C$. In this paper, our goal is to analyze methods for finding a *weakly efficient minimizer* of F with respect to $\preceq_C$, meaning a point $a \in \Omega$ such that there exists no $x \in \Omega$ satisfying $F(x) \prec_C F(a)$.

In [1], Bonnel, Iusem and Svaiter actually performed a similar extension for the case of the proximal point method for scalar-valued convex optimization. Let us give a brief description of this method. Given a Hilbert space X and a point-to-set (multivalued) operator $T : X \rightarrow 2^X$, the proximal point method, in its so-called exact version, is an iterative procedure for finding a zero of T , i.e., a point $x^* \in X$ such that $0 \in T(x^*)$. The method generates a sequence $\{x_n\} \subset X$, starting from an arbitrary $x_0 \in X$, through the following iteration: given a bounded exogenous sequence of positive real numbers $\{\alpha_n\}$ (called regularization parameters) and the current iterate x_n , the next iterate x_{n+1} is the unique vector in X such that $0 \in T_n(x_{n+1})$, where $T_n : X \rightarrow 2^X$ 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:

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

where $\{c_n\} \subset [c, \infty)$, $c > 0$, is a sequence of scalars. However, as pointed out in [5], the ideal form of 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 gave an inexact variant of the method:

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

where θ_{n+1} is regarded as an error sequence. This method is called an inexact proximal point algorithm. Rockafellar [6] proved in the setting of a finite-dimensional space $\mathcal{R}^n$ that if $\theta_n \rightarrow 0$

quickly enough such that $\sum_{n=1}^{\infty} \|\theta_n\| < \infty$, then $x_n \rightarrow z \in \mathcal{R}^n$ 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, the conditions ensuring that the error term being summable is an essential condition for the convergence of the method. In [6] and some sequel papers (e.g., [13]) a criterion for this is as follows:

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

In [5], Eckstein extended the method to Bregman-function-based inexact proximal methods and proved that the sequence $\{x_n\}$ generated by the algorithm converges to a root of T under the conditions

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

(see Eqs. (18) and (19) in [5]). Condition (4) is an assumption on the whole generated sequence $\{x_n\}$ and the error term sequence $\{\theta_n\}$, and thus seems to be slightly stronger, but it can be checked and enforced in practice more easily than those that existed earlier. Furthermore, da Silva e Silva et al. [14] and Solodov and Svaiter [15-17] very recently proposed some new accurate criteria for proximal point algorithms. Their criteria, rather than the imposed inequality (3), require only that $\sup_{n \geq 0} \sigma_n < 1$. On the other hand, He [9] gave another inexact criterion for the study of monotone general variational inequalities, which involves a relation between the error term and the residual function; in other words, the restriction $\sum_{n=0}^{\infty} \sigma_n < \infty$ in (3) is replaced by He's assumption $\sum_{n=0}^{\infty} \sigma_n^2 < \infty$. However, in [15-17] this comes at the cost of adding an additional projection or "extragradient" step to the algorithm, and the applicable portion of [14] is efficient only for convex minimization.

Now take a convex function $f : X \rightarrow \mathcal{R} \cup \{\infty\}$ and a closed and convex subset $K \subset X$. Then solutions to the convex optimization problem $\min f(x)$ subject to $x \in K$ are precisely the zeros of the maximal monotone operator $T = \partial(f + I_K)$, where ∂ denotes the subdifferential of a convex function and I_K is the indicator function of K defined as

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

Thus the proximal point method can be used to solve the above-mentioned optimization problem; in this case, it is easy to see that the iteration (1) has the form

$$x_{n+1} \in \operatorname{argmin}_{x \in K} \left\{ f(x) + \frac{\alpha_n}{2} \|x - x_n\|^2 \right\}. \quad (5)$$

Next we recall inexact version of the method, where x_{n+1} need not be the exact solution to the optimization subproblem in (2) but just an approximate solution of it. The following iteration was considered in [1] for the problem

$$\min f(x) \quad \text{subject to } x \in X,$$

with a convex function $f : X \rightarrow \mathcal{R} \cup \{\infty\}$: given x_n , take as x_{n+1} any vector $x \in X$ such that there exists $\varepsilon_n \in \mathcal{R}_+$ satisfying

$$0 \in \partial_{\varepsilon_n} f(x) + \alpha_n(x - x_n), \quad (6)$$

$$\varepsilon_n \leq \sigma \frac{\alpha_n}{2} \|x - x_n\|^2, \quad (7)$$

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

Recall that, for $\varepsilon \geq 0$,

$$\partial_\varepsilon f(x) := \{u \in X \mid f(y) - f(x) - \langle u, x - y \rangle \geq -\varepsilon, \forall y \in X\}. \quad (8)$$

The conditions in (6) and (7) simply mean that $\alpha_n(x_n - x_{n+1})$ lies in $\partial_{\varepsilon_n} f(x_{n+1})$, with the definition of ∂_{ε_n} given above and with ε_n satisfying inequality (7) for $x = x_{n+1}$.

The algorithm given in (6) and (7) reduces to the exact version for the optimization case as in (5) when $\sigma = 0$ (or, equivalently, $\varepsilon_n = 0$ for all n).

This algorithm was developed in [8, 11] for the more general problem of finding zeros of operators. The exact version of this method is the one presented in (1) where instead of $\partial_{\varepsilon_n} f_n$ with $f_n : x \mapsto f(x) + \frac{\alpha_n}{2} \|x - x_n\|^2$, an adequate enlargement of the operator T_n , related to ε_n , is used. This type of enlargement was introduced in [7]. The results in [8, 11] establish that the sequence generated by (6) and (7) converges to a minimizer of f in the weak topology of X under the same assumptions required in the exact case, namely: convexity of f and existence of solutions to the optimization problem.

Further, Bonnel, Iusem and Svaiter [1] considered extensions of both the exact proximal method (2) and its inexact counterpart (6) and (7) to the vector-valued optimization problem introduced at the beginning of this section. Basically, in the exact case the n th subproblem consists of finding weakly efficient minimizers of $F_n : X \rightarrow Y$ with

$$F_n(x) = F(x) + \alpha_n \|x - x_n\|^2 e_n, \quad (9)$$

restricted to the set $\Omega_n \subset X$ defined as $\Omega_n := \{x \in X : F(x) \preceq_C F(x_n)\}$, where e_n is an exogenously selected vector belonging to $\text{int}C$ and such that $\|e_n\| = 1$. On the other hand, for the inexact version they considered the topological dual space Y^* of Y , the positive polar cone $C^+ \subset Y^*$, given by $C^+ := \{z \in Y^* : \langle y, z \rangle \geq 0 \text{ for all } y \in C\}$, where $\langle \cdot, \cdot \rangle : Y \times Y^* \rightarrow \mathcal{R}$ is the duality pairing and the indicator function I_{Ω_n} of the set Ω_n defined as above. We take an exogenous sequence $\{\hbar_n\} \subset C^+$, with $\|\hbar_n\| = 1$ for all $n \geq 0$ and define, at iteration n , a function $f_n : X \rightarrow \mathcal{R} \cup \{\infty\}$ by

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

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

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

$$\varepsilon_n \leq \sigma \frac{\alpha_n}{2} \langle e_n, \hbar_n \rangle \|x_n - x\|^2, \quad (12)$$

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

In [1], Bonnel, Iusem and Svaiter established that any sequence 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 $F(a_{n+1}) \preceq_C F(a_n)$ for all $n \geq 0$ there exists $a \in X$ such that $F(a) \preceq_C F(a_n)$ for all $n \geq 0$.

In the absence of assumption (ii), they established weaker convergence results, namely, that the generated sequence is a minimizing one for the above vector-valued optimization problem.

Note that the vectorial proximal method is also discussed in section 4.2 of [18]. It is a generalization of algorithms for specific instances of a vector optimization problem (VOP): a particular control approximation problem in [19] and certain location problems in [20]. In the presentation given in [18], it deals with a problem more general than VOP, namely, the vector equilibrium problem (VEP). It can be seen that solutions to the scalarized equilibrium problem for a real bifunction f defined on $M \times M$, where M is a closed and convex subset of $X = \mathcal{R}^m$ (namely, the points $\bar{x} \in M$ such that $f(\bar{x}, x) \geq 0$ for all $x \in M$), are solutions to VEP. The authors proposed a scalar proximal method for this equilibrium problem. We refrain from making explicit the iterative formula of the method because, in the case of VOP, it ends up being just the standard scalar proximal point method, as given in (5), applied to the scalar function $\langle F(x), \bar{h}_n \rangle$, ($\bar{h}_n \in C^+$, $\|\bar{h}_n\| = 1$), except for the fact that a more general regularization term is used: the quadratic function $\|x - x_n\|^2$ of (5) is replaced by the Bregman distance $D_g(x, x_n)$ between x and x_n , induced by a Bregman function $g : X \rightarrow \mathcal{R} \cup \{\infty\}$ (see, e.g., [21] for the definition of Bregman functions and distances). The convergence analysis in [18] is also restricted to the finite-dimensional case. The fact that the method in [18] is essentially a scalar proximal method establishes a basic difference from that in [1]. Also, there is no doubt that the inexact version given in [1] is more suitable for numerical implementation.

Very recently, motivated by Bonnel, Iusem and Svaiter [1], Ceng and Yao [22] introduced and studied both the absolute approximate proximal method and the relative approximate proximal method for solving the vector-valued optimization problem introduced at the beginning of this section. Let $\{\alpha_n\}$ be a bounded sequence of positive real numbers. Basically, in the absolute case the n th subproblem consists of first finding weakly efficient minimizer $\tilde{x}_n$ of $\tilde{F}_n : X \rightarrow Y$ with

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

restricted to the set $\Omega_n \subset X$ defined as $\Omega_n = \{x \in X | F(x) \preceq_C F(x_n)\}$, 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

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

and e_n is an exogenously selected vector belonging to $\text{int}C$ and such that $\|e_n\| = 1$. On the other hand, for the relative version, they considered the topological dual space Y^* of Y , the positive polar cone $C^+ \subset Y^*$ and the indicator function I_{Ω_n} of the set Ω_n defined as above. We take an exogenous sequence $\{\tilde{h}_n\} \subset C^+$, with $\|\tilde{h}_n\| = 1$ for all $n \geq 0$ and define, at iteration n , the function $f_n : X \rightarrow \mathcal{R} \cup \{\infty\}$ as

$$f_n(x) := \langle F(x), \tilde{h}_n \rangle + I_{\Omega_n}(x). \quad (15)$$

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

$$0 \in \partial_{\varepsilon_n} f_n(x) + \alpha_n \langle e_n, \tilde{h}_n \rangle (x - x_n - \theta_n), \quad (16)$$

$$\varepsilon_n \leq \sigma \frac{\alpha_n}{2} \langle e_n, \tilde{h}_n \rangle \|x_n + \theta_n - x\|^2, \quad (17)$$

where $\{\theta_n\}$ is an error sequence in X satisfying condition (14), and $\sigma \in [0, 1]$ is again a measure of the relative error. They proved that any sequence generated by either their absolute or relative approximate proximal method converges in the weak topology of X to a weakly efficient minimizer of F under Bonnel, Iusem and Svaiter's two assumptions above.

We remind the reader of the fact that the exact case of Bonnel, Iusem and Svaiter's proximal method is indeed a particular case of their absolute approximate proximal method corresponding to choosing $\theta_n = 0$ and $\beta_n = 0$ for all n , and that the inexact case of Bonnel, Iusem and Svaiter's proximal method is actually a particular case of their relative approximate proximal method, corresponding to choosing $\theta_n = 0$ and $\beta_n = 0$ for all n . Moreover, the absolute form of the algorithm is a particular case of the relative one corresponding to choosing $\sigma = 0$, or, equivalently $\varepsilon_n = 0$ for all n , in the sense that any vector x_{n+1} satisfying (15)-(17) with $\sigma = 0$ is a weakly efficient minimizer of $\tilde{F}_n$ as defined in (12). Thus, a separate analysis of the absolute version might seem superfluous. However, both versions are presented somewhat differently: the subproblems of the absolute one are (vector-valued) optimization problems, whereas in each subproblem of the relative version they looked for zeros of approximate subdifferentials of scalar-valued convex functions.

On the other hand, let Ω be a nonempty closed convex subset of a real Hilbert space X and let P_Ω be the metric projection from X onto Ω . When $\{x_n\}$ is a sequence in X , then $x_n \rightarrow x$ (resp. $x_n \rightharpoonup x$) will denote strong (resp. weak) convergence of the sequence $\{x_n\}$ to x . Let A be a mapping from Ω into H . As usual, A is called *monotone* if

$$\langle Ax - Ay, x - y \rangle \geq 0, \quad \forall x, y \in \Omega.$$

A is called *α -inverse-strongly monotone* (see [23]) if there exists a positive constant α such that

$$\langle Ax - Ay, x - y \rangle \geq \alpha \|Ax - Ay\|^2, \quad \forall x, y \in \Omega.$$

A is called *k-Lipschitz-continuous* if there exists a positive constant k such that

$$\|Ax - Ay\| \leq k\|x - y\|, \quad \forall x, y \in \Omega.$$

It is easy to see that every α -inverse-strongly monotone mapping A under consideration is monotone and Lipschitz-continuous.

Let A be a mapping from Ω into X . The variational inequality problem is to find a $x \in \Omega$ such that

$$\langle Ax, y - x \rangle \geq 0, \quad \forall y \in \Omega.$$

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

In 1976, for finding a solution to the nonconstrained variational inequality problem in the finite-dimensional Euclidean space $\mathcal{R}^n$ under the assumption that a set $\Omega \subset \mathcal{R}^n$ is closed and convex and a mapping A of Ω into $\mathcal{R}^n$ is monotone and k -Lipschitz-continuous, Korpelevich [24] introduced the following so-called extragradient method:

$$\begin{cases} \text{choose } x_0 \in \Omega \text{ arbitrarily,} \\ \bar{x}_n = P(x_n - \lambda Ax_n), \\ x_{n+1} = P_C(x_n - \lambda A\bar{x}_n) \end{cases}$$

for all $n \geq 0$, where $\lambda \in (0, 1/k)$. She proved that if $VI(\Omega, A)$ is nonempty, then the sequence $\{x_n\}$ generated by the iterative scheme converges to an element of $VI(\Omega, A)$.

Very recently, utilizing a combination of the hybrid-type method [25] and the extragradient-type method, Nadezhkina and Takahashi [23] introduced an iterative process for finding a common element of the fixed-point set of a nonexpansive self-mapping of Ω and the set of solutions to the variational inequality problem for monotone, k -Lipschitz-continuous mapping. Subsequently, motivated by Nadezhkina and Takahashi [23], Ceng and Yao [26] introduced another iterative process for finding a common element of the common fixed-point set of N nonexpansive self-mappings of Ω and the set of solutions to the variational inequality problem for monotone, k -Lipschitz-continuous mapping by the combination of extragradient and approximate proximal methods.

In this paper inspired by Bonnel, Iusem and Svaiter [1], Nadezhkina and Takahashi [23], and Ceng and Yao [22, 26], we introduce and study both the absolute hybrid approximate proximal method and the relative hybrid approximate proximal method for solving the general vector-valued optimization problem considered at the beginning of this section. Let $\{\alpha_n\}$ be a bounded sequence of positive real numbers. Basically, in the absolute case the n th subproblem consists of first finding approximate solution z_n of the variational inequality problem for monotone, k -Lipschitz-continuous mapping A via

$$\begin{cases} y_n = P_\Omega(x_n - \lambda_n Ax_n), \\ z_n = \gamma_n x_n + (1 - \gamma_n)P_\Omega(x_n - \lambda_n Ay_n), \end{cases}$$

where $\{\lambda_n\} \subset (0, 1/k)$ and $\{\gamma_n\} \subset (0, 1)$, and then finding weakly efficient minimizer $\tilde{x}_n$ of $\tilde{F}_n : \Omega \rightarrow Y$ with

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

restricted to the set $\Omega_n \subset \Omega$ defined as $\Omega_n := \{x \in \Omega | F(x) \preceq_C F(x_n)\}$, and further computing the $(n+1)$ th iterate

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

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

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

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

On the other hand, for the relative version we consider the topological dual space Y^* of Y , the positive polar cone $C^+ \subset Y^*$, and the indicator function I_{Ω_n} of the set Ω_n , defined as above. We first find approximate solution z_n to the variational inequality problem for monotone, k -Lipschitz-continuous mapping A via

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

where $\{\lambda_n\} \subset (0, 1/k)$ and $\{\gamma_n\} \subset (0, 1)$, and then take an exogenous sequence $\{\hbar_n\} \subset C^+$, with $\|\hbar_n\| = 1$ for all $n \geq 0$, and define, at iteration n , the function $f_n : \Omega \rightarrow \mathcal{R} \cup \{\infty\}$ as

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

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

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

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

where $\{\theta_n\}$ is an error sequence in Ω satisfying the similar condition to (14), and where $\sigma \in [0, 1)$ is again a measure of the relative error.

It is shown in what follows that any sequence generated by either our absolute or relative hybrid approximate proximal method converges in the weak topology of X to a weakly efficient minimizer of F under quite mild assumptions.

The prototypical infinite-dimensional Banach spaces are the $\mathcal{L}^p$ spaces ($1 \leq p \leq \infty$), and the most relevant cones in them are the so-called positive cones consisting of all p -integrable functions, which are nonnegative almost everywhere. It is well known that these cones have empty interior, except for the case of $\mathcal{L}^\infty$ which is nonseparable and nonreflexive. The latter as well as as the space $C(K)$ (where K is a compact set in, e.g., R^n) provide good meaningful examples for us, where the order is induced by a cone in an infinite dimensional space. Observe that the approach to vector/multiobjective optimization problems developed by Mordukhovich [27] and based mainly on the *extremal principle* is applied in general multiobjective settings with possibly empty interior of the ordering cone C . Furthermore, the very recent paper by Bao and Mordukhovich [28] studies certain notions of *relaxed Pareto minimizers*, which is somehow close in spirit to weak minimizers while do not require 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 the relative Pareto notions studied in [28] from the viewpoints of existence theorems and necessary optimality conditions.

The paper is organized as follows. In Section 2 we formulate the problem and present some required preliminary material. The absolute version of the algorithm is analyzed in Section 3. Section 4 discusses an extension to Bregman-function-based hybrid approximate proximal algorithms, and Section 5 develops the relative version. We also adopt some notations in [1, 23].

2. Problems and Formulations

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

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

Observe that the embedding $Y \subset \bar{Y}$ 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 $\text{dom} F := \{x \in X \mid F(x) \neq \infty_C\}$. 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 $\bar{Y}$. For a set $U \subset \bar{Y}$, we denote its topological closure in the topological space $\bar{Y}$ by $\bar{U}$. We associate with a given set $U \subset \bar{Y}$ the following sets:

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

$$C\text{-INF}_p(U) := \{y \in \bar{U} \mid \exists K \subset Y \text{ pointed closed convex cone such that}$$

$$C \setminus \{0\} \subset \text{int} K, \ y \in K - \text{INF}(U)\}.$$

For a vector optimization problem

$$C\text{-MINIMIZE } G(x) \text{ subject to } x \in S,$$

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

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

Thus the sets of efficient (resp., weakly efficient or properly efficient) solutions, which are denoted by $C\text{-ARGMIN}\{G(x)|x \in S\}$ (resp., $C\text{-ARGMIN}_w\{G(x)|x \in S\}$ or $C\text{-ARGMIN}_p\{G(x)|x \in S\}$), 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 $d(y, U) := \inf\{\|y - z\| | z \in U \cap Y\}$.

Consider the following VOP:

$$C\text{-MINIMIZE}\{F(x)|x \in \Omega\},$$

where Ω is a nonempty, closed, and convex subset of X . The set of weakly efficient solutions of the VOP is denoted by $VO(\Omega, F)$. Let P_Ω be the *metric projection* from X onto Ω . As noted in [4], any constrained vector optimization problem (CVOP):

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

where $S \subset X$ is the feasible set and F_0 is a map from S to Y , is equivalent to the unconstrained extended-valued VOP:

$$C\text{-MINIMIZE}\{F(x)|x \in X\} \text{ with } F(x) = \begin{cases} F_0(x) & \text{if } x \in S, \\ \infty_C & \text{if } x \in X \setminus S. \end{cases}$$

The CVOP and VOP are equivalent in the sense that they have the same weakly efficient solutions and the same weakly infimal set. A map $G : X \rightarrow Y \cup \{\infty_C\}$ will be called *positively lower semicontinuous* if for each $z \in C^+$ the extended-valued scalar function $x \mapsto \langle G(x), z \rangle$ is lower semicontinuous.

Throughout this paper we consider a VOP, where the mapping F is C -convex. Such VOP is called a C -convex VOP.

Recall the following scalarization result (e.g., see [31]) known for finite-valued maps which can be easily generalized to extended-valued ones. Denote $C_s^+ := \{z \in Y^* | \langle y, z \rangle > 0 \text{ for all } y \in C \setminus \{0\}\}$.

Theorem 2.1. (c.f. [22]) If $S \subset X$ is a convex set and $G : S \rightarrow Y \cup \{\infty_C\}$ is a C -convex proper map, then

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

and

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

Remark 2.1. As pointed out in [4], the set $\operatorname{argmin}\{\langle G(x), z \rangle | x \in S\}$ in Theorem 2.1 may be empty for some $z \in C^+ \setminus \{0\}$.

Next we recall the following lemma, which is used in the sequel.

Lemma 2.1. (See [32]) Let X be a Hilbert space, let $\{\alpha_n\}$ be a sequence of real numbers such that $0 < a \leq \alpha_n \leq b < 1$ for every $n = 0, 1, 2, \dots$, and let $\{v_n\}$ and $\{w_n\}$ be sequences in X such that $\limsup_{n \rightarrow \infty} \|v_n\| \leq c$, $\limsup_{n \rightarrow \infty} \|w_n\| \leq c$, and $\lim_{n \rightarrow \infty} \|\alpha_n v_n + (1 - \alpha_n) w_n\| = c$, for some $c \geq 0$. Then $\lim_{n \rightarrow \infty} \|v_n - w_n\| = 0$.

Recall that if Ω is a nonempty, closed, 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 $\|x - P_\Omega x\| \leq \|x - y\|$ 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

$$\langle x - P_\Omega x, P_\Omega x - y \rangle \geq 0, \quad \forall x \in H, y \in \Omega; \quad (18)$$

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

$$\|x - y\|^2 \geq \|x - P_\Omega x\|^2 + \|y - P_\Omega x\|^2, \quad \forall x \in H, y \in \Omega. \quad (19)$$

Let A be a monotone mapping of Ω into X . In the context of the variational inequality problem under consideration the characterization of projection (18) implies

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

A mapping $T : \Omega \rightarrow \Omega$ is called *pseudocontractive* if for all $x, y \in \Omega$, we have

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

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

Recall that a set-valued mapping $T : X \rightarrow 2^X$ is monotone if for all $x, y \in X$, $f \in Tx$ and $g \in Ty$ imply $\langle x - y, f - g \rangle \geq 0$. A monotone mapping $T : X \rightarrow 2^X$ is maximal if its graph $\operatorname{Gr}(T)$ is not properly contained in the graph of any other monotone mapping. It is known that a monotone mapping T is *maximal* if and only if for $(x, f) \in H \times H$, $\langle x - y, f - g \rangle \geq 0$ for all $(y, g) \in \operatorname{Gr}(T)$ we have $f \in Tx$.

Let A be a monotone, k -Lipschitz-continuous mapping of Ω into H and let $N_\Omega v$ be the normal cone to Ω at $v \in \Omega$, i.e., $N_\Omega v = \{w \in X : \langle v - u, w \rangle \geq 0 \text{ for all } u \in \Omega\}$. Define

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

It is known that in this case T is maximal monotone, and $0 \in Tv$ if and only if $v \in VI(\Omega, A)$; see [33].

3. Absolute Hybrid Approximate Proximal Algorithm

For finding an element of $VO(\Omega, F)$, we first introduce the *absolute version of our hybrid approximate proximal algorithm*, which will be called Algorithm 1. It 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 real numbers $\{\alpha_n\}$ and $\{\sigma_n\}$, and a sequence $\{e_n\} \subset \text{int}C$ such that $\|e_n\| = 1$ for all n . Assume that $\Omega \cap \text{dom}(F) \neq \emptyset$. The method generates a sequence $\{x_n\} \subset \Omega$ in the following way:

Initialization: Choose $x_0 \in \Omega \cap \text{dom}(F)$.

Stopping rule: Given x_n , if $x_n \in C\text{-ARGMIN}_w\{F(x)|x \in \Omega\} = VO(\Omega, F)$, then let $x_{n+p} = x_n$ for all $p \geq 1$.

Iterative step: Given x_n , whenever $x_n \notin C\text{-ARGMIN}_w\{F(x)|x \in \Omega\} = VO(\Omega, F)$, we first compute

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

for every $n = 0, 1, 2, \dots$, where $\{\lambda_n\} \subset (0, 1)$ and $\{\gamma_n\} \subset [0, 1]$, and then takes as $\tilde{x}_n$ any vector u such that

$$u \in C\text{-ARGMIN}_w\{F(x) + \frac{\alpha_n}{2}\|x - z_n - \theta_n\|^2 e_n | x \in \Omega_n\}, \quad (21)$$

with $\Omega_n := \{x \in \Omega | F(x) \preceq_C F(x_n)\}$. Moreover, we further compute the next iterate

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

We make the following assumptions on the map F and the initial iterate x_0 :

(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 $n \geq 0$, it holds $F(u) \preceq_C F(a_n)$ for all $u \in VO(\Omega, F) \cap VI(\Omega, A)$ and all $n \geq 0$.

(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 each sequence $\{\tilde{h}_n\} \subset C^+$ converging weakly to some $\tilde{h} \in C^+$, we have for any sequence $\{y_n\} \subset \Omega$ that

$$\|x_n - y_n\| \rightarrow 0 \Rightarrow |\langle F(x_n) - F(y_n), \tilde{h}_n \rangle - \langle F(\hat{x}) - F(y_n), \tilde{h} \rangle| \rightarrow 0.$$

Now we prove the convergence of Algorithm 1 under condition (14) and assumptions (A) and (B).

Theorem 3.1. Let $F : X \rightarrow Y \cup \{\infty_C\}$ be a proper, C -convex, and positively lower semicontinuous map with $\Omega \cap \text{dom}(F) \neq \emptyset$. Let $A : \Omega \rightarrow X$ be a monotone and k -Lipschitz-continuous mapping such that $VO(\Omega, F) \cap VI(\Omega, A) \neq \emptyset$. Let $\{x_n\}$ be any sequence generated by Algorithm 1. Assume that condition (14), assumptions (A) and (B), and that the following conditions are satisfied:

- (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 the following hold:

(I) $\{x_n\}$ converges with respect to the weak topology of X to a weakly efficient solution of the VOP;

(II) $\{x_n\}$ converges with respect to the weak topology of X to an element of $VO(\Omega, F) \cap VI(\Omega, A)$ provided $x_n \notin C\text{-ARGMIN}_w\{F(x)|x \in \Omega\}$, $\forall n \geq 0$.

Proof. We divide the proof into several steps.

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

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

Indeed, put $t_n = P_\Omega(x_n - \lambda_n A y_n)$ for every $n = 0, 1, 2, \dots$. From (19), monotonicity of A , and $u \in VI(\Omega, A)$ we have

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

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

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

Thus we have

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

Therefore, from (23) and $z_n = \gamma_n x_n + (1 - \gamma_n) t_n$, we have

$$\begin{aligned} \|z_n - u\|^2 &= \|\gamma_n x_n + (1 - \gamma_n) t_n - u\|^2 \\ &= \|\gamma_n (x_n - u) + (1 - \gamma_n) (t_n - u)\|^2 \\ &\leq \gamma_n \|x_n - u\|^2 + (1 - \gamma_n) \|t_n - u\|^2 \\ &\leq \gamma_n \|x_n - u\|^2 + (1 - \gamma_n) (\|x_n - u\|^2 + (\lambda_n^2 k^2 - 1) \|x_n - y_n\|^2) \\ &= \|x_n - u\|^2 + (1 - \gamma_n) (\lambda_n^2 k^2 - 1) \|x_n - y_n\|^2 \\ &\leq \|x_n - u\|^2 \end{aligned} \tag{24}$$

for every $n = 0, 1, 2, \dots$

Step 2. Existence of iterates. Indeed, x_0 is chosen in the initialization step. Assuming that the algorithm reaches iteration n , we show next that an approximate x_{n+1} exists. By the stopping rule this is certainly the case if $x_n \in VO(\Omega, F)$. Otherwise, take any $z \in C^+ \setminus \{0\}$. Since $e_n \in \text{int}C$, it follows from the definition of C^+ that $\langle e_n, z \rangle > 0$. Define $\varphi_n : X \rightarrow \mathcal{R} \cup \{\infty\}$ by

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

C -convexity of F implies convexity of $\langle F(\cdot), z \rangle$ and convexity of Ω_n . Lower semicontinuity of F implies closedness of Ω_n ; thus $\langle F(\cdot), z \rangle + I_{\Omega_n}$ is convex and lower semicontinuous. Since $\langle e_n, z \rangle > 0$, φ_n is strongly convex, and the existence of minimizers of φ_n results from standard arguments for existence of iterates of the scalar-valued proximal method (c.f. [20]): the subdifferential of φ_n is maximal monotone and strongly monotone, and hence onto by *Minty's theorem*. Thus the subdifferential has some zero, which is a minimizer of φ_n . By Theorem 2.1 such a minimizer satisfies (21) and can be taken as $\tilde{x}_n$. Thus, via (22), we can compute an approximate x_{n+1} .

Step 3. Fejér convergence to the set of lower bounds of the initial section. Indeed, 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. Therefore, since $\tilde{x}_n$ solves the vector optimization problem in (21), by Theorem 2.1 there exists $\tilde{h}_n \in C^+ \setminus \{0\}$ such that $\tilde{x}_n$ solves the problem

$$\min \eta_n(x) \quad (26)$$

$$\text{subject to } x \in \Omega_n, \quad (27)$$

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

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

Since the solution to (26) and (27) is not altered through multiplication of $\tilde{h}_n$ by positive scalars, we can assume without loss of generality that $\|\tilde{h}_n\| = 1$ for all $n \geq 0$. Note that by definition, we have $\Omega_n \subset \text{dom}(\eta_n) = \text{dom}(F)$, so that $\emptyset \neq \text{dom}(I_{\Omega_n}) \subset \text{dom}(\eta_n)$. According to [34, Theorem 3.23], it follows 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

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

and

$$0 \leq \langle u_n, x - \tilde{x}_n \rangle, \quad (29)$$

for all $x \in \Omega_n$. Now define $\psi_n : X \rightarrow \mathcal{R} \cup \{\infty\}$ by

$$\psi_n(x) := \langle F(x), \tilde{h}_n \rangle. \quad (30)$$

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

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

for some

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

Now fix $u \in VO(\Omega, F) \cap VI(\Omega, A)$ arbitrarily. By condition (A), we have $u \in \Omega_n$ for all $n \geq 0$. Combining (29) with $x = u$ and (31) gives us

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

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

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

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

Furthermore, using the identity

$$\|x + y\|^2 = \|x\|^2 - \|y\|^2 + 2\langle y, x + y \rangle, \quad \forall x, y \in X,$$

we derive from (34) the relationships

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

Taking now $\sigma_n > 0$, observe that

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

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

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

Note that for all $x, y \in X$ and $0 \leq \lambda \leq 1$, 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

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

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

$$\|x_{n+1} - u\|^2 \leq \left(1 + \frac{2\sigma_n^2}{1 - 2\sigma_n^2}\right)\|x_n - u\|^2 - \frac{\delta}{2}\|\tilde{x}_n - z_n\|^2, \quad \forall n \geq N_0. \quad (38)$$

Also, from (22) we derive

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

and hence

$$\|\tilde{x}_n - x_n\| = \frac{1}{1 - \beta_n}\|x_{n+1} - x_n\| \geq \frac{1}{1 - \epsilon}\|x_{n+1} - x_n\|. \quad (39)$$

Step 4. Boundedness of the sequence and proximity of consecutive iterates. Indeed, we claim that for every $u \in VO(\Omega, F) \cap VI(\Omega, A)$, the sequence $\{\|x_n - u\|^2\}$ is convergent. In terms of (38) we have

$$\|x_{n+1} - u\|^2 \leq \left(1 + \frac{2\sigma_n^2}{1 - 2\sigma_n^2}\right)\|x_n - u\|^2, \quad \forall n \geq N_0. \quad (40)$$

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 that for all $n \geq N_0$ we have

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

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

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

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

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

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

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

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

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

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

$$\lim_{n \rightarrow \infty} \|\tilde{x}_n - z_n\| = 0. \quad (42)$$

Let $d := \lim_{n \rightarrow \infty} \|x_n - u\|$. Then from (24) and (37) we get

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

Moreover, observe that

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

Since $\{\beta_n\} \subset [\epsilon, 1 - \delta]$ for some $\epsilon, \delta \in (0, 1)$, we deduce from Lemma 2.1 that

$$\lim_{n \rightarrow \infty} \|x_n - \tilde{x}_n\| = 0, \quad (43)$$

and hence

$$\lim_{n \rightarrow \infty} \|x_{n+1} - x_n\| = 0. \quad (44)$$

Note that $\|x_n - z_n\| \leq \|x_n - \tilde{x}_n\| + \|\tilde{x}_n - z_n\|$. This together with (43) implies that

$$\lim_{n \rightarrow \infty} \|x_n - z_n\| = 0. \quad (45)$$

For every $u \in VO(\Omega, F) \cap VI(\Omega, A)$, from (24) we obtain

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

Therefore we have

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

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

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

Then, in contrast to (24), we get

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

and, rearranging as in (46),

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

Since $\lim_{n \rightarrow \infty} \|x_n - z_n\| = 0$ and the sequences $\{x_n\}$ and $\{z_n\}$ are bounded, we obtain $\lim_{n \rightarrow \infty} \|t_n - y_n\| = 0$. As A is k -Lipschitz-continuous, we get $\lim_{n \rightarrow \infty} \|Ay_n - At_n\| = 0$. Note that $\|x_n - t_n\| \leq \|x_n - y_n\| + \|y_n - t_n\|$. Thus we arrive at $\lim_{n \rightarrow \infty} \|x_n - t_n\| = 0$.

Step 5. Optimality of the weak cluster points of $\{x_n\}$. Indeed, since $\{x_n\}$ is bounded, it has weak cluster points. We will prove next that all of them lie in $VO(\Omega, F) \cap VI(\Omega, A)$.

Let $\hat{x}$ be a weak cluster point of $\{x_n\}$ and let $\{x_{k_n}\}$ be a subsequence weakly convergent to it. We can obtain that $\hat{x} \in VO(\Omega, F) \cap VI(\Omega, A)$. Let us first show $\hat{x} \in VI(\Omega, A)$. Since $\lim_{n \rightarrow \infty} \|x_n - t_n\| = 0$ and $\lim_{n \rightarrow \infty} \|x_n - y_n\| = 0$, we have $t_{k_n} \rightharpoonup \hat{x}$ and $y_{k_n} \rightharpoonup \hat{x}$. Denote

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

where $N_\Omega v$ is the normal cone to Ω at $v \in \Omega$. As already mentioned, in this case the mapping T is maximal monotone, and $0 \in Tv$ if and only if $v \in VI(\Omega, A)$; see [15]. Let $\text{Gr}(T)$ be the graph of T , and let $(v, w) \in \text{Gr}(T)$. Then, we have $w \in Tv = Av + N_\Omega v$ and hence $w - Av \in N_\Omega v$. This gives $\langle v - t, w - Av \rangle \geq 0$ for all $t \in \Omega$. On the other hand, from $t_n = P_\Omega(x_n - \lambda_n A y_n)$ and $v \in \Omega$ we have

$$\langle x_n - \lambda_n A y_n - t_n, t_n - v \rangle \geq 0$$

and hence

$$\langle v - t_n, \frac{t_n - x_n}{\lambda_n} + A y_n \rangle \geq 0.$$

From $\langle v - t, w - Av \rangle \geq 0$ for all $t \in \Omega$ and $t_{k_n} \in \Omega$, we further have

$$\begin{aligned} \langle v - t_{k_n}, w \rangle &\geq \langle v - t_{k_n}, Av \rangle \\ &\geq \langle v - t_{k_n}, Av \rangle - \langle v - t_{k_n}, \frac{t_{k_n} - x_{k_n}}{\lambda_{k_n}} + A y_{k_n} \rangle \\ &= \langle v - t_{k_n}, Av - A t_{k_n} \rangle + \langle v - t_{k_n}, A t_{k_n} - A y_{k_n} \rangle - \langle v - t_{k_n}, \frac{t_{k_n} - x_{k_n}}{\lambda_{k_n}} \rangle \\ &\geq \langle v - t_{k_n}, A t_{k_n} - A y_{k_n} \rangle - \langle v - t_{k_n}, \frac{t_{k_n} - x_{k_n}}{\lambda_{k_n}} \rangle. \end{aligned}$$

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

On the other hand, we define $\psi_z : X \rightarrow \mathcal{R}$ by $\psi_z(x) := \langle F(x), z \rangle$ and claim that

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

for all $z \in C^+$ and all $n \geq 0$. Indeed, since F is positively lower semicontinuous and C -convex, ψ_z is lower semicontinuous and convex, and so $\psi_z(\hat{x}) \leq \liminf_{n \rightarrow \infty} \psi_z(x_{k_n})$.

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

$$F(\hat{x}) \preceq_C F(x_n), \quad \forall n \geq 0. \tag{48}$$

Assume 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 $\bar{h}_n$ as chosen before (26). Since $\|\bar{h}_n\| = 1$ for all n , there exists, by Bourbaki-Alaoglu Theorem, a weak* cluster point of $\{\bar{h}_n\}$, say $\bar{h}$, which is a weak* limit of some subnet $\{\bar{h}_{j_n}\}$ of $\{\bar{h}_n\}$. We claim now that C^+ is weak* closed. Observe that $C^+ = \bigcap_{y \in C} \{z \in Y^* \mid \langle y, z \rangle \geq 0\}$. Since the linear forms $z \mapsto \langle y, z \rangle$ are weak* continuous for all $y \in Y$, we can represent C^+ as an intersection of weak* closed sets establishing the claim. It follows that $\bar{h} \in C^+$. Note that ψ_z is convex for each $z \in C^+$. Hence from (22) we get

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

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

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

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).

Using consequently (47), we conclude from (49) and (50) that

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

Note that $\lim_{n \rightarrow \infty} \|\tilde{x}_{j_n} - z_{j_n}\| = 0$ by (42), $\lim_{n \rightarrow \infty} \|x_{j_n} - \tilde{x}_{j_n}\| = 0$ by (51), and $\|\theta_{j_n}\| \leq \sigma_{j_n} \|\tilde{x}_{j_n} - z_{j_n}\|$ by (14). Now we take 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 $\|\bar{h}_n\| = \|e_n\| = 1$, we have that $\{\nu_n\}$ is bounded as well. Note again that $\{x_n\}$ is bounded so that $\{\|\bar{x} - \tilde{x}_{j_n}\|\}$ is also bounded. Finally, it is easy to see that $\lim_{n \rightarrow \infty} \|\tilde{x}_{j_n} - z_{j_n} - \theta_{j_n}\| = 0$, and

$$\lim_{n \rightarrow \infty} |\langle F(x_{j_n}) - F(\tilde{x}_{j_n}), \bar{h}_{j_n} \rangle - \langle F(\hat{x}) - F(\tilde{x}_{j_n}), \bar{h} \rangle| = 0$$

according to assumption (B). Since $\psi_{\bar{h}}$ is convex and lower semicontinuous, it is weakly lower semicontinuous. Thus it is clear that $\liminf_{n \rightarrow \infty} \psi_{\bar{h}}(\tilde{x}_{j_n}) \geq \psi_{\bar{h}}(\hat{x})$; i.e., $\liminf_{n \rightarrow \infty} (\psi_{\bar{h}}(\tilde{x}_{j_n}) -$

$\psi_h(\hat{x}) \geq 0$. Thus, for any given $\varepsilon > 0$, there exists an integer $N_0 \geq 1$ such that

$$\psi_h(\tilde{x}_{j_n}) - \psi_h(\hat{x}) \geq -\varepsilon, \quad \forall n \geq N_0.$$

This with together $\{\beta_n\} \subset [\varepsilon, 1 - \delta]$ imply that

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

Consequently, we have

$$\liminf_{n \rightarrow \infty} \beta_{j_n}(\psi_h(\tilde{x}_{j_n}) - \psi_h(\hat{x})) \geq -\varepsilon(1 - \delta).$$

Utilizing the arbitrariness of ε , we arrive at

$$\liminf_{n \rightarrow \infty} \beta_{j_n}(\psi_h(\tilde{x}_{j_n}) - \psi_h(\hat{x})) \geq 0$$

and we conclude that the lower limit of the rightmost expression in (51) as $n \rightarrow \infty$ is not less than zero. Since $\tilde{h}$ is the weak* limit of $\{\tilde{h}_{j_n}\}$, we get from (51) that

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

Next we claim that $\tilde{h} \neq 0$. Take $e \in \text{int}C$. It follows from Lemma 2.2 of [29] that $\langle e, \tilde{h}_n \rangle \geq d(e, Y \setminus C) > 0$ for all $n \geq 0$. Since $\tilde{h}$ is the weak* limit of $\{\tilde{h}_{j_n}\}$, we get that $\langle e, \tilde{h} \rangle > 0$, establishing the claim. Since $\tilde{h} \neq 0$, it is clear that (52) contradicts the fact that $\tilde{h}$ belongs to C^+ and the assumption that $F(\bar{x}) \prec_C F(\hat{x})$. Thus such an assumption is false, and $\hat{x}$ so is indeed weakly efficient for VOP.

Step 6. Uniqueness of the weak cluster point of $\{x_n\}$. This part of the proof, presented for the sake of completeness, is in the same line as the scalar-valued case in [6] using Brézis's uniqueness argument. By the same argument as in (48), both $\hat{x}$ and $\tilde{x}$ belong to $VO(\Omega, F) \cap VI(\Omega, A)$ and both $\{\|\hat{x} - x_n\|\}$ and $\{\|\tilde{x} - x_n\|\}$ converge as shown at the beginning of Step 4; i.e., there exist $\hat{\beta}, \tilde{\beta} \in \mathcal{R}_+$ such that

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

By the 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

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

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

$$\hat{\beta}^2 - \tilde{\beta}^2 = \|\tilde{x} - \hat{x}\|^2. \quad (55)$$

Reversing the roles of $\hat{x}$ and $\tilde{x}$, a similar reasoning leads to $\tilde{\beta}^2 - \hat{\beta}^2 = \|\tilde{x} - \hat{x}\|^2$. Combined this with (55), we get $\|\tilde{x} - \hat{x}\| = 0$, i.e., $\tilde{x} = \hat{x}$, which justifies the uniqueness of the weak cluster point of $\{x_n\}$. It follows that $\{x_n\}$ is weakly convergent to an element of $VO(\Omega, F) \cap VI(\Omega, A)$, and thus the proof is complete. $\square$

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

A lot of research during recent years has focused on nonlinear generations of recursion (1) based on *Bregman functions* defined in [5]. Suppose $h : \Omega \rightarrow \mathcal{R}$ is a strictly convex function that is Gâteaux differentiable on Ω . The *Bregman distance* between x and y is defined via the “*D*-function”

$$D_h(x, y) = h(x) - h(y) - \langle \nabla h(y), x - y \rangle, \quad (56)$$

where $x, y \in \Omega$. It follows from the strict convexity of h , that $D_h(x, y) \geq 0$ and $D_h(x, y) = 0$ if and only if $x = y$. If $h(x) = \frac{1}{2}\|x\|^2$, then $D_h(x, y) = \frac{1}{2}\|x - y\|^2$. In what follows, we use a class of functions that are presented 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 the definition of a Bregman function, so h is also a Bregman function. Thus for all $x, y \in \Omega$ we have

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

The absolute extension to Bregman-function-based hybrid approximate proximate algorithm, which is called Algorithm 2 below, 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 $\{\gamma_n\}$ in $[0, 1]$, and a sequence $\{e_n\} \subset \text{int}C$ such that $\|e_n\| = 1$ for all n . Assume that $\Omega \cap \text{dom}(F) \neq \emptyset$. The method generates a sequence $\{x_n\} \subset \Omega$ in the following way:

Initialization: Choose $x_0 \in \Omega \cap \text{dom}(F)$.

Stopping rule: Given x_n , if $x_n \in C\text{-ARGMIN}_w\{F(x)|x \in \Omega\} = VO(\Omega, F)$, then let $x_{n+p} = x_n$ for all $p \geq 1$.

Iterative step: Given x_n , if $x_n \notin C\text{-ARGMIN}_w\{F(x)|x \in \Omega\} = VO(\Omega, F)$, we first compute

$$\begin{cases} y_n = P_\Omega(x_n - \lambda_n A x_n), \\ z_n = \gamma_n x_n + (1 - \gamma_n) P_\Omega(x_n - \lambda_n A y_n) \end{cases} \quad (58)$$

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

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

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

In this algorithm, instead of condition (14), for every $u \in VO(\Omega, F) \cap VI(\Omega, A)$ we take

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

as the approximate criterion corresponding to recursion (59).

We make the following assumption on the map F and the initial iterate x_0 :

(A) The set $(F(x_0) - C) \cap F(\Omega)$ is C -quasicomplete for Ω , meaning 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 $n \geq 0$, it holds $F(u) \preceq_C F(a_n)$ for all $u \in VO(\Omega, F) \cap VI(\Omega, A)$ and all $n \geq 0$.

To prove the convergence of Algorithm 2, we need the following propositions, which can be found in Chen and Teboulle [13].

Proposition 4.1. For any $x, y, z \in X$ we have

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

Proposition 4.2. 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 assumption (A).

Theorem 4.1. Let $F : \Omega \rightarrow Y \cup \{\infty_C\}$ be a proper, C -convex, and positively lower semicontinuous map with $\Omega \cap \text{dom}(F) \neq \emptyset$. Let $\nabla h(\cdot)$ be uniformly continuous from the strong topology of X to the strong topology of X . Let $A : \Omega \rightarrow X$ be a monotone and k -Lipschitz-continuous mapping such that $VO(\Omega, F) \cap VI(\Omega, A) \neq \emptyset$. Let $\{x_n\}$ be any sequence generated by Algorithm 2. Assume that the fulfillment of condition (60), assumption (A), and the following conditions:

- (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 we have the conclusions:

(I) $\{x_n\}$ converges with respect to the weak topology of X to a weakly efficient solution of the VOP;

(II) $\{x_n\}$ converges with respect to the weak topology of X to an element of $VO(\Omega, F) \cap VI(\Omega, A)$ provided $x_n \notin C\text{-ARGMIN}_w \{F(x) | x \in \Omega\}, \forall n \geq 0$.

Proof. We divide the proof into several steps.

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

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

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

Step 2. Existence of iterates. This assertion can be proved by using the same argument as in Step 2 of the proof of Theorem 3.1 with $\varphi_n : X \rightarrow \mathcal{R} \cup \{\infty\}$ defined 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 Theorem 2.1 there exists $\tilde{h}_n \in C^+ \setminus \{0\}$ such that x_{n+1} solves the problem:

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

$$\text{subject to } x \in \Omega_n, \quad (63)$$

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

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

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

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

and

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

Now define $\psi_n : X \rightarrow \mathcal{R} \cup \{\infty\}$ as

$$\psi_n(x) := \langle F(x), \tilde{h}_n \rangle. \quad (66)$$

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

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

for some

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

Now take a fixed $u \in VO(\Omega, F) \cap VI(\Omega, A)$ arbitrarily. By condition (A), $u \in \Omega_n$ for all $n \geq 0$. Combining (65) with $x = u$ and (67), we have

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

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

Now define $\nu_n := \alpha_n \langle e_n, \tilde{h}_n \rangle$. Note that $\nu_n > 0$ by $\alpha_n > 0$, $e_n \in \text{int}C$, and $\tilde{h}_n \in C^+ \setminus \{0\}$. From (69) we obtain

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

which together with (60) imply that

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$D_h(u, x_{n+1}) \leq (1 + \frac{2\sigma_n^2}{1-2\sigma_n^2})D_h(u, x_n), \quad \forall n \geq N_0. \quad (77)$$

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

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

Observe that for all $n \geq N_0$ we have

$$\begin{aligned} D_h(u, x_{n+1}) &\leq (1 + \frac{2\sigma_n^2}{1-2\sigma_n^2})D_h(u, x_n) \\ &\leq (1 + \frac{2\sigma_n^2}{1-2\sigma_n^2})(1 + \frac{2\sigma_{n-1}^2}{1-2\sigma_{n-1}^2})D_h(u, x_{n-1}) \\ &\quad \vdots \\ &\leq \prod_{j=N_0}^n (1 + \frac{2\sigma_j^2}{1-2\sigma_j^2})D_h(u, x_{N_0}) \\ &\leq \prod_{j=N_0}^{\infty} (1 + \frac{2\sigma_j^2}{1-2\sigma_j^2})D_h(u, x_{N_0}) \\ &= K_1 D_h(u, x_{N_0}). \end{aligned}$$

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 $\{t_n\}$ and $\{z_n\}$ are bounded. Set $\widetilde{M} = \sup_{n \geq 0} D_h(u, x_n)$. Then from (77) we get

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

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

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

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

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

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

$$\frac{1}{2} D_h(x_{n+1}, x_n) \leq (1 + \frac{2\sigma_n^2}{1-2\sigma_n^2}) D_h(u, x_n) - D_h(u, x_{n+1}) \quad (78)$$

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

$$\lim_{n \rightarrow \infty} D_h(x_{n+1}, x_n) = 0. \quad (79)$$

Thus we arrive at $\lim_{n \rightarrow \infty} \|x_{n+1} - x_n\| = 0$ by (57).

On the other hand, since $\sum_{n=0}^{\infty} \sigma_n < \infty$, both $\{D_h(x_{n+1}, x_n)\}$ and $\{\|x_n - u\|\}$ are bounded, 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 $\lim_{n \rightarrow \infty} \frac{1}{2} \|u - x_n\|^2$ exists and so $\lim_{n \rightarrow \infty} \|u - x_n\|$ exists. Let $\tau := \lim_{n \rightarrow \infty} \|x_n - u\|$. Letting $n \rightarrow \infty$ we obtain from (76) that

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

and so

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

Note that (23) implies that $\limsup_{n \rightarrow \infty} \|u - t_n\| \leq \tau$. Utilizing Lemma 2.1, we have

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

which together with (58) imply that

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

Picking any $u \in VO(\Omega, F) \cap VI(\Omega, A)$, we obtain from (24) that

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

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

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

Then, in contrast to (24), the following hold:

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

Rearranging as in (80), we arrive at

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

Since $\lim_{n \rightarrow \infty} \|x_n - z_n\| = 0$ and the sequences $\{x_n\}$ and $\{z_n\}$ are bounded, we get $\lim_{n \rightarrow \infty} \|t_n - y_n\| = 0$. As A is k -Lipschitz-continuous, we further have $\lim_{n \rightarrow \infty} \|Ay_n - At_n\| = 0$.

Step 5. Optimality of the weak cluster points of $\{x_n\}$. Since $\{x_n\}$ is bounded, it has weak cluster points. We will prove next that all of them lie in $VO(\Omega, F) \cap VI(\Omega, A)$. Let $\hat{x}$ be a weak cluster point of $\{x_n\}$ and let $\{x_{k_n}\}$ be a subsequence weakly convergent to it. We can obtain that $\hat{x} \in VO(\Omega, F) \cap VI(\Omega, A)$. First, as in Step 5 of the proof of Theorem 3.1, we can show $\hat{x} \in VI(\Omega, A)$.

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

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

for all $z \in C^+$ and all $n \geq 0$. Indeed, since F is positively lower semicontinuous and C -convex, ψ_z is lower semicontinuous and convex so that $\psi_z(\hat{x}) \leq \lim_{n \rightarrow \infty} \psi_z(x_{k_n})$. Since $F(x_{n+1}) \preceq_C F(x_n)$ for all n , we have $\psi_z(x_{n+1}) \leq \psi_z(x_n)$ for all n so that $\lim_{n \rightarrow \infty} \psi_z(x_{k_n}) = \inf\{\psi_z(x_n)\}$. This shows that $\psi_z(\hat{x}) \leq \inf\{\psi_z(x_n)\}$, and hence (81) holds. It follows easily from (81) that

$$F(\hat{x}) \preceq_C F(x_n), \quad \forall n \geq 0.\tag{82}$$

Assume 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 $n \geq 0$. Obviously, Take $\bar{h}_n$ as chosen before (62). Since $\|\bar{h}_n\| = 1$ for all n , there exists, by the Bourbaki-Alaoglu

Theorem, a weak* cluster point of $\{\tilde{h}_{k_n}\}$, say $\tilde{h}$, which is a weak* limit of the subnet $\{\tilde{h}_{j_n}\}$ of $\{\tilde{h}_{k_n}\}$. We claim now that C^+ is weak* closed. Observe that $C^+ = \bigcap_{y \in C} \{z \in Y^* | \langle y, z \rangle \geq 0\}$. Since the linear forms $z \mapsto \langle y, z \rangle$ are weak* continuous for all $y \in Y$, we have written C^+ as an intersection of weak* closed sets, establishing the claim. It follows that $\tilde{h} \in C^+$. Thus

$$\begin{aligned}
\langle F(\bar{x}) - F(\hat{x}), \tilde{h}_{j_n} \rangle &\geq \langle F(\bar{x}) - F(x_{j_n+1}), \tilde{h}_{j_n} \rangle \\
&= \psi_{j_n}(\bar{x}) - \psi_{j_n}(x_{j_n+1}) \\
&\geq \langle v_{j_n}, \bar{x} - x_{j_n+1} \rangle \\
&= \langle u_{j_n}, \bar{x} - x_{j_n+1} \rangle - \nu_{j_n} \langle \nabla h(x_{j_n+1}) - \nabla h(x_{j_n}) - \theta_{j_n}, \bar{x} - x_{j_n+1} \rangle \\
&\geq -\nu_{j_n} \langle \nabla h(x_{j_n+1}) - \nabla h(x_{j_n}) - \theta_{j_n}, \bar{x} - x_{j_n+1} \rangle.
\end{aligned} \tag{83}$$

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

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

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

Now we take limits in the first and last expressions of (83). Regarding the rightmost one in (83), since $\{\alpha_n\}$ is bounded, and $\|\tilde{h}_n\| = \|e_n\| = 1$, we have that $\{\nu_n\}$ is bounded. Thus we conclude that the limit of the rightmost expression in (83) as $n \rightarrow \infty$ vanishes and so, since $\tilde{h}$ is the weak* limit of $\{\tilde{h}_{j_n}\}$, we get from (83) that

$$\langle F(\bar{x}) - F(\hat{x}), \tilde{h} \rangle \geq 0. \tag{84}$$

Now we claim that $\tilde{h} \neq 0$. Take $e \in \text{int}C$. It follows from Lemma 2.2 of [29] that $\langle e, \tilde{h}_n \rangle \geq d(e, Y \setminus C) > 0$ for all $n \geq 0$. Since $\tilde{h}$ is the weak* limit of $\{\tilde{h}_{j_n}\}$, we get that $\langle e, \tilde{h} \rangle > 0$, establishing the claim. Since $\tilde{h} \neq 0$, it is clear that (84) contradicts the fact that $\tilde{h}$ belongs to C^+ and the assumption that $F(\bar{x}) \prec_C F(\hat{x})$. Thus, such an assumption is false, and so $\hat{x}$ is indeed weakly efficient for VOP.

Step 6. Uniqueness of the weak cluster point of $\{x_n\}$. Indeed, this part of the proof, presented also for the sake of completeness again closely related to the scalar-valued case, as in [21], and it uses Brézis's uniqueness argument. With the same argument as in (82), both $\hat{x}$ and $\tilde{x}$ belong to $VO(\Omega, F) \cap VI(\Omega, A)$. As in the Step 6 of the proof of Theorem 3.1, we can prove that $\tilde{x} = \hat{x}$, establishing the uniqueness of the weak cluster point of $\{x_n\}$. It follows that $\{x_n\}$ is weakly convergent to an element of $VO(\Omega, F) \cap VI(\Omega, A)$ and the proof is complete. $\square$

Remark 4.1. In [21, Chapter 3], the scalar version of proximal point method may show us another way of using Bregman distances in order to produce weakly convergent algorithms. Meantime, the above Algorithm 2 is closely related to the algorithm in [7, Chapter 3] since the proximal point method discussed in [21, Chapter 3] is an exact scalar variant of the above Algorithm 2.

Remark 4.2. For "Stopping rule" in the above Algorithms 1 and 2, there is the requirement that $x_{n_p} = x_p$ for all $p \geq 1$ if for given x_n , $x_n \in C\text{-MINIMIZE}\{F(x) | x \in \Omega\}$. In

general, the requirement $x_{n+1} = x_n$ is sufficient as the usual stopping rule in scalar proximal point method. But, before the above Algorithms 1 and 2 are introduced, respectively, we specifically indicate and stress that “the method generates a sequence $\{x_n\}$ ”, i.e., an infinite sequence $\{x_n\}$. In this paper, the aim is to solve the VOP: $C\text{-MINIMIZE}\{F(x)|x \in \Omega\}$. In the proceeding of iterations we meet two possible cases.

Case (I). At each iteration step we have $x_n \notin C\text{-MINIMIZE}\{F(x)|x \in \Omega\}$. Hence the process of iteration continues infinitely producing an infinite sequence $\{x_n\}$. Under the conditions of Theorems 3.1 or 4.1, $\{x_n\}$ converges weakly to a solution of the VOP. This achieves our aim.

Case (II). There exists some iteration step such that we have $x_n \in C\text{-MINIMIZE}\{F(x)|x \in \Omega\}$. 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 $\{x_n\}$ converges weakly to a solution of the VOP.

5. Relative Hybrid Approximate Proximal Algorithm

In the last section we present the relative version of our hybrid approximate proximal algorithm, which is called Algorithm 3. It requires several exogenous sequences: the ones required by Algorithm 2, i.e., an error sequence $\{\theta_n\} \subset \Omega$, two bounded sequences of positive real numbers $\{\alpha_n\}$ and $\{\sigma_n\}$, a sequence $\{e_n\} \subset \text{int}(C)$ such that $\|e_n\| = 1$ for all n , and now also a sequence $\{\hbar_n\} \subset C^+$ such that $\|\hbar_n\| = 1$ for all $n \geq 0$. The method generates a sequence $\{x_n\} \subset \Omega$ in the following way:

Initialization: Choose $x_0 \in \Omega \cap \text{dom}(F)$.

Stopping rule: Given x_n , if $x_n \in C\text{-ARGMIN}_w\{F(x)|x \in \Omega\} = VO(\Omega, F)$, then let $x_{n+p} = x_n$ for all $p \geq 1$.

Iterative step: Given x_n , if $x_n \notin C\text{-ARGMIN}_w\{F(x)|x \in \Omega\} = VO(\Omega, F)$, we first compute

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

for every $n = 0, 1, 2, \dots$, 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

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

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

For this algorithm, instead of condition (14), we can take

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

as the approximate criterion corresponding to recursion (86).

Remark 5.1. We mention now that the difference between the presentation of the iteration steps in Algorithms 2 and 3 is not substantial. The subproblems of Algorithm 2 require finding weakly efficient points for a regularized vector-valued function; the subproblems of Algorithm 3 demand ε -subgradients of a scalar-valued one, which could be seen, in light of Theorem 2.1, as approximate weakly efficient points of the vector-valued one. We choose the presentation above in order to avoid the possibly cumbersome tasks of defining, for vector-valued maps, either some kind of approximate weakly efficient points or some kind of ε -subgradients. The subproblems of Algorithm 3, despite its scalar-valued presentation, in some cases are more suitably solved by algorithms specifically devised for vector optimization.

The convergence result for Algorithm 3 is the following.

Theorem 5.1. Let $F : \Omega \rightarrow Y \cup \{\infty_C\}$ be a proper, C -convex, and positively lower semicontinuous map with $\Omega \cap \text{dom}(F) \neq \emptyset$. Let $\nabla h(\cdot)$ be uniformly continuous from the strong topology of X to the strong topology of X . Let $A : \Omega \rightarrow X$ be a monotone and k -Lipschitz-continuous mapping such that $VO(\Omega, F) \cap VI(\Omega, A) \neq \emptyset$. Let $\{x_n\}$ be any sequence generated by Algorithm 3. Assume that condition (14), assumption (A), introduced just before the statement of Theorem 3.1, and the following conditions are satisfied:

- (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 following hold:

- (I) $\{x_n\}$ converges with respect to the weak topology of X to a weakly efficient solution of the VOP;
- (II) $\{x_n\}$ converges with respect to the weak topology of X to an element of $VO(\Omega, F) \cap VI(\Omega, A)$ provided $x_n \notin C\text{-ARGMIN}_w\{F(x)|x \in \Omega\}, \forall n \geq 0$.

Proof. We divide the proof into several steps.

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

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

Indeed, the arguments are similar to those in Step 1 of the proof of Theorem 3.1.

Step 2. Existence of the iterates. According to our definition given by (8) we have $\partial f(x) = \partial_0 f(x) \subset \partial f_\varepsilon(x)$ for all convex $f : X \rightarrow \mathcal{R} \cup \{\infty\}$, all $x \in X$, and all $\varepsilon \in \mathcal{R}_+$. Thus, assuming some z_n exists, the strongly convex function

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

has a subdifferential which has a zero x_{n+1} by Minty's theorem. Such a zero x_{n+1} satisfies the inclusion in (86), with $\varepsilon_n = 0$, which trivially satisfies (87).

Step 3. *Fejér convergence to the set of lower bounds of the initial section.* Indeed, let $\nu_n := \alpha_n \langle e_n, \hbar_n \rangle$ again. By (86) we have

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

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

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

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

$$f_n(u) = \langle F(u), \hbar_n \rangle \quad (91)$$

and also

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

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

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

Combining (90)-(93), we get

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

It follows from (94) that

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

Now for $\sigma_n > 0$, using (88) we get

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

Combining (87), (95) and (96), we obtain

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

Since $\sigma_n \rightarrow 0$, 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 for all $n \geq N_0$ we get

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

Step 4. Boundedness of $\{x_n\}$ and proximity of consecutive iterates. Indeed, as in Step 4 of the proof of Theorem 4.1, we can prove that $\lim_{n \rightarrow \infty} \|x_n - u\|$ exists and that

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

Utilizing Lemma 2.1, we have $\lim_{n \rightarrow \infty} \|x_n - t_n\| = 0$ and so $\lim_{n \rightarrow \infty} \|z_n - x_n\| = 0$ due to (85). Further, we can also obtain in this step the very similar results as in Step 4 of the proof of Theorem 4.1.

Steps 5-6. Optimality of the weak cluster points of $\{x_n\}$; Uniqueness of the weak cluster point of $\{x_n\}$. Because the remainder of the proof is very similar to the argument in Steps 5-6 in the proof of Theorem 4.1, we omit it. $\square$

References

- [1] H. BONNEL, A. N. IUSEM, AND B. F. SVAITER, Proximal methods in vector optimization, *SIAM J. Optim.*, 15 (4) (2005), pp. 953-970.
- [2] J. FLIEGE AND B. F. SVAITER, Steepest descent methods for multicriteria optimization, *Math. Methods Oper. Res.*, 51 (2000), pp. 479-494.
- [3] L. M. GRANA DRUMMOND AND B. F. SVAITER, A steepest descent method for vector optimization, *J. Comput. Appl. Math.*, 175 (2005), pp. 395-414.
- [4] L. M. GRANA DRUMMOND AND A. N. IUSEM, A projected gradient method for vector optimization problems, *Comput. Optim. Appl.*, 28 (2004), pp. 5-30.
- [5] J. ECKSTEIN, Approximate iterations in Bregman-function-based proximal algorithms, *Math. Programming*, 83 (1998), pp. 113-123.
- [6] R. T. ROCKAFELLAR, Monotone operators and the proximal point algorithm, *SIAM J. Control Optim.*, 14 (1976), pp. 877-898.
- [7] R. S. BURACHIK, A. N. IUSEM, AND B. F. SVAITER, Enlargement of monotone operators with applications to variational inequalities, *Set-Valued Anal.*, 5 (1997), pp. 159-180.
- [8] R. S. BURACHIK, S. SCHEIMBERG, AND B. F. SVAITER, Robustness of the hybrid extragradient proximal point algorithm, *J. Optim. Theory Appl.*, 111 (2001), pp. 117-136.
- [9] B. S. HE, Inexact implicit methods for monotone general variational inequalities, *Math. Programming*, 86 (1999), pp. 199-217.
- [10] J. JAHN, Theory of vector maximization: Various concepts of efficient solutions, in *Multicriteria Decision Making. Advances in MCDM Models, Algorithms, Theory, and Applications*, T. Gal, T. J. Stewart, and T. Hanne, eds., Kluwer, Boston, 1999, pp. 2-1-2-32.
- [11] M. V. SOLODOV AND B. F. SVAITER, An inexact hybrid extragradient-proximal point algorithm using the enlargement of a maximal monotone operator, *Set-Valued Anal.*, 7 (1999), pp. 323-345.

- [12]R. S. BURACHIK and A. N. IUSEM, Set-Valued Mappings and Enlargements of Monotone Operators, Springer Optimization and Its Applications, Vol. 8, 2008.
- [13]G. CHEN AND M. TEBOULLE, A proximal-based decomposition method for convex minimization problems, *Math. Programming*, 64 (1994), 81-100.
- [14]P. J. DA SILVA E SILVA, J. ECKSTEIN, AND C. HUMES, "Rescaling and Stepsize Selection in Proximal Methods Using Separable Generalized Distance," RUTCOR Research Report RRR 35-99, Rutgers University, 1999.
- [15]M. V. SOLODOV AND B. F. SVAITER, A hybrid projection-proximal point algorithm, *J. Convex Anal.*, 6 (1999), pp. 59-70.
- [16]M. V. SOLODOV AND B. F. SVAITER, Forcing strong convergence of proximal point iterations in a Hilbert space, *Math. Programming*, 87 (2000), 189-202.
- [17]M. V. SOLODOV AND B. F. SVAITER, An inexact hybrid generalized proximal point algorithm and some new results on the theory of Bregman functions, *Math. Oper. Res.*, 25 (2) (2000), pp. 214-230.
- [18]A. GOPFERT, H. RIAHI, C. TAMMER, AND C. ZALINESCU, *Variational Methods in Partially Ordered Spaces*, Springer, New York, 2003.
- [19]H. BENKER, A. HAMEL, AND C. TAMMER, An algorithm for vectorial control approximation problems, in *Multiple Criteria Decision Making*, G. Fandel and T. Gal, eds., Springer, Berlin, 1997, pp. 3-12.
- [20]M. KYONO AND M. FUKUSHIMA, Nonlinear proximal decomposition method for convex programming, *J. Optim. Theory Appl.*, 106 (2000), pp. 357-372.
- [21]D. BUTNARIU AND A. N. IUSEM, *Totally Convex Functions for Fixed Point Computation and Infinite Dimensional Optimization*, Kluwer, Dordrecht, 18 (7-8) (1997), 723-744.
- [22]L. C. CENG AND J. C. YAO, Approximate proximal methods in vector optimization, *European J. Oper. Res.*, 183 (1) (2007), pp. 1-19.
- [23]N. NADEZHKINA AND W. TAKAHASHI, Strong convergence theorem by a hybrid method for nonexpansive mappings and Lipschitz-continuous monotone mappings, *SIAM J. Optim.*, 16 (2006), pp. 1230-1241.
- [24]G. M. KORPELEVICH, The extragradient method for finding saddle points and other problems, *Matecon*, 12 (1976), pp. 747-756.
- [25]R. S. BURACHIK, J. O. LOPES, AND B. F. SVAITER, An outer approximation method for the variational inequality problem, *SIAM J. Control Optim.*, 43 (2005), pp. 2071-2088.
- [26]L. C. CENG AND J. C. YAO, An implicit iterative scheme for monotone variational inequalities and fixed point problems, *Nonlinear Analysis* (2007), doi:10.1016/j.na.2007.08.023.
- [27]B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation, II: Applications*. Springer, Berlin, 2006.
- [28]T. Q. BAO and B. S. MORDUKHOVICH, Relative Pareto minimizers for multiobjective problems: Existence and optimality conditions, *Math. Programming* (2009) (to appear).

- [29]S. BOLINTINEANU, Approximate efficiency and scalar stationarity in unbounded non-smooth convex vector optimization problems, J. Optim. Theory Appl., 106 (2000), pp. 265-296.
- [30]S. BOLINTINEANU, Vector variational principles; ε -efficiency and scalar stationarity, J. Convex Anal., 8 (2001), pp. 71-85.
- [31]T. D. LUC, Theory of Vector Optimization, Lecture Notes in Econom. and Math. Syst. 319, Springer, Berlin, 1989.
- [32]J. SCHU, Weak and strong convergence to fixed points of asymptotically nonexpansive mappings, Bull. Austral. Math. Soc., 43 (1991), pp. 153-159.
- [33]R. T. ROCKAFELLAR, On the maximality of sums of nonlinear monotone operators, Trans. Amer. Math. Soc., 149 (1970), pp. 75-88.
- [34]R. R. PHELPS, Convex Functions, Monotone Operators and Differentiability, Lecture Notes in Math. 1364, Springer, Berlin, 1988.