Variational Theory and Algorithms for a Class of Asymptotically Approachable Nonconvex Problems

Hanyang Li *   Ying Cui †

July 4, 2023 (Revision: December 26, 2023)

Abstract

We investigate a class of composite nonconvex functions, where the outer function is the sum of univariate extended-real-valued convex functions and the inner function is the limit of difference-of-convex functions. A notable feature of this class is that the inner function can be merely lower semicontinuous instead of continuously differentiable. It covers a range of important yet challenging applications, including the composite value functions of nonlinear programs and the value-at-risk constraints. We propose an asymptotic decomposition of the composite function that guarantees epi-convergence to the original function, leading to necessary optimality conditions for the corresponding minimization problem. The proposed decomposition also enables us to design a numerical algorithm such that any accumulation point of the generated sequence, if exists, satisfies the newly introduced optimality conditions. These results expand on the study of so-called amenable functions introduced by Poliquin and Rockafellar in 1992, which are compositions of convex functions with smooth maps, and the prox-linear methods for their minimization.

Keywords: epi-convergence; optimality conditions; nonsmooth analysis; difference-of-convex functions

1 Introduction.

We consider a class of composite optimization problems of the form:

$$\operatorname*{minimize}_{x\in\mathbb{R}^n} \quad \sum_{p=1}^{m}\Big[F_p(x) \triangleq \varphi_p\big(f_p(x)\big)\Big], \tag{CP$_0$}$$

where for each $p = 1,\cdots,m$, the outer function $\varphi_p : \mathbb{R} \to \mathbb{R}\cup\{+\infty\}$ is proper, convex, lower semicontinuous (lsc), and the inner function $f_p : \mathbb{R}^n \to \mathbb{R}$ can be merely lsc. Throughout the paper, we assume the existence of an optimal solution for (CP$_0$).

If each inner function f_p is continuously differentiable, then the objective in (CP$_0$) belongs to the family of *amenable functions* under a constraint qualification [20, 21]. For a thorough exploration

*Department of Industrial Engineering and Operations Research, University of California, Berkeley, Berkeley 94720 (hanyang_li@berkeley.edu)

†Department of Industrial Engineering and Operations Research, University of California, Berkeley, Berkeley 94720 (yingcui@berkeley.edu)

of the variational theory of amenable functions, readers are referred to [25, Chapter 10(F)]. The properties of amenable functions have also led to the development of prox-linear algorithms, where convex subproblems are constructed through the linearization of the inner smooth mapping [14, 3, 4, 15, 12].

However, there are various applications of composite optimization problem in the form of (CP_0) where the inner function f_p is nondifferentiable, or even discontinuous. In the following, we provide two such examples.

Example 1.1 (Composite value functions) For $p = 1, \cdots, m$, consider the optimal value function

$$f_p(x) \triangleq \inf_{y \in \mathbb{R}^{n_2}} \left\{ (c^p + C^p x)^\top y + \frac{1}{2}\, y^\top Q^p\, y \;\middle|\; A^p x + B^p y \leq b^p \right\} \qquad x \in \mathbb{R}^{n_1} \tag{1}$$

with appropriate dimensional vectors b^p and c^p, and matrices A^p, B^p, C^p and Q^p, where Q^p is symmetric and positive semidefinite. The function f_p is not smooth in general. The inverse (multi) optimal value problem [1, 19] finds a vector $x \in \mathbb{R}^{n_1}$ that minimizes the discrepancy between observed optimal values $\{v_p\}_{p=1}^m$ and true values $\{f_p\}_{p=1}^m$ based on a prescribed metric, such as the ℓ_1-error:

$$\underset{x \in \mathbb{R}^{n_1}}{\text{minimize}} \sum_{p=1}^{m} |v_p - f_p(x)| \,. \tag{2}$$

One can express problem (2) in the form of (CP_0) by defining the outer function $\varphi_p(t) = |v_p - t|$.

Example 1.2 (Optimal portfolios under a value-at-risk constraint) The *value-at-risk (VaR)* of a random variable Y at a confidence level $\alpha \in (0,1)$ is $\mathrm{VaR}_\alpha(Y) \triangleq \inf\{\gamma \in \mathbb{R} \mid \mathbb{P}(Y \leq \gamma) \geq \alpha\}$. Let Z be a random vector and $c(x, Z)$ represent the profit of investments parameterized by $x \in \mathbb{R}^n$. An agent's goal is to maximize the expected utility of $c(x, Z)$, denoted as $\mathbb{E}[\, u(c(x, Z))]$, while also controlling the risk via a constraint on $\mathrm{VaR}_\alpha[c(x, Z)]$ under a prescribed level r. Adapted from [30, Section 3.4], the model can be written as

$$\underset{x \in \mathbb{R}^n}{\text{maximize}} \quad \mathbb{E}\left[\, u\big(c(x, Z)\big)\right] \quad \text{subject to} \quad \mathrm{VaR}_\alpha[\, c(x, Z)] \leq r, \tag{3}$$

which can be put into the framework (CP_0) by defining $\varphi_1(t) = t$, $f_1(x) = -\mathbb{E}[\, u(c(x, Z))]$, $\varphi_2(t) = \delta_{(-\infty, r]}(t)$, and $f_2(x) = \mathrm{VaR}_\alpha[\, c(x, Z)]$. We note that the inner function $\mathrm{VaR}_\alpha[\, c(\cdot, Z)]$ is nonsmooth and can be discontinuous in general.

Due to the nondifferentiablity of the inner function f_p in (CP_0), the prox-linear algorithm is not applicable to solve this composite optimization problem. The present paper aims to develop an algorithmic framework for a subclass of (CP_0), where each inner function f_p, although nonsmooth, can be expressed as the limit of difference-of-convex (DC) functions. We refer to this class of functions as *approachable difference-of-convex (ADC) functions* (see section 2.1 for the formal definition). It is important to note that ADC functions are ubiquitous. In particular, the inner functions f_p in (2) and $\mathrm{VaR}_\alpha[\, c(\cdot, Z)]$ in (3) are instances of ADC functions. In fact, based on the result recently shown in [26], any lsc function is ADC.

With this new class of functions in hand, we have made a first step to understand the variational properties of the composite ADC minimization problem (CP_0), including an in-depth analysis of its necessary optimality conditions. The novel optimality conditions are defined through a handy approximation of the subdifferential ∂f_p that explores the ADC structure of f_p. Using a notion of *epi-convergence*, we further show that these optimality conditions are necessary conditions for any local solution of (CP_0). Additionally, we propose a double-loop algorithm for (CP_0), where the outer loop dynamically updates the DC functions approximating each f_p, and the inner loop finds an approximate stationary point of the resulting composite DC problem through successive convex approximations. It can be shown that any accumulation point of the sequence generated by our algorithm satisfies the newly introduced necessary optimality conditions.

Our strategy to handle the nonsmooth and possibly discontinuous inner function f_p through a sequence of DC functions shares certain similarities with the approximation frameworks in the existing literature. For instance, Ermoliev et al. [13] have designed smoothing approximations for lsc functions utilizing convolutions with bounded mollifier sequences, a technique akin to local "averaging". Research has sought to identify conditions that ensure gradient consistency for the smoothing approximation of composite nonconvex functions [8, 7, 5, 6]. Notably, Burke and Hoheisel [5] have emphasized the importance of epi-convergence for the approximating sequence, a less stringent requirement than the continuous convergence assumed in earlier works [8, 2]. In recent work, Royset [27] has studied the consistent approximation of the composite optimization in terms of the global minimizers and stationary solutions, where the inner function is assumed to be locally Lipschitz continuous. Our notion of subdifferentials and optimality conditions for (CP_0) takes inspiration from these works but adapts to accommodate nonsmooth approximating sequences that exhibit the advantageous property of being DC.

The rest of the paper is organized as follows. Section 2 presents a class of ADC functions and introduces a new associated notion of subdifferential. In section 3, we investigate the necessary optimality conditions for problem (CP_0). Section 4 is devoted to an algorithmic framework for solving (CP_0) and its convergence analysis to the newly introduced optimality conditions. The paper ends with a concluding section.

Notation and Terminology. We write $\mathbb{R}^n$ as the n-dimensional Euclidean space equipped with the inner product $\langle x, y\rangle = x^\top y$ and the induced norm $\|x\| \triangleq \sqrt{x^\top x}$. We use the symbol $\mathbb{B}(\bar{x}, \delta)$ to denote the Euclidean ball $\{x \in \mathbb{R}^n \mid \|x - \bar{x}\| \leq \delta\}$. The set of nonpositive, nonnegative and positive real numbers are denoted as $\mathbb{R}_-$, $\mathbb{R}_+$ and $\mathbb{R}_{++}$, respectively, and the set of nonnegative integers is denoted as $\mathbb{N}$. Notation $\{t^k\}$ is employed to abbreviate any sequence $\{t^k\}_{k\geq 0}$, wherein the elements may take the form of points, sets, or functions. By $t^k \to t$ and $t^k \to_N t$, we mean that the sequence $\{t^k\}$ and the subsequence $\{t^k\}_{k\in N}$ indexed by $N \subset \mathbb{N}$ converge to t, respectively. We further write $\mathbb{N}_\infty^\sharp \triangleq \{N \subset \mathbb{N} \mid N \text{infinite}\}$ and $\mathbb{N}_\infty \triangleq \{N \mid \mathbb{N} \setminus N \text{finite}\}$.

Given two sets A and B in $\mathbb{R}^n$ and a scalar $\lambda \in \mathbb{R}$, the Minkowski sum and the scalar multiple are defined as $A + B \triangleq \{a + b \mid a \in A, b \in B\}$ and $\lambda\, A \triangleq \{\lambda\, a \mid a \in A\}$. We also define $0 \cdot \emptyset = \{0\}$ and $\lambda \cdot \emptyset = \emptyset$ whenever $\lambda \neq 0$. When A and B are nonempty and closed, we define the one-sided deviation of A from B as $\mathbb{D}(A, B) \triangleq \sup_{x\in A} \operatorname{dist}(x, B)$, where $\operatorname{dist}(x, B) \triangleq \inf_{y\in B} \|y - x\|$. The Hausdorff distance between A and B is given by $\mathbb{H}(A, B) \triangleq \max\{\mathbb{D}(A, B), \mathbb{D}(B, A)\}$. The boundary and interior of A are denoted by $\operatorname{bdry}(A)$ and $\operatorname{int}(A)$. The topological closure and the convex hull of A are indicated by $\operatorname{cl}(A)$ and $\operatorname{con} A$. We let $\delta_A(x)$ be the indicator function of A, i.e., $\delta_A(x) = 0$ for $x \in A$ and $\delta_A(x) = +\infty$ for $x \notin A$.

For a sequence of sets $\{C^k\}$, we define its outer limit as

$$\operatorname*{Lim\,sup}_{k\to+\infty} C^k \triangleq \{u \mid \exists\, N \in \mathbb{N}_\infty^\sharp,\ u^k \to_N u \text{with } u^k \in C^k\},$$

and the horizon outer limit as

$$\operatorname*{Lim\,sup}_{k\to+\infty}{}^{\infty} C^k \triangleq \{0\} \cup \left\{ u \mid \exists\, N \in \mathbb{N}^\sharp_\infty,\ \lambda_k \downarrow 0,\ \lambda_k u^k \to_N u \text{with } u^k \in C^k \right\}.$$

The outer limit of a set-valued mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is defined as

$$\operatorname*{Lim\,sup}_{x\to\bar{x}} S(x) \triangleq \bigcup_{x^k\to\bar{x}} \operatorname*{Lim\,sup}_{k\to+\infty} S(x^k) = \{u \mid \exists\, x^k \to \bar{x},\ u^k \to u \text{with } u^k \in S(x^k)\} \quad \bar{x} \in \mathbb{R}^n.$$

We say S is outer semicontinuous (osc) at $\bar{x} \in \mathbb{R}^n$ if $\operatorname{Limsup}_{x\to\bar{x}} S(x) \subset S(\bar{x})$.

The regular normal cone and the limiting normal cone of a set $C \subset \mathbb{R}^n$ at $\bar{x} \in C$ are given by

$$\widehat{\mathcal{N}}_C(\bar{x}) \triangleq \left\{ v \,\middle|\, v^\top (x - \bar{x}) \le o(\|x - \bar{x}\|) \text{for all } x \in C \right\} \quad \text{and} \quad \mathcal{N}_C(\bar{x}) \triangleq \operatorname*{Lim\,sup}_{x(\in C)\to\bar{x}} \widehat{\mathcal{N}}_C(x).$$

The proximal normal cone of a set C at $\bar{x} \in C$ is defined as $\mathcal{N}^p_C(\bar{x}) \triangleq \{\lambda(x - \bar{x}) \mid \bar{x} \in P_C(x), \lambda \ge 0\}$, where P_C is the projection onto C that maps any x to the set of points in C that are closest to x.

For an extended-real-valued function $f : \mathbb{R}^n \to \overline{\mathbb{R}} \triangleq \mathbb{R} \cup \{\pm\infty\}$, we write its effective domain as $\operatorname{dom} f \triangleq \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$, and the epigraph as $\operatorname{epi} f \triangleq \{(x, \alpha) \in \mathbb{R}^{n+1} \mid \alpha \ge f(x)\}$. We say f is proper if $\operatorname{dom} f$ is nonempty and $f(x) > -\infty$ for all $x \in \mathbb{R}^n$. We adopt the common rules for extended arithmetic operations, including the lower and upper limits of a sequence of scalars in $\overline{\mathbb{R}}$ (cf. [25, Chapter 1(E)]).

Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper function. We write $x \to_f \bar{x}$, if $x \to \bar{x}$and $f(x) \to f(\bar{x})$. The regular subdifferential and the limiting subdifferential of f at $\bar{x} \in \operatorname{dom} f$ are respectively defined as

$$\widehat{\partial} f(\bar{x}) \triangleq \{v \mid f(x) \ge f(\bar{x}) + v^\top (x - \bar{x}) + o(\|x - \bar{x}\|) \text{for all } x\} \quad \text{and} \quad \partial f(\bar{x}) \triangleq \operatorname*{Lim\,sup}_{x\to_f\bar{x}} \widehat{\partial} f(x).$$

For any $\bar{x} \notin \operatorname{dom} f$, we set $\widehat{\partial} f(\bar{x}) = \partial f(\bar{x}) = \emptyset$. When f is locally Lipschitz continuous at $\bar{x}$, $\operatorname{con} \partial f(\bar{x})$ equals to the Clarke subdifferential $\partial_C f(\bar{x})$. We further say f is subdifferentially regular at $\bar{x} \in \operatorname{dom} f$ if f is lsc at $\bar{x}$ and $\widehat{\partial} f(\bar{x}) = \partial f(\bar{x})$. When f is proper and convex, $\widehat{\partial} f$, ∂f, and $\partial_C f$ coincide with the concept of the subdifferential in convex analysis.

Finally, we introduce the notion of function convergence. A sequence of functions $\{f^k : \mathbb{R}^n \to \overline{\mathbb{R}}\}$ is said to converge pointwise to $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, written $f^k \xrightarrow{\text{p}} f$, if $\lim_{k\to+\infty} f^k(x) = f(x)$ for any $x \in \mathbb{R}^n$. The sequence $\{f^k\}$ is said to epi-converge to f, written $f^k \xrightarrow{\text{e}} f$, if for any x, it holds

$$\begin{cases} \operatorname*{liminf}_{k\to+\infty} f^k(x^k) \ge f(x) & \text{for every sequence } x^k \to x, \\ \operatorname*{limsup}_{k\to+\infty} f^k(x^k) \le f(x) & \text{for some sequence } x^k \to x. \end{cases}$$

The sequence $\{f^k\}$ is said to converge continuously to f, written $f^k \xrightarrow{\text{c}} f$, if $\lim_{k\to+\infty} f^k(x^k) = f(x)$ for any x and any sequence $x^k \to x$.

2 Approachable difference-of-convex functions.

In this section, we formally introduce a class of functions that can be asymptotically approximated by DC functions. A new concept of the subdifferential that is defined through the approximating functions is proposed. At the end of this section, we provide several examples that demonstrate the introduced concepts.

2.1 Definitions and properties.

An extended-real-valued function can be approximated by a sequence of functions in various convergent notions, as comprehensively investigated in [25, Chapter 7(A-C)]. Among these approaches, epi-convergence has a notable advantage in its ability to preserve the minimizers [25, Theorem 7.31]. Our focus lies on a particular class of approximating functions, wherein each function exhibits a DC structure.

Definition 1. *A function f is said to be DC on its domain if there exist proper, lsc and convex functions $g, h : \mathbb{R}^n \to \overline{\mathbb{R}}$ such that $\operatorname{dom} f = [\operatorname{dom} g \cap \operatorname{dom} h]$ and $f(x) = g(x) - h(x)$ for any $x \in \operatorname{dom} f$.*

With this definition, we introduce the concept of ADC functions.

Definition 2 (ADC functions)**.** *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper function.*
(a) f is said to be pointwise approachable DC (p-ADC) if there exist proper functions $\{f^k : \mathbb{R}^n \to \overline{\mathbb{R}}\}$, DC on their respective domains, such that $f^k \xrightarrow{p} f$.
(b) f is said to be epigraphically approachable DC (e-ADC) if there exist proper functions $\{f^k : \mathbb{R}^n \to \overline{\mathbb{R}}\}$, DC on their respective domains, such that $f^k \xrightarrow{e} f$.
(c) f is said to be continuously approachable DC (c-ADC) if there exist proper functions $\{f^k : \mathbb{R}^n \to \overline{\mathbb{R}}\}$, DC on their respective domains, such that $f^k \xrightarrow{c} f$.
A function f confirming any of these properties are said to be an ADC function associated with $\{f^k\}$. By a slight abuse of notation, we denote the DC decompositions of each f^k as $f^k = g^k - h^k$, although the equality may only hold for $x \in \operatorname{dom} f^k$.

A p-ADC function may not be lsc. An example is given by $f(x) = \mathbf{1}_{\{0\}}(x) + 2 \cdot \mathbf{1}_{(0,+\infty)}(x)$, where for a set $C \subset \mathbb{R}^n$, we write $\mathbf{1}_C(x) = 1$ if $x \in C$ and $\mathbf{1}_C(x) = 0$ if $x \notin C$. In this case, f is not lsc at $x = 0$. However, f is p-ADC associated with $f^k(x) = \max(0, 2kx + 1) - \max(0, 2kx - 1)$. In contrast, any e-ADC function must be lsc [25, Proposition 7.4(a)], and any c-ADC function is continuous [25, Theorem 7.14].

The relationships among different notions of function convergence, including the unaddressed uniform convergence, have been thoroughly examined in [25]. Generally, pointwise convergence and epi-convergence do not imply one another, but they coincide when the sequence $\{f^k\}$ is asymptotically equi-lsc everywhere [25, Theorem 7.10]. In addition, f^k continuously converges to f if and only if both $f^k \xrightarrow{e} f$ and $(-f^k) \xrightarrow{e} (-f)$ are satisfied [25, Theorem 7.11]. While epi-convergence is often challenging to verify, it is simpler for a monotonic sequence $\{f^k\}$ that converges pointwise to f [25, Proposition 7.4(c-d)].

2.2 Subdifferentials of ADC functions.

Characterizing the limiting and Clarke subdifferentials can be challenging when dealing with functions that exhibit complex composite structures. Our focus in this subsection is on numerically computable approximations of the limiting subdifferentials.

Definition 3 (approximate subdifferentials)**.** *Consider an ADC function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ associated with $\{f^k = g^k - h^k\}$. The approximate subdifferential of f (associated with $\{f^k = g^k - h^k\}$) at*

$\bar{x} \in \mathbb{R}^n$ *is defined as*

$$\partial_A f(\bar{x}) \triangleq \bigcup_{x^k \to \bar{x}} \operatorname*{Lim\,sup}_{k \to +\infty} \left[\partial g^k(x^k) - \partial h^k(x^k)\right].$$

The approximate horizon subdifferential of f (associated with $\{f^k = g^k - h^k\}$) at $\bar{x} \in \mathbb{R}^n$ is defined as

$$\partial_A^\infty f(\bar{x}) \triangleq \bigcup_{x^k \to \bar{x}} \operatorname*{Lim\,sup^\infty}_{k \to +\infty} \left[\partial g^k(x^k) - \partial h^k(x^k)\right].$$

Unlike the limiting subdifferential which requires $x^k \to_f \bar{x}$, $\partial_A f(x)$ is defined using sequences $x^k \to \bar{x}$ without necessitating the convergence of function values. The following proposition establishes useful properties of the approximate (horizon) subdifferential mappings.

Proposition 1. *The following statements hold.*
(a) The mappings $x \mapsto \partial_A f(x)$ and $x \mapsto \partial_A^\infty f(x)$ are osc.
(b) Let $\bar{x} \notin \operatorname{dom} f$. Then $\partial_A f(\bar{x}) = \emptyset$ if for any sequence $x^k \to \bar{x}$, we have $x^k \notin \operatorname{dom} f^k$ for all k sufficiently large. The latter condition is particularly satisfied whenever $\operatorname{dom} f$ is closed and $\operatorname{dom} f^k \subset \operatorname{dom} f$ for all k sufficiently large.

Proof. The results in (a) follow directly from the definition of the approximate (horizon) subdifferential mappings. To show (b), note that for any $x^k \to \bar{x} \notin \operatorname{dom} f$, we have $[\partial g^k(x^k) - \partial h^k(x^k)] = \emptyset$ for all k sufficiently large due to $x^k \notin \operatorname{dom} f^k = [\operatorname{dom} g^k \cap \operatorname{dom} h^k]$. Thus, $\partial_A f(\bar{x}) = \emptyset$ for any $\bar{x} \notin \operatorname{dom} f$. The proof is then completed. □

Proposition 1(b) presents a sufficient condition for $\partial_A f(\bar{x}) = \partial f(\bar{x}) = \emptyset$ at any $\bar{x} \notin \operatorname{dom} f$. In the subsequent analysis, we restrict our attention to $\bar{x} \in \operatorname{dom} f$. Admittedly, the set $\partial_A f(\bar{x})$ depends on the approximating sequence $\{f^k\}$ and the DC decomposition of each f^k that may contain some irrelevant information concerning the local variational geometry of epi f. In fact, for a given ADC function f, we can make the set $\partial_A f(\bar{x})$ arbitrarily large by adding the same extra nonsmooth functions to both g^k and h^k. By Attouch's theorem (see for example [25, Theorem 12.35]), for proper, lsc, convex functions f and $\{f^k\}$, if $f^k \xrightarrow{e} f$, we immediately have $\partial_A f = \partial f$ when taking $g^k = f^k$ and $h^k = 0$. In what follows, we further explore the relationships among $\partial_A f$ and other commonly employed subdifferentials in the literature beyond the convex setting. As it turns out, with respect to an arbitrary DC decomposition of f^k that is lsc, $\partial_A f(\bar{x})$ contains the limiting subdifferential of f at any $\bar{x} \in \operatorname{dom} f$ whenever $f^k \xrightarrow{e} f$.

Theorem 1 (subdifferentials relationships)**.** *Consider an ADC function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. The following statements hold for any $\bar{x} \in \operatorname{dom} f$.*
(a) If f is e-ADC associated with $\{f^k\}$ and f^k is lsc, then $\partial f(\bar{x}) \subset \partial_A f(\bar{x})$ and $\partial^\infty f(\bar{x}) \subset \partial_A^\infty f(\bar{x})$.
(b) If f is locally Lipschitz continuous and bounded from below, then there exists a sequence of DC functions $\{f^k = g^k - h^k\}$ such that $f^k \xrightarrow{c} f$, $\partial f(\bar{x}) \subset \partial_A f(\bar{x}) \subset \partial_C f(\bar{x})$, and $\partial_A^\infty f(\bar{x}) = \{0\}$. Consequently, $\operatorname{con} \partial_A f(\bar{x}) = \partial_C f(\bar{x})$, the set $\partial_A f(\bar{x})$ is nonempty and bounded, and $\partial f(\bar{x}) = \partial_A f(\bar{x})$ when f is subdifferentially regular at $\bar{x}$.

Proof. (a) Since f is e-ADC, it must be lsc. By using epi-convergence of $\{f^k\}$ to f, we know from [25, Corollary 8.47(b)] and [25, Proposition 8.46(e)] that any element of $\partial f(\bar{x})$ can be generated as a limit of regular subgradients at x^k with $x^k \to_N \bar{x}$ and $f^k(x^k) \to_N f(\bar{x})$ for some $N \in \mathbb{N}_\infty$. Indeed, we can further restrict $x^k \in \operatorname{dom} f^k$ since $f^k(x^k) \to_N f(\bar{x})$ and $\bar{x} \in \operatorname{dom} f$. Then, we have

$$\partial f(\bar{x}) \subset \bigcup_{x^k(\in \operatorname{dom} f^k) \to \bar{x}} \operatorname*{Lim\,sup}_{k \to +\infty} \widehat{\partial} f^k(x^k) \subset \bigcup_{x^k(\in \operatorname{dom} f^k) \to \bar{x}} \operatorname*{Lim\,sup}_{k \to +\infty} \left[\partial g^k(x^k) - \partial h^k(x^k)\right] \subset \partial_A f(\bar{x}),$$

where the second inclusion can be verified as follows: Firstly, due to the lower semicontinuity of f^k and h^k, and $x^k \in \operatorname{dom} f^k \subset \operatorname{dom} g^k$, it follows from the sum rule of regular subdifferentials [25, Corollary 10.9] that $\widehat{\partial} g^k(x^k) \supset \widehat{\partial} f^k(x^k) + \widehat{\partial} h^k(x^k)$. Consequently, $\widehat{\partial} f^k(x^k) \subset \widehat{\partial} g^k(x^k) - \widehat{\partial} h^k(x^k) = \partial g^k(x^k) - \partial h^k(x^k)$ since g^k and h^k are proper and convex [25, Proposition 8.12]. Similarly, by [25, Corollary 8.47(b)], we have

$$\partial^\infty f(\bar{x}) \subset \bigcup_{x^k(\in \operatorname{dom} f^k) \to \bar{x}} \operatorname*{Lim\,sup^\infty}_{k \to +\infty} \widehat{\partial} f^k(x^k) \subset \bigcup_{x^k(\in \operatorname{dom} f^k) \to \bar{x}} \operatorname*{Lim\,sup^\infty}_{k \to +\infty} \left[\partial g^k(x^k) - \partial h^k(x^k)\right] \subset \partial_A^\infty f(\bar{x}).$$

(b) For a locally Lipschitz continuous function f, consider its Moreau envelope $e_\gamma f(x) \triangleq \inf_z\{f(z) + \|z - x\|^2/(2\gamma)\}$ and the set-valued mapping $P_{\gamma f}(x) \triangleq \operatorname{argmin}_z\{f(z) + \|z - x\|^2/(2\gamma)\}$. For any sequence $\gamma_k \downarrow 0$, we demonstrate in the following that $\{f^k \triangleq e_{\gamma_k} f\}$ is the desired sequence of approximating functions. Firstly, since f is bounded from below, it must be prox-bounded and, thus, each f^k is continuous and $f^k(\bar{x}) \uparrow f(\bar{x})$ for all $\bar{x}$ (cf. [25, Theorem 1.25]). By the continuity of f and f^k, we have $f^k \xrightarrow{\text{c}} f$ from [25, Proposition 7.4(c-d)]. It then follows from part (a) that $\partial f(\bar{x}) \subset \partial_A f(\bar{x})$. Consider the following DC decomposition for each f^k:

$$f^k(x) = \underbrace{\frac{\|x\|^2}{2\gamma_k}}_{\triangleq g^k(x)} - \underbrace{\sup_{z \in \mathbb{R}^n} \left\{-f(z) - \frac{\|z\|^2}{2\gamma_k} + \frac{z^\top x}{\gamma_k}\right\}}_{\triangleq h^k(x)} \qquad x \in \mathbb{R}^n.$$

It is clear that $f(z) + \|z\|^2/(2\gamma_k) + z^\top x/\gamma_k$ is level-bounded in z locally uniformly in x, since for any $r \in \mathbb{R}$ and any bounded set $X \subset \mathbb{R}^n$, the set

$$\left\{z \in \mathbb{R}^n \,\middle|\, x \in X, f(z) + \frac{\|z\|^2}{2\gamma_k} - \frac{z^\top x}{\gamma_k} \le r\right\} \subset \left\{z \in \mathbb{R}^n \,\middle|\, x \in X,\ \|z - x\|^2 \le \|x\|^2 + 2\gamma_k \left[r - \inf_z f(z)\right]\right\}$$

is bounded. Due to the level-boundedness condition, we can apply the subdifferential formula of the parametric minimization [25, Theorem 10.13] to get

$$\partial(-h^k)(x) \subset \bigcup_{z \in P_{\gamma_k f}(x)} \left\{y \,\middle|\, (0, y) \in \partial_{(z,x)} \left(f(z) + \frac{\|z\|^2}{2\gamma_k} - \frac{z^\top x}{\gamma_k}\right)\right\} \subset \bigcup_{z \in P_{\gamma_k f}(x)} \left\{\partial f(z) - \frac{x}{\gamma_k}\right\},$$

where the last inclusion is due to the calculus rules [25, Proposition 10.5 and Exercise 8.8(c)]. Since h^k is convex, we have $-\partial h^k(x) = \partial_C(-h^k)(x) = \operatorname{con} \partial(-h^k)(x)$ by [25, Theorem 9.61], which further yields that

$$\left[\partial g^k(x) - \partial h^k(x)\right] \subset \operatorname{con} \bigcup \{\partial f(z) \,|\, z \in P_{\gamma_k f}(x)\} \qquad \forall\, x \in \mathbb{R}^n,\ k \ge 0. \tag{4}$$

For any $x^k \to \bar{x}$ and any $z^k \in P_{\gamma_k f}(x^k)$, we have

$$\frac{1}{2\gamma_k}\|z^k - x^k\|^2 + \inf_x f(x) \le \frac{1}{2\gamma_k}\|z^k - x^k\|^2 + f(z^k) \le \frac{1}{2\gamma_k}\|\bar{x} - x^k\|^2 + f(\bar{x}).$$

Then, $\|z^k - x^k\| \leq \sqrt{\|\bar{x} - x^k\|^2 + 2\gamma_k[f(\bar{x}) - \inf_x f(x)]} \to 0$ due to the assumption that f is bounded from below and therefore $z^k \to \bar{x}$. By the locally Lipschitz continuity of f, it follows from [25, Theorem 9.13] that the mapping $\partial f : x \mapsto \partial f(x)$ is locally bounded at $\bar{x}$. Thus, there is a bounded set S such that $\bigcup\{\partial f(z^k) \mid z^k \in P_{\gamma_k f}(x^k)\} \subset S$ for all k sufficiently large. It follows directly from [25, Example 4.22] and the definition of the approximate horizon subdifferential that $\partial_A^\infty f(\bar{x}) = \{0\}$.

Next, we will prove the inclusion $\partial_A f(\bar{x}) \subset \partial_C f(\bar{x})$. For any $u \in \partial_A f(\bar{x})$, from (4), there exist sequences of vectors $x^k \to \bar{x}$ and $u^k \to u$ with each u^k taken from the convex hull of a bounded set $\bigcup\{\partial f(z^k) \mid z^k \in P_{\gamma_k f}(x^k)\}$. By Carathéodory's Theorem (see, e.g. [22, Theorem 17.1]), any point in the convex hull of a bounded set in $\mathbb{R}^n$ can be expressed as a convex combination of $(n+1)$ points in this set. Therefore, for each k, we have $u^k = \sum_{i=1}^{n+1} \lambda_{k,i}\, v^{k,i}$ for some nonnegative scalars $\{\lambda_{k,i}\}_{i=1}^{n+1}$ with $\sum_{i=1}^{n+1} \lambda_{k,i} = 1$ and a sequence $\left\{v^{k,i} \in \partial f(z^{k,i})\right\}_{i=1}^{n+1}$ with $\{z^{k,i} \in P_{\gamma_k f}(x^k)\}_{i=1}^{n+1}$. It is easy to see that the sequences $\{\lambda_{k,i}\}_{k\geq 0}$ and $\{v^{k,i}\}_{k\geq 0}$ are bounded for each i. We can then obtain convergent subsequences $\lambda_{k,i} \to_N \bar{\lambda}_i \geq 0$ with $\sum_{i=1}^{n+1} \bar{\lambda}_i = 1$ and $v^{k,i} \to_N \bar{v}^{\,i}$ for each i. Since $z^{k,i} \to \bar{x}$, we have $\bar{v}^{\,i} \in \partial f(\bar{x})$ by using the outer semicontinuity of ∂f. Thus, $u^k \to_N u = \sum_{i=1}^{n+1} \bar{\lambda}_i\, \bar{v}^{\,i} \in \operatorname{con} \partial f(\bar{x}) = \partial_C f(\bar{x})$. This implies $\partial_A f(\bar{x}) \subset \partial_C f(\bar{x})$. The rest statements in (b) follow from the fact that $\partial_C f(\bar{x})$ is nonempty and bounded whenever f is locally Lipschitz continuous [25, Theorem 9.61]. $\square$

Under suitable assumptions, Theorem 1(b) guarantees the existence of an ADC decomposition that has its approximate subdifferential contained in the Clarke subdifferential of the original function. Notably, this decomposition may not always be practically useful due to the necessity of computing the Moreau envelope for a generally nonconvex function. Another noteworthy remark is that the assumptions and results of Theorem 1 can be localized to any specific point $\bar{x}$. This can be accomplished by defining a notion of "local epi-convergence" at $\bar{x}$ and extending the result of [25, Corollary 8.47] accordingly.

2.3 Examples of ADC functions.

In this subsection, we provide examples of ADC functions, including functions that are discontinuous relative to their domains, with explicit and computationally tractable approximating sequences. Moreover, we undertake an investigation into the approximate subdifferentials of these ADC functions.

Example 2.1 (implicitly convex-concave functions) The concept of *implicitly convex-concave (icc) functions* is introduced in the monograph [11], and is further generalized to extended-real-valued functions in [16]. A proper function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is icc if there exists a lifted function $\overline{f} : \mathbb{R}^n \times \mathbb{R}^n \to \overline{\mathbb{R}}$ such that the following three conditions hold:
(i) $\overline{f}(z, x) = +\infty$ if $z \notin \operatorname{dom} f, x \in \mathbb{R}^n$, and $\overline{f}(z, x) = -\infty$ if $z \in \operatorname{dom} f, x \notin \operatorname{dom} f$;
(ii) $\overline{f}(\cdot, x)$ is convex for any fixed $x \in \operatorname{dom} f$, and $\overline{f}(z, \cdot)$ is concave for any fixed $z \in \operatorname{dom} f$;
(iii) $f(x) = \overline{f}(x, x)$ for any $x \in \operatorname{dom} f$.

A notable example of icc functions is the optimal value function f_p in (1), which is associated with the lifted function defined by (the subscripts/superscripts p are omitted for brevity):

$$\overline{f}(z, x) \triangleq \inf_{y \in \mathbb{R}^{n_2}} \left\{ (c + Cx)^\top y + \frac{1}{2} y^\top Q\, y \;\middle|\; Az + By \leq b \right\} \qquad (x, z) \in \operatorname{dom} f \times \operatorname{dom} f. \tag{5}$$

Let $\partial_1 \overline{f}(\cdot, x)$ and $\partial_2(-\overline{f})(z, \cdot)$ denote the subdifferentials of the convex functions $\overline{f}(\cdot, x)$ and $(-\overline{f})(z, \cdot)$, respectively, for any $(x, z) \in \operatorname{dom} f \times \operatorname{dom} f$. For any $\gamma > 0$, the *partial Moreau envelope* of an icc function f associated with $\overline{f}$ is given by

$$\inf_{z \in \mathbb{R}^{n_2}} \left\{ \overline{f}(z, x) + \frac{1}{2\gamma} \|z - x\|^2 \right\} = \underbrace{\frac{\|x\|^2}{2\gamma}}_{\triangleq g_\gamma(x)} - \underbrace{\sup_{z \in \mathbb{R}^{n_2}} \left\{ -\overline{f}(z, x) - \frac{\|z\|^2}{2\gamma} + \frac{z^\top x}{\gamma} \right\}}_{\triangleq h_\gamma(x)} \qquad x \in \operatorname{dom} f.$$

This decomposition, established in [16], offers computational advantages compared to the standard Moreau envelope, as the maximization problem defining h_γ is concave in z for any fixed x. In what follows, we present new results on the conditions under which the icc function f is e-ADC and c-ADC based on the partial Moreau envelope. Additionally, we explore a relationship between $\partial_A f(\bar{x})$ and $\partial_1 \overline{f}(\bar{x}, \bar{x}) - \partial_2(-\overline{f})(\bar{x}, \bar{x})$, where the latter is known to be an outer estimate of $\partial_C f(\bar{x})$ [11, Proposition 4.4.26]. We refer readers to Appendix A for the proof.

Proposition 2. *Let $f : \mathbb{R}^{n_1} \to \overline{\mathbb{R}}$ be a proper, lsc, icc function associated with $\overline{f}$, where $\operatorname{dom} f$ is closed and $\overline{f}$ is lsc on $\mathbb{R}^{n_1} \times \operatorname{dom} f$, bounded below on $\operatorname{dom} f \times \operatorname{dom} f$, and continuous relative to $\operatorname{int}(\operatorname{dom} f) \times \operatorname{int}(\operatorname{dom} f)$. Given a sequence of scalars $\{\gamma_k\} \downarrow 0$, we have:*
(a) f is e-ADC associated with $\{f^k\}$, where each $f^k(x) \triangleq g_{\gamma_k}(x) - h_{\gamma_k}(x) + \delta_{\operatorname{dom} f}(x)$. In addition, if $\operatorname{dom} f = \mathbb{R}^{n_1}$, then f is c-ADC associated with $\{f^k\}$.
(b) $\partial_A f(\bar{x}) \subset \partial_1 \overline{f}(\bar{x}, \bar{x}) - \partial_2(-\overline{f})(\bar{x}, \bar{x})$ and $\partial_A^\infty f(\bar{x}) = \{0\}$ for any $\bar{x} \in \operatorname{int}(\operatorname{dom} f)$.

Example 2.2 (VaR for continuous random variables) Given a random variable $Y : \Omega \to \mathbb{R}$, we consider its VaR mentioned in Example 1.2 and introduce the upper conditional VaR. The *upper conditional VaR* for Y at a confidence level $\alpha \in (0, 1)$ is defined as $\mathrm{CVaR}_\alpha^+(Y) \triangleq \mathbb{E}[Y \mid Y > \mathrm{VaR}_\alpha(Y)]$. Given a constant $\alpha \in (0, 1)$ and any $k > 1/\alpha$, we define

$$g^k(x) \triangleq [k(1-\alpha) + 1]\, \mathrm{CVaR}^+_{\alpha - 1/k}[\, c(x, Z)], \quad h^k(x) \triangleq k(1-\alpha)\, \mathrm{CVaR}^+_\alpha[\, c(x, Z)] \quad x \in \mathbb{R}^n. \tag{6}$$

The proof of the following proposition can be found in Appendix A.

Proposition 3. *Let $c : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be a lsc function. Suppose that $c(\cdot, z)$ is convex for any $z \in \mathbb{R}^m$, and $c(x, Z)$ is continuously distributed, induced by a random vector $Z : \Omega \to \mathbb{R}^m$, with $\mathbb{E}[\, |c(x, Z)|\,] < +\infty$ for any $x \in \mathbb{R}^n$. For any given constant $\alpha \in (0, 1)$, the following properties hold.*
(a) $VaR_\alpha[\, c(x, Z)]$ is lsc and e-ADC associated with $\{g^k - h^k\}$. Additionally, if $c(\cdot, \cdot)$ is continuous, then $VaR_\alpha[\, c(x, Z)]$ is continuous and c-ADC associated with $\{g^k - h^k\}$.
(b) If there exists a measurable function $\kappa : \mathbb{R}^m \to \mathbb{R}_+$ such that $\mathbb{E}[\, \kappa(Z)] < +\infty$ and $|c(x, z) - c(x', z)| \le \kappa(z) \|x - x'\|$ for all $x, x' \in \mathbb{R}^n$ and $z \in \mathbb{R}^m$, then for any $\bar{x} \in \mathbb{R}^n$,

$$\partial_A\, VaR_\alpha[\, c(\cdot, Z)](\bar{x}) = \bigcup_{x^k \to \bar{x}} \operatorname*{Lim\,sup}_{k \to +\infty} \mathbb{E}\left[\partial_x\, c(x^k, Z) \,\middle|\, VaR_{\alpha - 1/k}[\, c(x^k, Z)] < c(x^k, Z) < VaR_\alpha[\, c(x^k, Z)] \right],$$

where $\mathbb{E}[\, \mathcal{A}(x, Z)]$ represents the expectation of a random set-valued mapping $\mathcal{A}$, defined as the set of $\mathbb{E}[\, a(x, Z)]$ for all measurable selections $a(x, Z) \in \mathcal{A}(x, Z)$.

3 The convex composite ADC functions and minimization.

This section aims to derive necessary optimality conditions for (CP_0), particularly focusing on the inner function f_p that lacks local Lipschitz continuity. To prepare for it, we make the following assumption:

Assumption 1 For each p, we have

(a) f_p is an ADC function associated with $\{f_p^k = g_p^k - h_p^k\}_{k\geq 0}$, and $\mathrm{dom}\, g_p^k = \mathrm{dom}\, h_p^k = \mathbb{R}^n$;

(b) $-\infty < \liminf\limits_{x'\to x,\, k\to+\infty} f_p^k(x') \leq \limsup\limits_{x'\to x,\, k\to+\infty} f_p^k(x') < +\infty$ for all $x \in \mathbb{R}^n$;

(c) $\left[F_p^k \triangleq \varphi_p \circ f_p^k\right] \xrightarrow{\mathrm{e}} F_p$.

From Assumption 1(a), each f_p^k is locally Lipschitz continuous since any real-valued convex function is locally Lipschitz continuous. Obviously, $f_p^k \xrightarrow{\mathrm{c}} f_p$ is sufficient for Assumption 1(b) to hold. Under epi-convergence $f_p^k \xrightarrow{\mathrm{e}} f_p$, we have $\liminf_{x'\to x, k\to+\infty} f_p^k(x') \geq f_p(x) > -\infty$ for each p at any $x \in \mathbb{R}^n$. However, $\limsup_{x'\to x, k\to+\infty} f_p^k(x') < +\infty$ does not hold trivially. For example, consider a continuous function f and

$$f^k(x) = \begin{cases} f(x) + k^2 x + k & \text{if } x \in [-1/k, 0] \\ f(x) - k^2 x + k & \text{if } x \in (0, 1/k] \\ f(x) & \text{otherwise} \end{cases},$$

which results in $f^k \xrightarrow{\mathrm{e}} f$ but $\limsup_{k\to+\infty} f^k(0) = +\infty$. Additionally, Assumption 1(b) ensures that at each point x and for any sequence $x^k \to x$, the sequence $\{f_p^k(x^k)\}_{k\geq 0}$ must be bounded.

Sufficient conditions for Assumption 1(c) can be found in [25, Exercise 7.8(c)] and [27, Theorem 2.4]. Furthermore, Assumption 1(c) guarantees that each $F_p = \varphi_p \circ f_p$ is lsc, yet it doesn't necessarily lead to $\sum_{p=1}^m F_p^k \xrightarrow{\mathrm{e}} \sum_{p=1}^m F_p$. Let ε-$\arg\min f \triangleq \{x \mid f(x) \leq \inf f + \varepsilon\}$ be the set of points that minimize a function f to within ε. Hence, Assumption 1(c) alone may not be sufficient to ensure that every accumulation point of $\{x^k\}$ with $x^k \in \varepsilon_k$-$\arg\min \sum_{p=1}^m F_p^k$ for $\varepsilon_k \downarrow 0$ qualifies as a minimizer of $\sum_{p=1}^m F_p$. To maintain epi-convergence under addition of functions, one may refer to the sufficient conditions in [25, Theorem 7.46].

3.1 Asymptotic stationarity under epi-convergence.

In this subsection, we introduce a novel stationarity concept for problem (CP_0), grounded in a monotonic decomposition of univariate convex functions. We demonstrate that under certain constraint qualifications, epi-convergence of approximating functions ensures this stationarity concept as a necessary optimality condition. Alongside the fact that epi-convergence results in the convergence of global optimal solutions [25, Theorem 7.31(b)], this highlights the usefulness of epi-convergence as a tool for studying the approximation of the composite problem (CP_0).

The following lemma is an extension of [11, Lemma 6.1.1] from real-valued univariate convex functions to extended-real-valued univariate convex functions.

Lemma 1 (a monotonic decomposition of univariate convex functions)**.** *Let $\varphi : \mathbb{R} \to \overline{\mathbb{R}}$ be a proper, lsc and convex function. Then there exist a proper, lsc, convex and nondecreasing function $\varphi^{\uparrow}$, as well as a proper, lsc, convex and nonincreasing function $\varphi^{\downarrow}$, such that $\varphi = \varphi^{\uparrow} + \varphi^{\downarrow}$. In addition, if* $\operatorname{int}(\operatorname{dom}\varphi) \neq \emptyset$*, the following properties hold:*
(a) For any $z_0 \in \mathbb{R}$, there exists $\delta > 0$ such that $\mathcal{N}_{\operatorname{dom}\varphi^{\uparrow}}(z) = \{0\}$ for any $z \in \mathbb{B}(z_0, \delta)$, or $\mathcal{N}_{\operatorname{dom}\varphi^{\downarrow}}(z) = \{0\}$ for any $z \in \mathbb{B}(z_0, \delta)$.
(b) $\partial\varphi(z) = \partial\varphi^{\uparrow}(z) + \partial\varphi^{\downarrow}(z)$ and $\mathcal{N}_{\operatorname{dom}\varphi^{\uparrow}}(z) \bigcap \left[-\mathcal{N}_{\operatorname{dom}\varphi^{\downarrow}}(z)\right] = \{0\}$ for any $z \in \operatorname{dom}\varphi$. Consequently, $\mathcal{N}_{\operatorname{dom}\varphi}(z) = \mathcal{N}_{\operatorname{dom}\varphi^{\uparrow}}(z) + \mathcal{N}_{\operatorname{dom}\varphi^{\downarrow}}(z)$ for any $z \in \operatorname{dom}\varphi$.

Proof. From the convexity of φ, $\operatorname{dom}\varphi$ is an interval on $\mathbb{R}$, possibly unbounded. In fact, we can explicitly construct $\varphi^{\uparrow}$ and $\varphi^{\downarrow}$ in following two cases.

<u>Case 1.</u> If φ has no direction of recession, i.e., there does not exist $d \neq 0$ such that for any z, $\varphi(z + \lambda d)$ is a nonincreasing function of $\lambda > 0$, it follows from [22, Theorem 27.2] that φ attains its minimum at some $z^* \in \operatorname{dom}\varphi$. Define

$$\varphi^{\uparrow}(z) = \begin{cases} \varphi(z^*) & \text{if } z \leq z^* \\ \varphi(z) & \text{if } z > z^* \end{cases} \quad \text{and} \quad \varphi^{\downarrow}(z) = \begin{cases} \varphi(z) - \varphi(z^*) & \text{if } z \leq z^* \\ 0 & \text{if } z > z^* \end{cases}.$$

For any $z \neq z^*$, note that

$$\mathcal{N}_{\operatorname{dom}\varphi^{\uparrow}}(z) = \begin{cases} \{0\} & \text{if } z < z^* \\ \mathcal{N}_{\operatorname{dom}\varphi}(z) & \text{if } z > z^* \end{cases} \quad \text{and} \quad \mathcal{N}_{\operatorname{dom}\varphi^{\downarrow}}(z) = \begin{cases} \mathcal{N}_{\operatorname{dom}\varphi}(z) & \text{if } z < z^* \\ \{0\} & \text{if } z > z^* \end{cases}.$$

Thus, part (a) holds except at $z_0 = z^*$. When $z^* \in \operatorname{int}(\operatorname{dom}\varphi)$, there exists $\delta > 0$ such that $\mathcal{N}_{\operatorname{dom}\varphi^{\uparrow}}(z) = \mathcal{N}_{\operatorname{dom}\varphi^{\downarrow}}(z) = \{0\}$ for any $z \in \mathbb{B}(z^*, \delta)$. Next, consider the case of $z^* \in \operatorname{bdry}(\operatorname{dom}\varphi)$. If $\varphi(z) = +\infty$ for any $z < z^*$, then $\operatorname{dom}\varphi = [z^*, r)$ or $[z^*, r]$ for some $r \in (z^*, +\infty]$ due to the convexity of $\operatorname{dom}\varphi$ and $\operatorname{int}(\operatorname{dom}\varphi) \neq \emptyset$. Thus, $\varphi^{\uparrow}$ is finite-valued in a neighborhood $\mathbb{B}(z^*, \delta)$ of z^* with $\delta > 0$ and $\mathcal{N}_{\operatorname{dom}\varphi^{\uparrow}}(z) = \{0\}$ for any $z \in \mathbb{B}(z^*, \delta)$. Likewise, if $\varphi(z) = +\infty$ for any $z > z^*$, we have $\mathcal{N}_{\operatorname{dom}\varphi^{\downarrow}}(z) = \{0\}$ for any $z \in \mathbb{B}(z^*, \delta)$ with some $\delta > 0$. Combining the arguments for $z \neq z^*$, we conclude that (a) is true.

To show part (b), observe that $\emptyset \neq \operatorname{int}(\operatorname{dom}\varphi) \subset \left[\operatorname{int}(\operatorname{dom}\varphi^{\uparrow}) \cap \operatorname{int}(\operatorname{dom}\varphi^{\downarrow})\right]$. Consequently, from [22, Theorem 23.8], we have $\partial\varphi(z) = \partial\varphi^{\uparrow}(z) + \partial\varphi^{\downarrow}(z)$ for any $z \in \mathbb{R}$. The remaining results hold trivially if $\operatorname{dom}\varphi^{\uparrow} = \mathbb{R}$ or $\operatorname{dom}\varphi^{\downarrow} = \mathbb{R}$. Now we only need to consider the case where $\operatorname{dom}\varphi^{\uparrow} = (-\infty, p]$ and $\operatorname{dom}\varphi^{\downarrow} = [q, +\infty)$ for some $p > q$ ($p \neq q$ due to $\operatorname{int}(\operatorname{dom}\varphi) \neq \emptyset$), since the cases involving open domains can be derived similarly. It is evident that $\mathcal{N}_{(-\infty,p]}(z) \cap \left[-\mathcal{N}_{[q,+\infty)}(z)\right] = \{0\}$ for any $z \in \mathbb{R}$. By [25, Theorem 6.42] and $\operatorname{dom}\varphi = \operatorname{dom}\varphi^{\uparrow} \cap \operatorname{dom}\varphi^{\downarrow}$, it holds that $\mathcal{N}_{\operatorname{dom}\varphi}(z) = \mathcal{N}_{\operatorname{dom}\varphi^{\uparrow}}(z) + \mathcal{N}_{\operatorname{dom}\varphi^{\downarrow}}(z)$ for any $z \in \operatorname{dom}\varphi$.

<u>Case 2.</u> Otherwise, there exists $d \neq 0$ such that for any $z \in \mathbb{R}$, $\varphi(z + \lambda d)$ is a nonincreasing function of $\lambda > 0$. Consequently, $\operatorname{dom}\varphi$ must be an unbounded interval on $\mathbb{R}$. Let $d = 1$ (or -1) be such a recession direction, then φ is nonincreasing (or nondecreasing) on $\mathbb{R}$. We can set $\varphi^{\uparrow} = 0$ and $\varphi^{\downarrow} = \varphi$ (or $\varphi^{\uparrow} = \varphi$ and $\varphi^{\downarrow} = 0$). Since we have shown that φ is nondecreasing or nonincreasing in Case 2, the conclusions of (a) and (b) follow directly. The proof is thus completed. $\square$

In the subsequent analysis, we use $\varphi^{\uparrow}$ and $\varphi^{\downarrow}$ to denote the monotonic decomposition of any univariate, proper, lsc, and convex function φ constructed in the proof of Lemma 1 and, in particular, we take $\varphi^{\downarrow} = 0$ whenever φ is nondecreasing. We are now ready to present the definition of asymptotically stationary points.

Definition 4 (asymptotically stationary points)**.** *Let each f_p be an ADC function associated with $\{f_p^k = g_p^k - h_p^k\}_{k\geq 0}$. For each p, define*

$$T_p(x) \triangleq \left\{ t_p \,\middle|\, \exists\, N \in \mathbb{N}_\infty^\sharp, x^k \to x \, with \; f_p^k(x^k) \to_N t_p \right\} \qquad x \in \mathbb{R}^n. \tag{7}$$

We say that $\bar{x}$ is an asymptotically stationary (A-stationary) point of problem (CP$_0$) *if for each p, there exists $y_p \in \bigcup\{\partial\varphi_p(t_p) \mid t_p \in T_p(\bar{x})\}$ such that*

$$0 \in \sum_{p=1}^{m} \left(\{ y_p \, \partial_A f_p(\bar{x}) \} \cup \left[\pm \partial_A^\infty f_p(\bar{x}) \backslash \{0\} \right] \right). \tag{8}$$

We say that $\bar{x}$ is a weakly asymptotically stationary (weakly A-stationary) point of problem (CP$_0$) *if for each p, there exist $y_{p,1} \in \bigcup\{\partial\varphi_p^\uparrow(t_p) \mid t_p \in T_p(\bar{x})\}$ and $y_{p,2} \in \bigcup\{\partial\varphi_p^\downarrow(t_p) \mid t_p \in T_p(\bar{x})\}$ such that*

$$0 \in \sum_{p=1}^{m} \left(\{ y_{p,1} \, \partial_A f_p(\bar{x}) + y_{p,2} \, \partial_A f_p(\bar{x}) \} \cup \left[\pm \partial_A^\infty f_p(\bar{x}) \backslash \{0\} \right] \right).$$

Remark 1. *(i) Given that the approximate subdifferential $\partial_A f_p$ is determined by the approximating sequence $\{f_p^k\}_{k\geq 0}$ and their corresponding DC decompositions, the notion of (weak) A-stationarity also depends on these sequences and decompositions. (ii) It follows directly from Lemma 1(b) that an A-stationary point must be a weakly A-stationary point if* $\text{int}(\text{dom}\,\varphi_p) \neq \emptyset$ *for each $p = 1, \cdots, m$. (iii) When each φ_p is nondecreasing or nonincreasing, the concepts of weak A-stationarity and A-stationarity coincide. (iv) Given a point $\bar{x}$, we can rewrite* (8) *as*

$$0 \in \sum_{p \in I} \left[\pm \partial_A^\infty f_p(\bar{x}) \backslash \{0\} \right] + \sum_{p \in \{1,\cdots,m\} \backslash I} \{ y_p \, \partial_A f_p(\bar{x}) \}$$

for some index set $I \subset \{1, \cdots, m\}$ that is potentially empty. For each $p \in I$, although the scalar y_p does not explicitly appear in this inclusion, its existence implies that $\bigcup\{\partial\varphi_p(t_p) \mid t_p \in T_p(\bar{x})\} \neq \emptyset$, which plays a role in ensuring $\bar{x} \in \text{dom}(\varphi_p \circ f_p)$. For instance, if $f_p^k \xrightarrow{c} f_p$ for some $p \in I$, then $T_p(\bar{x}) = \{f_p(\bar{x})\}$, and the existence of $y_p \in \bigcup\{\partial\varphi_p(t_p) \mid t_p \in T_p(\bar{x})\} = \partial\varphi_p(f_p(\bar{x}))$ yields $\bar{x} \in \text{dom}(\varphi_p \circ f_p)$.

In the following, we take a detour to compare the A-stationarity with the stationarity defined in [27], where the author has focused on a more general composite problem

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \;\; \varphi\left(f(x)\right),$$

where $\varphi : \mathbb{R}^m \to \overline{\mathbb{R}}$ is proper, lsc, convex and $f \triangleq (f_1, \cdots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$ is a locally Lipschitz continuous mapping. Consider the special case where $\varphi(z) = \sum_{p=1}^m \varphi_p(z_p)$ with $z = (z_1, \cdots, z_m)$. Under this setting, a vector $\bar{x}$ is called a stationary point in [27] if there exists $\bar{y}$ and $\bar{z}$ such that

$$0 \in S(\bar{x}, \bar{y}, \bar{z}) \triangleq \left\{ (f_1(\bar{x}), \cdots, f_m(\bar{x})) - \bar{z} \right\} \times \left\{ \partial\varphi_1(\bar{z}_1) \times \cdots \times \partial\varphi_m(\bar{z}_m) - \bar{y} \right\} \times \left(\sum_{p=1}^{m} \bar{y}_p \, \partial_C f_p(\bar{x}) \right), \tag{9}$$

which can be equivalently written as

$$0 \in \sum_{p=1}^{m} \bar{y}_p \, \partial_C f_p(\bar{x}) \text{ for some } \; \bar{y}_p \in \partial \varphi_p(f_p(\bar{x})) \quad p = 1, \cdots, m.$$

For any fixed $k \geq 0$, the surrogate set-valued mapping S^k can be defined similarly as S in (9) by substituting f_p and φ_p with f_p^k and φ_p^k for each p. The cited paper provides sufficient conditions to ensure $\mathrm{Limsup}_{k\to+\infty}(\mathrm{gph}\, S^k) \subset \mathrm{gph}\, S$, which asserts that any accumulation point $(\bar{x}, \bar{y}, \bar{z})$ of a sequence $\{(x^k, y^k, z^k)\}$ with $0 \in S^k(x^k, y^k, z^k)$ yields a stationary point $\bar{x}$. Our study on the asymptotic stationarity differs from [27] in the following aspects:

1. Our outer convex function φ is assumed to have the separable form $\sum_{p=1}^m \varphi_p$, while [27] allows a general proper, lsc, convex function. In addition, each φ_p is fixed in our approximating problem while [27] considers a sequence of convex functions $\{\varphi_p^k\}_{k\geq 0}$ that epi-converges to φ_p.

2. We do not require the inner function f_p to be locally Lipschitz continuous.

If each f_p is locally Lipschitz continuous and bounded from below, it then follows from Proposition 1 that f_p is c-ADC associated with $\{f_p^k = g_p^k - h_p^k\}_{k\geq 0}$ such that $\partial f_p(x) \subset \partial_A f_p(x) \subset \partial_C f_p(x)$ and $\partial_A^\infty f_p(x) = \{0\}$ for any x. Moreover, by $f_p^k \xrightarrow{\mathrm{c}} f_p$, one has $T_p(x) = \{f_p(x)\}$. Thus, for any A-stationary point $\bar{x}$ induced by these ADC decompositions, there exists $\bar{y}_p \in \partial \varphi_p(f_p(\bar{x}))$ for each p such that

$$0 \in \sum_{p=1}^{m} \{\bar{y}_p \, \partial_A f_p(\bar{x})\} \;\subset\; \sum_{p=1}^{m} \{\bar{y}_p \, \partial_C f_p(\bar{x})\}\,. \tag{10}$$

Hence, $\bar{x}$ is also a stationary point defined in [27] satisfying $0 \in S(\bar{x}, \bar{y}, \bar{z})$. Indeed, A-stationarity here can be sharper than the latter one as the last inclusion in (10) may not hold with equality.

When f_p fails to be locally Lipschitz continuous for some p, it is not known if (9) is still a necessary condition for a local solution of (CP_0). This situation further complicates the fulfillment of conditions outlined in [27, Theorem 2.4], especially the requirement of $f_p^k \xrightarrow{\mathrm{c}} f_p$, due to the potential discontinuity of f_p. As will be shown in Theorem 2 below, despite these challenges, weak A-stationarity continues to be a necessary optimality condition under Assumption 1.

To proceed, for each p and any $x \in \mathrm{dom}(\varphi_p \circ f_p)$, we define $S_p(x)$ to be a collection of sequences:

$$S_p(x) \;\triangleq\; \left\{ \{x_p^k\}_{k\geq 0} \,\middle|\, x_p^k \to x \text{with } \varphi_p(f_p^k(x_p^k)) \to \varphi_p(f_p(x)) \right\}. \tag{11}$$

Theorem 2 (necessary conditions for optimality)**.** *Let $\bar{x} \in \bigcap_{p=1}^m \mathrm{dom}\, F_p$ be a local minimizer of problem (CP_0). Suppose that Assumption 1 and the following two conditions hold:*
(i) For each p and any sequence $\{x_p^k\}_{k\geq 0} \in S_p(\bar{x})$, there is a positive integer K such that

$$0 \notin \partial_C f_p^k(x_p^k) \quad \textit{or} \quad \mathcal{N}_{\mathrm{dom}\,\varphi_p}(f_p^k(x_p^k)) = \{0\} \quad \forall\, k \geq K, \tag{12}$$

and

$$\left[0 \in y_p \, \partial_A f_p(\bar{x}),\; y_p \in \bigcup \left\{ \mathcal{N}_{\mathrm{dom}\,\varphi_p}(t_p) \mid t_p \in T_p(\bar{x}) \right\} \right] \quad \Longrightarrow \quad y_p = 0, \quad p = 1, \cdots, m. \tag{13}$$

(ii) One has

$$\left[\sum_{p=1}^{m} w_p = 0,\ w_p \in \partial^{\infty}(\varphi_p \circ f_p)(\bar{x})\right] \quad \Longrightarrow \quad w_1 = \cdots = w_m = 0. \tag{14}$$

Then $\bar{x}$ is an A-stationary point of (CP$_0$)*. Additionally, $\bar{x}$ is a weakly A-stationary point of* (CP$_0$) *if* $\operatorname{int}(\operatorname{dom}\varphi_p) \neq \emptyset$ *for each* $p = 1, \cdots, m$.

Proof. By using Fermat's rule [25, Theorem 10.1] and the sum rule of the limiting subdifferentials [25, Corrollary 10.9] due to the condition (14), we have

$$\begin{aligned}
0 \ &\in \partial\left[\sum_{p=1}^{m}(\varphi_p \circ f_p)(\bar{x})\right] \subset \sum_{p=1}^{m}\partial(\varphi_p \circ f_p)(\bar{x}) \overset{\text{(i)}}{\subset} \sum_{p=1}^{m} \bigcup_{\{x_p^k\}_{k\geq 0}\in S_p(\bar{x})} \operatorname*{Lim\,sup}_{k\to+\infty}\ \partial(\varphi_p \circ f_p^k)(x_p^k)\\
&\overset{\text{(ii)}}{\subset} \sum_{p=1}^{m} \bigcup_{\{x_p^k\}_{k\geq 0}\in S_p(\bar{x})} \operatorname*{Lim\,sup}_{k\to+\infty} \bigcup\left\{\partial(y_p^k\, f_p^k)(x_p^k) \,\middle|\, y_p^k \in \partial\varphi_p(f_p^k(x_p^k))\right\}\\
&\overset{\text{(iii)}}{\subset} \sum_{p=1}^{m} \bigcup_{\{x_p^k\}_{k\geq 0}\in S_p(\bar{x})} \operatorname*{Lim\,sup}_{k\to+\infty}\ \left\{y_p^k\, v_p^k \,\middle|\, y_p^k \in \partial\varphi_p(f_p^k(x_p^k)),\ v_p^k \in \partial_C f_p^k(x_p^k)\right\}\\
&\overset{\text{(iv)}}{\subset} \sum_{p=1}^{m} \bigcup_{\{x_p^k\}_{k\geq 0}\in S_p(\bar{x})} \operatorname*{Lim\,sup}_{k\to+\infty}\ \left\{y_p^k\, v_p^k \,\middle|\, y_p^k \in \partial\varphi_p(f_p^k(x_p^k)),\ v_p^k \in \left[\partial g_p^k(x_p^k) - \partial h_p^k(x_p^k)\right]\right\}.
\end{aligned} \tag{15}$$

The inclusion (i) is due to approximation of subgradients under epi-convergence [25, Corollary 8.47] and [25, Proposition 8.46(e)]; (ii) follows from the nonsmooth Lagrange multiplier rule [25, Exercise 10.52] due to the locally Lipschitz continuity of f_p^k [25, Example 9.14] and the condition (12); (iii) and (iv) use the calculus rules of the Clarke subdifferential [10, Chapter 2.3]. For each p, any sequence $\{x_p^k\}_{k\geq 0} \in S_p(\bar{x})$ and any element

$$\bar{w}_p \in \operatorname*{Lim\,sup}_{k\to+\infty}\ \left\{y_p^k\, v_p^k \,\middle|\, y_p^k \in \partial\varphi_p(f_p^k(x_p^k)), v_p^k \in \left[\partial g_p^k(x_p^k) - \partial h_p^k(x_p^k)\right]\right\},$$

there is a subsequence $w_p^k \to_N \bar{w}_p$ with $w_p^k = y_p^k\, v_p^k$ for some $N \in \mathbb{N}_\infty^\sharp$. Next, we show the existence of $\bar{y}_p \in \bigcup\{\partial\varphi_p(t_p) \mid t_p \in T_p(\bar{x})\}$ for each p such that

$$\bar{w}_p \in \{\, \bar{y}_p\, \partial_A f_p(\bar{x})\,\} \,\cup\, \left[\, \pm\, \partial_A^\infty f_p(\bar{x})\backslash\{0\}\,\right]. \tag{16}$$

By Assumption 1(b), the subsequence $\{f_p^k(x_p^k)\}_{k\in N}$ is bounded. Taking a subsequence if necessary, we can suppose that $f_p^k(x_p^k) \to_N \bar{z}_p \in T_p(\bar{x})$. If $\{y_p^k\}_{k\in N}$ is unbounded, then $\{v_p^k\}_{k\in N}$ has a subsequence converging to 0 and, thus, $0 \in \partial_A f_p(\bar{x})$. Additionally, there exists $\widetilde{y}_p \neq 0$ such that

$$\frac{y_p^k}{|y_p^k|} \to_N \widetilde{y}_p \in \operatorname*{Lim\,sup}_{k(\in N)\to+\infty}{}^{\infty}\ \partial\varphi_p(f_p^k(x_p^k)) \overset{\text{(v)}}{=} \operatorname*{Lim\,sup}_{k(\in N)\to+\infty}{}^{\infty}\ \widehat{\partial}\varphi_p(f_p^k(x_p^k)) \overset{\text{(vi)}}{\subset} \partial^\infty\varphi_p(\bar{z}_p) \overset{\text{(vii)}}{=} \mathcal{N}_{\operatorname{dom}\varphi_p}(\bar{z}_p). \tag{17}$$

The equation (v) follows from [25, Proposition 8.12] by the convexity of φ_p. From $\{x_p^k\}_{k\geq 0} \in S_p(\bar{x})$ and $\bar{x} \in \operatorname{dom} F_p$, we must have $f_p^k(x_p^k) \in \operatorname{dom}\varphi_p$ for sufficiently large $k \in N$. Since φ_p is lsc, it holds that $\varphi_p(\bar{z}_p) \le \liminf_{k(\in N)\to+\infty} \varphi_p(f_p^k(x_p^k)) = \varphi_p(f_p(\bar{x}))$ and, thus, $\bar{z}_p \in \operatorname{dom}\varphi_p$. Also, notice that φ_p is continuous relative to its domain as it is univariate convex and lsc [22, Theorem 10.2]. This continuity implies $\varphi_p(f_p^k(x_p^k)) \to_N \varphi_p(\bar{z}_p)$. The inclusion (vi) follows directly from the definition of the horizon subdifferential. Lastly, (vii) is due to the lower semicontinuity of φ and [25, Proposition 8.12]. Therefore, we have $(0 \neq)\widetilde{y}_p \in \bigcup\{\mathcal{N}_{\operatorname{dom}\varphi_p}(t_p) \mid t_p \in T_p(\bar{x})\}$ with $0 \in \widetilde{y}_p\partial_A f_p(\bar{x})$ due to $0 \in \partial_A f_p(\bar{x})$, contradicting (13). So far, we conclude that $\{y_p^k\}_{k\in N}$ is a bounded sequence. Suppose that $y_p^k \to_N \bar{y}_p$ and, thus, $\bar{y}_p \in \partial\varphi_p(\bar{z}_p)$ by the outer semicontinuity of $\partial\varphi_p$ [25, Proposition 8.7].

<u>*Case 1.*</u> If $\bar{y}_p = 0$, inclusion (16) holds trivially for $\bar{w}_p = 0$, and for $\bar{w}_p \neq 0$ we can find a subsequence $\{|y_p^k|\}_{k\in N'} \downarrow 0$ such that $|y_p^k|\, v_p^k \to_{N'} \bar{w}_p$ or $-\bar{w}_p(\neq 0)$ with $v_p^k \in \left[\partial g_p^k(x_p^k) - \partial h_p^k(x_p^k)\right]$ for all $k \in N'$. Therefore, (16) follows from

$$\bar{w}_p \in \left[\Big(\pm \operatorname*{Lim\,sup}_{k\to+\infty}{}^{\infty} \left[\partial g_p^k(x_p^k) - \partial h_p^k(x_p^k)\right]\Big)\backslash\{0\}\right] \subset \left[\pm\partial_A^\infty f_p(\bar{x})\backslash\{0\}\right].$$

<u>*Case 2.*</u> Otherwise, $\|v_p^k\| \to_N \|\bar{w}_p\|/|\bar{y}_p|$. This means that $\{v_p^k\}_{k\in N}$ is bounded. Suppose $v_p^k \to_N \bar{v}_p$. Then, $\bar{v}_p \in \operatorname{Limsup}_{k\to+\infty}\left[\partial g_p^k(x_p^k) - \partial h_p^k(x_p^k)\right] \subset \partial_A f_p(\bar{x})$, and (16) is evident from $\bar{w}_p = \bar{y}_p\,\bar{v}_p$.

In either case, we have proved (16). Combining (15) with (16), for some $\bar{y}_p \in \bigcup\{\partial\varphi_p(t_p) \mid t_p \in T_p(\bar{x})\}$, we know that $\bar{x}$ is an A-stationary point of (CP_0). □

3.2 Examples of A-stationarity.

We present an example to illustrate the concept of A-stationarity and to study its relationship with other known optimality conditions.

Example 3.1 (bi-parametrized two-stage stochastic programs) Consider the following bi-parametrized two-stage stochastic program with fixed scenarios described in [17]:

$$\operatorname*{minimize}_{x\in\mathbb{R}^{n_1}} \;\theta(x) + \frac{1}{m_1}\sum_{p=1}^{m_1} f_p(x) \quad \text{subject to } \;\phi_p(x) \le 0, \quad p = 1,\cdots,m_2, \tag{18}$$

where $\theta, \phi_p : \mathbb{R}^{n_1} \to \mathbb{R}$ are convex, continuously differentiable for $p = 1,\cdots,m_2$, and f_p, as defined in (1), is finite for all $x \in \mathbb{R}^{n_1}$ and $p = 1,\cdots,m_1$. At $x = \bar{x}$, let $Y_p(\bar{x})$ and $\Lambda_p(\bar{x})$ represent the optimal solutions and multipliers for each second-stage problem (1). Suppose that $Y_p(\bar{x})$ and $\Lambda_p(\bar{x})$ are bounded. Note that θ and ϕ_p are ADC functions since they are convex. Example 2.1 shows that f_p is an ADC function, and therefore, problem (18) is a specific case of the composite model (CP_0). Given an A-stationary point $\bar{x}$ of (18), under the assumptions of Example 2.1, we have

$$\begin{aligned} 0 &\in \nabla\theta(\bar{x}) + \frac{1}{m_1}\sum_{p=1}^{m_1}\big(\{\partial_A f_p(\bar{x})\} \cup [\pm\partial_A^\infty f_p(\bar{x})\backslash\{0\}]\big) + \sum_{p=1}^{m_2}\bar{\mu}^{\,m_1+p}\,\nabla\phi_p(\bar{x}) \\ &\subset \nabla\theta(\bar{x}) + \frac{1}{m_1}\sum_{p=1}^{m_1}\left\{\partial_1\overline{f}_p(\bar{x},\bar{x}) - \partial_2(-\overline{f}_p)(\bar{x},\bar{x})\right\} + \sum_{p=1}^{m_2}\bar{\mu}^{\,m_1+p}\,\nabla\phi_p(\bar{x}), \end{aligned} \tag{19}$$

where $\bar{\mu}^{m_1+p} \in \mathcal{N}_{(-\infty,0]}(\phi_p(\bar{x}))$ for $p = 1,\cdots,m_2$ and $\overline{f}_p$ is defined in (5) for $p = 1,\cdots,m_1$. By assumptions, both $\Lambda_p(\bar{x})$ and $Y_p(\bar{x})$ are nonempty, bounded, and

$$\Lambda_p(\bar{x}) \times Y_p(\bar{x}) = \left\{(\bar{y}^p,\bar{\mu}^p)\,\middle|\, c^p + C^p\bar{x} + Q^p\,\bar{y}^p + (B^p)^\top\bar{\mu}^p = 0,\; 0 \le b^p - A^p\bar{x} - B^p\bar{y}^p \;\perp\; \bar{\mu}^p \ge 0\right\}.$$

It then follows from Danskin's Theorem [9, Theorem 2.1] that

$$\partial_1 \overline{f}_p(\bar{x},\bar{x}) = \mathrm{con}\left\{(A^p)^\top \bar{\mu}^p \,\middle|\, \bar{\mu}^p \in \Lambda_p(\bar{x})\right\} = \left\{(A^p)^\top \bar{\mu}^p \,\middle|\, \bar{\mu}^p \in \Lambda_p(\bar{x})\right\},$$

$$\partial_2(-\overline{f}_p)(\bar{x},\bar{x}) = \mathrm{con}\left\{-(C^p)^\top \bar{y}^p \,\middle|\, \bar{y}^p \in Y_p(\bar{x})\right\} = \left\{-(C^p)^\top \bar{y}^p \,\middle|\, \bar{y}^p \in Y_p(\bar{x})\right\}.$$

Combining these expressions with (19), we obtain

$$\begin{cases} 0 = \nabla\theta(\bar{x}) + \dfrac{1}{m_1}\displaystyle\sum_{p=1}^{m_1}\left[(C^p)^\top \bar{y}^p + (A^p)^\top \bar{\mu}^p\right] + \sum_{p=1}^{m_2} \bar{\mu}^{m_1+p}\, \nabla\phi_p(\bar{x}), \\ c^p + C^p\bar{x} + Q^p\,\bar{y}^p + (B^p)^\top \bar{\mu}^p = 0,\;\; 0 \le b^p - A^p\bar{x} - B^p\bar{y}^p \;\perp\; \bar{\mu}^p \ge 0, \quad p = 1,\cdots,m_1, \\ 0 \le \phi_p(\bar{x}) \;\perp\; \bar{\mu}^{m_1+p} \ge 0, \quad p = 1,\cdots,m_2, \end{cases}$$

which are the the Karush-Kuhn-Tucker (KKT) conditions for the deterministic equivalent of (18).

4 A computational algorithm.

In this section, we consider a double-loop algorithm for solving problem (CP_0). The inner loop finds an approximate stationary point of the perturbed composite optimization problem

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \;\; \sum_{p=1}^{m}\left[F_p^k(x) \triangleq \varphi_p(f_p^k(x))\right] \tag{20}$$

by solving a sequence of convex subproblems, while the outer loop drives $k \to +\infty$. It is important to note the potential infeasibility in (20) because $[F_p^k = \varphi_p \circ f_p^k] \xrightarrow{\text{e}} F_p$ in Assumption 1(c), together with $\mathrm{dom}(\varphi_p \circ f_p) \neq \emptyset$, does not guarantee $\mathrm{dom}(\varphi_p \circ f_p^k) \neq \emptyset$ for all $k \ge 0$. This can be seen from the example of $\varphi(t) = \delta_{(-\infty,0]}(t)$, $f(x) = \max\{x,0\} - 1/10$ and $f^k(x) = \max\{x,0\} + 1/k - 1/10$. Obviously $\mathrm{dom}(\varphi \circ f) = (-\infty, 1/10]$ and $\varphi \circ f^k \xrightarrow{\text{e}} \varphi \circ f$ by [27, Theorem 2.4(d)], but we have $\mathrm{dom}(\varphi \circ f^k) = \emptyset$ for $k = 1,\cdots,9$. Even though $\mathrm{dom}(\varphi_p \circ f_p^k) \neq \emptyset$ for all $k \ge 0$ and each p, this does not imply the feasibility of convex subproblems used in the inner loop to approximate (20).

For simplicity of the analysis, we assume that in problem (CP_0), φ_p is real-valued for $p = 1,\cdots,m_1$, and $\varphi_p = \delta_{(-\infty,0]}$ for $p = m_1+1,\cdots,m$. Namely, the problem takes the following form:

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \;\; \sum_{p=1}^{m_1}\left[F_p(x) = \varphi_p\big(f_p(x)\big)\right] \quad \text{subject to } \; f_p(x) \le 0,\;\; p = m_1+1,\cdots,m. \tag{CP_1}$$

4.1 Assumptions.

Firstly, we make an assumption to address the feasibility issue outlined at the start of this section. Let $\{\alpha_p^k\}_{k\ge0}$ be auxiliary sequences for $p = 1,\cdots,m$, where we set $\alpha_p^k \equiv 0$ for $p = 1,\cdots,m_1$, and for $p = m_1+1,\cdots,m$, we define

$$\alpha_p^k \triangleq \sup_{x\in X^k}\left[f_p^{k+1}(x) - f_p^k(x)\right]_+ \text{with } X^k \triangleq \left\{x \in \mathbb{R}^n \,\middle|\, f_p^k(x) \le 0,\, p = m_1+1,\cdots,m\right\}.$$

Based on these auxiliary sequences, we need an initial point x^0 that is strictly feasible to the constraints $f_p^0(x) \le 0$ for each $p = m_1+1,\cdots,m$.

Assumption 2 (**strict feasibility**) There exist x^0 and nonnegative sequences $\{\widehat{\alpha_p^k}\}_{k\geq 0}$ for $p = m_1+1,\cdots,m$, such that $\alpha_p^k \leq \widehat{\alpha_p^k}$ for all $k \geq 0$ and

$$\sum_{k=0}^{+\infty} \widehat{\alpha_p^k} < +\infty, \quad f_p^0(x^0) \leq -\sum_{k=0}^{+\infty} \widehat{\alpha_p^k} \qquad \forall\, p = m_1+1,\cdots,m.$$

Since the quantity α_p^k depends on the sequence $\{f_p^k\}_{k\geq 0}$, the above assumption posits a condition for this approximating sequence. As an example, consider f_p being icc associated with $\overline{f}_p$, where $\overline{f}_p(\cdot,x)$ is Lipschitz continuous with modulus L for any x. Using the sequence $\{f_p^k\}_{k\geq 0}$ in Example 2.1, we have

$$\alpha_p^k \leq \sup_{x\in\mathbb{R}^n} \Big[f_p^{k+1}(x) - f_p^k(x)\Big]_+ \leq \sup_{x\in\mathbb{R}^n} \Big[f_p(x) - f_p^k(x)\Big]_+ \leq \frac{\gamma_k L^2}{2} \triangleq \widehat{\alpha_p^k} \qquad \forall\, k \geq 0, \tag{21}$$

where the second inequality is due to $f_p^{k+1}(x) \leq f_p(x)$ for any x, and the last one uses the bound between the partial Moreau envelope and the original function [16, Lemma 3]. Thus, when $\{\gamma_k\}$ is summable, the sequence $\{\widehat{\alpha_p^k}\}_{k\geq 0}$ satisfies $\sum_{k=0}^{+\infty} \widehat{\alpha_p^k} < +\infty$.

Two more assumptions on the approximating sequences $\{f_p^k\}_{k\geq 0}$ are needed.

Assumption 3 (**smoothness of g_p^k or h_p^k**) For each $k \geq 0$, there exists $\ell_k > 0$ such that

$$\min\Big\{ \mathbb{H}\big(\partial g_p^k(x), \partial g_p^k(x')\big), \mathbb{H}\big(\partial h_p^k(x), \partial h_p^k(x')\big) \Big\} \leq \ell_k \|x'-x\| \quad \forall\, x, x' \in \mathbb{R}^n,\ p = 1,\cdots,m.$$

Assumption 4 (**level-boundedness**) For each $k \geq 0$, the function $H^k \triangleq \sum_{p=1}^m F_p^k$ is level-bounded, i.e., for any $r \in \mathbb{R}$, the set

$$\left\{ x \in \mathbb{R}^n \,\middle|\, \sum_{p=1}^{m_1} \varphi_p\left(f_p^k(x)\right) \leq r \right\} \cap X^k$$

is bounded.

Assumption 3 imposes conditions on the Lipschitz continuity of the subdifferential of either g_p^k or h_p^k, which will be used to determine the termination rule of the inner loop. A straightforward sufficient condition for this assumption is that, for each p and k, g_p^k or h_p^k is ℓ_k-smooth, i.e., $\|\nabla g_p^k(x) - \nabla g_p^k(x')\| \leq \ell_k\|x-x'\|$ or $\|\nabla h_p^k(x) - \nabla h_p^k(x')\| \leq \ell_k\|x-x'\|$ for any $x, x' \in \mathbb{R}^n$. Assumption 4 is a standard condition to ensure the boundedness of the generated sequences for each $k \geq 0$.

In addition, we need a technical assumption to ensure the boundedness of the multiplier sequences in our algorithm.

Assumption 5 (**an asymptotic constraint qualification**) For any $\bar{x} \in \cap_{p=1}^{m} \operatorname{dom} F_p$, if there exists $\{y_p\}_{p=1}^m$ satisfying $0 = \sum_{p=1}^m y_p\, v_p$ where for each p (with the definition of $T_p(\bar{x})$ in (7)),

$$(y_p,\, v_p) \in \Big(\bigcup \big\{\mathcal{N}_{\operatorname{dom}\varphi_p}(t_p) \mid t_p \in T_p(\bar{x})\big\} \times \operatorname{con} \partial_A f_p(\bar{x})\Big) \cup \Big(\mathbb{R} \times [\, \partial_A^{\infty} f_p(\bar{x}) \backslash \{0\}\,]\Big), \tag{22}$$

then we must have $y_1 = \cdots = y_m = 0$.

According to the definitions of $\partial_A f_p(\bar{x})$ and $\partial_A^\infty f_p(\bar{x})$, Assumption 5 depends on the approximating sequences $\{f_p^k\}_{k\geq 0}$ for $p = 1, \cdots, m$. It holds trivially if each φ_p is real-valued and $\partial_A^\infty f_p(\bar{x}) = \{0\}$. For Example 3.1, the assumption translates into

$$\left[\sum_{p=1}^{m_2} \lambda_p \nabla \phi_p(\bar{x}) = 0,\ \lambda_p \in \mathcal{N}_{(-\infty,0]}(\phi_p(\bar{x})),\ \ p = 1, \cdots, m_2\right] \quad \Longrightarrow \quad \lambda_1 = \cdots = \lambda_{m_2} = 0.$$

This is equivalent to the Mangasarian-Fromovitz constraint qualification (MFCQ) for problem (18) by [25, Example 6.40]; see also [23].

Furthermore, if each f_p is c-ADC associated with $\{f_p^k = g_p^k - h_p^k\}_{k\geq 0}$ such that $\operatorname{con} \partial_A f_p(\bar{x}) = \partial_C f_p(\bar{x})$, and $\partial_A^\infty f_p(\bar{x}) = \{0\}$, Assumption 5 states that

$$\left[0 \in \sum_{p=1}^{m} y_p\, \partial_C f_p(\bar{x}), \quad y_p \in \mathcal{N}_{\operatorname{dom}\varphi_p}(f_p(\bar{x})),\ \ p = 1, \cdots, m\right] \quad \Longrightarrow \quad y_1 = \cdots = y_m = 0.$$

This condition aligns with the constraint qualification for the composite optimization problem in [27, Proposition 2.1], and is stronger than the condition in the nonsmooth Lagrange multiplier rule [25, Exercise 10.52]. Finally, Assumption 5 implies the constraint qualifications in Theorem 2. We formally present this conclusion in the following proposition. Depending on whether φ_p is nondecreasing or not, we partition $\{1, \cdots, m\}$ into two categories:

$$I_1 \triangleq \{\, p \in \{1, \cdots, m\} \mid \varphi_p \text{nondecreasing}\} \quad \text{and} \quad I_2 \triangleq \{1, \cdots, m\} \backslash I_1. \tag{23}$$

Observe that $I_2 \subset \{1, \cdots, m_1\}$ for problem (CP$_1$). We do not specifically address the case where φ_p is nonincreasing, as one can always redefine $\widetilde{\varphi}_p(t) = \varphi_p(-t)$ and $\widetilde{f}_p(x) = -f_p(x)$, enabling the treatment of these indices in the same manner as those in I_1. The proof of Proposition 4 is in Appendix A.

Proposition 4 (consequences of Assumption 5)**.** *Suppose that Assumptions 1 and 5 hold and $f_p^k \xrightarrow{e} f_p$ for each p. If $\sup \varphi_p = +\infty$ for $p \in I_1$, and f_p is locally Lipschitz continuous for $p \in I_2$, then conditions* (12)*,* (13)*, and* (14) *hold at any feasible point $\bar{x}$ of* (CP$_1$)*. Consequently, any local solution of* (CP$_1$) *is a (weakly) A-stationary point of* (CP$_1$)*.*

4.2 The algorithmic framework and convergence analysis.

We now formalize the algorithm for solving (CP$_1$). For $p = m_1 + 1, \cdots, m$, recall the nonnegative sequences $\{\widehat{\alpha_p^k}\}_{k\geq 0}$ introduced in Assumption 2, and observe that $\sum_{k'=k}^{+\infty} \widehat{\alpha_p^{k'}} \to 0$ as $k \to +\infty$. For consistency of our notation, we also set $\widehat{\alpha_p^k} \equiv 0$ for all $k \geq 0$ and $p = 1, \cdots, m_1$. At the k-th outer iteration and for $p = 1, \cdots, m$, consider the upper and lower approximation of f_p^k at a point y by taking some $a_p^k \in \partial h_p^k(y)$, $b_p^k \in \partial g_p^k(y)$ and incorporating sequences $\{\widehat{\alpha_p^k}\}_{k \geq 0}$:

$$
\begin{aligned}
f_p^{k,\text{upper}}(x;y) &\triangleq \; g_p^k(x) - h_p^k(y) - (a_p^k)^\top (x-y) + \sum_{k'=k}^{+\infty} \widehat{\alpha_p^{k'}}, \\
f_p^{k,\text{lower}}(x;y) &\triangleq \; g_p^k(y) + (b_p^k)^\top (x-y) - h_p^k(x).
\end{aligned} \tag{24}
$$

Then, for $p = 1, \cdots, m_1$, a convex majorization of F_p^k at a point y can be constructed as

$$
\widehat{F_p^k}(x;y) \triangleq \varphi_p^\uparrow \left(f_p^{k,\text{upper}}(x;y) \right) + \varphi_p^\downarrow \left(f_p^{k,\text{lower}}(x;y) \right). \tag{25}
$$

For $p = m_1 + 1, \cdots, m$, we replace $f_p^k(x) \leq 0$ with a convex constraint $f_p^{k,\text{upper}}(x;y) \leq 0$. The proposed algorithm for solving (CP$_1$) is given below. In contrast to the prox-linear algorithm that is designed to minimize amenable functions and adopts complete linearization of the inner maps, the prox-ADC method retains more curvature information inherent in these maps; as illustrated in Figure 1.

Algorithm The prox-ADC method for solving (CP$_1$)

Input: Let $x = x^0$ be an initial point satisfying Assumption 2, and $\{\ell_k\}$ be the sequence satisfying Assumption 3. Set $\lambda > 0$, and a positive sequence $\{(\epsilon_k, \delta_k)\} \downarrow 0$ such that $\delta_k/\ell_k \to 0$.

Outer loop: Set $k = 0$.

1: Execute the inner loop with the initial point x^k, and parameters $(\varepsilon_k, \delta_k)$.

2: Set $k \leftarrow k + 1$ until a prescribed stopping criterion is satisfied.

Inner loop: Set $i = 0$ and $x^{k,0} = x^k$.

1: Take $\{a_p^{k,i} \in \partial\, g_p^k(x^{k,i})\}_{p=1}^m$, $\{b_p^{k,i} \in \partial h_p^k(x^{k,i})\}_{p=1}^{m_1}$ and solve the strongly convex subproblem

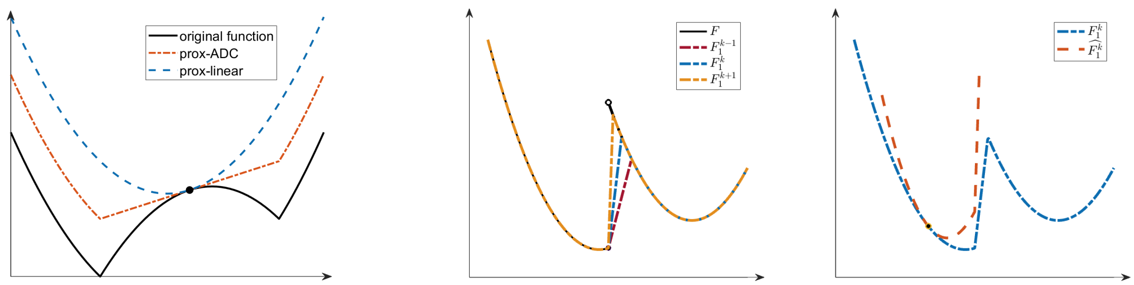
$$
x^{k,i+1} = \operatorname*{argmin}_{x \in \mathbb{R}^n} \left\{ \sum_{p=1}^{m_1} \widehat{F_p^k}(x; x^{k,i}) + \frac{\lambda}{2} \|x - x^{k,i}\|^2 \;\middle|\; f_p^{k,\text{upper}}(x; x^{k,i}) \leq 0,\, p = m_1 + 1, \cdots, m \right\}. \tag{26}
$$

2: If the following conditions hold

$$
\left\{
\begin{array}{rll}
f_p^{k,\text{upper}}(x^{k,i+1}; x^{k,i}) & \leq f_p^k(x^{k,i+1}) + \sum_{k'=k}^{+\infty} \widehat{\alpha_p^{k'}} + \epsilon_k, & p = 1, \cdots, m, \\
f_p^{k,\text{lower}}(x^{k,i+1}; x^{k,i}) & \geq f_p^k(x^{k,i+1}) - \epsilon_k, & p \in I_2, \\
\|x^{k,i+1} - x^{k,i}\| & \leq \delta_k/\ell_k, &
\end{array}
\right. \tag{27}
$$

break the inner loop and set $x^{k+1} = x^{k,i+1}$. Otherwise, set $i \leftarrow i + 1$ and return to step 1.

We emphasize that the prox-ADC method differs from [11, Algorithm 7.1.2] that is designed for solving a problem with a convex composite DC objective and DC constraints. Central to the prox-ADC method is the double-loop structure, where, in contrast to [11, Algorithm 7.1.2], the DC sequence f_p^k is dynamically updated in the outer loop rather than remaining the same. This adaptation necessitates specialized termination criteria (27) and the incorporation of $\widehat{\alpha_p^k}$ to maintain feasibility with each update of f_p^k. In the following, we demonstrate the well-definedness



(a) $F_1 = \varphi_1 \circ f_1$ for a convex φ_1 and a smooth f_1. (b) $F_1 = \varphi_1 \circ f_1$ for a convex nondecreasing φ_1 and a lsc f_1.

Figure 1: Illustrations of the prox-ADC method. (a): a comparison of the prox-ADC and the prox-linear method for minimizing an amenable function. (b): asymptotic approximations of a discontinuous composite function $F_1 = \varphi_1 \circ f_1$ that are constructed by an epi-convergent sequence $\{F_1^k = \varphi_1 \circ f_1^k\}$, and a convex majorization $\widehat{F_1^k}$ for F_1^k.

of the prox-ADC method. Specifically, we establish that for each iteration k, the criteria detailed in (27) are attainable within a finite number of steps.

Theorem 3 (convergence of the inner loop)**.** *Suppose that Assumptions 1-4 hold. Then the following statements hold.*
(a) Problem (26) *is feasible for any* $k, i \geq 0$.
(b) The stopping rule of the inner loop is achievable in finite steps, i.e., the smallest integer i *satisfying conditions* (27)*, denoted by* i_k*, is finite for any* $k \geq 0$.

Proof. We prove (a) and (b) by induction. For $k = 0$, notice from Assumption 2 that $f_p^{0,\text{upper}}(x^0; x^0) = f_p^0(x^0) + \sum_{k=0}^{+\infty} \widehat{\alpha_p^k} \leq 0$ for $p = m_1 + 1, \cdots, m$. Thus, problem (26) is feasible for $k = i = 0$. Assume that (26) is feasible for $k = 0$ and some $i = \bar{i}$ (≥ 0). Consequently, $x^{0,\bar{i}+1}$ is well-defined and for $p = m_1 + 1, \cdots, m$,

$$f_p^{0,\text{upper}}(x^{0,\bar{i}+1}; x^{k,\bar{i}+1}) = f_p^0(x^{0,\bar{i}+1}) + \sum_{k=0}^{+\infty} \widehat{\alpha_p^k} \;\leq\; f_p^{0,\text{upper}}(x^{0,\bar{i}+1}; x^{0,\bar{i}}) \leq 0,$$

which yields the feasibility of (26) for $k = 0$, $i = \bar{i}+1$. Hence, by induction, problem (26) is feasible for $k = 0$ and any $i \geq 0$. To proceed, recall the function H^k defined in Assumption 4. From the update of $x^{0,i+1}$, we have

$$H^0(x^{0,i+1}) = \sum_{p=1}^{m_1} F_p^0(x^{0,i+1}) \leq \sum_{p=1}^{m_1} \widehat{F_p^0}(x^{0,i+1}; x^{0,i}) \leq H^0(x^{0,i}) - \frac{\lambda}{2}\|x^{0,i+1} - x^{0,i}\|^2 \quad \forall\, i \geq 0. \tag{28}$$

Observe that H^0 is bounded from below by the continuity of $F_p^0 = \varphi_p \circ f_p^0$ for $p = 1, \cdots, m_1$ and the level-boundedness of H^0. Suppose for contradiction that the stopping rule of the inner loop is not achievable in finite steps. Then from (28), $\left\{H^0(x^{0,i})\right\}$ converges and $\sum_{i=0}^{\infty} \|x^{0,i+1} - x^{0,i}\|^2 < +\infty$. The latter further yields $\|x^{0,i+1} - x^{0,i}\| \to 0$ and thus the last condition in (27) is achievable in finite iterations. Next, to derive a contradiction, it suffices to prove that the first two conditions in (27) can also be achieved in finite number of steps. We only show the first one since the other can be done with similar arguments. By the level-boundedness of H^0, there exists a compact set S^0 containing the sequence $\{x^{0,i}\}$. For $p = 1, \cdots, m$, we have

$$0 \leq \; f_p^{0,\text{upper}}(x^{0,i+1}; x^{0,i}) - f_p^0(x^{0,i+1}) - \sum_{k=0}^{+\infty} \widehat{\alpha_p^k} \; = \; h_p^0(x^{0,i+1}) - h_p^0(x^{0,i}) - (a_p^{0,i})^\top (x^{0,i+1} - x^{0,i}) \; \longrightarrow \; 0,$$

because h_p^0 is uniformly continuous on S^0 and $\{a_p^{0,i}\}_{i\geq 0} \subset \bigcup \{\partial h_p^0(x) \mid x \in S^0\}$ is bounded by [22, Theorem 24.7]. Therefore, for a fixed $\epsilon_0 > 0$, there exists some i_0 such that $f_p^{0,\text{upper}}(x^{0,i_0+1}; x^{0,i_0}) \leq f_p^0(x^{0,i_0+1}) + \sum_{k=0}^{+\infty} \widehat{\alpha_p^k} + \epsilon_0$ holds for $p = 1, \cdots, m$. Thus, (a)-(b) hold for $k = 0$.

Now assume that (a)-(b) hold for some $k = \bar{k}$ (≥ 0) and, hence $i_{\bar{k}}$ is finite. It then follows from $x^{\bar{k}+1,0} = x^{\bar{k},i_{\bar{k}}+1} \in X^{\bar{k}}$ and $f_p^{\bar{k},\text{upper}}(x^{\bar{k},i_{\bar{k}}+1}; x^{\bar{k},i_{\bar{k}}}) \leq 0$ that for each p,

$$\begin{aligned} & f_p^{\bar{k}+1,\text{upper}}(x^{\bar{k}+1,0}; x^{\bar{k}+1,0}) = f_p^{\bar{k}+1}(x^{\bar{k}+1,0}) + \sum_{k=\bar{k}+1}^{+\infty} \widehat{\alpha_p^k} \\ \leq \;& f_p^{\bar{k}}(x^{\bar{k}+1,0}) + \sup_{x \in X^{\bar{k}}} \left[f_p^{\bar{k}+1}(x) - f_p^{\bar{k}}(x) \right]_+ + \sum_{k=\bar{k}+1}^{+\infty} \widehat{\alpha_p^k} \\ \leq \;& f_p^{\bar{k}}(x^{\bar{k}+1,0}) + \sum_{k=\bar{k}}^{+\infty} \widehat{\alpha_p^k} \; \leq \; f_p^{\bar{k},\text{upper}}(x^{\bar{k}+1}; x^{\bar{k},i_{\bar{k}}}) \leq 0. \end{aligned}$$

Thus, problem (26) is feasible for $k = \bar{k} + 1$ and any $i \geq 0$. Building upon this, we can now clearly see the validity of (b) for $k = \bar{k} + 1$, as we have shown similar results earlier in the case of $k = 0$. By induction, we complete the proof of (a)-(b). □

As we will see in the following lemma, the asymptotic constraint qualification (Assumption 5) implies the existence of the multipliers for problem (26).

Lemma 2. *Suppose that Assumptions 1-5 hold. Let $\{x^k\}$ be the sequence generated by the prox-ADC method and $\{x^k\}_{(k-1)\in N}$ be a subsequence converging to some $\bar{x}$. Then, for all $k \in N$, there exist $y_{p,1}^k \in \partial \varphi_p^{\uparrow}(f_p^{k,upper}(x^{k,i_k+1}; x^{k,i_k}))$, $y_{p,2}^k \in \partial \varphi_p^{\downarrow}(f_p^{k,lower}(x^{k,i_k+1}; x^{k,i_k}))$ for $p = 1, \cdots, m$ satisfying*

$$0 \in \sum_{p=1}^{m} \left[y_{p,1}^k \, \partial f_p^{k,upper}(x^{k,i_k+1}; x^{k,i_k}) + y_{p,2}^k \, \partial f_p^{k,lower}(x^{k,i_k+1}; x^{k,i_k}) \right] + \lambda(x^{k,i_k+1} - x^{k,i_k}). \tag{29}$$

Proof. Observe that $\{x^{k,i_k}\}_{k\in N}$ and $\{x^{k,i_k+1}\}_{k\in N}$ converge to $\bar{x}$ by the stopping conditions (27). By Theorem 3(a), we have $f_p^k(x^{k,i_k+1}) \leq 0$ for $p = m_1+1, \cdots, m$ and all $k \in N$. Due to epi-convergence in Assumption 1(c), we have

$$\delta_{(-\infty,0]}(f_p(\bar{x})) \leq \liminf_{k(\in N)\to+\infty} \delta_{(-\infty,0]}(f_p^k(x^{k,i_k+1})) = 0 \qquad \forall\, p = m_1 + 1, \cdots, m.$$

This means $f_p(\bar{x}) \leq 0$ for $p = m_1 + 1, \cdots, m$ and $\bar{x} \in \cap_{p=1}^m \operatorname{dom} F_p$. The conclusion is a direct consequence of the nonsmooth Lagrange multiplier rule [25, Exercise 10.52] for problem (26) if we can show that, for any $k \in N$, $y^k_{m_1+1} = \cdots = y^k_m = 0$ is the unique solution of the following system

$$0 \in \sum_{p=m_1+1}^{m} y^k_p \, \partial f_p^{k,\text{upper}}(x^{k,i_k+1}; x^{k,i_k}), \;\; y^k_p \in \mathcal{N}_{(-\infty,0]}(f_p^{k,\text{upper}}(x^{k,i_k+1}; x^{k,i_k})), \;\; p = m_1+1, \cdots, m. \tag{30}$$

Suppose that the above claim does not hold. Without loss of generality, take $\{y^k_p\}_{k\in N}$ for $p = m_1+1, \cdots, m$ satisfying (30) and $\sum_{p=m_1+1}^m |y^k_p| = 1$. For each p and $k \in N$, define

$$A^k_p \triangleq \left\{ y^k_p \, v^k_p \,\middle|\, v^k_p \in \left\{\partial g^k_p(x^{k,i_k}) - \partial h^k_p(x^{k,i_k})\right\} \cup \left\{\partial g^k_p(x^{k,i_k+1}) - \partial h^k_p(x^{k,i_k+1})\right\}\right\}.$$

Then, for all $k \in N$, we have

$$\begin{aligned}
&\operatorname{dist}\left(0, \sum_{p=m_1+1}^{m} A^k_p\right) \\
\overset{\text{(i)}}{\le}\; &\operatorname{dist}\left(0, \sum_{p=m_1+1}^{m} y^k_p \left[\partial g^k_p(x^{k,i_k+1}) - \partial h^k_p(x^{i,i_k})\right]\right) + \sum_{p=m_1+1}^{m} \mathbb{D}\left(y^k_p \left[\,\partial g^k_p(x^{k,i_k+1}) - \partial h^k_p(x^{i,i_k})\right], A^k_p\right) \\
\overset{\text{(ii)}}{\le}\; & 0 + \sum_{p=m_1+1}^{m} |y^k_p| \cdot \min\left\{\mathbb{H}(\partial g^k_p(x^{k,i_k+1}), \partial g^k_p(x^{k,i_k})), \, \mathbb{H}(\partial h^k_p(x^{k,i_k+1}), \partial h^k_p(x^{k,i_k}))\right\} \\
\overset{\text{(iii)}}{\le}\; & \sum_{p=m_1+1}^{m} |y^k_p| \cdot \ell_k \, \|x^{k,i_k+1} - x^{k,i_k}\| \overset{\text{(iv)}}{=} \delta_k,
\end{aligned}$$

where (i) uses the inequalities $\mathbb{D}(A,C) \le \mathbb{D}(A,B) + \mathbb{D}(B,C)$ and $\mathbb{D}(A+B, A'+B') \le \mathbb{D}(A,A') + \mathbb{D}(B,B')$; (ii) is due to (30) and the definition of A^k_p; (iii) is by Assumption 3; and (iv) is implied by conditions (27) and $\sum_{p=m_1+1}^m |y^k_p| = 1$. Equivalently, for all $k \in N$ and $p = m_1+1, \cdots, m$, there exist $y^k_p \in \mathcal{N}_{(-\infty,0]}\left(f_p^{k,\text{upper}}(x^{k,i_k+1}; x^{k,i_k})\right)$ with $\sum_{p=m_1+1}^m |y^k_p| = 1$ and

$$v^k_p \in \left\{\partial g^k_p(x^{k,i_k}) - \partial h^k_p(x^{k,i_k})\right\} \cup \left\{\partial g^k_p(x^{k,i_k+1}) - \partial h^k_p(x^{k,i_k+1})\right\}$$

such that $\|\sum_{p=m_1+1}^m y^k_p \, v^k_p\| \le \delta_k$. Taking a subsequence if necessary and using conditions (27), we can assume that $f_p^{k,\text{upper}}(x^{k,i_k+1}; x^{k,i_k})$ and $f^k_p(x^{k,i_k+1})$ converge to the same limit point $\bar z_p \in T_p(\bar x)$ as $k\ (\in N) \to +\infty$ for each $p = m_1+1, \cdots, m$. Notice that, for each p, $\bar z_p$ must satisfy $\bar z_p \le 0$, and $f_p^{k,\text{upper}}(x^{k,i_k+1}; x^{i,i_k}) \le 0$ for all $k \in N$ from Theorem 3(a). Suppose that $y^k_p \to_N \bar y_p$ for each p. Then, by the outer semicontinuity of the normal cone [25, Proposition 6.6],

$$\bar y_p \in \mathcal{N}_{(-\infty,0]}(\bar z_p) \subset \bigcup \left\{\mathcal{N}_{\operatorname{dom}\varphi_p}(t_p) \mid t_p \in T_p(\bar x)\right\}, \quad p = m_1+1, \cdots, m.$$

Obviously, $\sum_{p=m_1+1}^m |\bar y_p| = 1$, and $\{\bar y_p\}_{p=m_1+1}^m$ has at least one nonzero element. Consider two cases.

Case 1. If $\{v^k_p\}_{k\in N}$ is bounded for $p = m_1+1, \cdots, m$, then there exist vectors $\{\bar v_p\}_{p=m_1+1}^m$ with $\bar v_p \in \partial_A f_p(\bar x)$ such that $v^k_p \to_N \bar v_p$ and $0 = \sum_{p=m_1+1}^m \bar y_p \, \bar v_p \in \sum_{p=m_1+1}^m \bar y_p \, \partial_A f_p(\bar x)$, contradicting Assumption 5 since $\bar y_{m_1+1}, \cdots, \bar y_m$ are not all zeros.

Case 2. Otherwise, there exists some p such that $\{v^k_p\}_{k\in N}$ is unbounded, define the index sets

$$I_{\text{ub}} \triangleq \left\{ p \in \{m_1+1, \cdots, m\} \,\middle|\, \{v^k_p\}_{k\in N} \text{unbounded} \right\} \ (\ne \emptyset) \quad \text{and} \quad I_{\text{b}} \triangleq \{m_1+1, \cdots, m\} \backslash I_{\text{ub}}.$$

Notice that $\{\sum_{p\in I_{\rm b}} y_p^k\, v_p^k\}_{k\in N}$ is bounded. Without loss of generality, assume that this sequence converges to some $\bar{w}$ and, thus, $\sum_{p\in I_{\rm ub}} y_p^k\, v_p^k \to_N (-\bar{w})$.

Step 1: Next we prove by contradiction that, for each $p \in I_{\rm ub}$, the sequence $\{y_p^k\, v_p^k\}_{k\in N}$ is bounded. Suppose that the boundedness fails and $\sum_{p\in I_{\rm ub}} \|y_p^k\, v_p^k\| \to_N +\infty$ by passing to a subsequence. Consider $\widetilde{w}_p^k \triangleq y_p^k\, v_p^k / \sum_{p\in I_{\rm ub}} \|y_p^k\, v_p^k\|$ for $p \in I_{\rm ub}$. Then $\sum_{p\in I_{\rm ub}} \widetilde{w}_p^k \to_N 0$. Since $\sum_{p\in I_{\rm ub}} \|\widetilde{w}_p^k\| = 1$ for all $k \in N$, we can assume that there exist $p_1 \in I_{\rm ub}$ and $\widetilde{w}_{p_1} \neq 0$ such that $\widetilde{w}_{p_1}^k \to_N \widetilde{w}_{p_1}$. It then follows from the construction of $\widetilde{w}_p^k$ that $\{\widetilde{w}_p^k\}_{k\in N}$ has a subsequence converging to some element of $\pm\partial_A^\infty f_p(\bar{x})$ for each $p \in I_{\rm ub}$ and, in particular, $\widetilde{w}_{p_1} \in \big[\pm\partial_A^\infty f_{p_1}(\bar{x})\backslash\{0\}\big]$. From $\sum_{p\in I_{\rm ub}} \widetilde{w}_p^k \to_N 0$, we obtain

$$0 \in [\,\pm\partial_A^\infty f_{p_1}(\bar{x})\backslash\{0\}\,] + \sum_{p\in I_{\rm ub}\backslash\{p_1\}} [\,\pm\partial_A^\infty f_p(\bar{x})\,],$$

contradicting Assumption 5 since the coefficient of the term $[\,\pm\partial_A^\infty f_{p_1}(\bar{x})\backslash\{0\}\,]$ is nonzero. So far, we have shown the boundedness of $\{y_p^k\, v_p^k\}_{k\in N}$ for each $p \in I_{\rm ub}$.

Step 2: Now suppose that $y_p^k\, v_p^k \to_N \bar{w}_p$ for each $p \in I_{\rm ub}$ with $\sum_{p\in I_{\rm ub}} \bar{w}_p = -\bar{w}$. Thus $y_p^k \to_N 0$ and $\bar{w}_p \in [\,\pm\partial_A^\infty f_p(\bar{x})\,]$ for each $p \in I_{\rm ub}$. Since $\sum_{p=m_1+1}^m |\bar{y}_p| = 1$, there exists $p_2 \in I_{\rm b}$ such that $\bar{y}_{p_2} \neq 0$. Then $\sum_{p=m_1+1}^m y_p^k\, v_p^k \to_N 0$ implies

$$0 \in \bar{y}_{p_2}\, \partial_A f_{p_2}(\bar{x}) + \sum_{p\in I_{\rm b}\backslash\{p_2\}} \bar{y}_p\, \partial_A f_p(\bar{x}) + \sum_{p\in I_{\rm ub}} [\,\pm\partial_A^\infty f_p(\bar{x})\,],$$

which leads to a contradiction to Assumption 5 and therefore completes the proof. $\square$

The main convergence result of the prox-ADC method follows. Recall the definitions of I_1 and I_2 in (23). An additional assumption is the boundedness of the set $\partial_A f_p(\bar{x})$ for $p \in I_2$, ensured by assuming $\partial_A^\infty f_p(\bar{x}) = \{0\}$. There are some sufficient conditions for $\partial_A^\infty f_p(\bar{x}) = \{0\}$ to hold: (i) If f_p is locally Lipschitz continuous and bounded below, from Proposition 1(b), we have $\partial_A^\infty f_p(x) = \{0\}$ at any $x \in \operatorname{dom} f_p$ for the approximating sequence generated by the Moreau envelope. (ii) If f_p is icc associated with $\overline{f}_p$ satisfying all assumptions in Example 2.1, then $\partial_A^\infty f_p(x) = \{0\}$ still holds at any $x \in \operatorname{int}(\operatorname{dom} f_p)$ for the approximating sequence based on the partial Moreau envelope. It is worth mentioning that the icc function f_p under condition (ii) is not necessarily locally Lipschitz continuous.

Theorem 4. *Suppose that Assumptions 1-5 hold, and the sequence $\{x^k\}$ generated by the prox-ADC method has an accumulation point $\bar{x}$. Suppose in addition that $\partial_A^\infty f_p(\bar{x}) = \{0\}$ for $p \in I_2$. Then $\bar{x}$ is a weakly A-stationary point of* (CP$_1$). *Moreover, if for each $p \in I_2$, the functions g_p^k and h_p^k are ℓ_k-smooth for all $k \geq 0$, i.e., there exists a sequence $\{\ell_k\}$ such that for all $k \geq 0$,*

$$\max\Big\{\, \big\|\nabla g_p^k(x) - \nabla g_p^k(x')\big\|,\, \big\|\nabla h_p^k(x) - \nabla h_p^k(x')\big\| \,\Big\} \leq \ell_k \|x' - x\| \quad \forall\, x, x' \in \mathbb{R}^n,\ p \in I_2, \tag{31}$$

then $\bar{x}$ is also an A-stationary point of (CP$_1$).

Proof. Let $\{x^k\}_{(k-1)\in N}$ be a subsequence converging to $\bar{x}$. Similar to Lemma 2, we also have $x^{k,i_k} \to_N \bar{x}$, $x^{k,i_k+1} \to_N \bar{x}$, and $\bar{x} \in \cap_{p=1}^m \operatorname{dom} F_p$. By Lemma 2, for all $k \in N$, we have

$$0 \in \sum_{p=1}^{m} \Big[y_{p,1}^k \left(\partial g_p^k(x^{k,i_k+1}) - \partial\, h_p^k(x^{k,i_k})\right) + y_{p,2}^k \left(\partial g_p^k(x^{k,i_k}) - \partial\, h_p^k(x^{k,i_k+1})\right) \Big] + \lambda(x^{k,i_k+1} - x^{k,i_k}), \tag{32}$$

where $y_{p,1}^k \in \partial\varphi_p^{\uparrow}(f_p^{k,\text{upper}}(x^{k,i_k+1}; x^{k,i_k}))$, $y_{p,2}^k \in \partial\varphi_p^{\downarrow}(f_p^{k,\text{lower}}(x^{k,i_k+1}; x^{k,i_k}))$ for $p = 1, \cdots, m$. Due to Assumption 3 and similar arguments in Lemma 2, the optimality condition (32) implies

$$\left\{ \begin{array}{l} \left\| \displaystyle\sum_{p=1}^{m} \left(y_{p,1}^k\, v_{p,1}^k + y_{p,2}^k\, v_{p,2}^k\right) \right\| \leq \lambda\, \delta_k/\ell_k + \displaystyle\sum_{p=1}^{m} \left(|y_{p,1}^k| + |y_{p,2}^k|\right)\delta_k, \\ \left[\begin{array}{c} v_{p,1}^k \in \partial g_p^k(x^{k,i_k}) - \partial h_p^k(x^{k,i_k}) \\ v_{p,2}^k \in \partial g_p^k(x^{k,i_k+1}) - \partial h_p^k(x^{k,i_k+1}) \end{array} \right] \text{ or } \left[\begin{array}{c} v_{p,1}^k \in \partial g_p^k(x^{k,i_k+1}) - \partial h_p^k(x^{k,i_k+1}) \\ v_{p,2}^k \in \partial g_p^k(x^{k,i_k}) - \partial h_p^k(x^{k,i_k}) \end{array} \right], \, p = 1, \cdots, m. \end{array} \right. \tag{33}$$

Recall that, for $p \in I_1$, φ_p is nondecreasing, i.e., $\varphi_p^{\downarrow} = 0$. Then $y_p^k = 0$ for all $k \in N$ and $p \in I_2$, and the first inequality of (33) is equivalent to

$$\left\| \sum_{p\in I_1} y_{p,1}^k\, v_{p,1}^k + \sum_{p\in I_2} \left(y_{p,1}^k\, v_{p,1}^k + y_{p,2}^k\, v_{p,2}^k\right) \right\| \leq \frac{\lambda\, \delta_k}{\ell_k} + \left(\sum_{p\in I_1} |y_{p,1}^k| + \sum_{p\in I_2} \left(|y_{p,1}^k| + |y_{p,2}^k|\right) \right) \delta_k. \tag{34}$$

<u>Step 1:</u> To start with, we prove the boundedness of the multiplier subsequences along $k \in N$. Similarly as in the proof of Lemma 2, assume that $f_p^{k,\text{upper}}(x^{k,i_k+1}; x^{k,i_k})$, $f_p^{k,\text{lower}}(x^{k,i_k+1}; x^{k,i_k})$ and $f_p^k(x^{k,i_k+1})$ converge to the same limit point $\bar{z}_p \in T_p(\bar{x})$ as k $(\in N) \to +\infty$ for each p.

For $p \in I_2 \subset \{1, \cdots, m_1\}$, given $\varphi_p^{\uparrow}$ is convex, real-valued, and $f_p^{k,\text{upper}}(x^{k,i_k+1}; x^{k,i_k}) \to_N \bar{z}_p$, we can invoke [22, Theorem 24.7] to deduce the boundedness of $\{y_{p,1}^k\}_{k\in N}$. A parallel reasoning applies to demonstrate the boundedness of $\{y_{p,2}^k\}_{k\in N}$ for $p \in I_2$. Note that $\{v_{p,1}^k\}_{k\in N}$, $\{v_{p,2}^k\}_{k\in N}$ must also be bounded for $p \in I_2$, otherwise we could assume $\|v_{p,1}^k\| \to_N +\infty$ and then every accumulation point of unit vectors $\{v_{p,1}^k/\|v_{p,1}^k\|\}_{k\in N}$ would be in the set $\partial_A^\infty f_p(\bar{x})$, contradicting our assumption that $\partial_A^\infty f_p(\bar{x}) = \{0\}$ for each $p \in I_2$.

For $p \in I_1$, suppose for contradiction that $\big\{ \sum_{p\in I_1} |y_{p,1}^k| \big\}_{k\in N}$ is unbounded and $\sum_{p\in I_1} |y_{p,1}^k| \to_N +\infty$ by passing to a subsequence. Consider the normalized subsequence $\{\widetilde{y}_{p,1}^k \triangleq y_{p,1}^k / \sum_{p\in I_1} |y_{p,1}^k|\}_{k\in N}$ for each p. Consequently, $\widetilde{y}_{p,1}^k \to_N 0$ for $p \in I_2$. By the triangle inequality and (34), we have

$$\begin{aligned} &\left\| \sum_{p\in I_1} \widetilde{y}_{p,1}^k\, v_{p,1}^k \right\| - \left\| \sum_{p\in I_2} \left(\widetilde{y}_{p,1}^k\, v_{p,1}^k + \widetilde{y}_{p,2}^k\, v_{p,2}^k\right) \right\| \leq \left\| \sum_{p\in I_1} \widetilde{y}_{p,1}^k\, v_{p,1}^k + \sum_{p\in I_2} \left(\widetilde{y}_{p,1}^k\, v_{p,1}^k + \widetilde{y}_{p,2}^k\, v_{p,2}^k\right) \right\| \\ \leq\;& \frac{\lambda\, \delta_k}{\ell_k} \cdot \frac{1}{\sum_{p\in I_1} |y_{p,1}^k|} + \left(1 + \frac{\sum_{p\in I_2} \left(|y_{p,1}^k| + |y_{p,2}^k|\right)}{\sum_{p\in I_1} |y_{p,1}^k|} \right) \delta_k \;\longrightarrow_N\; 0, \end{aligned}$$

which further implies $\| \sum_{p\in I_1} \widetilde{y}_{p,1}^k\, v_{p,1}^k \| \to_N 0$ by the boundedness of $\{v_{p,1}^k\}_{k\in N}$ and $\{v_{p,2}^k\}_{k\in N}$ for $p \in I_2$. Now suppose that $\widetilde{y}_{p,1}^k \to_N \widetilde{y}_{p,1}$ for $p \in I_1$. Then from a similar reasoning in (17), for $p \in I_1$,

$$\widetilde{y}_{p,1} \in \operatorname*{Lim\,sup}_{k(\in N)\to+\infty}{}^{\infty}\, \partial\varphi_p^{\uparrow} \left(f_p^{k,\text{upper}}(x^{k,i_k+1}; x^{k,i_k}) \right) \subset \partial^\infty \varphi_p^{\uparrow}(\bar{z}_p) = \mathcal{N}_{\operatorname{dom} \varphi_p^{\uparrow}}(\bar{z}_p),$$

and obviously $\sum_{p\in I_1} |\widetilde{y}_{p,1}| = 1$. The remaining argument to derive a contradiction to Assumption 5 is actually the same as the proof in Lemma 2 for two cases, except changing the index set $\{m_1+1,\cdots,m\}$ to I_1. Therefore, we establish the boundedness of $\{y^k_{p,1}\}_{k\in N}$ for $p \in I_1$ and $\{y^k_{p,2}\}_{k\in N}$ for $p \in I_1 \cup I_2$. Suppose that $y^k_{p,1} \to_N \bar{y}_{p,1}$ for $p \in I_1$ and $y^k_{p,2} \to_N \bar{y}_{p,2}$ for $p \in I_1 \cup I_2$. Then, by the outer semicontinuity of $\partial\varphi^{\uparrow}_p$, we have $\bar{y}_{p,1} \in \partial\varphi^{\uparrow}_p(\bar{z}_p)$ for $p \in I_1$. Similarly, $\bar{y}_{p,2} \in \partial\varphi^{\downarrow}_p(\bar{z}_p)$ for $p \in I_1 \cup I_2$.

Step 2: To proceed, we prove by contradiction that the sequence $\{y^k_{p,1}\, v^k_{p,1}\}_{k\in N}$ is bounded for $p \in I_1$. Suppose that $\sum_{p\in I_1} \|y^k_{p,1}\, v^k_{p,1}\| \to_N +\infty$. Based on step 1, assume that

$$\sum_{p\in I_2} \left(y^k_{p,1}\, v^k_{p,1} + y^k_{p,2}\, v^k_{p,2}\right) \to_N \ \bar{w} \quad \left(\in \sum_{p\in I_2} (\bar{y}_{p,1}\, \partial_A f_p(\bar{x}) + \bar{y}_{p,2}\, \partial_A f_p(\bar{x})) \right),$$

and thus $\sum_{p\in I_1} y^k_{p,1}\, v^k_{p,1} \to_N (-\bar{w})$. Consider $\widetilde{w}^k_p \triangleq y^k_{p,1}\, v^k_{p,1} / \sum_{p\in I_1} \|y^k_{p,1}\, v^k_{p,1}\|$ for $p \in I_1$, and then $\sum_{p\in I_1} \widetilde{w}^k_p \to_N 0$. Given $\sum_{p\in I_1} \|\widetilde{w}^k_p\| = 1$ for all $k \in N$, there must exist $p_1 \in I_1$ such that $\widetilde{w}^k_{p_1} \to_N \widetilde{w}_{p_1} \neq 0$. For each $p \in I_1$, it then follows from $y^k_{p,1} / \sum_{p\in I_1} \|y^k_{p,1} v^k_{p,1}\| \to_N 0$ that $\{\widetilde{w}^k_p\}_{k\in N}$ has a subsequence converging to some element in $\partial^\infty_A f_p(\bar{x})$. In particular, $\widetilde{w}_{p_1} \in \partial^\infty_A f_{p_1}(\bar{x})\backslash\{0\}$. Since $\sum_{p\in I_1} \widetilde{w}^k_p \to_N 0$, this implies that

$$0 \in [\,\partial^\infty_A f_{p_1}(\bar{x})\backslash\{0\}\,] + \sum_{p\in I_1\backslash\{p_1\}} \partial^\infty_A f_p(\bar{x}),$$

which contradicts Assumption 5. Hence, $\{y^k_{p,1}\, v^k_{p,1}\}_{k\in N}$ is bounded for $p \in I_1$.

Step 3: We are now ready to prove that $\bar{x}$ is a weakly A-stationary point. Suppose that $y^k_{p,1}\, v^k_{p,1} \to_N \bar{w}_p$ for $p \in I_1$ with $\sum_{p\in I_1} \bar{w}_p = -\bar{w}$. It remains to show that for each $p \in I_1$, there exists $\bar{y}_{p,1} \in \bigcup\{\partial\varphi^{\uparrow}_p(t_p) \mid t_p \in T_p(\bar{x})\}$ such that

$$\bar{w}_p \in \{\, \bar{y}_{p,1}\, \partial_A f_p(\bar{x})\,\} \cup [\,\partial^\infty_A f_p(\bar{x})\backslash\{0\}\,],$$

which can be derived similarly as the proof of (16) in Theorem 2. Summarizing these arguments, we conclude that $\bar{x}$ is a weakly A-stationary point of (CP_1).

Under the additional assumption of the theorem, there exist $y^k_{p,1} \in \partial\varphi^{\uparrow}_p(f^{k,\text{upper}}_p(x^{k,i_k+1}; x^{k,i_k}))$, $y^k_{p,2} \in \partial\varphi^{\downarrow}_p(f^{k,\text{lower}}_p(x^{k,i_k+1}; x^{k,i_k}))$, and

$$v^k_{p,1} \in \left\{\partial g^k_p(x^{k,i_k}) - \partial h^k_p(x^{k,i_k})\right\} \cup \left\{\partial g^k_p(x^{k,i_k+1}) - \partial h^k_p(x^{k,i_k+1})\right\}$$

such that

$$
\begin{aligned}
&\left\| \sum_{p \in I_1} y_{p,1}^k \, v_{p,1}^k + \sum_{p \in I_2} (y_{p,1}^k + y_{p,2}^k) \left[\nabla g_p^k(x^{k,i_k}) - \nabla h_p^k(x^{k,i_k}) \right] \right\| \\
&\overset{(\mathrm{i}')}{\leq} \lambda \|x^{k,i_k+1} - x^{k,i_k}\| + \sum_{p \in I_1} |y_{p,1}^k| \cdot \min \left\{ \left\| \nabla g_p^k(x^{k,i_k+1}) - \nabla g_p^k(x^{k,i_k}) \right\|, \left\| \nabla h_p^k(x^{k,i_k+1}) - \nabla h_p^k(x^{k,i_k}) \right\| \right\} \\
&\quad + \sum_{p \in I_2} \Big(|y_{p,1}^k| \cdot \|\nabla g_p^k(x^{k,i_k}) - \nabla g_p^k(x^{k,i_k+1})\| + |y_{p,2}^k| \cdot \|\nabla h_p^k(x^{k,i_k+1}) - \nabla h_p^k(x^{k,i_k})\| \Big) \\
&\overset{(\mathrm{ii}')}{\leq} \lambda \|x^{k,i_k+1} - x^{k,i_k}\| + \left(\sum_{p \in I_1} |y_{p,1}^k| + \sum_{p \in I_2} \big(|y_{p,1}^k| + |y_{p,2}^k| \big) \right) \ell_k \, \|x^{k,i_k+1} - x^{k,i_k}\| \\
&\overset{(\mathrm{iii}')}{\leq} \lambda \, \delta_k / \ell_k + \left(\sum_{p \in I_1} |y_{p,1}^k| + \sum_{p \in I_2} \big(|y_{p,1}^k| + |y_{p,2}^k| \big) \right) \delta_k \qquad \forall \, k \geq 0,
\end{aligned}
$$

where (i′) is implied by the optimality condition (32), (ii′) employs (31), and (iii′) follows from conditions (27). This inequality is a tighter version of (34) in the sense that, for each $p \in I_2$ and $k \geq 0$, $v_{p,1}^k$ and $v_{p,2}^k$ are elements taken from the single-valued mapping $\nabla g_p^k(\cdot) - \nabla h_p^k(\cdot)$ evaluated at the same point x^{k,i_k}. A straightforward adaptation of the preceding argument confirms that $\bar{x}$ is an A-stationary point of (CP_1). □

The algorithm in [16] for solving the bi-parameterized two-stage stochastic program with fixed scenarios can be viewed as a special application of the prox-ADC algorithm. Since each outer function φ_p in [16] is real-valued, the algorithm in the cited paper simplifies the stopping criteria of the inner loop by dropping the first two conditions in (27).

5 Conclusions.

In this paper, we have introduced a new class of composite functions that broadens the scope of the well-established class of amenable functions. Our principal objective has been to demonstrate that when the outer convex function is separable across each coordinate, and the inner function is ADC, the resulting composite function retains computational amenability. Despite the theoretical advances we have achieved, the practical implementation of this framework to address real-world applications is yet to be explored. Future work should aim to bridge this gap, translating the theoretical aspects of our findings into tangible computational solutions.

Acknowledgments.

The authors are partially supported by the National Science Foundation under grants CCF-2153352 and DMS-2309729, and the National Institutes of Health under grant 1R01CA287413-01. The authors are grateful to the associate editor and two reviewers for their careful reading and constructive comments that have substantially improved the paper.

References

[1] Shabbir Ahmed and Yongpei Guan. The inverse optimal value problem. *Mathematical Programming*, 102(1):91–110, 2005.

[2] Amir Beck and Marc Teboulle. Smoothing and first order methods: a unified framework. *SIAM Journal on Optimization*, 22(2):557–580, 2012.

[3] James V Burke. Descent methods for composite nondifferentiable optimization problems. *Mathematical Programming*, 33(3):260–279, 1985.

[4] James V Burke and Michael C Ferris. A Gauss-Newton method for convex composite optimization. *Mathematical Programming*, 71(2):179–194, 1995.

[5] James V Burke and Tim Hoheisel. Epi-convergent smoothing with applications to convex composite functions. *SIAM Journal on Optimization*, 23(3):1457–1479, 2013.

[6] James V Burke and Tim Hoheisel. Epi-convergence properties of smoothing by infimal convolution. *Set-Valued and Variational Analysis*, 25(1):1–23, 2017.

[7] James V Burke, Tim Hoheisel, and Christian Kanzow. Gradient consistency for integral-convolution smoothing functions. *Set-Valued and Variational Analysis*, 21(2):359–376, 2013.

[8] Xiaojun Chen. Smoothing methods for nonsmooth, nonconvex minimization. *Mathematical Programming*, 134(1):71–99, 2012.

[9] Frank H Clarke. Generalized gradients and applications. *Transactions of the American Mathematical Society*, 205:247–262, 1975.

[10] Frank H Clarke. *Optimization and Nonsmooth Analysis*. Society for Industrial and Applied Mathematics, 1990.

[11] Ying Cui and Jong-Shi Pang. *Modern Nonconvex Nondifferentiable Optimization*. SIAM Publications, Philadelphia, 2021.

[12] Dmitriy Drusvyatskiy and Courtney Paquette. Efficiency of minimizing compositions of convex functions and smooth maps. *Mathematical Programming*, 178:503–558, 2019.

[13] Yuri M Ermoliev, Vladimir I Norkin, and Roger JB Wets. The minimization of semicontinuous functions: mollifier subgradients. *SIAM Journal on Control and Optimization*, 33(1):149–167, 1995.

[14] Roger Fletcher. A model algorithm for composite nondifferentiable optimization problems. *Nondifferential and Variational Techniques in Optimization*, pages 67–76, 1982.

[15] Adrian S Lewis and Stephen J Wright. A proximal method for composite minimization. *Mathematical Programming*, 158(1-2):501–546, 2016.

[16] Hanyang Li and Ying Cui. A decomposition algorithm for two-stage stochastic programs with nonconvex recourse. *arXiv:2204.01269*, 2022.

[17] Junyi Liu, Ying Cui, Jong-Shi Pang, and Suvrajeet Sen. Two-stage stochastic programming with linearly bi-parameterized quadratic recourse. *SIAM Journal on Optimization*, 30(3):2530–2558, 2020.

[18] Boris S Mordukhovich. Generalized differential calculus for nonsmooth and set-valued mappings. *Journal of Mathematical Analysis and Applications*, 183(1):250–288, 1994.

[19] Giuseppe Paleologo and Samer Takriti. Bandwidth trading: A new market looking for help from the or community. *AIRO News VI*, 3:1–4, 2001.

[20] RA Poliquin and R Tyrrell Rockafellar. Amenable functions in optimization. *Nonsmooth Optimization: Methods and Applications (Erice, 1991)*, pages 338–353, 1992.

[21] RA Poliquin and R Tyrrell Rockafellar. A calculus of epi-derivatives applicable to optimization. *Canadian Journal of Mathematics*, 45(4):879–896, 1993.

[22] R Tyrrell Rockafellar. *Convex Analysis*, volume 18. Princeton University Press, 1970.

[23] R Tyrrell Rockafellar. Lagrange multipliers and optimality. *SIAM Review*, 35(2):183–238, 1993.

[24] R Tyrrell Rockafellar and Stanislav Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–42, 2000.

[25] R Tyrrell Rockafellar and Roger JB Wets. *Variational Analysis*, volume 317. Springer Science & Business Media, 2009.

[26] Johannes O Royset. Approximations of semicontinuous functions with applications to stochastic optimization and statistical estimation. *Mathematical Programming*, 184(1):289–318, 2020.

[27] Johannes O Royset. Consistent approximations in composite optimization. *Mathematical Programming*, pages 1–34, 2022.

[28] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on Stochastic Programming: Modeling and Theory (Third Edition)*. SIAM Publications, Philadelphia, 2021.

[29] Wim van Ackooij. A discussion of probability functions and constraints from a variational perspective. *Set-Valued and Variational Analysis*, 28(4):585–609, 2020.

[30] Pengyu Wei. Risk management with weighted VaR. *Mathematical Finance*, 28(4):1020–1060, 2018.

Appendix A. Proofs of Proposition 2 and Proposition 3

Proof of Proposition 2. (a) We first generalize the convergence result of the classical Moreau envelopes when $\gamma_k \downarrow 0$ (see, e.g., [25, Theorem 1.25]) to the partial Moreau envelopes. Fixing any $\gamma_0 > 0$, we consider the function $\psi(z, x, \gamma) \triangleq \overline{f}(z, x) + \delta_{\operatorname{dom} f}(x) + \psi_0(z, x, \gamma)$ with

$$\psi_0(z, x, \gamma) \triangleq \begin{cases} \|z - x\|^2/(2\gamma) & \text{if } \gamma \in (0, \gamma_0], \\ 0 & \text{if } \gamma = 0, z = x, \\ \infty & \text{otherwise.} \end{cases}$$

Notice that $f^k(x) = g_{\gamma_k}(x) - h_{\gamma_k}(x) + \delta_{\mathrm{dom}\, f}(x) = \inf_z \psi(z, x, \gamma_k)$. It is easy to verify that ψ is proper and lsc based on our assumptions. Under the assumption that $\overline{f}$ is bounded from below on $\mathrm{dom}\, f \times \mathrm{dom}\, f$, we can also show by contradiction that $\psi(z, x, \gamma)$ is level-bounded in z locally uniformly in (x, γ). Consequently, it follows from [25, Theorem 1.17] that $f^k(x) = \inf_z \psi(z, x, \gamma_k) \uparrow f(x)$ for any fixed x and each f^k is lsc.

Hence, $f^k \xrightarrow{\text{e}} f$ is a direct consequence of [25, Proposition 7.4(d)] by $f^k(x) \uparrow f(x)$ for all x and the lower semicontinuity of f^k. If $\mathrm{dom}\, f = \mathbb{R}^{n_1}$, then f is continuous, and thus $f^k \xrightarrow{\text{c}} f$ by [25, Proposition 7.4(c-d)]. We then complete the proof of (a).

(b) For any $\bar{x} \in \mathrm{int}(\mathrm{dom}\, f)$,

$$\begin{aligned}
\partial_A f(\bar{x}) = &\bigcup_{x^k \to \bar{x}} \mathop{\mathrm{Lim\,sup}}_{k \to +\infty} \left\{ \partial g^k(x^k) - \partial h^k(x^k) \right\} \\
\overset{\text{(i)}}{=} &\bigcup_{x^k \to \bar{x}} \mathop{\mathrm{Lim\,sup}}_{k \to +\infty} \left\{ \frac{x^k}{\gamma_k} - \partial_2(-\overline{f})(z^k, x^k) - \frac{z^k}{\gamma_k} \;\middle|\; z^k = \mathop{\arg\min}_{z \in \mathbb{R}^{n_1}} \left[\overline{f}(z, x^k) + \frac{1}{2\gamma_k} \|z - x^k\|^2 \right] \right\} \\
\overset{\text{(ii)}}{\subset} &\bigcup_{(x^k, z^k) \to (\bar{x}, \bar{x})} \mathop{\mathrm{Lim\,sup}}_{k \to +\infty} \left[\partial_1 \overline{f}(z^k, x^k) - \partial_2(-\overline{f})(z^k, x^k) \right] \\
\overset{\text{(iii)}}{=} &\ \partial_1 \overline{f}(\bar{x}, \bar{x}) - \partial_2(-\overline{f})(\bar{x}, \bar{x}),
\end{aligned}$$

where (i) follows from the convexity of $(-\overline{f})(z, \cdot)$ for any $z \in \mathrm{dom}\, f$ and Danskin's Theorem [9, Theorem 2.1]; (ii) is due to the optimality condition for z^k, and $z^k \to \bar{x}$ is obtained by similar arguments in the proof of Theorem 1(b) due to our assumption that $\overline{f}$ is bounded from below on $\mathrm{dom}\, f \times \mathrm{dom}\, f$; and (iii) uses the outer semicontinuity of $\partial_1 \overline{f}$ and $\partial_2(-\overline{f})$ at $(\bar{x}, \bar{x})$ [16, Lemma 5]. Therefore, for any $\bar{x} \in \mathrm{int}(\mathrm{dom}\, f)$, $\partial f(\bar{x}) \subset \partial_A f(\bar{x}) \subset \partial_1 \overline{f}(\bar{x}, \bar{x}) - \partial_2(-\overline{f})(\bar{x}, \bar{x})$. Moreover, due to the local boundedness of the mappings $\partial_1 \overline{f}$ and $\partial_2(-\overline{f})$ at $(\bar{x}, \bar{x})$ [16, Lemma 5], it follows from [25, Example 4.22] that $\partial_A^\infty f(\bar{x}) = \{0\}$. □

Proof of Proposition 3. (a) Note that for any $x \in \mathbb{R}^n$, $\mathrm{CVaR}_\alpha^+[\, c(x, Z)\,]$ is well-defined and takes finite value due to $\mathbb{E}[\, |c(x, Z)|\,] < +\infty$. Since $c(x, Z)$ follows a continuous distribution for any $x \in \mathbb{R}^n$, we know that

$$\mathrm{CVaR}_\alpha^+[\, c(x, Z)\,] = \inf_{t \in \mathbb{R}} \left\{ t + \frac{1}{1 - \alpha} \mathbb{E}\left[\max\{c(x, Z) - t, 0\} \right] \right\} = \frac{1}{1 - \alpha} \int_\alpha^1 \mathrm{VaR}_t[\, c(x, Z)\,]\, \mathrm{d}t,$$

and $\mathrm{CVaR}_\alpha^+[\, c(\cdot, Z)\,]$ is convex by the convexity of $c(\cdot, z)$ for any fixed $z \in \mathbb{R}^m$ (cf. [24, Theorem 2]). Therefore, both g^k and h^k defined in (6) are convex. For any $x \in \mathbb{R}^n$, we have

$$g^k(x) - h^k(x) = k \int_{\alpha - 1/k}^1 \mathrm{VaR}_t[\, c(x, Z)\,]\, \mathrm{d}t - k \int_\alpha^1 \mathrm{VaR}_t[\, c(x, Z)\,]\, \mathrm{d}t = k \int_{\alpha - 1/k}^\alpha \mathrm{VaR}_t[\, c(x, Z)\,]\, \mathrm{d}t. \tag{35}$$

Thus, $\mathrm{VaR}_{\alpha - 1/k}[\, c(x, Z)\,] \le g^k(x) - h^k(x) \le \mathrm{VaR}_\alpha[\, c(x, Z)\,]$ for any $x \in \mathbb{R}^n$ and $k > 1/\alpha$. Since $\mathrm{VaR}_t(Z)$ as a function of t on $(0, 1)$ is left-continuous, it follows that $[g^k(x) - h^k(x)] \uparrow \mathrm{VaR}_\alpha[\, c(x, Z)\,]$ for all x. Observe that

$$\{x \mid \mathrm{VaR}_\alpha[\, c(x, Z)\,] \le r\} = \{x \mid \mathbb{P}(c(x, Z) \le r) \ge \alpha\}.$$

Based on our assumptions and [29, Proposition 2.2], for any $r \in \mathbb{R}$, the probability function $x \mapsto -\mathbb{P}(c(x, Z) \leq r)$ is lsc, which implies the closedness of the level set $\{x \mid \mathbb{P}(c(x, Z) \leq r) \geq \alpha\}$ for any $(r, \alpha) \in \mathbb{R} \times (0, 1)$. Hence, $\mathrm{VaR}_\alpha[\, c(x, Z)\,]$ is lsc for any given $\alpha \in (0, 1)$ and is continuous if $c(\cdot, \cdot)$ is further assumed to be continuous. Then (a) follows from [25, Proposition 7.4(c-d)].

(b) We use $\mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ to denote the space of all random variables $\phi : \Omega \to \mathbb{R}$ with $\mathbb{E}[\,|\phi(\omega)|] < +\infty$. According to [28, Example 6.19], the function $\mathrm{CVaR}^+_\alpha : \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ is subdifferentiable (see [28, (9.281)] for the definition). Consider any fixed $x \in \mathbb{R}^n$. Given that $c(x, Z)$ is a continuous random variable, it follows from [28, (6.81)] that the subdifferential of $\mathrm{CVaR}^+_\alpha[\cdot]$ at $c(x, Z)$ is:

$$\partial \left(\mathrm{CVaR}^+_\alpha[\,\cdot\,]\right)\, [c(x, Z)] = \left\{ \phi : \Omega \to \mathbb{R}_+ \;\middle|\; \begin{array}{ll} \phi(\omega) = (1-\alpha)^{-1} & \text{if } c(x, Z(\omega)) > \mathrm{VaR}_\alpha[\, c(x, Z)] \\ \phi(\omega) = 0 & \text{if } c(x, Z(\omega)) < \mathrm{VaR}_\alpha[\, c(x, Z)] \end{array} \right\}.$$

Let $\mathbb{P}_Z$ denote the probability measure associated with Z. By using [28, Theorem 6.14]), we obtain the subdifferential of the convex function $\mathrm{CVaR}^+_\alpha[\, c(\cdot, Z)\,]$ at x:

$$\partial \left(\mathrm{CVaR}^+_\alpha[\, c(\cdot, Z)\,]\right)(x) = \mathrm{cl}\left(\bigcup_{\phi \in \partial(\mathrm{CVaR}^+_\alpha[\,\cdot\,])\,[c(x,Z)]} \int \partial_x\, c(x, Z(\omega))\, \phi(\omega)\, \mathrm{d}\mathbb{P}_Z(\omega) \right).$$

By the convexity of $c(\cdot, z)$ for any fixed $z \in \mathbb{R}^m$ and the existence of a measurable function κ, it follows from [9, Theorem 2.7.2] that the set $\int \partial_x c(x, Z(\omega))\, \phi(\omega)\, \mathrm{d}\mathbb{P}_Z(\omega) = \partial \int c(x, Z(\omega))\, \phi(\omega)\, \mathrm{d}\mathbb{P}_Z(\omega)$ is closed. Then, for any $k > 1/\alpha$, $\{\partial g^k(x) - \partial h^k(x)\}$ can be written as

$$\left\{ \int \partial_x\, c(x, Z(\omega))\, \phi(\omega)\, \mathrm{d}\mathbb{P}_Z(\omega) \;\middle|\; \phi(\omega) = k \text{if } \mathrm{VaR}_{\alpha - 1/k}[\, c(x, Z)] < c(x, Z(\omega)) < \mathrm{VaR}_\alpha[\, c(x, Z)] \right\}.$$

We then complete the proof by the definition of the approximate subdifferential. □

Appendix B. The proof of Proposition 4

We start with the chain rules for $\partial(\varphi \circ f)$ and $\partial^\infty(\varphi \circ f)$ where the inner function f is merely lsc. These results are extensions of the nonlinear rescaling [25, Proposition 10.19(b)] to the case where φ may lack the strictly increasing property at a given point. One can also derive the same results through a general chain rule of the coderivative for composite set-valued mappings [18, Theorem 5.1]. However, to avoid the complicated computations accompanied by the introduction of the coderivative, we give an alternative proof below that is more straightforward. We will present the proof of Proposition 4 after this lemma.

Lemma 3 (chain rules for the limiting subdifferential)**.** *Let $\varphi : \mathbb{R} \to \overline{\mathbb{R}}$ be proper, lsc, convex, and nondecreasing with $\sup \varphi = +\infty$, and $f : \mathbb{R}^n \to \mathbb{R}$ be lsc. Consider $\bar{x} \in \mathrm{dom}(\varphi \circ f)$. If the only scalar $y \in \mathrm{Limsup}_{x \to_{(\varphi \circ f)} \bar{x}}\, \mathcal{N}_{\mathrm{dom}\, \varphi}(f(x))$ with $0 \in y \cdot \mathrm{Limsup}_{x \to \bar{x}}\, \partial f(x)$ is $y = 0$, then*

$$\partial(\varphi \circ f)(\bar{x}) \subset \bigcup \left\{ y \cdot \operatorname*{Lim\,sup}_{x \to \bar{x}} \partial f(x) \;\middle|\; y \in \operatorname*{Lim\,sup}_{x \to_{(\varphi \circ f)}\, \bar{x}} \partial \varphi(f(x)) \right\} \cup \left[\operatorname*{Lim\,sup}_{x \to \bar{x}}{}^{\infty} \partial f(x) \backslash \{0\} \right],$$

$$\partial^\infty(\varphi \circ f)(\bar{x}) \subset \bigcup \left\{ y \cdot \operatorname*{Lim\,sup}_{x \to \bar{x}} \partial f(x) \;\middle|\; y \in \operatorname*{Lim\,sup}_{x \to_{(\varphi \circ f)}\, \bar{x}} \mathcal{N}_{\mathrm{dom}\, \varphi}(f(x)) \right\} \cup \left[\operatorname*{Lim\,sup}_{x \to \bar{x}}{}^{\infty} \partial f(x) \backslash \{0\} \right].$$

Proof of Lemma 3. The basic idea is to rewrite $\varphi \circ f$ as a parametric minimization problem and apply [25, Theorem 10.13]. Note that $\varphi\big(f(x)\big) = \inf_\alpha\, [g(x,\alpha) \triangleq \delta_{\mathrm{epi}\, f}(x,\alpha) + \varphi(\alpha)]$ for $x \in \mathrm{dom}(\varphi \circ f)$. Define the corresponding set of optimal solutions as $\Lambda(x)$ for any $x \in \mathrm{dom}(\varphi \circ f)$. Then, we have $f(\bar{x}) \in \Lambda(\bar{x})$ and $\varphi(\alpha) = \varphi(f(\bar{x}))$ for any $\alpha \in \Lambda(\bar{x})$. By our assumptions, it is obvious that $\mathrm{dom}\,\varphi \in \{(-\infty, b), (-\infty, b]\}$ for some $b \in \mathbb{R} \cup \{+\infty\}$. Based on our assumption that $\sup \varphi = +\infty$ and f is lsc, it is easy to verify that g is proper, lsc, and level-bounded in α locally uniformly in x. Then we apply [25, Theorem 10.13] to obtain

$$\partial(\varphi \circ f)(\bar{x}) \subset \{\, v \mid (v,0) \in \partial g(\bar{x},\bar{\alpha}), \bar{\alpha} \in \Lambda(\bar{x}) \,\}\,, \quad \partial^\infty(\varphi \circ f)(\bar{x}) \subset \{\, v \mid (v,0) \in \partial^\infty g(\bar{x},\bar{\alpha}), \bar{\alpha} \in \Lambda(\bar{x}) \,\}\,. \tag{36}$$

<u>*Step 1:*</u> We will show that for any $\bar{\alpha} \in \Lambda(\bar{x})$,

$$\mathcal{N}_{\mathrm{epi}\, f}(\bar{x},\bar{\alpha}) \cap \big(\, \{0\} \times [-\mathcal{N}_{\mathrm{dom}\,\varphi}(\bar{\alpha})] \,\big) = \{0\}. \tag{37}$$

We divide the proof of (37) into two cases.
<u>*Case 1.*</u> If $\Lambda(\bar{x})$ is a singleton $\{f(\bar{x})\}$, we can characterize $\mathcal{N}_{\mathrm{epi}\, f}(\bar{x}, f(\bar{x}))$ by using the result in [25, Theorem 8.9]. Since $\partial f(\bar{x}) \subset \mathrm{Limsup}_{x \to \bar{x}}\, \partial f(x)$ and $\mathcal{N}_{\mathrm{dom}\,\varphi}(f(\bar{x})) \subset \mathrm{Limsup}_{x \to_{(\varphi \circ f)} \bar{x}}\, \mathcal{N}_{\mathrm{dom}\,\varphi}(f(x))$, it follows from our assumption that either $0 \notin \partial f(\bar{x})$ or $\mathcal{N}_{\mathrm{dom}\,\varphi}(f(\bar{x})) = \{0\}$. Hence, (37) is satisfied.
<u>*Case 2.*</u> Otherwise, there exists $\bar{\alpha}_{\max} \in (f(\bar{x}), +\infty)$ such that $\Lambda(\bar{x}) = [f(\bar{x}), \bar{\alpha}_{\max}]$ since φ is lsc, nondecreasing and $\sup \varphi = +\infty$. Thus,

$$\begin{aligned} \partial(\varphi \circ f)(\bar{x}) &\subset \big[\{v \mid (v,0) \in \partial g(\bar{x}, f(\bar{x}))\} \cup \{v \mid (v,0) \in \partial g(\bar{x},\bar{\alpha}), f(\bar{x}) < \bar{\alpha} \le \bar{\alpha}_{\max}\}\big], \\ \partial^\infty(\varphi \circ f)(\bar{x}) &\subset \big[\{v \mid (v,0) \in \partial^\infty g(\bar{x}, f(\bar{x}))\} \cup \{v \mid (v,0) \in \partial^\infty g(\bar{x},\bar{\alpha}), f(\bar{x}) < \bar{\alpha} \le \bar{\alpha}_{\max}\}\big]. \end{aligned} \tag{38}$$

Let $\Lambda_1(\bar{x}) \triangleq \big\{\bar{\alpha} \in (f(\bar{x}), \bar{\alpha}_{\max}] \,\big|\, \exists\, x^k \to \bar{x} \text{with } f(x^k) \to \bar{\alpha}\big\}$ and $\Lambda_2(\bar{x}) \triangleq \Lambda(\bar{x}) \backslash \Lambda_1(\bar{x})$. In the following, we characterize $\mathcal{N}_{\mathrm{epi}\, f}(\bar{x},\bar{\alpha})$ and verify (37) separately for $\bar{\alpha} \in \Lambda_1(\bar{x})$ and $\bar{\alpha} \in \Lambda_2(\bar{x})$.
<u>*Case 2.1.*</u> For any $\bar{\alpha} \in \Lambda_1(\bar{x})$, we first prove the inclusion:

$$\mathcal{N}_{\mathrm{epi}\, f}(\bar{x},\bar{\alpha}) \subset \left[\left\{ \lambda(v,-1) \,\middle|\, v \in \operatorname*{Lim\,sup}_{x \to \bar{x}} \partial f(x), \lambda > 0 \right\} \cup \left\{ (v,0) \,\middle|\, v \in \operatorname*{Lim\,sup}_{x \to \bar{x}}{}^{\infty} \partial f(x) \right\} \right]. \tag{39}$$

Observe that for any $\bar{\alpha} \in \Lambda_1(\bar{x})$, it holds that

$$\mathcal{N}_{\mathrm{epi}\, f}(\bar{x},\bar{\alpha}) \subset \operatorname*{Lim\,sup}_{(x,\alpha)(\in \mathrm{epi}\, f) \to (\bar{x},\bar{\alpha})} \mathcal{N}^p_{\mathrm{epi}\, f}(x,\alpha) \subset \operatorname*{Lim\,sup}_{x \to \bar{x}} \mathcal{N}^p_{\mathrm{epi}\, f}(x, f(x)) \subset \operatorname*{Lim\,sup}_{x \to \bar{x}} \mathcal{N}_{\mathrm{epi}\, f}(x, f(x)), \tag{40}$$

where the first inclusion is because any normal vector is a limit of proximal normals at nearby points [25, Exercise 6.18]; the second one uses the fact that, for any fixed $\alpha > f(x)$, any proximal normal to $\mathrm{epi}\, f$ at (x,α) is also a proximal normal to $\mathrm{epi}\, f$ at $(x, f(x))$; the last inclusion follows directly from the definition of proximal normals. Based on the the result of [25, Theorem 8.9] that

$$\mathcal{N}_{\mathrm{epi}\, f}(x, f(x)) = \{\lambda(v,-1) \,|\, v \in \partial f(x), \lambda > 0\} \cup \{(v,0) \,|\, v \in \partial^\infty f(x)\}\,,$$

we conclude that $\mathcal{N}_{\mathrm{epi}\, f}(\bar{x},\bar{\alpha}) \subset \mathbb{R}^n \times \mathbb{R}_-$ for any $\bar{\alpha} \in \Lambda_1(\bar{x})$. For any $(v,-1) \in \mathcal{N}_{\mathrm{epi}\, f}(\bar{x},\bar{\alpha})$ with $\bar{\alpha} \in \Lambda_1(\bar{x})$, there exist $x^k \to \bar{x}$, $v^k \to v$ with $v^k \in \partial f(x^k)$. Then $v \in \mathrm{Limsup}_{x \to \bar{x}}\, \partial f(x)$.

To prove (39), it remains to show that $v \in \mathrm{Limsup}^\infty_{x \to \bar{x}}\, \partial f(x)$ whenever $(v,0) \in \mathcal{N}_{\mathrm{epi}\, f}(\bar{x},\bar{\alpha})$. It follows from (40) that $(v,0)$ is a limit of proximal normals of $\mathrm{epi}\, f$ at $(x^k, f(x^k))$ for some sequence

$x^k \to \bar{x}$. (i) First consider the case $(v^k, 0) \to (v, 0)$ with $(v^k, 0) \in \mathcal{N}^p_{\operatorname{epi} f}(x^k, f(x^k))$. Following the argument in the proof of [25, Theorem 8.9], we can derive $v^k \in \partial^\infty f(x^k)$. Therefore,

$$v \in \operatorname*{Lim\,sup}_{k\to+\infty} \partial^\infty f(x^k) \subset \operatorname*{Lim\,sup}_{k\to+\infty} \left(\bigcup_{x^{k,i} \to_f x^k} \operatorname*{Lim\,sup}^\infty_{i\to+\infty} \partial f(x^{k,i}) \right) \subset \bigcup_{x^j \to \bar{x}} \operatorname*{Lim\,sup}^\infty_{j\to+\infty} \partial f(x^j),$$

where the first inclusion is due to the definition of the horizon subdifferential, and the last inclusion follows from a standard diagonal extraction procedure. (ii) In the other case, we have $\lambda_k(v^k, -1) \to (v, 0)$ with $\lambda_k \downarrow 0$ and $v^k \in \partial f(x^k)$ for all $k \geq 0$. It is easy to see $v \in \operatorname{Limsup}^\infty_{x\to\bar{x}} \partial f(x)$. So far, we obtain inclusion (39). Since $\bar{\alpha} \in \Lambda_1(\bar{x})$, we have $\mathcal{N}_{\operatorname{dom}\varphi}(\bar{\alpha}) \subset \operatorname{Limsup}_{x \to_{(\varphi\circ f)} \bar{x}} \mathcal{N}_{\operatorname{dom}\varphi}(f(x))$, and our assumption implies that $\lambda = 0$ is the unique solution satisfying $0 \in \lambda \cdot \operatorname{Limsup}_{x\to\bar{x}} \partial f(x)$ with $\lambda \in \mathcal{N}_{\operatorname{dom}\varphi}(\bar{\alpha})$. Thus, (37) is satisfied.

Case 2.2. For any $\bar{\alpha} \in \Lambda_2(\bar{x})$, consider any sequence $\{(x^k, \alpha^k)\} \subset \operatorname{epi} f$ converging to $(\bar{x}, \bar{\alpha})$. Then $\alpha^k > f(x^k)$ for all k sufficiently large since $\bar{\alpha} \notin \Lambda_1(\bar{x})$. It is easy to see that $\mathcal{N}^p_{\operatorname{epi} f}(x^k, \alpha^k) \subset \mathbb{R}^n \times \{0\}$, which gives us $\mathcal{N}_{\operatorname{epi} f}(x^k, \alpha^k) \subset \mathbb{R}^n \times \{0\}$. By following a similar pattern as the final part of Case 2.1, it is not difficult to obtain, for any $\bar{\alpha} \in \Lambda_2(\bar{x})$,

$$\mathcal{N}_{\operatorname{epi} f}(\bar{x}, \bar{\alpha}) \subset \left\{ (v, 0) \,\middle|\, v \in \operatorname*{Limsup}^\infty_{x\to\bar{x}} \partial f(x) \right\}. \tag{41}$$

In this case, (37) holds trivially. Hence, we have verified (37) for cases 1 and 2.

Step 2: Based on (37) in step 1, we can now apply the sum rule [25, Corollary 10.9] for $\partial g(\bar{x}, \bar{\alpha})$ to obtain

$$\partial g(\bar{x}, \bar{\alpha}) \subset \mathcal{N}_{\operatorname{epi} f}(\bar{x}, \bar{\alpha}) + \{0\} \times \partial\varphi(\bar{\alpha}), \qquad \partial^\infty g(\bar{x}, \bar{\alpha}) \subset \mathcal{N}_{\operatorname{epi} f}(\bar{x}, \bar{\alpha}) + \{0\} \times \mathcal{N}_{\operatorname{dom}\varphi}(\bar{\alpha}). \tag{42}$$

Case 1. For $\Lambda(\bar{x}) = \{f(\bar{x})\}$, by combining (42) with (36), we can derive the stated results for $\partial(\varphi \circ f)(\bar{x})$ and $\partial^\infty(\varphi \circ f)(\bar{x})$ based on the observations that $\partial\varphi(f(\bar{x})) \subset \operatorname{Limsup}_{x\to_{(\varphi\circ f)}\bar{x}} \varphi(f(x))$ and $\partial^\infty f(\bar{x}) \subset \operatorname{Limsup}^\infty_{x\to\bar{x}} \partial f(x)$.

Case 2. Otherwise, by (42), we have

$$\begin{aligned}
&\{v \mid (v, 0) \in \partial g(\bar{x}, \bar{\alpha}), f(\bar{x}) < \bar{\alpha} \leq \bar{\alpha}_{\max}\} \\
\overset{(39)(41)}{\subset}\ & \bigcup \left\{ y \cdot \operatorname*{Lim\,sup}_{x\to\bar{x}} \partial f(x) \,\middle|\, y \in \partial\varphi(\bar{\alpha}), \bar{\alpha} \in \Lambda_1(\bar{x}) \right\} \cup \left\{ \operatorname*{Lim\,sup}^\infty_{x\to\bar{x}} \partial f(x) \,\middle|\, 0 \in \partial\varphi(\bar{\alpha}), f(\bar{x}) < \bar{\alpha} \leq \bar{\alpha}_{\max} \right\} \\
\subset\ & \bigcup \left\{ y \cdot \operatorname*{Lim\,sup}_{x\to\bar{x}} \partial f(x) \,\middle|\, y \in \operatorname*{Lim\,sup}_{x\to_{(\varphi\circ f)}\bar{x}} \partial\varphi(f(x)) \right\} \cup \left[\operatorname*{Lim\,sup}^\infty_{x\to\bar{x}} \partial f(x) \backslash \{0\} \right],
\end{aligned}$$

where the last inclusion is because 0 will be included in the first set if $0 \in \partial\varphi(\bar{\alpha})$ for some $\bar{\alpha} \in (f(\bar{x}), \bar{\alpha}_{\max}]$ and the second set will be empty otherwise. Similarly,

$$\begin{aligned}
&\{v \mid (v, 0) \in \partial g^\infty(\bar{x}, \bar{\alpha}), f(\bar{x}) < \bar{\alpha} \leq \bar{\alpha}_{\max}\} \\
\subset\ & \bigcup \left\{ y \cdot \operatorname*{Lim\,sup}_{x\to\bar{x}} \partial f(x) \,\middle|\, y \in \operatorname*{Lim\,sup}_{x\to_{(\varphi\circ f)}\bar{x}} \mathcal{N}_{\operatorname{dom}\varphi}(f(x)) \right\} \cup \left[\operatorname*{Lim\,sup}^\infty_{x\to\bar{x}} \partial f(x) \backslash \{0\} \right].
\end{aligned}$$

We then complete the proof by using the inclusions in (38). $\square$

Equipped with the chain rules, we are now ready to prove Proposition 4.

Proof of Proposition 4. Let $\bar{x}$ be any feasible point, i.e., $\bar{x} \in \cap_{p=1}^m \operatorname{dom} F_p$. Suppose for contradiction that (12) does not hold at $\bar{x}$. Thus, there exist $p_1 \in \{1, \cdots, m\}$, $\{x^k\} \in S_{p_1}(\bar{x})$ and an index set $N \in \mathbb{N}_\infty^\sharp$ such that $0 \in \partial_C f_{p_1}^k(x^k)$ and $\mathcal{N}_{\operatorname{dom} \varphi_{p_1}}\left(f_{p_1}^k(x^k)\right) \neq \{0\}$ for all $k \in N$. Take an arbitrary nonzero scalar $y^k \in \mathcal{N}_{\operatorname{dom} \varphi_{p_1}}\left(f_{p_1}^k(x^k)\right)$ for all $k \in N$. Let $\widetilde{y}$ be any accumulation point of the unit scalars $\{y^k/|y^k|\}_{k \in N}$. Then, we have $(0 \neq)\widetilde{y} \in \bigcup\{\mathcal{N}_{\operatorname{dom} \varphi_{p_1}}(t_{p_1}) \mid t_{p_1} \in T_{p_1}(\bar{x})\}$ and $0 \in \operatorname{con} \partial_A f_{p_1}(\bar{x})$ by Proposition 1(a), contradicting Assumption 5. This proves condition (12).

For any fixed $p = 1, \cdots, m$, let $y_{p'} = 0$ for any $p' \in \{1, \cdots, m\} \backslash \{p\}$ in Assumption 5. Then the only scalar $y_p \in \bigcup\left\{\mathcal{N}_{\operatorname{dom} \varphi_p}(t_p) \mid t_p \in T_p(\bar{x})\right\}$ with $0 \in y_p \operatorname{con} \partial_A f_p(\bar{x})$ is $y_p = 0$, which completes the proof of (13).

To derive the constraint qualification (14), we consider two cases.

<u>*Case 1.*</u> For $p \in I_2$, we have $\mathcal{N}_{\operatorname{dom} \varphi_p}(f_p(\bar{x})) \subset \bigcup\left\{\mathcal{N}_{\operatorname{dom} \varphi_p}(t_p) \mid t_p \in T_p(\bar{x})\right\}$ due to $f_p^k \xrightarrow{e} f_p$ and $\partial(y f_p)(\bar{x}) \subset y\, \partial_C f_p(\bar{x}) \subset y \cdot \operatorname{con} \partial_A f_p(\bar{x})$ for any y by Theorem 1. Together with Assumption 5, we deduce that the only scalar $y \in \mathcal{N}_{\operatorname{dom} \varphi_p}(f_p(\bar{x}))$ with $0 \in \partial(y f_p)(\bar{x})$ is $y = 0$. From this condition and the local Lipschitz continuity of f_p for $p \in I_2$, we can apply the chain rule [25, Theorem 10.49] to get

$$\partial^\infty(\varphi_p \circ f_p)(\bar{x}) \subset \bigcup\left\{y \cdot \operatorname{con} \partial_A f_p(\bar{x}) \mid y \in \mathcal{N}_{\operatorname{dom} \varphi_p}(t_p), t_p \in T_p(\bar{x})\right\}. \tag{43}$$

<u>*Case 2.*</u> For $p \in I_1$, to utilize the chain rules (Proposition 3) for $\partial^\infty(\varphi_p \circ f_p)$, we must first confirm the validity of the condition:

$$\left[0 \in y \cdot \operatorname*{Lim\,sup}_{x \to \bar{x}} \partial f_p(x), \quad y \in \operatorname*{Lim\,sup}_{x \to_{F_p} \bar{x}} \mathcal{N}_{\operatorname{dom} \varphi_p}(f_p(x))\right] \implies y = 0. \tag{44}$$

Indeed, it suffices to consider the case of $\operatorname{dom} \varphi_p^\uparrow = (-\infty, r_p)$ or $(-\infty, r_p]$ for some $r_p \in \mathbb{R}$, because the statement holds trivially when $\varphi_p^\uparrow$ is real-valued. For any element $\bar{y} \in \operatorname{Limsup}_{x \to_{F_p} \bar{x}} \mathcal{N}_{\operatorname{dom} \varphi_p}(f_p(x))$, there exist $(x^k, y_k) \to (\bar{x}, \bar{y})$ with $y_k \in \mathcal{N}_{\operatorname{dom} \varphi_p}(f_p(x^k))$ and $F_p(x^k) \to F_p(\bar{x})$. Since $\bar{x} \in \operatorname{dom} F_p$, we must have $x^k \in \operatorname{dom} F_p$ for k sufficiently large, i.e., $f_p(x^k) \in \operatorname{dom} \varphi_p^\uparrow$, and $\{f_p(x^k)\}_{k \geq 0}$ is bounded from above due to $\operatorname{dom} \varphi_p^\uparrow = (-\infty, r_p)$ or $(-\infty, r_p]$. It follows immediately from the lower semicontinuity of f_p that $\{f_p(x^k)\}_{k \geq 0}$ is bounded. Assume that this sequence converges to some $\bar{z}_p$. Note that $\bar{z}_p \in \operatorname{dom} \varphi_p$ due to $F_p(\bar{x}) = \liminf_{k \to +\infty} \varphi_p(f_p(x^k)) \geq \varphi_p(\bar{z}_p)$. Thus, by the outer semicontinuity, $y_k \to \bar{y} \in \mathcal{N}_{\operatorname{dom} \varphi_p}(\bar{z}_p)$. By $f_p^k \xrightarrow{e} f_p$, each $f_p(x^k)$ can be expressed as the limit of a sequence $\{f_p^i(x^{k,i})\}_{i \geq 0}$ with $x^{k,i} \to x^k$ for any fixed $k \geq 0$. Using a standard diagonal extraction procedure, one can extract a subsequence $f_p^{i_k}(x^{k,i_k}) \to \bar{z}_p$ with $x^{k,i_k} \to \bar{x}$. Hence, $\bar{z}_p \in T_p(\bar{x})$ and

$$\operatorname*{Limsup}_{x \to_{F_p} \bar{x}} \mathcal{N}_{\operatorname{dom} \varphi_p}(f_p(x)) \subset \bigcup\{\mathcal{N}_{\operatorname{dom} \varphi_p}(t_p) \mid t_p \in T_p(\bar{x})\}. \tag{45}$$

Using the subdifferentials relationships in Proposition 1 and the outer semicontinuity of $\partial_A f_p$ in Proposition 1(a), we have

$$\operatorname*{Lim\,sup}_{x \to \bar{x}} \partial f_p(x) \subset \operatorname*{Lim\,sup}_{x \to \bar{x}} \partial_A f_p(x) = \partial_A f_p(\bar{x}). \tag{46}$$

By (45), (46) and Assumption 5, we immediately get (44). Thus, we can apply the chain rule in Proposition 3, and use (45), (46) again to obtain

$$\begin{aligned} \partial^{\infty}(\varphi_p \circ f_p)(\bar{x}) \quad &\subset \bigcup \left\{ y\, \partial_A f_p(x) \,\middle|\, y \in \mathcal{N}_{\operatorname{dom} \varphi_p}(t_p),\ t_p \in T_p(\bar{x}) \right\} \cup \left[\operatorname*{Lim\,sup}_{x \to \bar{x}}{}^{\infty} f_p(x) \backslash \{0\} \right] \\ &\subset \bigcup \left\{ y\, \partial_A f_p(x) \,\middle|\, y \in \mathcal{N}_{\operatorname{dom} \varphi_p}(t_p),\ t_p \in T_p(\bar{x}) \right\} \cup \left[\partial_A^{\infty} f_p(\bar{x}) \backslash \{0\} \right]. \end{aligned} \tag{47}$$

For the last inclusion, we use $\operatorname*{Lim\,sup}_{x \to \bar{x}}{}^{\infty}\, \partial f_p(x) \subset \operatorname*{Lim\,sup}_{x \to \bar{x}}{}^{\infty}\, \partial_A f_p(x) \subset \partial_A^{\infty} f_p(\bar{x})$ by Theorem 1(a) and using the diagonal extraction procedure again. Combining inclusions (43), (47) for two cases with Assumption 5, we derive (14) and complete the proof. □