Characterizations of ε-duality gap statements for composed optimization problems

Horaţiu-Vasile Boncea* Sorin-Mihai Grad†

Abstract. In this paper we present different regularity conditions that equivalently characterize ε-duality gap statements for optimization problems consisting of minimizing the sum of a function with the precomposition of a cone-increasing function to a vector function. These regularity conditions are formulated by using epigraphs and ε-subdifferentials. Taking $\varepsilon = 0$ one can rediscover recent results on stable strong and total duality and zero duality gap from the literature. Moreover, as byproducts we deliver ε-optimality conditions and (ε, η)-saddle point statements for the mentioned type of problems, and ε-Farkas statements involving the sum of a function with the precomposition of a cone-increasing function to a vector function.

2000 Mathematics Subject Classification: 49N15, 90C25, 90C34

Key words and phrases: Conjugate functions, $\varepsilon-$duality gap, constraint qualifications, Fenchel-Lagrange dual

1 Introduction and preliminaries

Different important combinations of functions like the precomposition of a vector function with a cone-increasing function or the sum of a function

*Faculty of Mathematics and Computer Science, Babeş-Bolyai University, 1 M. Kogălniceanu Street, 400084 Cluj-Napoca, Romania, e-mail: horatiuboncea@yahoo.com. Research made during the stay of the author in the Summer Semester 2012 at Chemnitz University of Technology as a guest of the Chair of Applied Mathematics (Approximation Theory). The author wishes to thank for the financial support provided from programs co-financed by The Sectoral Operational Program for Human Resources Development, Contract POSDRU /88/1.5/S/60185 - “Innovative doctoral studies in a knowledge based society”

†Department of Mathematics, Chemnitz University of Technology, D-09107 Chemnitz, Germany, e-mail: grad@mathematik.tu-chemnitz.de. Research partially supported by DFG (German Research Foundation), project WA 922/1-3.

with another one postcomposed with a linear continuous operator can be seen as special cases of the sum of a function and the precomposition with a cone-increasing function of a vector function. Moreover, the optimization problem consisting in minimizing a function subject to both geometric and cone-inequality constraints can be recovered as a special case of the problem of minimizing the sum of a function and the precomposition with a cone-increasing function of a vector function, too. Consequently, there developed a growing interest in investigating the latter problem and its objective function by means of conjugate optimization, materialized in works like [2, 4, 5, 7, 11, 13–16] and the references therein. Nevertheless, the mentioned problem was approached in [12] from the point of view of approximation theory, ε-optimality conditions being delivered for it.

In our paper we extend some of the results in [3, 4] by delivering general characterizations of ε-duality gap statements for optimization problems consisting of minimizing the sum of a function with the precomposition of a cone-increasing function to a vector function by means of epigraphs and ε-subdifferentials. Taking $\varepsilon = 0$ and adding appropriate convexity and topological hypotheses to the functions involved, one can rediscover different results from [2–4], which, further particularized, can lead to recent statements on stable strong and total duality and zero duality gap from papers like [8–10] and the references therein. Moreover, as byproducts we deliver ε-optimality conditions and (ε, η)-saddle point statements for the mentioned type of problems, and ε-Farkas statements involving the sum of a function with the precomposition of a cone-increasing function to a vector function, which can easily lead to rediscovering statements from [5,11–13]. In this way one can say that the results we provide in this paper may be seen as a a kind of umbrella for different recent results from the literature on optimization problems involving compositions of functions.

Now let us present the necessary preliminaries (following [2,4,5,7,8,20]) in order to make the paper as self-contained as possible.

Consider two separated locally convex vector spaces X and Y and their continuous dual spaces X^* and Y^*, endowed with the weak* topologies $w(X^*, X)$ and $w(Y^*, Y)$ respectively. Some of the following notions and results as well as most of the statements we deliver in this paper can be given in the more general framework of linear spaces, but in order to avoid juggling with spaces we decided to consider only locally convex spaces. Let the nonempty convex cone $C \subseteq Y$, i.e. $\lambda C \subseteq C$ for all $\lambda \geq 0$, and its dual cone $C^* = \{y^* \in Y^* : \langle y^*, y\rangle \geq 0 \ \forall y \in C\}$ be given, where we denote by $\langle y^*, y\rangle = y^*(y)$ the value at y of the continuous linear functional y^*. On Y we consider the partial ordering "$\leqq_C$" induced by C, defined by $z \leqq_C y \Leftrightarrow y - z \in C$, $z, y \in Y$. To Y we attach a greatest element with respect to "$\leqq_C$" denoted by ∞_C which does not belong to Y and let $Y^\bullet = Y \cup \{\infty_C\}$. Then for any $y \in Y^\bullet$ one has $y \leqq_C \infty_C$ and we consider on $Y^\bullet$ the following operations: $y + \infty_C = \infty_C + y = \infty_C$ and $t\infty_C = \infty_C$ for all $y \in Y$ and all $t \geq 0$. Moreover, for $y^* \in C^*$ we set $\langle y^*, \infty_C \rangle = +\infty$. By $\operatorname{cl} U$ we denote the *closure* of the set $U \subseteq X$ in the corresponding topology. A set $U \subseteq X$ is said to be *closed regarding the set* $Z \subseteq X$ if $U \cap Z = (\operatorname{cl} U) \cap Z$. We extend this definition as follows, noting that the notion of a ε-*closed set* was considered in the literature in different instances that have nothing in common with our research, see for instance [1, 17].

Definition 1.1 *Let $\varepsilon \geq 0$. A set $U \subseteq X \times Y$ is said to be $(0, \varepsilon)$-closed regarding the set $Z \subseteq X \times Y$ if $(\operatorname{cl} U) \cap Z \subseteq (U \cap Z) - (0, \varepsilon)$.*

For a function $f : X \to \overline{\mathbb{R}}$ we have its *domain* and *epigraph* defined by $\operatorname{dom} f = \{x \in X : f(x) < +\infty\}$ and $\operatorname{epi} f = \{(x, r) \in X \times \mathbb{R} : f(x) \leq r\}$, respectively. We say that f is *proper* if $f(x) > -\infty$ for all $x \in X$ and $\operatorname{dom} f \neq \emptyset$. The lower semicontinuous envelope of f is the function $\operatorname{cl} f : X \to \overline{\mathbb{R}}$, whose epigraph is $\operatorname{cl} \operatorname{epi} f$. The (classical/convex/Fenchel-Moreau) *conjugate function of* f is the function $f^* : X^* \to \overline{\mathbb{R}}$, given by $f^*(x^*) = \sup\{\langle x^*, x \rangle - f(x) : x \in X\}$. The function $f : X \to \overline{\mathbb{R}}$ and its conjugate fulfill the *Fenchel-Young inequality*, namely $f^*(x^*) + f(x) \geq \langle x^*, x \rangle$ for all $x \in X$ and all $x^* \in X^*$. Let $f : X \to \overline{\mathbb{R}}$, $x \in X$ with $f(x) \in \mathbb{R}$ and $\varepsilon \geq 0$. The set $\partial_\varepsilon f(x) = \{x^* \in X^* : f(u) - f(x) + \varepsilon \geq \langle x^*, u - x \rangle \ \forall u \in X\}$ is called the ε-*subdifferential* of f at x. When $f(x) \notin \mathbb{R}$ we take by convention $\partial_\varepsilon f(x) = \emptyset$. The set $R(\partial_\varepsilon f) = \{x^* \in X^* : \exists x \in X \text{ s.t. } x^* \in \partial_\varepsilon f(x)\}$ is called the *range* of the ε-subdifferential of the function f. Given a proper function $f : X \to \overline{\mathbb{R}}$, for all $\varepsilon \geq 0$, $x \in X$ and $x^* \in X^*$ one has $x^* \in \partial_\varepsilon f(x) \Leftrightarrow f^*(x^*) + f(x) \leq \langle x^*, x \rangle + \varepsilon$. For $\varepsilon = 0$, the ε-subdifferential of f becomes the classical (convex) subdifferential, denoted ∂f. A function $f : X \to \overline{\mathbb{R}}$ is said to be *lower semicontinuous regarding the set* $Z \subseteq X$ if $\operatorname{epi} f \cap (Z \times \mathbb{R}) = (\operatorname{cl} \operatorname{epi} f) \cap (Z \times \mathbb{R})$, i.e. $\operatorname{epi} f$ is closed regarding $Z \times \mathbb{R}$. Given two proper functions f, $g : X \to \overline{\mathbb{R}}$, their *infimal convolution* is $f \Box g : X \to \overline{\mathbb{R}}$, $f \Box g(a) = \inf\{f(x) + g(a - x) : x \in X\}$, and it is called *exact* at some $a \in X$ when there is an $x \in X$ such that $f \Box g(a) = f(x) + g(a - x)$.

There are notions given for functions with extended real values that can be formulated also for vector functions as follows. We say that $h : X \to Y^\bullet$ is *proper* if its *domain* $\operatorname{dom} h = \{x \in X : h(x) \in Y\}$ is nonempty and, respectively, C-*convex* if $h(tx + (1-t)y) \leqq_C th(x) + (1-t)h(y)$ for all $x, y \in X$ and all $t \in [0, 1]$. For $\lambda \in C^*$, we define $(\lambda h) : X \to \overline{\mathbb{R}}$, $(\lambda h)(x) = \langle \lambda, h(x) \rangle$

for all $x \in X$. When C is closed, we say that h is C*-epi-closed* if its *epigraph* $\operatorname{epi}_C h = \{(x,y) \in X \times Y : y \in h(x) + C\}$ is closed. A function $g : Y \to \overline{\mathbb{R}}$ is called C*-increasing* if for $x, y \in Y$ such that $y \leqq_C x$, it follows $g(y) \leq g(x)$.

Let $f : X \to \overline{\mathbb{R}}$ be a proper function, $g : Y \to \overline{\mathbb{R}}$ be a proper function, which is also C-increasing and $h : X \to Y^\bullet$ a proper vector function fulfilling $\operatorname{dom} g \cap (h(\operatorname{dom} f) + C) \neq \emptyset$. Unless otherwise stated, these hypotheses remain valid through the entire paper. Consider the optimization problem

$$(P^C) \qquad \inf_{x \in X} [f(x) + (g \circ h)(x)],$$

and, for $x^* \in X^*$, the linearly perturbed optimization problem

$$(P^C_{x^*}) \qquad \inf_{x \in X} \left[f(x) + (g \circ h)(x) - \langle x^*, x \rangle \right].$$

To $(P^C_{x^*})$ one can assign several dual problems. We attach to it two Fenchel-Lagrange-type duals. If f and (λh) are taken together one gets

$$(D^C_{x^*}) \qquad \sup_{\lambda \in C^*} \{-g^*(\lambda) - (f + (\lambda h))^*(x^*)\},$$

while if f and (λh) are separated, we have the following dual

$$(\overline{D^C_{x^*}}) \qquad \sup_{\substack{\lambda \in C^* \\ \beta \in X^*}} \{-g^*(\lambda) - f^*(\beta) - (\lambda h)^*(x^* - \beta)\}.$$

We denote by $v(P^C)$ the optimal objective value of the optimization problem (P^C). Note that $v(\overline{D^C_{x^*}}) \leq v(D^C_{x^*}) \leq v(P^C_{x^*})$ for all $x^* \in X^*$. When $x^* = 0$ these duals to (P^C) are denoted simply by (D^C) and $(\overline{D^C})$, respectively. For (P^C) and its duals one always has *weak duality*, i.e. $v(P^C) \geq v(D^C)$, respectively, $v(P^C) \geq v(\overline{D^C})$. When $v(P^C) = v(D^C)$ we say that there is *zero duality gap* for (P^C) and (D^C) and if (D^C) has moreover an optimal solution, the situation is called *strong duality.* If $v(P^C) - v(D^C) \leq \varepsilon$, with $\varepsilon \geq 0$, we have an ε*-duality gap* for (P^C) and (D^C). If one of these situations holds for $(P^C_{x^*})$ and $(D^C_{x^*})$ for all $x^* \in X^*$, it will be called *stable.* In the following we write min(max) instead of inf(sup) when the corresponding infimum (supremum) is attained.

Remark 1 Consider an arbitrary $x^* \in X^*$. By using the Fenchel-Young inequality one can show that for all $\lambda \in C^*$ and for all $\beta \in X^*$ the inequalities

$$(f + g \circ h)^*(x^*) \leq g^*(\lambda) + (f + (\lambda h))^*(x^*) \leq g^*(\lambda) + f^*(\beta) + (\lambda h)^*(x^* - \beta) \tag{1.1}$$

are always fulfilled. Under some additional hypotheses (see for instance [4]) the existence of some $\lambda \in C^*$ and $\beta \in X^*$ for which the inequalities in (1.1) are fulfilled as equalities is secured.

2 ε-duality gap statements using epigraphs

Let $\varepsilon \geq 0$. Consider the regularity conditions
(RC)
$$\left| \begin{array}{l} \{(x^*,0,r) : (x^*,r) \in \operatorname{epi}(f + g \circ h)^*\} \subseteq [\{0\} \times \operatorname{epi}(g^*) + \bigcup\limits_{\lambda \in C^*} \{(a,-\lambda,r) : \\ (a,r) \in \operatorname{epi}((f + (\lambda h))^*)\}] \cap (X^* \times \{0\} \times \mathbb{R}) - (0,0,\varepsilon) \end{array} \right.$$

and
$(\overline{RC})$
$$\left| \begin{array}{l} \{(x^*,0,r) : (x^*,r) \in \operatorname{epi}(f + g \circ h)^*\} \subseteq [\{0\} \times \operatorname{epi}(g^*) + \{(x^*,0,r) : \\ (x^*,r) \in \operatorname{epi}(f^*)\} + \bigcup\limits_{\lambda \in C^*} \{(a,-\lambda,r) : (a,r) \in \operatorname{epi}((\lambda h)^*)\}] \cap \\ (X^* \times \{0\} \times \mathbb{R}) - (0,0,\varepsilon) \end{array} \right.$$

They are inspired by the closedness type regularity conditions from [4], but unlike there, we do not use convexity and topological hypotheses for most of the proven statements. However, adding such hypotheses one can give proper generalizations of some results from [4,8] extended in this section by making use of the notion of a set which is $(0,\varepsilon)$-closed regarding another set. In order not to overcomplicate the paper we will not give these results, leaving most of our statements of algebraical nature.

Theorem 2.1 *The condition (RC) is fulfilled if and only if for any $x^* \in X^*$ there exists a $\overline{\lambda} \in C^*$ such that*

$$(f + g \circ h)^*(x^*) \geq g^*(\overline{\lambda}) + (f + (\overline{\lambda} h))^*(x^*) - \varepsilon. \tag{2.1}$$

Proof. "$\Rightarrow$" Let $x^* \in X^*$. If $(f+g\circ h)^*(x^*) = +\infty$, (2.1) holds. Otherwise, $(x^*, (f+g\circ h)^*(x^*)) \in \operatorname{epi}((f+g\circ h)^*)$. From (RC) we have that $(x^*, 0, (f+g\circ h)^*(x^*)) \in [\{0\} \times \operatorname{epi}(g^*) + \cup_{\lambda \in C^*}\{(a,-\lambda,r) : (a,r) \in \operatorname{epi}((f+(\lambda h))^*)\}] \cap (X^* \times \{0\} \times \mathbb{R}) - (0,0,\varepsilon)$. Therefore there exist some $\overline{\lambda} \in C^*$ and $\eta \geq 0$ such that

$$(x^*, 0, (f + g \circ h)^*(x^*)) = (0, \overline{\lambda}, g^*(\overline{\lambda})) + (x^*, -\overline{\lambda}, (f + (\overline{\lambda} h))^*(x^*))$$

$$+(0,0,\eta) - (0,0,\varepsilon).$$

Thus we get $(f + g \circ h)^*(x^*) \geq g^*(\overline{\lambda}) + (f + (\overline{\lambda} h))^*(x^*) - \varepsilon$.

"$\Leftarrow$" Let $(x^*, r) \in \operatorname{epi}(f + g \circ h)^*$. This means that $(f + g \circ h)^*(x^*) \leq r$. From (2.1) there exists $\overline{\lambda} \in C^*$ such that $g^*(\overline{\lambda}) + (f + (\overline{\lambda} h))^*(x^*) - \varepsilon \leq r$, thus $(f + (\overline{\lambda} h))^*(x^*) \leq r + \varepsilon - g^*(\overline{\lambda})$. But $(x^*, 0, r) = (0, \overline{\lambda}, g^*(\overline{\lambda}) - \varepsilon) + (x^*, -\overline{\lambda}, r - g^*(\overline{\lambda}) + \varepsilon)$, where the first term in the right-hand side belongs to $\{0\} \times \operatorname{epi}(g^*) - (0,0,\varepsilon)$ and the second to $\cup_{\lambda \in C^*}\{(a,-\lambda,r) : (a,r) \in \operatorname{epi}((f+(\lambda h))^*)\}$. Consequently $(x^*,0,r) \in [\{0\} \times \operatorname{epi}(g^*) + \cup_{\lambda \in C^*}\{(a,-\lambda,r) : (a,r) \in \operatorname{epi}((f+(\lambda h))^*)\}] \cap (X^* \times \{0\} \times \mathbb{R}) - (0,0,\varepsilon)$. Therefore, (RC) is fulfilled. $\square$

Remark 2 In the left-hand side of (2.1) one can easily recognize $-v(P^C_{x^*})$. The quantity in the right-hand side of (2.1) is not necessarily $-v(D^C_{x^*}) - \varepsilon$, as the supremum in $(D^C_{x^*})$ is not shown to be attained at $\overline{\lambda}$. Though, (2.1) implies $v(P^C_{x^*}) \leq v(D^C_{x^*}) + \varepsilon$, which actually means that for $(P^C_{x^*})$ and $(D^C_{x^*})$ there is ε-duality gap. Thus, (RC) yields that there is stable ε-duality gap for (P^C) and (D^C). Note also that the $\overline{\lambda} \in C^*$ obtained in Theorem 2.1 is an ε-optimal solution of $(D^C_{x^*})$.

For $\varepsilon = 0$ we obtain the following consequence of Theorem 2.1.

Corollary 2.2 *The regularity condition*

$$\left| \begin{array}{l} \{(x^*,0,r) : (x^*,r) \in \operatorname{epi}(f + g \circ h)^*\} = [\{0\} \times \operatorname{epi}(g^*) + \bigcup\limits_{\lambda \in C^*} \{(a,-\lambda,r) : \\ (a,r) \in \operatorname{epi}((f+(\lambda h))^*)\}] \cap (X^* \times \{0\} \times \mathbb{R}) \end{array} \right. \tag{2.2}$$

is fulfilled if and only if for any $x^* \in X^*$ it holds

$$(f + g \circ h)^*(x^*) = \min_{\lambda \in C^*} [g^*(\lambda) + (f + (\lambda h))^*(x^*)].$$

Proof. Because of (1.1), the right-hand side of (2.2) is included in the left-hand one, thus we get that (RC) turns in case $\varepsilon = 0$ into an equality. Further, using again (1.1) it follows that in case $\varepsilon = 0$ (2.1) turns into an equality, too, where the infimum regarding $\lambda \in C^*$ of the sum in the right-hand side is attained at the $\overline{\lambda} \in C^*$ obtained in Theorem 2.1. $\square$

Adding to Corollary 2.2 the necessary convexity and topological hypotheses one can rediscover [4, Theorem 3.3a)] as follows.

Corollary 2.3 *Let $f : X \to \overline{\mathbb{R}}$ be a proper convex lower semicontinuous function, $g : Y \to \overline{\mathbb{R}}$ a proper convex lower semicontinuous function which is also C-increasing and $h : X \to Y^\bullet$ a proper C-convex C-epi-closed vector function. The regularity condition*

$$\left| \begin{array}{l} \{0\} \times \operatorname{epi}(g^*) + \bigcup\limits_{\lambda \in C^*} \{(a,-\lambda,r) : (a,r) \in \operatorname{epi}((f+(\lambda h))^*)\} \\ \textit{is closed regarding the subspace } X^* \times \{0\} \times \mathbb{R} \end{array} \right.$$

is fulfilled if and only if for any $x^* \in X^*$ one has

$$(f + g \circ h)^*(x^*) = \min_{\lambda \in C^*} \{g^*(\lambda) + (f + (\lambda h))^*(x^*)\}.$$

If we take $x^* = 0$ we obtain from Theorem 2.1 the following result.

Corollary 2.4 *The regularity condition*
(RC^0)

$$\left| \begin{array}{l} \{(0,0,r) : (0,r) \in \operatorname{epi}(f + g \circ h)^*\} \subseteq [\{0\} \times \operatorname{epi}(g^*) + \bigcup\limits_{\lambda \in C^*} \{(a, -\lambda, r) : \\ (a, r) \in \operatorname{epi}((f + (\lambda h))^*)\}] \cap (\{0\} \times \{0\} \times \mathbb{R}) - (0, 0, \varepsilon) \end{array} \right.$$

is fulfilled if and only if there exists a $\overline{\lambda} \in C^*$ such that

$$\inf_{x \in X} (f(x) + g \circ h(x)) \leq -g^*(\overline{\lambda}) - (f + (\overline{\lambda} h))^*(0) + \varepsilon. \tag{2.3}$$

Remark 3 The quantity in the right-hand side of (2.3) is not necessarily $v(D^C) + \varepsilon$, as the supremum in (D^C) is not shown to be attained at $\overline{\lambda}$, while in the left-hand side of (2.3) we have $v(P^C)$. Though, (2.3) implies $v(P^C) \leq v(D^C) + \varepsilon$, which actually means that for (P^C) and (D^C) there is ε-duality gap and thus (RC^0) is a regularity condition that guarantees this result. Note also that the $\overline{\lambda} \in C^*$ obtained in Corollary 2.4 is an ε-optimal solution of (D^C).

Similar results can be obtained for $(\overline{D^C})$ by using $(\overline{RC})$ as follows.

Theorem 2.5 *$(\overline{RC})$ is fulfilled if and only if for any $x^* \in X^*$ there exist some $\overline{\lambda} \in C^*$ and $\overline{\beta} \in X^*$ such that*

$$(f + g \circ h)^*(x^*) \geq g^*(\overline{\lambda}) + f^*(\overline{\beta}) + (\overline{\lambda} h)^*(x^* - \overline{\beta}) - \varepsilon. \tag{2.4}$$

Proof. "$\Rightarrow$" Let $x^* \in X^*$. If $(f + g \circ h)^*(x^*) = +\infty$, (2.4) holds. Otherwise, $(x^*, (f + g \circ h)^*(x^*)) \in \operatorname{epi}((f + g \circ h)^*)$. From $(\overline{RC})$ we have that $(x^*, 0, (f + g \circ h)^*(x^*)) \in [\{0\} \times \operatorname{epi}(g^*) + \{(x^*, 0, r) : (x^*, r) \in \operatorname{epi}(f^*)\} + \cup_{\lambda \in C^*} \{(a, -\lambda, r) : (a, r) \in \operatorname{epi}((\lambda h)^*)\}] \cap (X^* \times \{0\} \times \mathbb{R}) - (0, 0, \varepsilon)$. Therefore there exist some $\overline{\lambda} \in C^*, \overline{\beta} \in X^*$ and $\eta \geq 0$ such that

$$(x^*, 0, (f + g \circ h)^*(x^*)) = (0, \overline{\lambda}, g^*(\overline{\lambda})) + (\overline{\beta}, 0, f^*(\overline{\beta})) +$$

$$(x^* - \overline{\beta}, -\overline{\lambda}, (\overline{\lambda} h)^*(x^* - \overline{\beta})) + (0, 0, \eta) - (0, 0, \varepsilon),$$

which implies $(f + g \circ h)^*(x^*) \geq g^*(\overline{\lambda}) + f^*(\overline{\beta}) + (\overline{\lambda} h)^*(x^* - \overline{\beta}) - \varepsilon$.

"$\Leftarrow$" Let $(x^*, r) \in \text{epi}(f + g \circ h)^*$. This means that $(f + g \circ h)^*(x^*) \leq r$. From (2.4) there exists $\overline{\lambda} \in C^*$ and $\overline{\beta} \in X^*$ such that $g^*(\overline{\lambda}) + f^*(\overline{\beta}) + (\overline{\lambda}h)^*(x^* - \overline{\beta}) - \varepsilon \leq r$, thus $(\overline{\lambda}h)^*(x^* - \overline{\beta}) \leq r - g^*(\overline{\lambda}) - f^*(\overline{\beta}) + \varepsilon$. But $(x^*, 0, r) = (0, \overline{\lambda}, g^*(\overline{\lambda}) - \varepsilon) + (\overline{\beta}, 0, f^*(\overline{\beta})) + (x^* - \overline{\beta}, -\overline{\lambda}, r - g^*(\overline{\lambda}) - f^*(\overline{\beta}) + \varepsilon)$, where the first term in the right-hand side belongs to $\{0\} \times \text{epi}(g^*) - (0, 0, \varepsilon)$, the second to $\{(a, 0, r) : (a, r) \in \text{epi}(f^*)\}$ and the third one to $\cup_{\lambda \in C^*}\{(a, -\lambda, r) : (a, r) \in \text{epi}((\lambda h)^*)\}$. Consequently $(x^*, 0, r) \in [\{0\} \times \text{epi}(g^*) + \{(x^*, 0, r) : (x^*, r) \in \text{epi}(f^*)\} + \cup_{\lambda \in C^*}\{(a, -\lambda, r) : (a, r) \in \text{epi}((\lambda h)^*)\}] \cap (X^* \times \{0\} \times \mathbb{R}) - (0, 0, \varepsilon)$. Therefore, $(\overline{RC})$ is fulfilled. $\square$

Remark 4 In the left-hand side of (2.4) one can easily recognize $-v(P^C_{x^*})$. The quantity in the right-hand side of (2.4) is not necessarily $-v(\overline{D^C_{x^*}}) - \varepsilon$, as the supremum in $(\overline{D^C_{x^*}})$ is not shown to be attained at $\overline{\lambda}$ and $\overline{\beta}$. Though, (2.4) implies $v(P^C_{x^*}) \leq v(\overline{D^C_{x^*}}) + \varepsilon$, which actually means that for $(P^C_{x^*})$ and $(\overline{D^C_{x^*}})$ there is ε-duality gap. Thus $(\overline{RC})$ guarantees stable ε-duality gap for (P^C) and $(\overline{D^C})$ and, moreover, also for (P^C) and (D^C). Note also that the pair $(\overline{\lambda}, \overline{\beta}) \in C^* \times X^*$ obtained in Theorem 2.5 is an ε-optimal solution of $(\overline{D^C_{x^*}})$.

For $\varepsilon = 0$ we obtain the following consequence of Theorem 2.5.

Corollary 2.6 *The regularity condition*

$$\left| \begin{array}{l} \{(x^*, 0, r) : (x^*, r) \in \text{epi}(f + g \circ h)^*\} = [\{0\} \times \text{epi}(g^*) + \{(x^*, 0, r) : (x^*, r) \\ \in \text{epi}(f^*)\} + \bigcup\limits_{\lambda \in C^*} \{(a, -\lambda, r) : (a, r) \in \text{epi}((\lambda h)^*)\}] \cap (X^* \times \{0\} \times \mathbb{R}) \end{array} \right.$$

is fulfilled if and only if for any $x^* \in X^*$ it holds

$$(f + g \circ h)^*(x^*) = \min_{\substack{\lambda \in C^* \\ \beta \in X^*}} [g^*(\lambda) + f^*(\beta) + (\lambda h)^*(x^* - \beta)].$$

Adding to Corollary 2.6 the necessary convexity and topological hypotheses, one can rediscover [4, Theorem 3.8a)] as follows.

Proposition 2.7 *Let $f : X \to \overline{\mathbb{R}}$ be a proper convex lower semicontinuous function, $g : Y \to \overline{\mathbb{R}}$ a proper convex lower semicontinuous function which is also C-increasing and $h : X \to Y^\bullet$ a proper C-convex C-epi-closed vector function. The regularity condition*

$$\left| \begin{array}{l} \{0\} \times \text{epi}(g^*) + \{(x^*, 0, r) : (x^*, r) \in \text{epi}(f^*)\} + \bigcup\limits_{\lambda \in C^*} \{(a, -\lambda, r) : \\ (a, r) \in \text{epi}((\lambda h)^*)\} \textit{ is closed regarding the subspace } X^* \times \{0\} \times \mathbb{R} \end{array} \right.$$

is fulfilled if and only if for any $x^* \in X^*$ one has

$$(f + g \circ h)^*(x^*) = \min_{\substack{\lambda \in C^* \\ \beta \in X^*}} \{g^*(\lambda) + f^*(\beta) + (\lambda h)^*(x^* - \beta)\}.$$

If we take $x^* = 0$ we obtain from Theorem 2.5 the following result.

Corollary 2.8 *The regularity condition*

$$(\overline{RC}^0) \quad \left| \begin{array}{l} \{(0,0,r) : (0,r) \in \mathrm{epi}(f + g \circ h)^*\} \subseteq [\{0\} \times epi(g^*) + \{(0,0,r) : \\ (0,r) \in \mathrm{epi}(f^*)\} + \bigcup\limits_{\lambda \in C^*} \{(a, -\lambda, r) : (a,r) \in \mathrm{epi}((\lambda h)^*)\}] \cap \\ (\{0\} \times \{0\} \times \mathbb{R}) - (0,0,\varepsilon) \end{array} \right.$$

is fulfilled if and only if there exist some $\overline{\lambda} \in C^*$ and $\overline{\beta} \in X^*$ such that

$$\inf_{x \in X} (f(x) + g \circ h(x)) \leq -g^*(\overline{\lambda}) - f^*(\overline{\beta}) - (\overline{\lambda} h)^*(-\overline{\beta}) + \varepsilon. \tag{2.5}$$

Remark 5 The quantity in the right-hand side of (2.5) is not necessarily $v(\overline{D^C}) + \varepsilon$, as the supremum in $(\overline{D^C})$ is not shown to be attained at $(\overline{\lambda}, \overline{\beta})$, while in the left-hand side of (2.5) we have $v(P^C)$. Though, (2.5) implies $v(P^C) \leq v(\overline{D^C}) + \varepsilon$, which actually means that for (P^C) and $(\overline{D^C})$ there is ε-duality gap and thus $(\overline{RC}^0)$ is a regularity condition that garantees this result. Note also that the pair $(\overline{\lambda}, \overline{\beta}) \in C^* \times X^*$ obtained in Corollary 2.8 is an ε-optimal solution of $(\overline{D^C})$. Moreover, $\overline{\lambda}$ is an ε-optimal solution of (D^C) and $(\overline{RC}^0)$ guarantees ε-duality gap for (P^C) and (D^C), too.

Remark 6 The intersection with $X^* \times \{0\} \times \mathbb{R}$ in the right-hand side of the inclusion is not necessary in (RC) or $(\overline{RC})$ for Theorem 2.1, Theorem 2.5, Corollary 2.4 or Corollary 2.8, respectively, but we need it in Corollary 2.2 and Corollary 2.6 to ensure that the elements in the right-hand side of the corresponding inclusions have that form.

In order to characterize formulae similar to (2.1) and (2.4), where appear actually the optimal values of (D^C) and $(\overline{D^C})$, let us consider the following regularity conditions

$$(RCI) \qquad \mathrm{epi}(f + g \circ h)^* \subseteq \mathrm{epi} \inf_{\lambda \in C^*} [g^*(\lambda) + (f + (\lambda h))^*(\cdot)] - (0, \varepsilon)$$

and

$$(\overline{RCI}) \quad \mathrm{epi}(f + g \circ h)^* \subseteq \mathrm{epi} \inf_{\substack{\lambda \in C^* \\ \beta \in X^*}} [g^*(\lambda) + f^*(\beta) + (\lambda h)^*(\cdot - \beta)] - (0, \varepsilon).$$

Theorem 2.9 *(RCI) is fulfilled if and only if for any $x^* \in X^*$ we have*

$$(f + g \circ h)^*(x^*) \geq \inf_{\lambda \in C^*}[g^*(\lambda) + (f + (\lambda h))^*(x^*)] - \varepsilon. \tag{2.6}$$

Proof. "$\Rightarrow$" Let $x^* \in X^*$. If $(f+g\circ h)^*(x^*) = +\infty$, (2.6) holds. Otherwise, it is clear that $(x^*, (f+g\circ h)^*(x^*)) \in \operatorname{epi}((f+g\circ h)^*)$. From ($RCI$) we get $(f+g\circ h)^*(x^*) \geq \inf_{\lambda\in C^*}[g^*(\lambda) + (f + (\lambda h))^*(x^*)] - \varepsilon$.

"$\Leftarrow$" Let $(x^*, r) \in \operatorname{epi}(f + g\circ h)^*$. This means that $(f+g\circ h)^*(x^*) \leq r$. From (2.6) we have that $\inf_{\lambda\in C^*}[g^*(\lambda) + (f + (\lambda h))^*(x^*)] \leq r + \varepsilon$. This means that $(x^*, r) \in \operatorname{epi}\inf_{\lambda\in C^*}[g^*(\lambda) + (f + (\lambda h))^*(x^*)] - (0, \varepsilon)$. $\square$

Remark 7 Relation (2.6) means actually $v(P^C_{x^*}) \leq v(D^C_{x^*}) + \varepsilon$, i.e. we have stable ε-duality gap for (P^C) and (D^C).

Taking $\varepsilon = 0$ in Theorem 2.9, we get the following obvious statement.

Corollary 2.10 *The regularity condition*

$$\operatorname{epi}(f + g\circ h)^* = \operatorname{epi}\inf_{\lambda\in C^*}[g^*(\lambda) + (f + (\lambda h))^*(\cdot)]$$

is fulfilled if and only if for any $x^* \in X^*$ we have

$$(f + g\circ h)^*(x^*) = \inf_{\lambda\in C^*}[g^*(\lambda) + (f + (\lambda h))^*(x^*)].$$

Nevertheless, if we take $x^* = 0$ in Theorem 2.9 we obtain the following ε-duality gap statement for (P^C) and (D^C).

Corollary 2.11 *The condition*
(RCI^0)
$$(\operatorname{epi}(f+g\circ h)^*)\cap(\{0\}\times\mathbb{R}) \subseteq (\operatorname{epi}\inf_{\lambda\in C^*}[g^*(\lambda)+(f+(\lambda h))^*(\cdot)])\cap(\{0\}\times\mathbb{R}) - (0,\varepsilon)$$

is fulfilled if and only if we have

$$\inf_{x\in X}(f(x) + g\circ h(x)) \leq \sup_{\lambda\in C^*}[-g^*(\lambda) - (f + (\lambda h))^*(0)] + \varepsilon.$$

Theorem 2.12 *($\overline{RCI}$) is fulfilled if and only if for any $x^* \in X^*$ we have*

$$(f + g\circ h)^*(x^*) \geq \inf_{\substack{\lambda\in C^*\\ \beta\in X^*}}[g^*(\lambda) + f^*(\beta) + (\lambda h)^*(x^* - \beta)] - \varepsilon. \tag{2.7}$$

If we take $x^* = 0$ we obtain the following ε-duality gap statement for (P^C) and $(\overline{D^C})$.

Corollary 2.13 *The condition*

$$(\overline{RCI}^0) \quad (epi(f + g \circ h)^*) \cap (\{0\} \times \mathbb{R}) \subseteq \big(\text{epi} \inf_{\substack{\lambda \in C^* \\ \beta \in X^*}} [g^*(\lambda) + f^*(\beta) \\ +(\lambda h)^*(\cdot - \beta)]\big) \cap (\{0\} \times \mathbb{R}) - (0, \varepsilon)$$

is fulfilled if and only if we have

$$\inf_{x \in X} (f(x) + g \circ h(x)) \leq \sup_{\substack{\lambda \in C^* \\ \beta \in X^*}} [-g^*(\lambda) - f^*(\beta) - (\lambda h)^*(-\beta)] + \varepsilon.$$

Remark 8 Taking into consideration Theorem 2.1, Theorem 2.5, Theorem 2.9 and Theorem 2.12 we get the following implications: $(\overline{RC}) \Rightarrow (RC) \Rightarrow (RCI)$ and $(\overline{RC}) \Rightarrow (\overline{RCI}) \Rightarrow (RCI)$. Using, for instance, [4, Example 3.10], one can construct examples that show that the opposite implications are not valid in general.

Remark 9 Note also that (RCI) together with the condition

$$\left\{(x^*, 0, r) : \inf_{\lambda \in C^*} [g^*(\lambda) + (f + (\lambda h))^*(x^*)] \leq r \right\} =$$

$$[\{0\} \times \text{epi}(g^*) + \bigcup_{\lambda \in C^*} \{(a, -\lambda, r) : (a, r) \in \text{epi}((f + (\lambda h))^*)\}] \cap (X^* \times \{0\} \times \mathbb{R})$$

implies (RC) and, analogously, $(\overline{RCI})$ together with the condition

$$\left\{(x^*, 0, r) : \inf_{\substack{\lambda \in C^* \\ \beta \in X^*}} [g^*(\lambda) + f^*(\beta) + (\lambda h)^*(x^* - \beta)] \leq r \right\} = (X^* \times \{0\} \times \mathbb{R}) \cap [\{0\} \times$$

$$\text{epi}(g^*) + \{(x^*, 0, r) : (x^*, r) \in \text{epi}(f^*)\} + \bigcup_{\lambda \in C^*} \{(a, -\lambda, r) : (a, r) \in \text{epi}((\lambda h)^*)\}]$$

yields $(\overline{RC})$.

Now let us give some statements regarding $(\varepsilon + \eta)$-duality gap for (P^C) and its duals, where $\eta > 0$.

Theorem 2.14 *If (RCI^0) is fulfilled then for each $\eta > 0$ there exists $\overline{\lambda}_\eta \in C^*$ such that*

$$v(P^C) \leq -g^*(\overline{\lambda}_\eta) - (f + (\overline{\lambda}_\eta h))^*(0) + \varepsilon + \eta.$$

Proof. From (RCI^0) we get $v(P^C) = -(f+g\circ h)^*(0) \leq \sup_{\lambda\in C^*}[-g^*(\lambda) - (f+(\lambda h))^*(0)] + \varepsilon$. For each $\eta > 0$ there exists $\overline{\lambda}_\eta \in C^*$ such that $\sup_{\lambda\in C^*}[-g^*(\lambda) - (f+(\lambda h))^*(0)] \leq -g^*(\overline{\lambda}_\eta) - (f+(\overline{\lambda}_\eta h))^*(0) + \eta$. So, $v(P^C) \leq -g^*(\overline{\lambda}_\eta) - (f+(\overline{\lambda}_\eta h))^*(0) + \varepsilon + \eta$. □

Analogously, one can show the following result.

Theorem 2.15 *If $(\overline{RCI}^0)$ is fulfilled then for each $\eta > 0$ there exist $\overline{\lambda}_\eta \in C^*$ and $\overline{\beta}_\eta \in X^*$ such that*

$$v(P^C) \leq -g^*(\overline{\lambda}_\eta) - f^*(\overline{\beta}_\eta) - (\overline{\lambda}_\eta h)^*(-\overline{\beta}_\eta) + \varepsilon + \eta.$$

Remark 10 Note that when η goes towards 0 it does not follow that the sequence $(\overline{\lambda}_\eta)_{\eta>0}$ obtained in Theorem 2.14 converges towards an ε-optimal solution to (D^C) since (RCI^0) is equivalent only to ε-duality gap for (P^C) and (D^C), with no guarantee that the dual has an ε-optimal solution. Analogously, since $(\overline{RCI}^0)$ is equivalent to ε-duality gap for (P^C) and $(\overline{D^C})$, in general the sequence $(\overline{\lambda}_\eta, \overline{\beta}_\eta)_{\eta\geq 0}$ obtained in Theorem 2.15 does not converge towards an ε-optimal solution of $(\overline{D^C})$, when η converges towards 0.

For the following result we consider some additional topological and convexity hypotheses on the involved functions.

Theorem 2.16 *Let the functions f and g be moreover convex lower semicontinuous and h also proper C-convex C-epi-closed. If*

$$(RCD) \qquad \inf_{\lambda\in C^*}[g^*(\lambda) + (f+(\lambda h))^*(\cdot)] \text{ is lower semicontinuous,}$$

then for each $\eta > 0$ there exists $\overline{\lambda}_\eta \in C^*$ such that

$$v(P^C_{x^*}) \leq -g^*(\overline{\lambda}_\eta) - (f+(\overline{\lambda}_\eta h))^*(x^*) + \eta \text{ for all } x^* \in X^*.$$

Proof. As (cf. [6, Theorem 3.1]) we have that $(f+g\circ h)^* = \operatorname{cl}\inf_{\lambda\in C^*}[g^*(\lambda) + (f+(\lambda h))^*(\cdot)]$, the hypotheses of the theorem imply $(f+g\circ h)^*(x^*) = \inf_{\lambda\in C^*}[g^*(\lambda) + (f+(\lambda h))^*(x^*)]$ for all $x^* \in X^*$. Thus for each $\eta > 0$ there exists $\overline{\lambda}_\eta \in C^*$ such that $g^*(\overline{\lambda}_\eta) + (f+(\overline{\lambda}_\eta h))^*(x^*) \leq (f+g\circ h)^*(x^*) + \eta$. So, we get $v(P^C_{x^*}) = -(f+g\circ h)^*(x^*) \leq -g^*(\overline{\lambda}_\eta) - (f+(\overline{\lambda}_\eta h))^*(x^*) + \eta$. □

Remark 11 Like in the proof of [4, Theorem 4.3], it can be shown that $(f+g\circ h)^*(x^*) = \inf_{\lambda\in C^*}[g^*(\lambda) + (f+(\lambda h))^*(x^*)]$ is also equivalent to $\inf_{\lambda\in C^*+q}[g^*(\lambda) + (f+((\lambda-q)h))^*(\cdot)]$ is lower semicontinuous regarding $\{0\}\times X^*$. This condition can be used instead of (RCD) in Theorem 2.16.

Analogously one can show the following statement.

Theorem 2.17 *Let the functions f and g be moreover convex lower semicontinuous and h also C-convex C-epi-closed. If*

$$(\overline{RCD}) \quad \left| \begin{array}{l} \inf\{g^*(\lambda) + f^*(\beta) + (\lambda h)^*(\cdot - \beta) : \lambda \in C^*, \beta \in X^*\} \textit{ is} \\ \textit{lower semicontinuous,} \end{array} \right.$$

then for each $\eta > 0$ there exist $\overline{\lambda}_\eta \in C^*$ and $\overline{\beta}_\eta \in X^*$ such that

$$v(P^C_{x^*}) \leq -g^*(\overline{\lambda}_\eta) - f^*(\overline{\beta}_\eta) - (\overline{\lambda}_\eta h)^*(x^* - \overline{\beta}_\eta) + \eta \textit{ for all } x^* \in X^*.$$

Remark 12 Note that when η goes towards 0 it does not follow that the sequence $(\overline{\lambda}_\eta)_{\eta>0}$ obtained in Theorem 2.16 converges towards an optimal solution to (D^C) since (RCD) is equivalent only to the zero duality gap for $(P^C_{x^*})$ and $(D^C_{x^*})$ for all $x^* \in X^*$, with no guarantee that the duals have optimal solutions. Analogously, since $(\overline{RCD})$ is equivalent to the zero duality gap for $(P^C_{x^*})$ and $(\overline{D^C_{x^*}})$ for all $x^* \in X^*$, in general the sequence $(\overline{\lambda}_\eta, \overline{\beta}_\eta)_{\eta\geq 0}$ obtained in Theorem 2.17 does not converge towards an optimal solution of the corresponding $(\overline{D^C_{x^*}})$, when η converges towards 0.

3 ε-duality gap statements using subdifferentials

In this section we show that the relations (2.1)–(2.7) can be characterized by regularity conditions involving ε-subdifferentials, too.

Theorem 3.1 *One has*

$$(RCSC) \qquad \partial(f + g \circ h)(x) \subseteq \bigcap_{\eta>0} \bigcup_{\substack{\varepsilon_{1,2}\geq 0 \\ \varepsilon_1+\varepsilon_2=\varepsilon+\eta \\ \lambda\in C^*\cap\partial_{\varepsilon_2} g(h(x))}} \partial_{\varepsilon_1}(f + (\lambda h))(x)$$

for all $x \in X$ if and only if (2.6) *holds for all $x^* \in R(\partial(f + g \circ h))$.*

Proof. "$\Rightarrow$" Let $x^* \in R(\partial(f + g \circ h))$. Then exists $x \in X$ such that $x^* \in \partial(f + g \circ h)(x)$. This means that $(f + g \circ h)(x) + (f + g \circ h)^*(x^*) = \langle x^*, x\rangle$. Because the condition $(RCSC)$ is satisfied, for each $\eta > 0$ there are some $\overline{\lambda}_\eta \in C^* \cap \partial_{\varepsilon_2} g(h(x))$ and $\varepsilon_1, \varepsilon_2 \geq 0$ with $\varepsilon_1 + \varepsilon_2 = \varepsilon + \eta$, such that $(f + (\overline{\lambda}_\eta h))^*(x^*) + (f + (\overline{\lambda}_\eta h))(x) \leq \langle x^*, x\rangle + \varepsilon_1$ and $g(h(x)) + g^*(\overline{\lambda}_\eta) \leq (\overline{\lambda}_\eta h)(x) + \varepsilon_2$. It follows that $g^*(\overline{\lambda}_\eta) + (f + (\overline{\lambda}_\eta h))^*(x^*) \leq \langle x^*, x\rangle - (f + g \circ$

$h)(x) + \varepsilon + \eta = (f + g \circ h)^*(x^*) + \varepsilon + \eta$. Consequently, $(f + g \circ h)^*(x^*) \geq \inf_{\lambda \in C^*}[g^*(\lambda) + (f + (\lambda h))^*(x^*)] - \varepsilon - \eta$.

Letting η converge towards 0, (2.6) follows.

"$\Leftarrow$" Taking $x^* \in \partial(f + g \circ h)(x)$ we have $(f + g \circ h)(x) + (f + g \circ h)^*(x^*) = \langle x^*, x \rangle$. For each $\eta > 0$ there is some $\overline{\lambda}_\eta \in C^*$ such that $\inf_{\lambda \in C^*}[g^*(\lambda) + (f + (\lambda h))^*(x^*)] \geq g^*(\overline{\lambda}_\eta) + (f + (\overline{\lambda}_\eta h))^*(x^*) - \eta$. Using (2.6) we get $g^*(\overline{\lambda}_\eta) + (f + (\overline{\lambda}_\eta h))^*(x^*) - \eta - \varepsilon \leq \langle x^*, x \rangle - (f + g \circ h)(x)$. This is equivalent to $g^*(\overline{\lambda}_\eta) + g(h(x)) + (f + (\overline{\lambda}_\eta h))^*(x^*) + (f + (\overline{\lambda}_\eta h))(x) \leq \langle x^*, x \rangle + (\overline{\lambda}_\eta h)(x) + \eta + \varepsilon$. Using the Young-Fenchel inequality, it follows that there exist $\varepsilon_1, \varepsilon_2 \geq 0$ with $\varepsilon_1 + \varepsilon_2 = \varepsilon + \eta$ such that $g^*(\overline{\lambda}_\eta) + g(h(x)) \leq (\overline{\lambda}_\eta h)(x) + \varepsilon_2$ and $(f + (\overline{\lambda}_\eta h))^*(x^*) + (f + (\overline{\lambda}_\eta h))(x) \leq \langle x^*, x \rangle + \varepsilon_1$. So, we get that $\overline{\lambda}_\eta \in \partial_{\varepsilon_2} g(h(x))$ and $x^* \in \partial_{\varepsilon_1}(f + (\lambda h))(x)$, which means that $(RCSC)$ holds. □

The assertion of Theorem 3.1 can be refined as follows.

Corollary 3.2 *Let $x \in X$. Then $(RCSC)$ holds for x if and only if for all $x^* \in \partial(f + g \circ h)(x)$ one has*

$$-(f+g\circ h)(x)+\langle x^*, x\rangle = (f+g\circ h)^*(x^*) \geq \inf_{\lambda \in C^*}[g^*(\lambda)+(f+(\lambda h))^*(x^*)]-\varepsilon.$$

Remark 13 The inequality in Corollary 3.2 is nothing but $v(P^C_{x^*}) \leq v(D^C_{x^*}) + \varepsilon$, i.e. ε-duality gap for the pair of problems $(P^C_{x^*})$ and $(D^C_{x^*})$, when x is an optimal solution of $(P^C_{x^*})$.

Remark 14 One can analogously show that relation (2.6) implies

$$\partial_\nu(f + g \circ h)(x) \subseteq \bigcap_{\eta > 0} \bigcup_{\substack{\varepsilon_{1,2} \geq 0 \\ \varepsilon_1 + \varepsilon_2 = \varepsilon + \eta + \nu \\ \lambda \in C^* \cap \partial_{\varepsilon_2} g(h(x))}} \partial_{\varepsilon_1}(f + (\lambda h))(x), \tag{3.1}$$

for all $x \in X$ and $\nu > 0$. Viceversa, for $\nu > 0$ (3.1) implies

$$(f + g \circ h)^*(x^*) \geq \inf_{\lambda \in C^*}[g^*(\lambda) + (f + (\lambda h))^*(x^*)] - \varepsilon - \nu,$$

for all $x^* \in R(\partial_\nu(f + g \circ h))$.

Remark 15 If $\varepsilon = 0$, relation (3.1) becomes

$$\partial_\nu(f + g \circ h)(x) \subseteq \bigcap_{\eta > 0} \bigcup_{\substack{\varepsilon_{1,2} \geq 0 \\ \varepsilon_1 + \varepsilon_2 = \eta + \nu \\ \lambda \in C^* \cap \partial_{\varepsilon_2} g(h(x))}} \partial_{\varepsilon_1}(f + (\lambda h))(x) \tag{3.2}$$

and Remark 14 yields for $\nu > 0$ that, for $x \in X$, (3.2) implies

$$(f + g \circ h)^*(x^*) \geq \inf_{\lambda \in C^*}[g^*(\lambda) + (f + (\lambda h))^*(x^*)] - \nu \ \forall x^* \in \partial_\nu(f + g \circ h)(x).$$

Thus, (3.2) holds for all $x \in X$ and all $\nu > 0$ if and only if $(f + g \circ h)^*(x^*) = \inf_{\lambda \in C^*}[g^*(\lambda) + (f + (\lambda h))^*(x^*)]$ for all $x^* \in \cap_{\nu > 0} R(\partial_\nu(f + g \circ h)) = R(\partial f + g \circ h))$. As the set in the right-hand side of (3.2) is always a subset of the one in the left-hand side, adding the necessary topological and convexity hypotheses on the involved functions, one can rediscover [3, Proposition 3.2].

In the following result we give another characterization for relation (2.1), this time by making use of ε-subdifferentials.

Theorem 3.3 *One has*

$$(RCLC) \qquad \partial(f + g \circ h)(x) \subseteq \bigcup_{\substack{\varepsilon_{1,2} \geq 0 \\ \varepsilon_1 + \varepsilon_2 = \varepsilon \\ \lambda \in C^* \cap \partial_{\varepsilon_2} g(h(x))}} \partial_{\varepsilon_1}(f + (\lambda h))(x)$$

for all $x \in X$ if and only if for each $x^* \in R(\partial(f + g \circ h))$ there exists $\overline{\lambda} \in C^*$ such that (2.1) *holds.*

Proof. "$\Rightarrow$" Let $x^* \in \partial(f + g \circ h)(x)$. This means that $(f + g \circ h)(x) + (f + g \circ h)^*(x^*) = \langle x^*, x \rangle$. Because the condition $(RCLC)$ is satisfied there are some $\overline{\lambda} \in C^* \cap \partial_{\varepsilon_2} g(h(x))$ and $\varepsilon_1, \varepsilon_2 \geq 0$ with $\varepsilon_1 + \varepsilon_2 = \varepsilon$, such that $(f + (\overline{\lambda} h))^*(x^*) + (f + (\overline{\lambda} h))(x) \leq \langle x^*, x \rangle + \varepsilon_1$ and $g(h(x)) + g^*(\overline{\lambda}) \leq (\overline{\lambda} h)(x) + \varepsilon_2$. It follows that $g^*(\overline{\lambda}) + (f + (\overline{\lambda} h))^*(x^*) \leq \langle x^*, x \rangle - (f + g \circ h)(x) + \varepsilon = (f + g \circ h)^*(x^*) + \varepsilon$. So, (2.1) holds.

"$\Leftarrow$" Taking $x^* \in \partial(f + g \circ h)(x)$ we have $(f + g \circ h)(x) + (f + g \circ h)^*(x^*) = \langle x^*, x \rangle$. Using (2.1) we get that there exists $\overline{\lambda} \in C^*$ such that $g^*(\overline{\lambda}) + (f + (\overline{\lambda} h))^*(x^*) - \varepsilon \leq \langle x^*, x \rangle - (f + g \circ h)(x)$. This is equivalent to $g^*(\overline{\lambda}) + g(h(x)) + (f + (\overline{\lambda} h))^*(x^*) + (f + (\overline{\lambda} h))(x) \leq \langle x^*, x \rangle + (\overline{\lambda} h)(x) + \varepsilon$. Using the Young-Fenchel inequality, it follows that there exist $\varepsilon_1, \varepsilon_2 \geq 0$ with $\varepsilon_1 + \varepsilon_2 = \varepsilon$ such that $g^*(\overline{\lambda}) + g(h(x)) \leq (\overline{\lambda} h)(x) + \varepsilon_2$ and $(f + (\overline{\lambda} h))^*(x^*) + (f + (\overline{\lambda} h))(x) \leq \langle x^*, x \rangle + \varepsilon_1$. So, we get that $\overline{\lambda} \in \partial_{\varepsilon_2} g(h(x))$ and $x^* \in \partial_{\varepsilon_1}(f + (\lambda h))(x)$, which means that $(RCLC)$ holds. $\Box$

Remark 16 An observation similar to Remark 14 can be made in the sense that relation (2.1) implies a formula like (3.1) but without η. Investigating further, a similar analysis to the one in Remark 15 can be given, too, with [4, Theorem 3.3 b)] as the rediscovered result.

The following result characterizes relation (2.7).

Theorem 3.4 *One has*

$$(\overline{RCSC}) \qquad \partial(f+g\circ h)(x) \subseteq \bigcap_{\eta>0} \bigcup_{\substack{\varepsilon_{1,2}\geq 0 \\ \varepsilon_1+\varepsilon_2+\varepsilon_3=\varepsilon+\eta \\ \lambda\in C^*\cap\partial_{\varepsilon_3} g(h(x))}} \partial_{\varepsilon_1} f(x) + \partial_{\varepsilon_2}(\lambda h)(x)$$

for all $x \in X$ if and only if for all $x^* \in R(\partial(f+g\circ h))$, (2.7) holds.

The assertion of Theorem 3.4 can be refined as follows.

Corollary 3.5 *Let $x \in X$. Then $(\overline{RCSC})$ holds for x if and only if for all $x^* \in \partial(f+g\circ h)(x)$ one has*

$$-(f+g\circ h)(x)+\langle x^*, x\rangle = (f+g\circ h)^*(x^*) \geq \inf_{\substack{\lambda\in C^* \\ \beta\in X^*}} [g^*(\lambda)+f^*(\beta)+(\lambda h)^*(x^*-\beta)]-\varepsilon.$$

Remark 17 The inequality in Corollary 3.5 is nothing but $v(P^C_{x^*}) \leq v(\overline{D^C_{x^*}})+\varepsilon$, i.e. ε-duality gap for the pair of problems $(P^C_{x^*})$ and $(\overline{D^C_{x^*}})$, when x is an optimal solution of $(P^C_{x^*})$. Consequently, $(\overline{RCSC})$ yields ε-duality gap for the pair of problems $(P^C_{x^*})$ and $(D^C_{x^*})$, too.

Remark 18 Relation (2.7) implies

$$\partial_\nu(f+g\circ h)(x) \subseteq \bigcap_{\eta>0} \bigcup_{\substack{\varepsilon_{1,2}\geq 0 \\ \varepsilon_1+\varepsilon_2=\varepsilon+\eta+\nu \\ \lambda\in C^*\cap\partial_{\varepsilon_2} g(h(x))}} \partial_{\varepsilon_1} f(x) + \partial_{\varepsilon_2}(\lambda h)(x), \tag{3.3}$$

for all $x \in X$ and $\nu > 0$. Viceversa, for $\nu > 0$ (3.3) implies

$$(f+g\circ h)^*(x^*) \geq \inf_{\substack{\lambda\in C^* \\ \beta\in X^*}} [g^*(\lambda)+f^*(\beta)+(\lambda h)^*(x^*-\beta)]-\varepsilon-\nu,$$

for all $x^* \in R(\partial_\nu(f+g\circ h))$.

Remark 19 If $\varepsilon = 0$, relation (3.3) becomes

$$\partial_\nu(f+g\circ h)(x) \subseteq \bigcap_{\eta>0} \bigcup_{\substack{\varepsilon_{1,2}\geq 0 \\ \varepsilon_1+\varepsilon_2=\eta+\nu \\ \lambda\in C^*\cap\partial_{\varepsilon_2} g(h(x))}} \partial_{\varepsilon_1} f(x) + \partial_{\varepsilon_2}(\lambda h)(x) \tag{3.4}$$

and Remark 18 yields for $\nu > 0$ that, for $x \in X$, (3.4) implies

$$(f+g\circ h)^*(x^*) \geq \inf_{\substack{\lambda\in C^*,\\ \beta\in X^*}} [g^*(\lambda)+f^*(\beta)+(\lambda h)^*(x^*-\beta)]-\nu \; \forall x^* \in R(\partial_\nu(f+g\circ h)).$$

Thus, (3.4) holds for all $x \in X$ and for all $\nu > 0$ if and only if $(f+g\circ h)^*(x^*) = \inf_{\lambda\in C^*,\beta\in X^*}[g^*(\lambda) + f^*(\beta) + (\lambda h)^*(x^* - \beta)]$ for all $x^* \in R(\partial(f + g \circ h))$. As the set in the right-hand side of (3.4) is always a subset of the one in the left-hand side, adding the necessary topological and convexity hypotheses on the involved functions, one can rediscover [3, Proposition 3.3].

In the following result we characterize relation (2.4).

Theorem 3.6 *One has*

$$(\overline{RCLC}) \qquad \partial(f + g \circ h)(x) \subseteq \bigcup_{\substack{\varepsilon_{1,2}\geq 0\\ \varepsilon_1+\varepsilon_2+\varepsilon_3=\varepsilon+\eta\\ \lambda\in C^*\cap\partial_{\varepsilon_3}g(h(x))}} \partial_{\varepsilon_1}f(x) + \partial_{\varepsilon_2}(\lambda h)(x)$$

for all $x \in X$ if and only if for each $x^* \in R(\partial(f + g \circ h))$ there exist $\overline{\lambda} \in C^*$ and $\overline{\beta} \in X^*$ such that (2.4) *holds.*

Remark 20 An observation similar to Remark 18 can be made in the sense that relation (2.4) implies a formula like (3.3) but without η. Investigating further, a similar analysis to the one in Remark 19 can be given, too, with [4, Theorem 3.8b)] as the rediscovered result.

Remark 21 Looking at the conditions (RC) and $(RCLC)$ one can observe that (RC) is equivalent to the validity of (2.1) for all $x^* \in X^*$, while $(RCLC)$ holds if and only if (2.1) is satisfied only for all $x^* \in R(\partial(f + g \circ h))$. This yields that (RC) implies $(RCLC)$. Analogously, $(\overline{RC})$ means the satisfaction of (2.4) for all $x^* \in X^*$, while $(\overline{RCLC})$ is equivalent to the validity of (2.4) for all $x^* \in R(\partial(f + g \circ h))$, consequently, $(\overline{RC})$ implies $(\overline{RCLC})$.

4 Byproducts: ε-optimality conditions, ε-Farkas statements and (ε, η)-saddle points

From the results presented in the previous sections one can derive other useful statements concerning ε-optimality conditions, ε-Farkas assertions and characterizations for (ε, η)-saddle points as follows. We begin with the ε-optimality conditions.

Theorem 4.1 *(a) Let $\varepsilon, \eta \geq 0$. Suppose that the condition (RCI^0) is fulfilled. If $\overline{x}$ is an ε-optimal solution of the problem (P^C), then there exist $\varepsilon_1, \varepsilon_2 \geq 0$, and $\overline{\lambda} \in C^*$ such that*

(i) $g^*(\overline{\lambda}) + g(h(\overline{x})) \leq (\overline{\lambda}h)(\overline{x}) + \varepsilon_2$,

(ii) $(f + (\overline{\lambda}h))^*(0) + (f + (\overline{\lambda}h))(\overline{x}) \leq \varepsilon_1$,

(iii) $\varepsilon_1 + \varepsilon_2 = \varepsilon + \eta$.

Moreover, $\overline{\lambda}$ is an $(\varepsilon + \eta)$-optimal solution of the problem (D^C).

(b) If there exist $\varepsilon_1, \varepsilon_2 \geq 0$ and $\overline{\lambda} \in C^*$ such that the relations $(i) - (iii)$ hold for $\overline{x} \in X$ and $\overline{\lambda} \in C^*$ then $\overline{x}$ is an $(\varepsilon + \eta)$-optimal solution of the problem (P^C). Moreover, $\overline{\lambda}$ is an $(\varepsilon + \eta)$-optimal solution of the problem (D^C).

Proof. (a) As $\overline{x}$ is an ε-optimal solution of the problem (P^C) we have that $0 \in \partial_\varepsilon(f + g \circ h)(\overline{x})$. By relation (3.1) written for $x^* = 0$, i.e.

$$0 \in \partial_\varepsilon(f + g \circ h)(\overline{x}) \Rightarrow 0 \in \bigcap_{\nu > 0} \bigcup_{\substack{\varepsilon_{1,2} \geq 0 \\ \varepsilon_1 + \varepsilon_2 = \varepsilon + \nu \\ \lambda \in C^* \cap \partial_{\varepsilon_2} g(h(\overline{x}))}} \partial_{\varepsilon_1}(f + (\lambda h))(\overline{x}),$$

there exist $\varepsilon_1, \varepsilon_2 \geq 0$ and $\overline{\lambda} \in C^*$ such that $\varepsilon_1 + \varepsilon_2 = \varepsilon + \eta$, $\overline{\lambda} \in C^* \cap \partial_{\varepsilon_2} g(h(\overline{x}))$ and $0 \in \partial_{\varepsilon_1}(f + (\overline{\lambda}h))(\overline{x})$. As $\overline{\lambda} \in \partial_{\varepsilon_2} g(h(\overline{x}))$, the assertion (i) arises directly. From $0 \in \partial_{\varepsilon_1}(f + (\overline{\lambda}h))(\overline{x})$ the assertion (ii) can be deduced from the definition of ε-subdifferential. Further, from relations $(i) - (ii)$ and taking into consideration the relation (iii), we get that

$$f(\overline{x}) + g(h(\overline{x})) \leq -g^*(\overline{\lambda}) - (f + (\overline{\lambda}h))^*(0) + \varepsilon + \eta. \tag{4.1}$$

We know that weak duality always holds, i.e. $v(D^C) \leq v(P^C)$ and, since $v(P^C) \leq f(\overline{x}) + (g \circ h)(\overline{x})$, one gets $v(D^C) \leq -g^*(\overline{\lambda}) - (f + (\overline{\lambda}h))^*(0) + \varepsilon + \eta$, which means that $\overline{\lambda}$ is an $(\varepsilon + \eta)$-optimal solution of the problem (D^C).

(b) By summing the relations (i) and (ii) and taking into consideration the relation (iii), we get that

$$g^*(\overline{\lambda}) + g(h(\overline{x})) - (\overline{\lambda}h)(\overline{x}) + (f + (\overline{\lambda}h))^*(0) + (f + (\overline{\lambda}h))(\overline{x}) \leq \varepsilon_1 + \varepsilon_2 = \varepsilon + \eta. \tag{4.2}$$

By (1.1) we get $(f + g \circ h)^*(0) + (f + g \circ h)(\overline{x}) \leq \varepsilon + \eta$, thus $0 \in \partial_{\varepsilon+\eta}(f + g \circ h)(\overline{x})$, i.e. $\overline{x}$ is an $(\varepsilon + \eta)$-optimal solution of the problem (P^C). On the other hand, (4.2) implies (4.1), so $\overline{\lambda}$ is an $(\varepsilon + \eta)$-optimal solution of (D^C). □

The similar statement for $(\overline{D^C})$ can be proven analogously.

Theorem 4.2 *(a) Let $\varepsilon, \eta \geq 0$. Suppose that the condition $(\overline{RCI}^0)$ is fulfilled. If $\overline{x}$ is an ε-optimal solution of the problem (P^C), then there exist $\varepsilon_1, \varepsilon_2, \varepsilon_3 \geq 0$, $\overline{\lambda} \in C^*$ and $\overline{\beta} \in X^*$ such that*

(i) $g^*(\overline{\lambda}) + g(h(\overline{x})) \leq (\overline{\lambda}h)(\overline{x}) + \varepsilon_3$,

(ii) $f^*(\overline{\beta}) + f(\overline{x}) \leq \langle \overline{\beta}, \overline{x} \rangle + \varepsilon_1$,

(iii) $(\overline{\lambda}h)^*(-\overline{\beta}) + (\overline{\lambda}h)(\overline{x}) \leq \langle -\overline{\beta}, \overline{x} \rangle + \varepsilon_2$,

(iv) $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 = \varepsilon + \eta$.

Moreover, $(\overline{\lambda}, \overline{\beta})$ is an $(\varepsilon + \eta)$-optimal solution of the problem $(\overline{D^C})$.

(b) If there exist $\varepsilon_1, \varepsilon_2, \varepsilon_3 \geq 0$, $\overline{\lambda} \in C^*$ and $\overline{\beta} \in X^*$ such that the relations $(i) - (iv)$ hold for $\overline{x} \in X$, $\overline{\lambda} \in C^*$ and $\overline{\beta} \in X^*$ then $\overline{x}$ is an $(\varepsilon + \eta)$-optimal solution of the problem (P^C). Moreover, $(\overline{\lambda}, \overline{\beta})$ is an $(\varepsilon + \eta)$-optimal solution of the problem $(\overline{D^C})$.

Remark 22 Similar optimality conditions to the ones in Theorem 4.2 were obtained also in [12, Theorem 4] but under convexity and topological hypotheses, with some other regularity conditions and without involving η. These can be rediscovered, too, as consequences of our results by using in Theorem 4.2 the regularity condition $(\overline{RC}^0)$. Note also that employing (RC^0) as regularity condition in Theorem 4.1 one obtains similar ε-optimality conditions, but without involving η, which can be found in the literature under additional convexity and topological hypotheses.

In the following we give ε-Farkas-type results for (P^C) and its duals, too.

Theorem 4.3 *(a) Suppose that (RC^0) holds. If $f(x)+(g\circ h)(x) \geq \varepsilon/2$ for all $x \in X$ then there exists $\overline{\lambda} \in C^*$ such that $g^*(\overline{\lambda})+(f+\overline{\lambda}h)^*(0) \leq \varepsilon/2$.*

(b) If there exists $\overline{\lambda} \in C^*$ such that $g^*(\overline{\lambda}) + (f + \overline{\lambda}h)^*(0) \leq -\varepsilon/2$, then $f(x) + (g \circ h)(x) \geq \varepsilon/2$ for all $x \in X$.

Proof. (a) From (RC^0) we have that there exists $\overline{\lambda} \in C^*$ fulfilling $\inf_{x\in X}[f(x)+(g\circ h)(x)] \leq -g^*(\overline{\lambda}) - (f+\overline{\lambda}h)^*(0)+\varepsilon$. Then $-g^*(\overline{\lambda}) - (f+\overline{\lambda}h)^*(0) \geq -\varepsilon/2$ and the conclusion follows.

(b) As we can find some $\overline{\lambda} \in C^*$ fulfilling $-g^*(\overline{\lambda}) - (f + \overline{\lambda}h)^*(0) \geq \varepsilon/2$, it follows from weak duality that $f(x) + (g \circ h)(x) \geq \varepsilon/2$. □

Analogously, one can prove the following statements for (P^C) and $(\overline{D^C})$, too.

Theorem 4.4 *(a) Suppose that* $(\overline{RC}^0)$ *holds. If* $f(x) + (g \circ h)(x) \geq \varepsilon/2$ *for all* $x \in X$ *then there exist* $\overline{\lambda} \in C^*$ *and* $\overline{\beta} \in X^*$ *such that* $f^*(\overline{\beta}) + g^*(\overline{\lambda}) + (\overline{\lambda}h)^*(-\overline{\beta}) \leq \varepsilon/2$.

(b) If there exist $\overline{\lambda} \in C^*$ *and* $\overline{\beta} \in X^*$ *such that* $f^*(\overline{\beta})+g^*(\overline{\lambda})+(\overline{\lambda}h)^*(-\overline{\beta}) \leq -\varepsilon/2$*, then* $f(x) + (g \circ h)(x) \geq \varepsilon/2$ *for all* $x \in X$.

Remark 23 Taking $\varepsilon = 0$ and adding convexity and topological assumptions, one can rediscover, via Theorem 4.3 and Theorem 4.4, the Farkas-type results for composed convex functions from [11] (see also [13]).

We can give ε-Farkas type statements for the other regularity conditions, too.

Theorem 4.5 *(a) Suppose* (RCI^0) *holds. If* $f(x) + (g \circ h)(x) \geq \varepsilon/2$ *for all* $x \in X$ *then* $\inf_{\lambda \in C^*}[g^*(\lambda) + (f + \lambda h)^*(0)] \leq \varepsilon/2$,

(b) If $\inf_{\lambda \in C^*}[g^*(\lambda) + (f + \lambda h)^*(0)] \leq -\varepsilon/2$*, then* $f(x) + (g \circ h)(x) \geq \varepsilon/2$ *for all* $x \in X$.

Theorem 4.6 *(a) Suppose* $(\overline{RCI}^0)$ *holds. If* $f(x) + (g \circ h)(x) \geq \varepsilon/2$ *for all* $x \in X$ *then* $\inf\{f^*(\overline{\beta})+g^*(\overline{\lambda})+(\overline{\lambda}h)^*(-\overline{\beta}) : \lambda \in C^*, \beta \in X^*\} \leq \varepsilon/2$,

(b) If $\inf\{f^*(\overline{\beta}) + g^*(\overline{\lambda}) + (\overline{\lambda}h)^*(-\overline{\beta}) : \lambda \in C^*, \beta \in X^*\} \leq -\varepsilon/2$*, then* $f(x) + (g \circ h)(x) \geq \varepsilon/2$ *for all* $x \in X$.

Nevertheless, one can extend the investigations from this paper also towards generalized saddle points.

The *Lagrangian* function assigned to $(P^C) - (D^C)$ is $L^C : X \times Y^* \to \overline{\mathbb{R}}$, defined by (cf. [5])

$$L^C(x, \lambda) = \begin{cases} f(x) + (\lambda h)(x) - g^*(\lambda), & \text{if } \lambda \in C^*, \\ -\infty, & \text{otherwise.} \end{cases}$$

Let $\eta \geq 0$. We say that $(\overline{x}, \overline{\lambda}) \in X \times Y^*$ is an (η, ε)-*saddle point* of the Lagrangian L^C if

$$L^C(\overline{x}, \lambda) - \eta \leq L^C(\overline{x}, \overline{\lambda}) \leq L^C(x, \overline{\lambda}) + \varepsilon \text{ for all } (x, \lambda) \in X \times Y^*.$$

Remark 24 The notion of a ε-saddle point of a function with two variables, where $\varepsilon \geq 0$ was already considered in the literature, see for instance [18,19]. However, we are not aware of any work dealing with (η, ε)-saddle points as introduced above.

Slightly weakening the properness hypothesis of g and adding to it convexity and topological assumptions, one obtains the following statement connecting the (η, ε)-saddle points of L^C with the $(\varepsilon + \eta)$-duality gap for the problems (P^C) and (D^C), and the existence of some $(\varepsilon + \eta)$-optimal solutions to them.

Theorem 4.7 *Assume that g is a convex and lower semicontinuous function fulfilling $g(y) > -\infty$ for all $y \in Y$. If $(\overline{x}, \overline{\lambda})$ is an (η, ε)-saddle point of L^C then $\overline{x} \in X$ is an $(\varepsilon + \eta)$-optimal solution to (P^C), $\overline{\lambda} \in C^*$ is an $(\varepsilon + \eta)$-optimal solution to (D^C) and there is $(\varepsilon + \eta)$-duality gap for the pair of problems (P^C) and (D^C), i.e. $v(P^C) \leq (D^C) + \varepsilon + \eta$.*

Proof. If $(\overline{x}, \overline{\lambda})$ is a (η, ε)-saddle point of L^C, we get that

$$f(\overline{x}) + (\lambda h)(\overline{x}) - g^*(\lambda) - \eta \leq f(\overline{x}) + (\overline{\lambda} h)(\overline{x}) - g^*(\overline{\lambda}) \leq$$
$$f(x) + (\overline{\lambda} h)(x) - g^*(\overline{\lambda}) + \varepsilon \text{ for all } (x, \lambda) \in X \times Y^\bullet \tag{4.3}$$

If $\overline{\lambda} \notin C^*$, the second and the third terms from (4.3) are $-\infty$, while the first one takes also real values, so $\overline{\lambda} \in C^*$. The first inequality from (4.3) yields that $\sup_{\lambda \in Y^\bullet} [(\lambda h)(\overline{x}) - g^*(\lambda)] - \eta \leq (\overline{\lambda} h)(\overline{x}) - g^*(\overline{\lambda})$, which is equivalent to $g^{**}(h(\overline{x})) - \eta \leq (\overline{\lambda} h)(\overline{x}) - g^*(\overline{\lambda})$. But $g^{**}(x) = g(x)$, so we get

$$g(h(\overline{x})) - \eta \leq (\overline{\lambda} h)(\overline{x}) - g^*(\overline{\lambda}). \tag{4.4}$$

The second inequality from (4.3) yields, via Fenchel-Young inequality, that $f(\overline{x}) + (\overline{\lambda} h)(\overline{x}) - g^*(\overline{\lambda}) \leq f(x) + g(h(x)) + \varepsilon$, for all $x \in X$, thus

$$f(\overline{x}) + (\overline{\lambda} h)(\overline{x}) - g^*(\overline{\lambda}) \leq \inf_{x \in X} [f(x) + g(h(x))] + \varepsilon = v(P^C) + \varepsilon.$$

From the latter inequality and using (4.4) we get $f(\overline{x}) + g(h(\overline{x})) - \eta \leq v(P^C) + \varepsilon$, which means that $\overline{x}$ is an $(\varepsilon + \eta)$-optimal solution to (P^C).

On the other hand, the second inequality from (4.3) can be rewritten as

$$f(\overline{x}) + (\overline{\lambda} h)(\overline{x}) - g^*(\overline{\lambda}) \leq -(f + (\overline{\lambda} h))^*(0) - g^*(\overline{\lambda}) + \varepsilon \leq v(D^C) + \varepsilon.$$

So, $v(P^C) \leq (f + g \circ h)(\overline{x}) \leq (f + (\overline{\lambda} h))(\overline{x}) - g^*(\overline{\lambda}) + \eta \leq v(D^C) + \varepsilon + \eta$, which yields that there is $(\varepsilon + \eta)$-duality gap for the pair of problems (P^C) and (D^C).

We know that weak duality always holds we get that $v(D^C) - \eta \leq v(P^C) - \eta \leq -(f + (\overline{\lambda} h))^*(0) - g^*(\overline{\lambda}) + \varepsilon$, which yields that $\overline{\lambda}$ is an $(\varepsilon + \eta)$-optimal solution to (D^C). $\square$

An analogous result with Theorem 4.7 can be formulated for the pair of problems (P^C) and $(\overline{D^C})$ with the corresponding Lagrangian function given by (cf. [5]) $\overline{L^C} : X \times X^* \times Y^* \to \overline{\mathbb{R}}$

$$\overline{L^C}(x, \beta, \lambda) = \begin{cases} \langle \beta, x \rangle + (\lambda h)(x) - f^*(\beta) - g^*(\lambda), & \text{if } \lambda \in C^* \\ -\infty, & \text{otherwise.} \end{cases}$$

Theorem 4.8 *Assume that g is a convex and lower semicontinuous function fulfilling $g(y) > -\infty$ for all $y \in Y$. If $(\overline{x}, \overline{\lambda}, \overline{\beta})$ is an (η, ε)-saddle point of $\overline{L^C}$ then $\overline{x} \in X$ is an $(\varepsilon + \eta)$-optimal solution to (P^C), $(\overline{\lambda}, \overline{\beta}) \in C^* \in X^*$ is an $(\varepsilon + \eta)$-optimal solution to $(\overline{D^C})$ and there is $(\varepsilon + \eta)$-duality gap for the pair of problems (P^C) and $(\overline{D^C})$, i.e. $v(P^C) \leq (\overline{D^C}) + \varepsilon + \eta$.*

Remark 25 One can formulate also reverse statements for Theorem 4.7 and Theorem 4.8, which together with these collapse in case $\varepsilon = \eta = 0$ to [5, Theorem 3.4.3] and [5, Theorem 3.4.7], respectively. Moreover, it may be worth trying to see if the regularity conditions we introduced in this paper guarantee the existence of some (η, ε)-saddle points of the considered Lagrangians, in the sense of [5, Corollary 3.4.4] and [5, Corollary 3.4.8], respectively.

Remark 26 The results we gave for composed functions can be particularized for combinations of functions that appear often in both theoretical and practical problems. One of the most important such particular cases is obtained when one takes $f(x) = 0$ for all $x \in X$, when different characterizations and statements involving the function $g \circ h$ and the optimization problem of minimizing it can be derived. Another important combination of functions often met in optimization problems is $f + g \circ A$, where $A : X \to Y$ is a linear continuous mapping, and it can be recovered as a special instance of $f + g \circ h$ by taking $h(x) = Ax$ for all $x \in X$. In both these special cases, due to the fact that either f and, respectively, h, are taken to be continuous functions, the duals and the conditions obtained when f and h appear separated coincide with their counterparts where they are taken together. Consequently, in each of these cases we obtain only a dual problem and a single set of results. An interesting special case of the problem of minimizing a composition of functions as investigated in this paper is the general constrained minimization problem (see, for instance, [8]). Consequently, one can use the results delivered in this paper for recovering and extending statements from papers dealing with this problem, such as [8–10].

References

[1] Y. Anbo, *Nonstandard arguments and the characterization of independence in generic structures*, RIMS Kôkyûroku 1646, 4–17, 2009.

[2] R.I. Boţ, *Conjugate duality in convex optimization*, Lecture Notes in Economics and Mathematical Systems 637, Springer-Verlag, Berlin Heidelberg, 2010.

[3] R.I. Boţ, S.-M. Grad, *Lower semicontinuous type regularity conditions for subdifferential calculus*, Optimization Methods and Software 25(1), 37–48, 2010.

[4] R.I. Boţ, S.-M. Grad, G. Wanka, *A new constraint qualification for the formula of the subdifferential of composed convex functions infinite dimensional spaces*, Mathematische Nachrichten 281(8), 1088–1107, 2008.

[5] R.I. Boţ, S.-M. Grad, G. Wanka, *Duality in vector optimization*, Springer-Verlag, Berlin Heidelberg, 2009.

[6] R.I. Boţ, S.-M. Grad, G. Wanka, *Generalized Moreau-Rockafellar results for composed convex functions*, Optimization 58(7), 917–933, 2009.

[7] R.I. Boţ, S.-M. Grad, G. Wanka, *New constraint qualification and conjugate duality for composed convex optimization problems*, Journal of Optimization Theory and Applications 135(2), 241–255, 2007.

[8] R.I. Boţ, S.-M. Grad, G. Wanka, *New regularity conditions for Lagrange and Fenchel-Lagrange duality in infinite dimensional spaces*, Mathematical Inequalities & Applications 12(1), 171–189, 2009.

[9] R.I. Boţ, S.-M. Grad, G. Wanka: *New regularity conditions for strong and total Fenchel-Lagrange duality in infinite dimensional spaces*, Nonlinear Analysis: Theory, Methods & Applications 69(1), 323–336, 2008.

[10] R.I. Boţ, S.-M. Grad, G. Wanka: *On strong and total Lagrange duality for convex optimization problems*, Journal of Mathematical Analysis and Applications 337(2), 1315–1325, 2008.

[11] R.I. Boţ, I.B. Hodrea, G. Wanka, *Composed convex programming: duality and Farkas-type results*, in: Z. Kasa, G. Kassay, J. Kolumban (Eds.), "Proceedings of the International Conference in Memoriam Gyula Farkas", Cluj University Press, Cluj-Napoca, 22–35, 2006.

[12] R.I. Boţ, I.B. Hodrea, G. Wanka, *$\varepsilon-$optimality conditions for composed convex optimization problems*, Journal of Approximation Theory 153(1), 108–121, 2008.

[13] R.I. Boţ, I.B. Hodrea, G. Wanka, *Farkas-type results for inequality systems with composed convex functions via conjugate duality*, Journal of Mathematical Analysis and Applications 322(1), 316–328, 2006.

[14] C. Combari, M. Laghdir, L. Thibault, *A note on subdifferentials of convex composite functionals*, Archiv der Mathematik 67(3), 239–252, 1996.

[15] C. Combari, M. Laghdir, L. Thibault, *On subdifferential calculus for convex functions defined on locally convex spaces*, Annales des sciences mathématiques du Québec 23(1), 23–36, 1999.

[16] C. Combari, M. Laghdir, L. Thibault, *Sous-différentiels de fonctions convexes composées*, Annales des sciences mathématiques du Québec 18(2), 119–148, 1994.

[17] H.M. Friedman, *A way out*, in: Godehard Link (Ed.), "One Hundred Years of Russell's Paradox", de Gruyter, 49–86, 2004.

[18] G.S. Kim, G.M. Lee, *On ε-approximate solutions for convex semidefinite optimization problems*, Taiwanese Journal of Mathematics 11(3), 765–784, 2007.

[19] M.M. Tidball, O. Pourtallier, E. Altman, *Approximations in dynamic zero-sum games*, SIAM Journal on Optimization 35(6), 2101-2117, 2006.

[20] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific, Singapore, 2002.