Duality and epsilon-optimality conditions for multi-composed optimization problems with applications to fractional and entropy optimization

S.-M. Grad, G. Wanka, O. Wilfer

Preprint 2016-4

Preprintreihe der Fakultät für Mathematik
ISSN 1614-8835

Duality and ε-optimality conditions for multi-composed optimization problems with applications to fractional and entropy optimization

Sorin-Mihai Grad* Gert Wanka† Oleg Wilfer‡

Abstract: We introduce a closedness type regularity condition that characterizes the stable strong duality for convex constrained optimization problems with multi-composed objective functions and guarantees a formula for the ε-subdifferential of a multi-composed function, that is employed for delivering necessary and sufficient ε-optimality conditions that characterize ε-optimality solutions to multi-composed optimization problems. As a byproduct, a formula of the conjugate function of a multi-composed function is provided under a regularity condition weaker than known in the literature. We also present two possible applications of our investigations in fractional programming and entropy optimization, respectively.

Key words: convex functions, composed functions, regularity conditions, conjugate functions, fractional programming, entropy optimization

1 Introduction and preliminaries

Motivated by possible applications in fractional programming and entropy optimization, that are discussed in the last section, as well as in other fields briefly mentioned later, we present in this paper some investigations on duality and optimality as well as corresponding ε-subdifferential formulae for convex constrained optimization problems with multi-composed objective functions. To the best of our knowledge such functions were considered in similar contexts only in [20], where investigations on Lagrange duality for the mentioned class of problems were presented, strong duality being delivered under some interiority type regularity conditions. However, the term "multi-composed" can be found in different research fields in connection to (mechanical) systems, materials, substances or images, and an eventual mathematical modelling of such problems may contain multi-compositions of functions. We introduce a closedness type regularity condition that characterizes the stable strong duality for convex constrained optimization problems with multi-composed objective functions and we show that it also guarantees a formula for the ε-subdifferential of a multi-composed function, where $\varepsilon \geq 0$. The latter is then employed

*Faculty of Mathematics, Chemnitz University of Technology, D-09107 Chemnitz, Germany, telephone: 0371 531-36002, fax number: 0371 531-22409, e-mail: grad@mathematik.tu-chemnitz.de

†Faculty of Mathematics, Chemnitz University of Technology, D-09107 Chemnitz, Germany, telephone: 0371 531-38560, fax number: 0371 531-22409, e-mail: gert.wanka@mathematik.tu-chemnitz.de

‡Faculty of Mathematics, Chemnitz University of Technology, D-09107 Chemnitz, Germany, telephone: 0371 531-39521, fax number: 0371 531-22409, e-mail: oleg.wilfer@mathematik.tu-chemnitz.de

for delivering necessary and sufficient ε-optimality conditions that characterize ε-optimality solutions to multi-composed optimization problems. As a byproduct, a formula of the conjugate function of a multi-composed function is provided under a closedness type regularity condition that is weaker than the one given in [20]. Different results involving composed functions from the literature can be recovered as special cases of the statements we provide in this paper. In all the new formulae we deliver, the functions involved in the original chain of compositions appear alone, allowing thus a separate processing. This might prove to be of advantage when concretely solving such problems by means of numerical algorithms, for instance by employing primal-dual splitting type methods. However, such investigations remain subject to future research.
In the following we present the framework we work in and some preliminary notions and results needed later in our investigations.

Let X be a Hausdorff locally convex space and X^* its topological dual space endowed with the weak* topology $w(X^*, X)$. For $x \in X$ and $x^* \in X^*$, let $\langle x^*, x\rangle := x^*(x)$ be the value of the linear continuous functional x^* at x.
A set $D \subseteq X$ is said to be closed regarding the subspace $T \subseteq X$ if $D \cap T = \operatorname{cl} D \cap T$, where $\operatorname{cl} D$ denotes the closure of D. Consider a convex cone $K \subseteq X$, which induces on X a partial ordering relation "$\leqq_K$", defined by $\leqq_K := \{(x, y) \in X \times X : y - x \in K\}$, i.e. for $x, y \in X$ it holds $x \leqq_K y \Leftrightarrow y - x \in K$. Note that we assume that all cones we consider contain the origin. Further, we attach to X a greatest element with respect to "$\leqq_K$", denoted by $+\infty_K$, which does not belong to X and denote $\overline{X} = X \cup \{+\infty_K\}$. Then it holds $x \leqq_K +\infty_K$ for all $x \in \overline{X}$. We write $x \leq_K y$ if and only if $x \leqq_K y$ and $x \neq y$. Further, we write $\leqq_{\mathbb{R}_+} =: \leq$ and $\leq_{\mathbb{R}_+} =: <$.
On $\overline{X}$ we consider the following operations and conventions: $x + (+\infty_K) = (+\infty_K) + x := +\infty_K \ \forall x \in X \cup \{+\infty_K\}$ and $\lambda \cdot (+\infty_K) := +\infty_K \ \forall \lambda \in [0, +\infty]$. Further, $K^* := \{x^* \in X^* : \langle x^*, x\rangle \geq 0, \ \forall x \in K\}$ is the dual cone of K and we take by convention $\langle x^*, +\infty_K\rangle := +\infty$ for all $x^* \in K^*$. By a slight abuse of notation we denote the extended real space $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ and consider on it the following operations and conventions: $\lambda + (+\infty) = (+\infty) + \lambda := +\infty \ \forall \lambda \in [-\infty, +\infty]$, $\lambda + (-\infty) = (-\infty) + \lambda := -\infty \ \forall \lambda \in [-\infty, +\infty)$, $\lambda \cdot (+\infty) := +\infty \ \forall \lambda \in [0, +\infty]$, $\lambda \cdot (+\infty) := -\infty \ \forall \lambda \in [-\infty, 0)$, $\lambda \cdot (-\infty) := -\infty \ \forall \lambda \in (0, +\infty]$, $\lambda \cdot (-\infty) := +\infty \ \forall \lambda \in [-\infty, 0)$, and $0(-\infty) := 0$. For a subset $A \subseteq X$, its indicator function $\delta_A : X \to \overline{\mathbb{R}}$ is

$$\delta_A(x) := \begin{cases} 0, & \text{if } x \in A, \\ +\infty, & \text{otherwise.} \end{cases}$$

For a given function $f : X \to \overline{\mathbb{R}}$ we consider its effective domain $\operatorname{dom} f := \{x \in X : f(x) < +\infty\}$ as well as its graph $\operatorname{gra} f := \{(x, f(x)) : x \in \operatorname{dom} f\}$, and call it f proper if $\operatorname{dom} f \neq \emptyset$ and $f(x) > -\infty$ for all $x \in X$. The epigraph of f is $\operatorname{epi} f = \{(x, r) \in X \times \mathbb{R} : f(x) \leq r\}$. The conjugate function of f with respect to the non-empty subset $S \subseteq X$ is defined by $f_S^* : X^* \to \overline{\mathbb{R}}$, $f_S^*(x^*) = \sup_{x \in S}\{\langle x^*, x\rangle - f(x)\}$. In the case $S = X$, f_S^* turns into the classical Fenchel-Moreau conjugate function of f denoted by f^*. Recall that a function $f : X \to \overline{\mathbb{R}}$ is called convex if $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda) f(y)$ for all $x, y \in X$ and all $\lambda \in [0, 1]$. A function $f : X \to \overline{\mathbb{R}}$ is called lower semicontinuous at $\overline{x} \in X$ if $\liminf_{x \to \overline{x}} f(x) \geq f(\overline{x})$ and when this function is lower semicontinuous at all $x \in X$, then we call it lower semicontinuous (l.s.c. for short). Let $W \subseteq X$ be a non-empty set, then a function $f : X \to \overline{\mathbb{R}}$ is called K-increasing on W, if from $x \leqq_K y$ follows $f(x) \leq f(y)$ for all $x, y \in W$. When $W = X$, then we call the function f K-increasing.

Remark 1.1. *Note that for a proper function f it holds* $\operatorname{epi} f = \{(x,r) \in X \times \mathbb{R} : k \geq 0,\ r = f(x) + k\} = \operatorname{gra} f + \{0_X\} \times \mathbb{R}_+$.

If we take for a proper function $f : X \to \overline{\mathbb{R}}$ an arbitrary $x \in X$ such that $f(x) \in \mathbb{R}$, then we call the set $\partial_\varepsilon f(x) := \{x^* \in X^* : f(y) - f(x) \geq \langle x^*, y - x\rangle - \varepsilon,\ \forall y \in X\}$ for $\varepsilon \geq 0$ the ε-subdifferential of f at x. Moreover, for $\varepsilon = 0$ we write $\partial f(x) = \partial_0 f(x)$ and we say that f is subdifferentiable at x if $\partial f(x) \neq \emptyset$. Additionally, we make the convention that $\partial_\varepsilon f(x) := \emptyset$ if $f(x) \notin \mathbb{R}$. It is well-known that (see [9])

$$f(x) + f^*(x^*) \leq \langle x^*, x\rangle + \varepsilon \Leftrightarrow x^* \in \partial_\varepsilon f(x). \tag{1}$$

Let Z be another Hausdorff locally convex space partially ordered by the convex cone $Q \subseteq Z$ and Z^* its topological dual space endowed with the weak* topology $w(Z^*, Z)$. The domain of a vector function $F : X \to \overline{Z} = Z \cup \{+\infty_Q\}$ is $\operatorname{dom} F := \{x \in X : F(x) \neq +\infty_Q\}$. F is called proper if $\operatorname{dom} F \neq \emptyset$. When $F(\lambda x + (1-\lambda)y) \leqq_Q \lambda F(x) + (1-\lambda)F(y)$ holds for all $x, y \in X$ and all $\lambda \in [0,1]$ the function F is said to be Q-convex. The Q-epigraph of a vector function F is $\operatorname{epi}_Q F = \{(x,z) \in X \times Z : F(x) \leqq_Q z\}$ and when Q is closed we say that F is Q-epi closed if $\operatorname{epi}_Q F$ is a closed set. For a $z^* \in Q^*$ we define the function $(z^*F) : X \to \overline{\mathbb{R}}$ by $(z^*F)(x) := \langle z^*, F(x)\rangle$. Then $\operatorname{dom}(z^*F) = \operatorname{dom} F$. Moreover, it is easy to see that if F is Q-convex, then (z^*F) is convex for all $z^* \in Q^*$. The vector function F is called positively Q-lower semicontinuous at $x \in X$ if (z^*F) is lower semicontinuous at x for all $z^* \in Q^*$. The function F is called positively Q-lower semicontinuous if it is positively Q-lower semicontinuous at every $x \in X$. Note that if F is positively Q-lower semicontinuous, then it is also Q-epi closed, while the inverse statement is not true in general (see: [4, Proposition 2.2.19]). Let us mention that in the case $Z = \mathbb{R}$ and $Q = \mathbb{R}_+$, the notion of Q-epi closedness falls into the classical notion of lower semicontinuity. $F : X \to \overline{Z}$ is called (K,Q)-increasing on W, if from $x \leqq_K y$ follows $F(x) \leqq_Q F(y)$ for all $x, y \in W$. When $W = X$, we call this function (K,Q)-increasing.
Given an optimization problem (P), we denote its optimal objective value by $v(P)$.

We give now some statements that will be useful later in our presentation, beginning with one whose proof is straightforward.

Lemma 1.1. *Let V be a Hausdorff locally convex space partially ordered by the convex cone U, $F : X \to \overline{Z}$ be a proper and Q-convex function and $G : Z \to \overline{V}$ be an U-convex and (Q,U)-increasing function on $F(\operatorname{dom} F) \subseteq \operatorname{dom} G$ with the convention $G(+\infty_Q) = +\infty_U$. Then the function $(G \circ F) : X \to \overline{V}$ is U-convex.*

Lemma 1.2. *Let Y be a Hausdorff locally convex space, Q also closed, $h : X \times Y \to \overline{Z}$ and $F : X \to \overline{Z}$ proper vector functions and $G : Y \to Z$ a continuous vector functions, where h is defined by $h(x,y) := F(x) + G(y)$. Then F is Q-epi closed if and only if h is Q-epi closed.*

Proof. "$\Rightarrow$": Let $(x_\alpha, y_\alpha, z_\alpha)_\alpha \subseteq \operatorname{epi}_Q h$ such that $(x_\alpha, y_\alpha, z_\alpha) \to (\overline{x}, \overline{y}, \overline{z})$. Then $F(x_\alpha) + G(y_\alpha) \leq z_\alpha$ for any α, followed by $(x_\alpha, z_\alpha - G(y_\alpha))_\alpha \subseteq \operatorname{epi}_Q F$ and $(y_\alpha, G(y_\alpha))_\alpha \subseteq \operatorname{epi}_Q G$. Because G is continuous and $y_\alpha \to \overline{y}$, it follows that $G(y_\alpha) \to G(\overline{y})$. Then $(x_\alpha, z_\alpha - G(y_\alpha)) \to (\overline{x}, \overline{z} - G(\overline{y})) \in \operatorname{epi}_Q F$, because this set is closed. One has then $F(\overline{x}) \leqq_Q \overline{z} - G(\overline{y})$, i.e. $(\overline{x}, \overline{y}, \overline{z}) \in \operatorname{epi}_Q h$. As the convergent nets $(x_\alpha)_\alpha$, $(y_\alpha)_\alpha$ and $(z_\alpha)_\alpha$ were arbitrarily chosen, it follows that $\operatorname{epi}_Q h$ is closed, i.e. h is Q-epi closed.
"$\Leftarrow$": Let $(x_\alpha, z_\alpha)_\alpha \subseteq \operatorname{epi}_Q F$ such that $(x_\alpha, z_\alpha) \to (\overline{x}, \overline{z})$. Take also $(y_\alpha)_\alpha \subseteq Y$ such that $y_\alpha \to \overline{y}$. Because G is continuous, one has $G(y_\alpha) \to G(\overline{y})$. Then $(x_\alpha, y_\alpha, z_\alpha + G(y_\alpha))_\alpha \subseteq \operatorname{epi}_Q h$,

which is closed, consequently $(\overline{x}, \overline{y}, \overline{z} + G(\overline{y})) \in \operatorname{epi}_Q h$, i.e. $F(\overline{x}) + G(\overline{y}) \leqq_Q \overline{z} + G(\overline{y})$. Therefore $F(\overline{x}) \leqq_Q \overline{z}$, i.e. $(\overline{x}, \overline{z}) \in \operatorname{epi}_Q F$. As the convergent nets $(x_\alpha)_\alpha$ and $(z_\alpha)_\alpha$ were arbitrarily chosen, it follows that $\operatorname{epi}_Q F$ is closed, i.e. F is Q-epi closed. $\square$

Remark 1.2. *Note that a continuous proper vector function $G : Y \to \overline{Z}$, where Y is a Hausdorff locally convex space, has a full domain, thus one can directly take $G : Y \to Z$ in this situation. The question whether the equivalence in Lemma 1.2 remains valid if one considers a proper vector function $G : Y \to \overline{Z}$ that is not necessarily continuous is still open.*

Remark 1.3. *If we set $Y = Z$ and $G(y) = -y$, $\forall y \in Y$, then Lemma 1.2 says that F is Q-epi closed if and only if the vector function defined by $(x, y) \in X \times Y \mapsto F(x) - y$ is Q-epi closed. For this special case a similar statement can be found in [20, Lemma 2.1], but under the additional hypothesis* $\operatorname{int} Q \neq \emptyset$.

Let $X_0, \ldots, X_n$ be Hausdorff locally convex spaces and consider the functions $f_i : X_i \to \overline{\mathbb{R}}$, $i = 0, \ldots, n$ and $\phi : X_0 \times \ldots \times X_n \to \overline{\mathbb{R}}$ defined by $\phi(y^0, \ldots, y^n) = \sum_{i=0}^n f_i(y^i)$. It can easily be verified that $\operatorname{dom} \phi = \prod_{i=0}^n \operatorname{dom} f_i$. Furthermore, letting $\mathcal{T}^n_{X_i} : X_i \times \mathbb{R} \to \{0_{X_0}\} \times \ldots \times \{0_{X_{i-1}}\} \times X_i \times \{0_{X_{i+1}}\} \times \ldots \times \{0_{X_n}\} \times \mathbb{R}$ be defined by $\mathcal{T}^n_{X_i}(x^i, r) := (0_{X_0}, \ldots, 0_{X_{i-1}}, x^i, 0_{X_{i+1}}, \ldots, 0_{X_n}, r)$ for all $x^i \in X_i$, $i = 0, \ldots, n$, (with the usual conventions, i.e. when $i = 0$ there is no X_{i-1}) and $r \in \mathbb{R}$, one gets the following statement.

Lemma 1.3. *Let $f_i : X_i \to \overline{\mathbb{R}}$ be a proper function, $i = 0, \ldots, n$, then it holds*

$$\operatorname{epi} \phi = \sum_{i=0}^{n} \mathcal{T}^n_{X_i} (\operatorname{epi} f_i).$$

Proof. Using Remark 1.1, one gets

$$\begin{aligned}
&\operatorname{epi} \phi = \{(y^0, \ldots, y^n, r) : \phi(y^0, \ldots, y^n) \leq r\} \\
= \ & \{(y^0, \ldots, y^n, \phi(y^0, \ldots, y^n)) : (y^0, \ldots, y^n) \in \operatorname{dom} \phi\} + \{0_{X_0}\} \times \ldots \times \{0_{X_n}\} \times \mathbb{R}_+ \\
= \ & \{(y^0, \ldots, y^n, f_0(y^0) + \ldots + f_n(y^n)) : y^i \in \operatorname{dom} f_i, \ i = 0, \ldots, n\} + \{0_{X_0}\} \times \ldots \times \{0_{X_n}\} \times \mathbb{R}_+ \\
= \ & \sum_{i=0}^{n} (\{(0_{X_0}, \ldots, 0_{X_{i-1}}, y^i, 0_{X_{i+1}}, \ldots, 0_{X_n}, f_i(y^i)) : y^i \in \operatorname{dom} f_i\} + \{0_{X_0}\} \times \ldots \times \{0_{X_n}\} \times \mathbb{R}_+) \\
= \ & \sum_{i=0}^{n} \{(0_{X_0}, \ldots, 0_{X_{i-1}}, y^i, 0_{X_{i+1}}, \ldots, 0_{X_n}, r_i) : f_i(y_i) \leq r_i\} = \sum_{i=0}^{n} \mathcal{T}^n_{X_i} (\operatorname{epi}(f_i)). \quad \square
\end{aligned}$$

We also consider the operator $\widetilde{\mathcal{T}}^n_{X_i} : X_i \times \mathbb{R} \times X_{i-1} \to \{0_{X_0}\} \times \ldots \times \{0_{X_{i-2}}\} \times X_{i-1} \times X_i \times \{0_{X_{i+1}}\} \times \ldots \times \{0_{X_n}\} \times \mathbb{R}$ defined by $\widetilde{\mathcal{T}}^n_{X_i}(x^i, r, x^{i-1}) := (0_{X_0}, \ldots, 0_{X_{i-2}}, x^{i-1}, x^i, 0_{X_{i+1}}, \ldots, 0_{X_n}, r)$ for all $x^i \in X_i$, $i = 1, \ldots, n$, and $r \in \mathbb{R}$, where it is easy to see that

$$\begin{aligned}
& \mathcal{T}^n_{X_i}(x^i, r) + (0_{X_0}, \ldots, 0_{X_{i-2}}, x^{i-1}, 0_{X_i}, \ldots, 0_{X_n}, 0) \\
= \ & (0_{X_0}, \ldots, 0_{X_{i-2}}, x^{i-1}, x^i, 0_{X_{i+1}}, \ldots, 0_{X_n}, r) = \widetilde{\mathcal{T}}^n_{X_i}(x^i, r, x^{i-1})
\end{aligned} \tag{2}$$

for all $x^i \in X_i$, $i = 1, \ldots, n$, and $r \in \mathbb{R}$.

Remark 1.4. *Note that the operators $\mathcal{T}^n_{X_i}$, $i = 1, \ldots, n$, are homeomorphisms. This means that for a non-empty subset $P_i \subseteq X_i \times \mathbb{R}$ the set $\mathcal{T}^n_{X_i}(P_i)$ is compact if and only if the subset P_i is compact. The same holds also for the function $\widetilde{\mathcal{T}}^n_{X_i}$, $i = 1, \ldots, n$.*

2 Lagrange duality for multi-composed optimization problems

The starting point of our research is for a fixed $x^* \in X_n^*$ the following multi-composed problem

$$
\begin{aligned}
(P^C_{x^*}) \qquad & \inf_{x \in \mathcal{A}} \{(f \circ F^1 \circ \ldots \circ F^n)(x) - \langle x^*, x\rangle\},\\
& \mathcal{A} = \{x \in S : g(x) \in -Q\},
\end{aligned}
$$

where X_i, $i = 0, \ldots, n$, are Hausdorff locally convex spaces such that X_j is partially ordered by the convex cone $K_j \subseteq X_j$ for $j = 0, \ldots, n-1$. Moreover, $S \subseteq X_n$ is a non-empty set, $f : X_0 \to \overline{\mathbb{R}}$ is a proper and K_0-increasing function on $\operatorname{dom} f$ and $F^1(\operatorname{dom} F^1) \subseteq \operatorname{dom} f$, $F^i : X_i \to \overline{X}_{i-1} = X_{i-1} \cup \{+\infty_{K_{i-1}}\}$ is a proper and (K_i, K_{i-1})-increasing function on $\operatorname{dom} F^i$ and $F^{i+1}(\operatorname{dom} F^{i+1}) \subseteq \operatorname{dom} F^i$ for $i = 1, \ldots, n-2$, $F^{n-1} : X_{n-1} \to \overline{X}_{n-2} \cup \{+\infty_{K_{n-2}}\}$ is a proper and (K_{n-1}, K_{n-2})-increasing function on $\operatorname{dom} F^{n-1}$ and $F^n(\operatorname{dom} F^n \cap \mathcal{A}) \subseteq \operatorname{dom} F^{n-1}$, $F^n : X_n \to \overline{X}_{n-1} = X_{n-1} \cup \{+\infty_{K_{n-1}}\}$ is a proper function and $g : X_n \to \overline{Z}$ is a proper function fulfilling $S \cap g^{-1}(-Q) \cap ((F^n)^{-1} \circ \ldots \circ (F^1)^{-1})(\operatorname{dom} f) \cap \operatorname{dom} F^n \neq \emptyset$. We also make the following conventions: $f(+\infty_{K_0}) = +\infty$ and $F^i(+\infty_{K_i}) = +\infty_{K_{i-1}}$, extending thus the involved functions as follows $f : \overline{X}_0 \to \overline{\mathbb{R}}$ and $F^i : \overline{X}_i \to \overline{X}_{i-1}$, $i = 1, \ldots, n-1$.
When $x^* = 0_{X_n^*}$ the problem $(P^C_{x^*})$ collapses to

$$
\begin{aligned}
(P^C) \qquad & \inf_{x \in \mathcal{A}} \{(f \circ F^1 \circ \ldots \circ F^n)(x)\},\\
& \mathcal{A} = \{x \in S : g(x) \in -Q\},
\end{aligned}
$$

that was considered in [20]. One can notice that $(P^C_{x^*})$ is a linearly perturbed problem of (P^C). In order to approach $(P^C_{x^*})$, for fixed $x^* \in X_n^*$, by means of Lagrange duality consider the following optimization problem

$$
\begin{aligned}
(\widetilde{P}^C_{x^*}) \qquad & \inf_{(y^0, \ldots, y^n) \in \widetilde{\mathcal{A}}} \{\widetilde{f}(y^0, \ldots, y^n) - \langle x^*, y^n\rangle\}\\
& \widetilde{\mathcal{A}} = \{(y^0, \ldots, y^{n-1}, y^n) \in X_0 \times \ldots \times X_{n-1} \times S :\\
& \qquad g(y^n) \in -Q,\ h^i(y^i, y^{i-1}) \in -K_{i-1},\ i = 1, \ldots, n\},
\end{aligned}
$$

where $\widetilde{f} : X_0 \times \ldots \times X_n \to \overline{\mathbb{R}}$ and $h^i : X_i \times X_{i-1} \to \overline{X}_{i-1}$ are defined as

$$
\widetilde{f}(y^0, \ldots, y^n) = f(y^0) \text{ and } h^i(y^i, y^{i-1}) = F^i(y^i) - y^{i-1} \text{ for } i = 1, \ldots, n.
$$

Its Lagrange dual problem is

$$
\begin{aligned}
(\widetilde{D}^{C_L}_{x^*}) \qquad & \sup_{\substack{z^{n*} \in Q^*,\ z^{i*} \in K_i^* \\ i=0,\ldots,n-1}} \inf_{\substack{y^n \in S,\ y^i \in X_i \\ i=0,\ldots,n-1}} \Bigg\{ \widetilde{f}(y^0, \ldots, y^n) - \langle x^*, y^n\rangle + \sum_{i=1}^{n} \langle z^{(i-1)*}, h^i(y^i, y^{i-1})\rangle \\
& \qquad + \langle z^{n*}, g(y^n)\rangle \Bigg\}.
\end{aligned}
$$

As $v(P^C_{x^*}) = v(\widetilde{P}^C_{x^*})$ (cf. [20, Theorem 2]), we use $(\widetilde{D}^{C_L}_{x^*})$ to assign the following Lagrange dual problem to $(P^C_{x^*})$

$$
\begin{aligned}
(D^{C_L}_{x^*}) \qquad & \sup_{\substack{z^{n*} \in Q^*,\ z^{i*} \in K_i^* \\ i=0,\ldots,n-1}} \inf_{\substack{y^n \in S,\ y^i \in X_i \\ i=0,\ldots,n-1}} \Bigg\{ f(y^0) - \langle x^*, y^n\rangle + \langle z^{(n-1)*}, F^n(y^n) - y^{n-1}\rangle + \\
& \qquad \langle z^{n*}, g(y^n)\rangle + \sum_{i=1}^{n-1} \langle z^{(i-1)*}, F^i(y^i) - y^{i-1}\rangle \Bigg\},
\end{aligned}
$$

that can be equivalently written as

$$(D^{C_L}_{x^*})\quad \sup_{\substack{z^{n*}\in Q^*,\ z^{i*}\in K_i^*,\\ i=0,\dots,n-1}} \Bigg\{ -\sup_{y^n\in S}\{\langle x^*,y^n\rangle - \langle z^{(n-1)*},F^n(y^n)\rangle - \langle z^{n*},g(y^n)\rangle\} - \sup_{y^0\in X_0}\{\langle z^{0*},y^0\rangle - f(y^0)\} - \sum_{i=1}^{n-1}\sup_{\substack{y^i\in X_i,\\ i=1,\dots,n-1}}\{\langle z^{i*},y^i\rangle - \langle z^{(i-1)*},F^i(y^i)\rangle\}\Bigg\}$$

and even simplified to

$$(D^{C_L}_{x^*})\qquad \sup_{\substack{z^{n*}\in Q^*,\ z^{i*}\in K_i^*\\ i=0,\dots,n-1}} \Big\{ -f^*(z^{0*}) - \sum_{i=1}^{n-1}(z^{(i-1)*}F^i)^*(z^{i*}) - ((z^{(n-1)*}F^n)+(z^{n*}g))^*_S(x^*)\Big\}.$$

Remark 2.1. *For $x^*=0$ $(D^{C_L}_{x^*})$ turns out to be the Lagrange dual problem to (P^C) which was introduced in [20] and will be denoted further by (D^{C_L}). Additionally, note that the weak duality for $(\widetilde{P}^C_{x^*})$ and $(\widetilde{D}^{C_L}_{x^*})$ is always fulfilled, i.e. $v(\widetilde{P}^C_{x^*}) \geq v(\widetilde{D}^{C_L}_{x^*})$. Thus one has $v(P^C_{x^*}) = v(\widetilde{P}^C_{x^*}) \geq v(\widetilde{D}^{C_L}_{x^*}) = v(D^{C_L}_{x^*})$ and, hence, for any $x^*\in X_n^*$, it holds*

$$\sup_{x\in X_n}\{\langle x^*,x\rangle - (f\circ F^1\circ\ldots\circ F^n)(x) - \delta_{\mathcal{A}}(x)\} \leq \inf_{\substack{z^{n*}\in Q^*,\ z^{i*}\in K_i^*,\\ i=0,\dots,n-1}} \Big\{ f^*(z^{0*}) + \sum_{i=1}^{n-1}(z^{(i-1)*}F^i)^*(z^{i*}) + ((z^{(n-1)*}F^n)+(z^{n*}g))^*_S(x^*)\Big\},$$

i.e. for all $z^{n*}\in Q^*$ and $z^{i*}\in K_i^*$, $i=0,\dots,n-1$, one has the inequality

$$(f\circ F^1\circ\ldots\circ F^n)^*_{\mathcal{A}}(\cdot) \leq f^*(z^{0*}) + \sum_{i=1}^{n-1}(z^{(i-1)*}F^i)^*(z^{i*}) + ((z^{(n-1)*}F^n)+(z^{n*}g))^*_S(\cdot). \tag{3}$$

In order to achieve strong duality for the primal-dual pair $(P^C_{x^*})$-$(D^{C_L}_{x^*})$, that is actually stable strong duality for (P^C)-(D^{C_L}) and also corresponds to the equality case in (3), one needs additional hypotheses. To this end we employ the following regularity condition, considered in [16] for guaranteeing stable strong Lagrange duality,

$$(RC'_L)\ \Big|\ M' := \bigcup_{\widetilde{z}^*\in\widetilde{K}^*} \operatorname{epi}((\widetilde{f} + (\widetilde{z}^*\widetilde{h}) + \delta_{\widetilde{S}})^*) \text{ is closed regarding the subspace } \mathcal{U},$$

where $\mathcal{U} := \{0_{X_0^*}\}\times\ldots\times\{0_{X_{n-1}^*}\}\times X_n^*\times\mathbb{R}$, $\widetilde{y} := (y^0,\dots,y^n)\in\widetilde{X} := X_0\times\ldots\times X_n$, $\widetilde{K} := K_0\times\ldots\times K_{n-1}\times Q$, $\widetilde{S} := X_0\times\ldots\times X_{n-1}\times S$, $\widetilde{Z} := X_0\times\ldots\times X_{n-1}\times Z$, $\widetilde{X}^* := X_0^*\times\ldots\times X_n^*$, $\widetilde{z}^* := (z^{0*},\dots,z^{(n-1)*},z^{n*})\in\widetilde{K}^* := K_0^*\times\ldots\times K_{n-1}^*\times Q^*$ and $\widetilde{h}:\widetilde{X}\to\overline{\widetilde{Z}} = \widetilde{Z}\cup\{+\infty_{\widetilde{K}}\}$ defined as

$$\widetilde{h}(\widetilde{y}) := \begin{cases} (h^1(y^1,y^0),\dots,h^n(y^n,y^{n-1}),g(y^n)), & \text{if } (y^i,y^{i-1})\in\operatorname{dom}h^i,\ i=1,\dots,n,\ y^n\in\operatorname{dom}g,\\ +\infty_{\widetilde{K}}, & \text{otherwise.}\end{cases} \tag{4}$$

In order to formulate the regularity condition only by means of the originally considered functions and sets we have the following statement.

Lemma 2.1. *The set M' can equivalently be expressed as*

$$M = \mathcal{T}^n_{X_0^*}(\text{epi}(f^*)) + \bigcup_{\substack{z^{n*}\in Q^*,\ z^{i*}\in K_i^*,\\ i=0,\dots,n-1}} \Bigg(\sum_{i=1}^{n-1} \widetilde{\mathcal{T}}^n_{X_i^*} \left(\text{epi}((z^{(i-1)*}F^i)^*) \times \{-z^{(i-1)*}\}\right)$$
$$+\widetilde{\mathcal{T}}^n_{X_n^*} \left(\text{epi}(((z^{(n-1)*}F^n) + (z^{n*}g))^*_S) \times \{-z^{(n-1)*}\}\right) \Bigg).$$

Proof. For fixed $z^{n*} \in Q^*$, $z^{i*} \in K_i^*$, $i = 0, \dots, n-1$, and $\widetilde{y}^* = (y^{0*}, \dots, y^{n*}) \in X_0^* \times \dots \times X_n^*$, we have

$$\begin{aligned}
&(\widetilde{f} + (\widetilde{z}^*\widetilde{h}) + \delta_{\widetilde{S}})^*(\widetilde{y}^*) = \sup_{\widetilde{y}\in\widetilde{S}} \{\langle \widetilde{y}^*, \widetilde{y}\rangle - \widetilde{f}(\widetilde{y}) - \langle \widetilde{z}^*, \widetilde{h}(\widetilde{y})\rangle\} \\
= &\sup_{\substack{y^i\in X_i,\ i=0,\dots,n-1\\ y^n\in S}} \left\{ \sum_{i=0}^{n} \langle y^{i*}, y^i\rangle - f(y^0) - \sum_{i=1}^{n} \langle z^{(i-1)*}, F^i(y^i) - y^{i-1}\rangle - \langle z^{n*}, g(y^n)\rangle \right\} \\
= &\sup_{y^0\in X_0} \{\langle y^{0*} + z^{0*}, y^0\rangle - f(y^0)\} + \sup_{y^n\in S} \{\langle y^{n*}, y^n\rangle - \langle z^{(n-1)*}, F^n(y^n)\rangle - \langle z^{n*}, g(y^n)\rangle\} + \\
&\sum_{i=1}^{n-1} \sup_{y^i\in X_i,\ i=1,\dots,n-1} \left\{ \langle y^{i*} + z^{i*}, y^i\rangle - \langle z^{(i-1)*}, F^i(y^i)\rangle \right\} \\
= &f^*(y^{0*} + z^{0*}) + ((z^{(n-1)*}F^n) + (z^{n*}g))^*_S(y^{n*}) + \sum_{i=1}^{n-1} (z^{(i-1)*}F^i)^*(y^{i*} + z^{i*}).
\end{aligned} \tag{5}$$

Moreover, one has

$$\begin{aligned}
&(\widetilde{y}^*, r) \in \bigcup_{\widetilde{z}^*\in\widetilde{K}^*} \text{epi}(\widetilde{f} + (\widetilde{z}^*\widetilde{h}) + \delta_{\widetilde{S}})^* \\
\Leftrightarrow\ &\exists (z^{0*}, \dots, z^{(n-1)*}, z^{n*}) \in K_0^* \times K_1^* \times \dots \times K_{n-1}^* \times Q^* \text{ such that} \\
&f^*(y^{0*} + z^{0*}) + ((z^{(n-1)*}F^n) + (z^{n*}g))^*_S(y^{n*}) + \sum_{i=1}^{n-1} (z^{(i-1)*}F^i)^*(y^{i*} + z^{i*}) \le r.
\end{aligned} \tag{6}$$

Employing Lemma 1.3 and (2), this is further equivalent to

$$\begin{aligned}
&\exists (z^{0*}, \dots, z^{(n-1)*}, z^{n*}) \in K_0^* \times K_1^* \times \dots \times K_{n-1}^* \times Q^* \text{ such that} \\
&(y^{0*}, \dots, y^{n*}, r) \in \mathcal{T}^n_{X_0^*}(\text{epi}(f^*)) + \sum_{i=1}^{n-1} \widetilde{\mathcal{T}}^n_{X_i^*} \left(\text{epi}((z^{(i-1)*}F^i)^*) \times \{-z^{(i-1)*}\} \right) \\
&+\widetilde{\mathcal{T}}^n_{X_n^*} \left(\text{epi}(((z^{(n-1)*}F^n) + (z^{n*}g))^*_S) \times \{-z^{(n-1)*}\} \right),
\end{aligned}$$

which actually means that

$$(y^{0*}, \dots, y^{n*}, r) \in \mathcal{T}^n_{X_0^*}(\text{epi}(f^*)) + \bigcup_{\substack{z^{n*}\in Q,\ z^{i*}\in K_i^*,\\ i=0,\dots,n-1}} \Bigg(\sum_{i=1}^{n-1} \widetilde{\mathcal{T}}^n_{X_i^*} \left(\text{epi}((z^{(i-1)*}F^i)^*) \times \{-z^{(i-1)*}\} \right)$$
$$+\widetilde{\mathcal{T}}^n_{X_n^*} \left(\text{epi}(((z^{(n-1)*}F^n) + (z^{n*}g))^*_S) \times \{-z^{(n-1)*}\} \right) \Bigg). \qquad \square$$

The regularity condition (RC'_L) introduced above can be thus reformulated as

$$(RC_L) \quad \Big| \quad M \text{ is closed regarding the subspace } \mathcal{U}.$$

In order to show the stable strong duality statement for (P^C) and (D^{C_L}), we also need to impose some convexity and topological hypotheses on the sets and functions. We assume for the rest of this paper that $S \subseteq X_n$ is a closed and convex set, f is a convex and lower semicontinuous function, F^i is a K_{i-1}-convex and K_{i-1}-epi closed vector function for $i = 1, \dots, n$ and g is a Q-convex and Q-epi closed vector function.

Theorem 2.1. *The regularity condition (RC_L) is fulfilled if and only if*

$$(f \circ F^1 \circ \dots \circ F^n)^*_{\mathcal{A}}(x^*) =$$
$$\min_{\substack{z^{n*} \in Q^*,\ z^{i*} \in K_i^*,\\ i=0,\dots,n-1}} \left\{ f^*(z^{0*}) + \sum_{i=1}^{n-1} (z^{(i-1)*} F^i)^*(z^{i*}) + ((z^{(n-1)*} F^n) + (z^{n*} g))^*_S(x^*) \right\} \forall x^* \in X_n^*,$$

i.e. there is stable strong duality for (P^C) and (D^{C_L}).

Proof. According to the previous considerations, it holds

$$v(P^C_{x*}) = v(\widetilde{P}^C_{x*}) \geq v(\widetilde{D}^{C_L}_{x*}) = v(D^{C_L}_{x*}).$$

The convexity and topological hypotheses imposed above guarantee, via Lemma 1.1 and Lemma 1.2, that $\widetilde{f}$ is convex and lower semicontinuous and h^i, $i = 1, \dots, n$, are K_{i-1}-convex and K_{i-1}-epi closed. Then, by [16, Theorem 2.7 and Corollary 2.3] one obtains for any $x^* \in X_n^*$ the existence of $\overline{z}^{n*} \in Q^*$ and $\overline{z}^{i*} \in K_i^*$, $i = 0, \dots, n-1$ such that

$$v(\widetilde{P}^C_{x*}) = \inf_{(y^0,\dots,y^n) \in \mathcal{A}} \{\widetilde{f}(y^0, \dots, y^n) - \langle x^*, y^n \rangle\}$$

$$= \sup_{\substack{z^{n*} \in Q^*,\ z^{i*} \in K_i^*\\ i=0,\dots,n-1}} \inf_{\substack{y^n \in S,\ y^i \in X_i\\ i=0,\dots,n-1}} \left\{ \widetilde{f}(y^0, \dots, y^n) - \langle x^*, y^n \rangle + \sum_{i=1}^{n} \langle z^{(i-1)*}, h^i(y^i, y^{i-1}) \rangle + \langle z^{n*}, g(y^n) \rangle \right\}$$

$$= \inf_{\substack{y^n \in S,\ y^i \in X_i\\ i=0,\dots,n-1}} \left\{ \widetilde{f}(y^0, \dots, y^n) - \langle x^*, y^n \rangle + \sum_{i=1}^{n} \langle \overline{z}^{(i-1)*}, h^i(y^i, y^{i-1}) \rangle + \langle \overline{z}^{n*}, g(y^n) \rangle \right\} = v(\widetilde{D}^{C_L}_{x*}).$$

Consequently, $v(P^C_{x*}) = v(D^{C_L}_{x*})$. □

Remark 2.2. *Since one has via [4, Theorem 3.5.9] stable strong duality for $(\widetilde{P}^C)$ and its Lagrange dual problem if and only if M' is closed in the topology $w(\widetilde{X}^*, \widetilde{X}) \times \mathbb{R}$, the regularity condition*

$$(RC_L^T) \quad \Big| \quad M \text{ is closed in the topology } w(\widetilde{X}^*, \widetilde{X}) \times \mathbb{R}$$

is a sufficient condition to have stable strong duality for (P^C) and (D^{C_L}).

Remark 2.3. *Alternatively to the Lagrange duality approach, one can consider the Fenchel-Lagrange type dual problem for (P^C), by employing the following perturbation function*

$$\Phi(x, y^0, \dots, y^{n+1}) := \begin{cases} f(F^1(\dots F^{n-1}(F^n(x + y^n) + y^{n-1}) \dots) + y^0), & \text{if } g(x) \in y^{n+1} - Q, \\ +\infty, & \text{otherwise,} \end{cases}$$

where $(y^0,\dots,y^n,y^{n+1}) \in X_0\times\ldots\times X_n\times Z$ *are the dual variables. By simple calculations one derives the following conjugate function of* Φ,

$$\Phi^*(x^*,y^{0*},\dots,y^{(n+1)*}) =$$
$$f^*(y^{0*})+\sum_{i=0}^{n-1}(y^{i*}F^{i+1})^*(y^{(i+1)*})+(-y^{(n+1)*}g)^*_S(x^*-y^{n*})+\delta_{Q^*}(-y^{(n+1)*}),$$

and hence, the following Fenchel-Lagrange type dual problem is assigned to (P^C),

$$(D^{C_{FL}})\qquad \sup_{(y^{0*},\dots,y^{n*},y^{(n+1)*})\in X_0^*\times\ldots\times X_n^*\times Z^*}\left\{-\Phi^*\left(0_{X_n^*},y^{0*},\dots,y^{n*},y^{(n+1)*}\right)\right\}$$
$$=\sup_{\substack{y^{i*}\in X_i^*, i=0,\dots,n,\\ y^{(n+1)*}\in Q^*}}\left\{-f^*(y^{0*})-\sum_{i=0}^{n-1}(y^{i*}F^{i+1})^*(y^{(i+1)*})-(y^{(n+1)*}g)^*_S(-y^{n*})\right\}.$$

Notice that different to (D^{C_L}), *in* $(D^{C_{FL}})$ *all the involved functions appear separately. This might be useful for computational reasons, for instance when employing splitting type methods. Moreover, we can formulate an associated closedness type condition ensuring strong duality between* (P^C) *and* $(D^{C_{FL}})$. *For this purpose, we have to ensure that* $\mathrm{Pr}_{X_n^*\times\mathbb{R}}(\mathrm{epi}\,\Phi^*)$ *is closed in the topology* $w(X_n^*,X_n)\times\mathbb{R}$ *(see [2, 4]). It is an easy exercise to observe that*

$$\mathrm{Pr}_{X_n^*\times\mathbb{R}}(\mathrm{epi}\,\Phi^*)=\bigcup_{y^{(n+1)*}\in Q^*}\mathrm{epi}(y^{(n+1)*}g)^*_S+$$
$$\bigcup_{\substack{y^{i*}\in X_i^*,\\ i=0,\dots,n-1}}\left(\mathrm{epi}(y^{(n-1)*}F^n)^*+\left(0_{X_n^*},f^*(y^{0*})+\sum_{i=0}^{n-2}(y^{i*}F^{i+1})^*(y^{(i+1)^*})\right)\right).$$

Recall that when $\mathrm{Pr}_{X_n^*\times\mathbb{R}}(\mathrm{epi}\,\Phi^*)$ *is closed, there is actually strong duality for* $(P^C_{x^*})$ *and its corresponding Fenchel-Lgrange type dual problem for all* $x^*\in X_n^*$, *i.e. stable strong duality for* (P^C) *and* $(D^{C_{FL}})$.

From Theorem 2.1 one can also derive a formula for the conjugate function of a multi-composed function and a characterization via epigraph inclusions for it.

Corollary 2.1. *It holds*

$$(f\circ F^1\circ\ldots\circ F^n)^*(x^*)=\min_{\substack{z^{n*}\in Q^*,\ z^{i*}\in K_i^*,\\ i=0,\dots,n-1}}\left\{f^*(z^{0*})+\sum_{i=1}^{n-1}(z^{(i-1)*}F^i)^*(z^{i*})+(z^{(n-1)*}F^n)^*(x^*)\right\}\quad(7)$$

for all $x^*\in X_n^*$ *if and only if*

$$M_0=\mathcal{T}^n_{X_0^*}(\mathrm{epi}(f^*))+\bigcup_{\substack{z^{n*}\in Q^*,\ z^{i*}\in K_i^*,\\ i=0,\dots,n-1}}\left(\sum_{i=1}^{n-1}\widetilde{\mathcal{T}}^n_{X_i^*}\left(\mathrm{epi}((z^{(i-1)*}F^i)^*)\times\{-z^{(i-1)*}\}\right)\right.$$
$$\left.+\widetilde{\mathcal{T}}^n_{X_n^*}\left(\mathrm{epi}((z^{(n-1)*}F^n)^*)\times\{-z^{(n-1)*}\}\right)\right)$$

is closed regarding $\mathcal{U}$.

Remark 2.4. *The regularity condition* M_0 *closed regarding* $\mathcal{U}$ *is equivalent to the formula* (7), *thus a natural sufficient condition in order to guarantee* (7) *is to ask* M_0 *to be closed. The formula* (7) *can be found also in [20], but under a regularity condition of interiority type that is stronger than the ones just mentioned.*

3 ε-subdifferential formulae and ε-optimality conditions

In this section we give a formula for the ε-subdifferential of the function $f \circ F^1 \circ \ldots \circ F^n + \delta_{\mathcal{A}}$, to the best of our knowledge the first one in the literature for such a multi-composed function, that is subsequently employed for deriving necessary and sufficient ε-optimality conditions for characterizing the ε-optimal solutions to the problem (P^C), where $\varepsilon \geq 0$. As a special case, a formula for the ε-subdifferential of the multi-composed function $f \circ F^1 \circ \ldots \circ F^n$ is derived, too. Moreover, we briefly discuss how can one obtain different duality and optimality statements concerning composed optimization problems from the literature (see for example [2–4, 9]) as special cases of our approach.

Theorem 3.1. *The regularity condition (RC_L) is fulfilled if and only if for all $x \in X^n$ and for all $\varepsilon \geq 0$ it holds*

$$
\begin{aligned}
&\partial_\varepsilon((f \circ F^1 \circ \ldots \circ F^n) + \delta_{\mathcal{A}})(x) = \\
&\bigcup_{\substack{\sum_{i=0}^{n+1} \varepsilon_i = \varepsilon,\ \varepsilon_i \geq 0, \\ i=0,\ldots,n+1}} \Big\{ \partial_{\varepsilon_n}((z^{(n-1)*}F^n) + (z^{n*}g) + \delta_S)(x) : z^{0*} \in K_0^* \cap \partial_{\varepsilon_0} f((F^1(\ldots F^n(x)))), \\
&\qquad z^{i*} \in K_i^* \cap \partial_{\varepsilon_i}(z^{(i-1)*}F^i)(F^{i+1}(\ldots F^n(x))),\ i = 1,\ldots,n-1, \\
&\qquad z^{n*} \in Q^* \text{ and } 0 \leq \langle z^{n*}, g(x)\rangle + \varepsilon_{n+1} \Big\}.
\end{aligned}
$$

Proof. "$\Rightarrow$": If $x \notin S \cap g^{-1}(-Q) \cap ((F^n)^{-1}(\ldots(F^1)^{-1}))(\operatorname{dom} f) \cap \operatorname{dom} F^n$ then both sides of the equality we have to prove are empty sets, so we take further an arbitrary $x \in S \cap g^{-1}(Q) \cap ((F^n)^{-1}(\ldots(F^1)^{-1})(\operatorname{dom} f)) \cap \operatorname{dom} F^n$ and an $\varepsilon \geq 0$.
"$\subseteq$": For $x^* \in \partial_\varepsilon(f \circ F^1 \circ \ldots \circ F^n + \delta_{\mathcal{A}})(x)$ by (1) it holds

$$
((f \circ F^1 \circ \ldots \circ F^n) + \delta_{\mathcal{A}})^*(x^*) + (f \circ F^1 \circ \ldots \circ F^n)(x) + \delta_{\mathcal{A}}(x) \leq \langle x^*, x\rangle + \varepsilon. \tag{8}
$$

Following Theorem 2.1, (RC_L) implies the existence of $\overline{z}^{n*} \in Q^*$ and $\overline{z}^{i*} \in K_i^*,\ i = 0,\ldots,n-1$, such that

$$
\begin{aligned}
&f^*(\overline{z}^{0*}) + \sum_{i=1}^{n-1}(\overline{z}^{(i-1)*}F^i)^*(\overline{z}^{i*}) + ((\overline{z}^{(n-1)*}F^n) + (\overline{z}^{n*}g))_S^*(x^*) \\
&+(f \circ F^1 \circ \ldots \circ F^n)(x) + \langle \overline{z}^{n*}, g(x)\rangle(x) - \langle \overline{z}^{n*}, g(x)\rangle \leq \langle x^*, x\rangle + \varepsilon.
\end{aligned} \tag{9}
$$

Further, the inequality in (9) can be written as

$$
\begin{aligned}
&[f^*(\overline{z}^{0*}) + f((F^1 \circ \ldots \circ F^n)(x)) - \langle \overline{z}^{0*}, (F^1 \circ \ldots \circ F^n)(x)\rangle] + \\
&\sum_{i=1}^{n-1}[(\overline{z}^{(i-1)*}F^i)^*(\overline{z}^{i*}) + (\overline{z}^{(i-1)*}F^i)((F^{i+1} \circ \ldots \circ F^n)(x)) - \langle \overline{z}^{i*}, (F^{i+1} \circ \ldots \circ F^n)(x)\rangle] \\
&+[((\overline{z}^{(n-1)*}F^n) + (\overline{z}^{n*}g))_S^*(x^*) + (\overline{z}^{(n-1)*}F^n)(x) + (\overline{z}^{n*}g)(x) - \langle x^*, x\rangle] \\
&-\langle \overline{z}^{n*}, g(x)\rangle \leq \varepsilon.
\end{aligned} \tag{10}
$$

Now, we define $\widetilde{\varepsilon}_0 := f^*(\overline{z}^{0*}) + f((F^1 \circ \ldots \circ F^n)(x)) - \langle \overline{z}^{0*}, (F^1 \circ \ldots \circ F^n)(x)\rangle$, $\overline{\varepsilon}_i := (\overline{z}^{(i-1)*}F^i)^*(\overline{z}^{i*}) + (\overline{z}^{(i-1)*}F^i)((F^{i+1} \circ \ldots \circ F^n)(x)) - \langle \overline{z}^{i*}, (F^{i+1} \circ \ldots \circ F^n)(x)\rangle$, $i = 1,\ldots,n-1$, $\overline{\varepsilon}_n := ((\overline{z}^{(n-1)*}F^n) +$

$(\overline{z}^{n*}g))^*_S(x^*) + (\overline{z}^{(n-1)*}F^n)(x) + (\overline{z}^{n*}g)(x) - \langle x^*, x\rangle$ and $\varepsilon_{n+1} := -\langle \overline{z}^{n*}, g(x)\rangle$. By the Young-Fenchel inequality it is clear that $\widetilde{\varepsilon}_0 \geq 0$ and $\overline{\varepsilon}_i \geq 0$, $i = 1, \dots, n$ and as $\overline{z}^{n*} \in Q^*$ and $g(x) \in -Q$ it follows that $\varepsilon_{n+1} \geq 0$. Moreover, (10) yields $\widetilde{\varepsilon}_0 + \sum_{i=1}^{n+1} \overline{\varepsilon}_i \leq \varepsilon$. Setting $\overline{\varepsilon}_0 := \varepsilon - \sum_{i=1}^{n+1} \overline{\varepsilon}_i > \widetilde{\varepsilon}_0$, it holds $\overline{z}^{0*} \in \partial_{\overline{\varepsilon}_0} f(F^1(\dots F^n(x)))$, $\overline{z}^{i*} \in \partial_{\overline{\varepsilon}_i}(\overline{z}^{(i-1)*}F^i)(F^{i+1}(\dots F^n(x)))$, $i = 1, \dots, n-1$ and $x^* \in \partial_{\overline{\varepsilon}_n}((\overline{z}^{(n-1)*}F^n) + (\overline{z}^{n*}g) + \delta_S)(x)$. Therefore, we have

$$\begin{aligned}
&x^* \in \partial_{\overline{\varepsilon}_n}((\overline{z}^{(n-1)*}F^n) + (\overline{z}^{n*}g) + \delta_S)(x) \\
&\subseteq \bigcup_{\substack{\sum_{i=0}^{n+1} \varepsilon_i = \varepsilon,\ \varepsilon_i \geq 0, \\ i=0,\dots,n+1}} \Big\{ \partial_{\varepsilon_n}((\overline{z}^{(n-1)*}F^n) + (\overline{z}^{n*}g) + \delta_S)(x) : \overline{z}^{0*} \in K_0^* \cap \partial_{\varepsilon_0} f(F^1(\dots F^n(x))), \\
&\quad \overline{z}^{i*} \in K_i^* \cap \partial_{\varepsilon_i}(\overline{z}^{(i-1)*}F^i)(F^{i+1}(F^{i+2}(\dots F^n(x)))),\ i = 1, \dots, n-1,\ \overline{z}^{n*} \in Q^* \\
&\quad \text{and } 0 \leq \langle \overline{z}^{n*}, g(x)\rangle + \varepsilon_{n+1} \Big\}.
\end{aligned}$$

"$\supseteq$": Let us take an arbitrary

$$\begin{aligned}
x^* \in &\bigcup_{\substack{\sum_{i=0}^{n+1} \varepsilon_i = \varepsilon,\ \varepsilon_i \geq 0, \\ i=0,\dots,n+1}} \Big\{ \partial_{\varepsilon_n}((z^{(n-1)*}F^n) + (z^{n*}g) + \delta_S)(x) : z^{0*} \in K_0^* \cap \partial_{\varepsilon_0} f(F^1(\dots F^n(x))), \\
&\quad z^{i*} \in K_i^* \cap \partial_{\varepsilon_i}(z^{(i-1)*}F^i)(F^{i+1}(F^{i+2}(\dots F^n(x)))),\ i = 1, \dots, n-1, \\
&\quad z^{n*} \in Q^* \text{ and } 0 \leq \langle z^{n*}, g(x)\rangle + \varepsilon_{n+1} \Big\}.
\end{aligned}$$

Therefore, there exist $\varepsilon_i \geq 0$, $i = 0, \dots, n+1$, $x^* \in \partial_{\varepsilon_n}((z^{(n-1)*}F^n) + (z^{n*}g) + \delta_S)(x)$, $z^{0*} \in K_0^* \cap \partial_{\varepsilon_0}(f(F^1 \circ \dots \circ F^n)(x))$, $z^{i*} \in K_i^* \cap \partial_{\varepsilon_i}(z^{(i-1)*}F^i)(F^{i+1}(\dots F^n(x)))$, $i = 1, \dots, n-1$, and $z^{n*} \in Q^*$ such that $\sum_{i=0}^{n+1} \varepsilon_i = \varepsilon$, $f^*(z^{0*}) + f((F^1 \circ \dots \circ F^n)(x)) \leq \langle z^{0*}, (F^1 \circ \dots \circ F^n)(x)\rangle + \varepsilon_0$, $(z^{(i-1)*}F^i)^*(z^{i*}) + (z^{(i-1)*}F^i)((F^{i+1} \circ \dots \circ F^n)(x)) \leq \langle z^{i*}, (F^{i+1} \circ \dots \circ F^n)(x)\rangle + \varepsilon_i$, $i = 1, \dots, n-1$, $((z^{(n-1)*}F^n) + (z^{n*}g))^*_S(x^*) + (z^{(n-1)*}F^n)(x) + (z^{n*}g)(x) \leq \langle x^*, x\rangle + \varepsilon_n$ and $0 \leq \langle z^{n*}, g(x)\rangle + \varepsilon_{n+1}$. By taking the sum we obtain

$$\begin{aligned}
&f^*(z^{0*}) + f((F^1 \circ \dots \circ F^n)(x)) + \sum_{i=1}^{n-1} [(z^{(i-1)*}F^i)^*(z^{i*}) + (z^{(i-1)*}F^i)((F^{i+1} \circ \dots \circ F^n)(x))] + \\
&((z^{(n-1)*}F^n) + (z^{n*}g))^*_S(x^*) + (z^{(n-1)*}F^n)(x) + (z^{n*}g)(x) - \langle z^{n*}, g(x)\rangle \\
&\leq \langle z^{0*}, (F^1 \circ \dots \circ F^n)(x)\rangle + \varepsilon_0 + \sum_{i=1}^{n-1} [\langle z^{i*}, (F^{i+1} \circ \dots \circ F^n)(x)\rangle + \varepsilon_i] + \langle x^*, x\rangle + \varepsilon_n + \varepsilon_{n+1} \\
&= \sum_{i=0}^{n-1} \langle z^{i*}, (F^{i+1} \circ \dots \circ F^n)(x)\rangle + \sum_{i=0}^{n+1} \varepsilon_i + \langle x^*, x\rangle,
\end{aligned}$$

which is equivalent to

$$f^*(z^{0*}) + f((F^1 \circ \dots \circ F^n)(x)) + \sum_{i=1}^{n-1} (z^{(i-1)*}F^i)^*(z^{i*}) + ((z^{(n-1)*}F^n) + (z^{n*}g))^*_S(x^*) \leq \langle x^*, x\rangle + \varepsilon.$$

By using (3) we get

$$(f \circ F^1 \circ \ldots \circ F^n)(x) + ((f \circ F^1 \circ \ldots \circ F^n) + \delta_{\mathcal{A}})^*(x^*) \leq \langle x^*, x\rangle + \varepsilon$$

i.e. $x^* \in \partial_\varepsilon((f \circ F^1 \circ \ldots \circ F^n) + \delta_{\mathcal{A}})(x)$.
"$\Leftarrow$": For the trivial case $(f \circ F^1 \circ ... \circ F^n)^*_{\mathcal{A}} = +\infty$ the statement is obviously fulfilled. Let us now assume that $x \in \mathcal{A} \cap \operatorname{dom} f$ an denote

$$\varepsilon := (f \circ F^1 \circ ... \circ F^n)^*_{\mathcal{A}}(x^*) + (f \circ F^1 \circ ... \circ F^n)(x) - \langle x^*, x\rangle \geq 0, \tag{11}$$

which in turn implies that $x^* \in \partial_\varepsilon(f \circ F^1 \circ ... \circ F^n + \delta_{\mathcal{A}})(x)$ and so, there exist $\varepsilon_i \geq 0$, $i = 0, ..., n+1$, $z^{i*} \in K_i^*$, $i = 0, ..., n-1$, and $z^{n*} \in Q^*$ such that

$$\begin{aligned}
&(z^{(n-1)*}F^n + z^{n*}g + \delta_S)^*(x^*) + (z^{(n-1)*}F^n + z^{n*}g)(x) \leq \langle x^*, x\rangle + \varepsilon_n,\\
&(z^{(i-1)*}F^i)^*(z^{i*}) + (z^{(i-1)*}F^i)(F^{i+1}(...(F^n(x)))) \leq \langle z^{i*}, F^{i+1}(...(F^n(x)))\rangle + \varepsilon_i,\ \ i = 1, ..., n-1,\\
&0 \leq (z^{n*}g)(x) + \varepsilon_{n+1},\\
&f(F^1(...(F^n(x)))) + f^*(z^{0*}) \leq \langle z^{0*}, F^1(F^2(...(F^n(x))))\rangle + \varepsilon_0.
\end{aligned}$$

Summing up these inequalities leads to

$$\begin{aligned}
&(z^{(n-1)*}F^n + z^{n*}g + \delta_S)^*(x^*) + (z^{(n-1)*}F^n + z^{n*}g)(x) + \sum_{i=1}^{n-1}(z^{(i-1)*}F^i)^*(z^{i*})\\
&+\sum_{i=1}^{n-1}(z^{(i-1)*}F^i)(F^{i+1}(...(F^n(x)))) + f(F^1(...(F^n)))(x) + f^*(z^{0*})\\
&\leq \langle x^*, x\rangle + \sum_{i=0}^{n+1}\varepsilon_i + \sum_{i=1}^{n}(z^{(i-1)*}F^i)(F^{i+1}(...(F^n(x)))) + (z^{n*}g)(x)
\end{aligned}$$

and by using $\varepsilon = \sum_{i=0}^{n+1}\varepsilon_i$ and (11) this is equivalent to

$$\begin{aligned}
&(z^{(n-1)*}F^n + z^{n*}g + \delta_S)^*(x^*) + \sum_{i=1}^{n-1}(z^{(i-1)*}F^i)^*(z^{i*}) + f(F^1(...(F^n)))(x) + f^*(z^{0*})\\
&\leq \langle x^*, x\rangle + (f \circ F^1 \circ ... \circ F^n)^*_{\mathcal{A}}(x^*) + (f \circ F^1 \circ ... \circ F^n)(x) - \langle x^*, x\rangle\\
\Leftrightarrow\ \ &(z^{(n-1)*}F^n + z^{n*}g + \delta_S)^*(x^*) + \sum_{i=1}^{n-1}(z^{(i-1)*}F^i)^*(z^{i*}) + f^*(z^{0*}) \leq (f \circ F^1 \circ ... \circ F^n)^*_{\mathcal{A}}(x^*).
\end{aligned}$$

Finally, Theorem 2.1 provides the desired statement. □

Remark 3.1. *Note that no regularity condition is needed for proving the inclusion "$\supseteq$" in the Theorem 3.1.*

When $\varepsilon = 0$, $S = X_n$ and g is identical zero, the previous statement delivers the formula for the subdifferential of a multi-composed function.

Corollary 3.1. *Let M_0 be closed regarding $\mathcal{U}$. Then for all $x \in X^n$ it holds*

$$\partial(f \circ F^1 \circ \ldots \circ F^n)(x) = \bigcup_{\substack{z^{0*} \in K_0^* \cap \partial f((F^1(...F^n(x)))),\\ z^{i*} \in K_i^* \cap \partial(z^{(i-1)*}F^i)(F^{i+1}(...F^n(x))),\\ i=1,...,n-1}} \partial(z^{(n-1)*}F^n)(x).$$

We employ now the result of Theorem 3.1 for giving necessary and sufficient ε-optimality conditions for (P^C). Recall that $\overline{x} \in \mathcal{A}$ is an ε-optimal solution of problem (P^C) if

$$(f \circ F^1 \circ \ldots \circ F^n)(\overline{x}) \leq \inf_{x \in \mathcal{A}} \{(f \circ F^1 \circ \ldots \circ F^n)(x)\} + \varepsilon,$$

which happens if and only if $0_{X_n^*} \in \partial_\varepsilon(f \circ F^1 \circ \ldots \circ F_n + \delta_{\mathcal{A}})(\overline{x})$.

Theorem 3.2. *(a) Assume that the regularity condition (RC_L) is fulfilled and let $\varepsilon \geq 0$. If $\overline{x} \in \mathcal{A}$ is an ε-optimal solution to (P^C), then there exist $\varepsilon_i \geq 0$, $i = 0, \ldots, n+1$, $\overline{z}^{i*} \in K_i^*$, $i = 0, \ldots, n-1$, and $\overline{z}^{n*} \in Q^*$ such that $(\overline{z}^{0*}, \ldots, \overline{z}^{n*})$ is an ε-optimal solution to (D^{C_L}) fulfilling*

(i) $0 \leq f(F^1(\ldots F^n(\overline{x}))) + f^*(\overline{z}^{0*}) - \langle \overline{z}^{0*}, F^1(\ldots F^n(\overline{x}))\rangle \leq \varepsilon_0$,

(ii) $0 \leq (\overline{z}^{(i-1)*} F^i)(F^{i+1}(\ldots F^n(\overline{x}))) + (\overline{z}^{(i-1)*} F^i)^*(\overline{z}^{i*}) - \langle \overline{z}^{i*}, F^{i+1}(\ldots F^n(\overline{x}))\rangle \leq \varepsilon_i$
$\forall\, i = 1, \ldots, n-1$,

(iii) $0 \leq (\overline{z}^{(n-1)*} F^n)(\overline{x}) + (\overline{z}^{n*} g)(\overline{x}) + ((\overline{z}^{(n-1)*} F^n) + (\overline{z}^{n*} g))_S^*(0_{X_n^*}) \leq \varepsilon_n$,

(iv) $0 \leq -\langle \overline{z}^{n*}, g(\overline{x})\rangle \leq \varepsilon_{n+1}$,

(v) $\sum\limits_{i=0}^{n+1} \varepsilon_i = \varepsilon$.

(b) If there exist $\varepsilon_i \geq 0$, $i = 0, \ldots, n+1$, $\overline{z}^{i*} \in K_i^*$, $i = 0, \ldots, n-1$ and $\overline{z}^{n*} \in Q^*$ such that (i)-(v) are fulfilled for some $\overline{x} \in \mathcal{A}$, then $\overline{x}$ is an ε-optimal solution to (P^C), $(\overline{z}^{0*}, \ldots, \overline{z}^{n*})$ an ε-optimal solution to (D^{C_L}) and $v(P^C) \leq v(D^{C_L}) + \varepsilon$.

Proof. Since $\overline{x}$ is an ε-optimal solution of the problem (P^C), it holds $0_{X_n^*} \in \partial_\varepsilon(f \circ F^1 \circ \ldots \circ F^n + \delta_{\mathcal{A}})(\overline{x})$ and by Theorem 3.1 that there exist $\varepsilon_i \geq 0$, $i = 0, \ldots, n+1$, $\overline{z}^{i*} \in K_i^*$, $i = 0, \ldots, n-1$ and $\overline{z}^{n*} \in Q^*$ such that $\overline{z}^{0*} \in K_i^* \cap \partial_{\varepsilon_0} f(F^1 \ldots F^n(\overline{x}))$, $\overline{z}^{i*} \in K_i^* \cap \partial_{\varepsilon_i}(\overline{z}^{(i-1)*} F^i)(F^{i+1} \ldots F^n(\overline{x}))$, $i = 1, \ldots, n-1$, $0_{X_n^*} \in \partial_{\varepsilon_n}((\overline{z}^{(n-1)*} F^n) + (\overline{z}^{n*} g) + \delta_S)(\overline{x})$, $0 \leq \langle \overline{z}^{n*}, g(\overline{x})\rangle + \varepsilon_{n+1}$ and $\sum_{i=0}^{n+1} \varepsilon_i = \varepsilon$. Using (1) one obtains (i)-(iii).
(b) The sum of relations (i)-(iv) yields

$$(f \circ F^1 \circ \ldots \circ F^n)(\overline{x}) + f^*(\overline{z}^{0*}) - \langle \overline{z}^{0*}, (F^1 \circ \ldots \circ F^n)(\overline{x})\rangle +$$

$$\sum_{i=1}^{n-1} [(\overline{z}^{(i-1)*} F^i)((F^{i+1} \circ \ldots \circ F^n)(\overline{x})) + (\overline{z}^{(i-1)*} F^i)^*(\overline{z}^{i*}) - \langle \overline{z}^{i*}, (F^{i+1} \circ \ldots \circ F^n)(\overline{x})\rangle] +$$

$$(\overline{z}^{(n-1)*} F^n)(\overline{x}) + (\overline{z}^{n*} g)(\overline{x}) + ((\overline{z}^{(n-1)*} F^n) + (\overline{z}^{n*} g))_S^*(0_{X_n^*}) - \langle \overline{z}^{n*}, g(\overline{x})\rangle \leq \sum_{i=0}^{n+1} \varepsilon_i \Leftrightarrow$$

$$(f \circ F^1 \circ \ldots \circ F^n)(\overline{x}) + f^*(\overline{z}^{0*}) + \sum_{i=1}^{n-1} (\overline{z}^{(i-1)*} F^i)^*(\overline{z}^{i*}) + ((z^{(n-1)*} F^n) + (\overline{z}^{n*} g))_S^*(0_{X_n^*}) \leq \sum_{i=0}^{n+1} \varepsilon_i.$$

Employing relation (v), the last inequality yields the conclusion. □
When $\varepsilon = 0$, the previous statement delivers the following necessary and sufficient optimality conditions for characterizing the optimal solution to (P^C), providing thus weaker hypotheses for the similar statement [20, Theorem 4.2].

Corollary 3.2. *(a) Assume that the regularity condition (RC_L) is fulfilled. If $\overline{x} \in \mathcal{A}$ is an optimal solution to (P^C), then there exist $\overline{z}^{i*} \in K_i^*$, $i = 0, \ldots, n-1$ and $\overline{z}^{n*} \in Q^*$ such that $(\overline{z}^{0*}, \ldots, \overline{z}^{n*})$ is an optimal solution to (D^{C_L}) fulfilling*

(i) $f(F^1(\ldots F^n(\overline{x}))) + f^*(\overline{z}^{0*}) = \langle \overline{z}^{0*}, F^1(\ldots F^n(\overline{x})\rangle$,

(ii) $(\overline{z}^{(i-1)*} F^i)(F^{i+1}(\ldots F^n(\overline{x}))) + (\overline{z}^{(i-1)*} F^i)^*(\overline{z}^{i*}) - \langle \overline{z}^{i*}, F^{i+1}(\ldots F^n(\overline{x}))\rangle = 0 \ \forall i = 1, \ldots, n-1$,

(iii) $(\overline{z}^{(n-1)*} F^n)(\overline{x}) + (\overline{z}^{n*} g)(\overline{x}) + ((\overline{z}^{(n-1)*} F^n) + (\overline{z}^{n*} g))^*_S(0_{X_n^*}) = 0$,

(iv) $\langle \overline{z}^{n*}, g(\overline{x})\rangle = 0$.

(b) If there exist $\overline{z}^{i*} \in K_i^*$, $i = 0, \ldots, n-1$ and $\overline{z}^{n*} \in Q^*$ such that (i)-(iv) are fulfilled for some $\overline{x} \in \mathcal{A}$, then $\overline{x}$ is an optimal solution to (P^C), $(\overline{z}^{0*}, \ldots, \overline{z}^{n*})$ one to (D^{C_L}) and there is strong duality for the primal-dual pair of problems $(P^C) - (D^{C_L})$.

Remark 3.2. *The classical composed optimization problem*

$$(P) \quad \inf_{x \in X} \{V(x) + (G \circ H)(x)\},$$

where Y is a Hausdorff locally convex space partially ordered by the convex cone C, $V : X \to \overline{\mathbb{R}}$ is a proper and convex function, $G : \overline{Y} \to \overline{\mathbb{R}}$ is a proper, convex and C-increasing function on $\operatorname{dom} G$ *and* $H(\operatorname{dom} H) \subseteq \operatorname{dom} G$ *and $H : X \to \overline{Y}$ is a proper and C-convex function, can be obtained as a special case of (P^C) by taking $X_0 = \mathbb{R} \times Y$ partially ordered by $K_0 = \mathbb{R}_+ \times C$, $X_1 = X$,*

$$f : \overline{\mathbb{R} \times Y} \to \overline{\mathbb{R}}, \ f(y_1^0, y_2^0) := y_1^0 + G(y_2^0) \ \textit{with} \ (y_1^0, y_2^0) \in \mathbb{R} \times Y$$

and

$$F^1 : X \to \overline{\mathbb{R} \times Y}, \ F^1(y_1^1, y_2^1) := \begin{cases} (V(x), H(x)), & \text{if } x \in \operatorname{dom} V \cap \operatorname{dom} H, \\ +\infty_{\mathbb{R}_+ \times C}, & \text{otherwise.} \end{cases}$$

Moreover, different results on stable strong Lagrange duality, ε-subdifferential and ε-optimality conditions involving (P^C) from [3, 9, 15, 16] can be recovered then as special cases of our statements. This shows that the introduced concept of multi-composed optimization problems combines several approaches to give formulae for the characterization of strong and total Lagrange duality. Moreover, since the formulae for the conjugate functions of $(V + (G \circ H))$ can also be received by using the perturbation theory (see [2, 4]), the introduced concept can also be interpreted as an umbrella for several perturbations.

4 Applications

In this section we discuss two possible directions where our main results can be applied, fractional programming and entropy optimization.

4.1 ε-optimality conditions for convex fractional problems

Consider the following convex fractional optimization problem

$$
\begin{aligned}
(P^F) \qquad & \inf_{x\in\mathcal{A}} \left\{ \Phi\left(c_1 \frac{[h_1(x)]^2}{l_1(x)}, \ldots, c_n \frac{[h_n(x)]^2}{l_n(x)} \right) \right\}, \\
& \mathcal{A} = \{x \in S : g(x) \in -Q\},
\end{aligned}
$$

where c_i are positive numbers for $i = 1, \ldots, n$. In order to deal with the problem (P^F) by means of duality, let us in the following assume that $X_0 = \mathbb{R}^n$ is partially ordered by $K_0 = \mathbb{R}^n_+$, $X_1 = \mathbb{R}^n \times \mathbb{R}^n$ is partially ordered by $K_1 = \mathbb{R}^n_+ \times \mathbb{R}^n_+$, $X_2 = X$ and Z is partially ordered by the convex cone Q, where X and Z are locally convex Hausdorff spaces. In addition we assume that S is a non-empty, closed and convex subset of X, $g : X \to \overline{Z}$ is a proper, Q-convex and Q-epi closed function and $\Phi : \overline{\mathbb{R}^n} \to \overline{\mathbb{R}}$ is a proper, convex, $\mathbb{R}^n_+$-increasing on $\mathbb{R}^n_+$ and lower semicontinuous function. Further, let $h_i : X \to \overline{\mathbb{R}}$ be a proper, convex and lower semicontinuous function fulfilling $h_i(x) \geq 0$ for all $x \in X$ and $l_i : X \to \mathbb{R} \cup \{+\infty\}$ be a proper, concave and upper semicontinuous function fulfilling $l_i(x) > 0$ for all $x \in X$, $i = 1, \ldots, n$. As the function Φ is defined in a general way, this fractional problem covers a broad class of optimization problems, with applications in many areas such as finance, economics and engineering. Examples for the function Φ are $\Phi(y^0) = \max\{y^0_1, \ldots, y^0_n\}$, $\Phi(y^0) = \sum_{i=1}^n y^0_i$ or $\Phi(y^0) = \|y^0 - a\|$, where $\|\cdot\|$ is the Euclidean norm and $a \in \mathbb{R}^n$ a suitably chosen point (see, for instance, [18, 19]).
To write the problem (P^F) as a special case of the problem (P^C) we introduce the following functions:

- $f : \overline{\mathbb{R}^n} \to \overline{\mathbb{R}}$ defined by $f(y^0) := \Phi(y^0),\ y^0 = (y^0_1, \ldots, y^0_n)^T \in \mathbb{R}^n$,
- $F^1 : \overline{\mathbb{R}^n \times \mathbb{R}^n} \to \overline{\mathbb{R}^n}$ defined by
$$F^1(y^1, \widetilde{y}^1) := \begin{cases} \left(-c_1 \frac{[y^1_1]^2}{\widetilde{y}^1_1}, \ldots, -c_n \frac{[y^1_n]^2}{\widetilde{y}^1_n}\right)^T, & \text{if } y^1_i \geq 0,\ \widetilde{y}^1_i < 0\ \forall i = 1, \ldots, n, \\ +\infty_{\mathbb{R}^n_+}, & \text{otherwise,} \end{cases}$$

and

- $F^2 : X \to \overline{\mathbb{R}^n \times \mathbb{R}^n}$ defined by
$$F^2(x) = \begin{cases} (h_1(x), \ldots, h_n(x), -l_1(x), \ldots, -l_n(x)), & \text{if } x \in \bigcap\limits_{i=1}^n (\operatorname{dom} h_i \cap \operatorname{dom}(-l_i)), \\ +\infty_{\mathbb{R}^n_+ \times \mathbb{R}^n_+}, & \text{otherwise.} \end{cases}$$

The problem (P^F) can be written as

$$
\begin{aligned}
(P^F) \qquad & \inf_{x\in\mathcal{A}} \{(f \circ F^1 \circ F^2)(x)\}, \\
& \mathcal{A} = \{x \in S : g(x) \in -Q\}.
\end{aligned}
$$

It is worth noting that the functions f, F^1 and F^2 fulfill the conditions considered in the previous sections, namely f is proper, convex, lower semicontinuous and $\mathbb{R}^n_+$-increasing on $\mathbb{R}^n_+$, and $F^1(\operatorname{dom} F^1) = \mathbb{R}^n_+$, F^1 is proper, $\mathbb{R}^n_+$-convex and $(\mathbb{R}^n_+ \times \mathbb{R}^n_+, \mathbb{R}^n_+)$-increasing on $\operatorname{dom} F^1 = \mathbb{R}^n_+ \times \operatorname{int}(-\mathbb{R}^n_+) = F^2(\operatorname{dom} F^2)$ and $\mathbb{R}^n_+$-epi closed. Note that the convexity of F^1 follows as $-[y^1_i]^2/\widetilde{y}^1_i$ is convex for all $y^1_i \geq 0$ and $y^1_i < 0$ and moreover, as $-[y^1_i]^2/\widetilde{y}^1_i$ is $\mathbb{R}^2_+$-increasing for all $y^1_i \geq 0$ and $y^1_i < 0$, $i = 1, \ldots, n$, we can guarantee that F^1 is $(\mathbb{R}^n_+ \times \mathbb{R}^n_+, \mathbb{R}^n_+)$-increasing on $\mathbb{R}^n_+ \times \operatorname{int}(-\mathbb{R}^n_+)$. The $\mathbb{R}^n_+$-epi closedness of F^1 follows by the continuity of $-[y^1_i]^2/\widetilde{y}^1_i$ for all $y^1_i \geq 0$

and $y_i^1 < 0$, $i = 1, \ldots, n$. Further, it is not hard to see that F^2 is proper, $\mathbb{R}^n_+ \times \mathbb{R}^n_+$-convex and $\mathbb{R}^n_+ \times \mathbb{R}^n_+$-epi closed.

In the next step we want to determine the conjugate functions of $(z^{0*}F^1)$ and $(((z^{1*}, \widetilde{z}^{1*})F^2) + (z^{2*}g))^*_S$, where $z^{0*} \in \mathbb{R}^n_+$, $(z^{1*}, \widetilde{z}^{1*}) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+$ and $z^{2*} \in Q^*$. The one of $(z^{0*}F^1)$ is (cf. [11])

$$(z^{0*}F^1)^*(z^{1*}, \widetilde{z}^{1*}) = \begin{cases} 0, \text{ if } \widetilde{z}_i^{1*} - \frac{[z_i^{1*}]^2}{4z_i^{0*}c_i} \geq 0, \; i = 1, \ldots, n, \\ +\infty, \text{ otherwise.} \end{cases}$$

and for the last one we obtain

$$\begin{aligned} (((z^{1*}, \widetilde{z}^{1*})F^2) + (z^{2*}g))^*_S(x^*) &= \sup_{x \in S} \left\{ \langle x^*, x \rangle - \sum_{i=1}^n z_i^{1*} h_i(x) + \sum_{i=1}^n \widetilde{z}_i^{1*} l_i(x) - (z^{2*}g)(x) \right\} \\ &= \left(\sum_{i=1}^n (z_i^{1*} h_i - \widetilde{z}_i^{1*} l_i) + (z^{2*}g) \right)^* (x^*). \end{aligned}$$

To give a formula for the closedness type regularity condition one also needs the epigraph of $(z^{0*}F^1)^*$ for $z^{0*} \in \mathbb{R}^n_+$, that can be expressed as $\operatorname{epi}((z^{0*}F^1)^*) = N \times \mathbb{R}_+$, where

$$N := \left\{ (z^{1*}, \widetilde{z}^{1*}) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ : \widetilde{z}_i^{1*} - \frac{[z_i^{1*}]^2}{4z_i^{0*}c_i} \geq 0, \; i = 1, \ldots, n \right\}.$$

The corresponding closedness type regularity condition for the problem (P^F) is then

$$(RC_F) \quad \left| \begin{array}{l} \mathcal{T}^2_{X_0^*}(\operatorname{epi}(\Phi^*)) + \bigcup\limits_{\substack{z^{0*} \in \mathbb{R}^n_+, \; (z^{1*}, \widetilde{z}^{1*})^T \in \mathbb{R}^{n \times n}_+ \\ z^{2*} \in Q^*}} \left(\widetilde{\mathcal{T}}^2_{X_1^*}(N \times \mathbb{R}_+ \times \{-z^{0*}\}) \right. \\ \left. + \widetilde{\mathcal{T}}^2_{X_2^*} \left(\operatorname{epi} \left(\left(\sum\limits_{i=1}^n (z_i^{1*} h_i - \widetilde{z}_i^{1*} l_i) + (z^{2*}g) \right)^* \right) \times \{-(z^{1*}, \widetilde{z}^{1*})\} \right) \right) \\ \text{is closed regarding the subspace } \{0_{X_0^*}\} \times \{0_{X_1^*}\} \times X^* \times \mathbb{R}. \end{array} \right.$$

Theorem 2.1 implies the following stable strong duality statement for problem (P^F).

Theorem 4.1. *The regularity condition (RC_F) is fulfilled if and only if*

$$\begin{aligned} &\sup_{x \in \mathcal{A}} \left\{ \langle x^*, x \rangle - \Phi \left(c_1 \frac{[h_1(x)]^2}{l_1(x)}, \ldots, c_n \frac{[h_n(x)]^2}{l_n(x)} \right) \right\} \\ = &\min_{(z^{0*}, z^{1*}, \widetilde{z}^1, z^{2*}) \in \mathcal{B}} \left\{ \Phi^*(z^{0*}) + \left(\sum_{i=1}^n (z_i^{1*} h_i - \widetilde{z}_i^{1*} l_i) + (z^{2*}g) \right)^* (x^*) \right\} \end{aligned}$$

for all $x^* \in X^*$, where

$$\mathcal{B} := \left\{ (z^{0*}, z^{1*}, \widetilde{z}^{1*}, z^{2*}) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ \times \mathbb{R}^n_+ \times Q^* : \widetilde{z}_i^{1*} - \frac{[z_i^{1*}]^2}{4z_i^{0*}c_i} \geq 0, \; i = 1, \ldots, n \right\}.$$

One can also provide necessary and sufficient optimality conditions for the ε-optimal solutions of the problem (P^F) via Theorem 3.2.

Theorem 4.2. *(a) Assume that the regularity condition (RC_F) is fulfilled and let $\varepsilon \geq 0$. If $\overline{x} \in X$ is an ε-optimal solution of the problem (P^F), then there exist $\varepsilon_i \geq 0$, $i = 0, \dots, 3$ and $(\overline{z}^{0*}, \overline{z}^{1*}, \overline{\widetilde{z}}^{1*}, \overline{z}^{2*}) \in \mathcal{B}$ such that*

(i) $0 \leq \Phi\left(c_1 \frac{[h_1(\overline{x})]^2}{l_1(x)}, \dots, c_n \frac{[h_n(\overline{x})]^2}{l_n(x)}\right) + \Phi^*(\overline{z}^{0*}) - \sum\limits_{i=1}^n \overline{z}_i^{0*} c_i \frac{[h_i(\overline{x})]^2}{l_i(\overline{x})} \leq \varepsilon_0,$

(ii) $0 \leq \sum\limits_{i=1}^n \left(\overline{z}_i^{0*} c_i \frac{[h_i(\overline{x})]^2}{l_i(\overline{x})} - \overline{z}_i^{1*} h_i(\overline{x}) + \overline{\widetilde{z}}_i^{1*} l_i(\overline{x})\right) \leq \varepsilon_1$

(iii) $0 \leq \sum\limits_{i=1}^n ((\overline{z}_i^{1*} h_i)(\overline{x}) - \overline{\widetilde{z}}_i^{1*} l_i(\overline{x})) + (z^{2*} g)(x) + \left(\sum\limits_{i=1}^n (z_i^{1*} h_i - \widetilde{z}_i^{1*} l_i) + (z^{2*} g)\right)^* (0_{X^*}) \leq \varepsilon_2,$

(iv) $0 \leq -\langle \overline{z}^{n*}, g(\overline{x}) \rangle \leq \varepsilon_3,$

(v) $\varepsilon_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3 = \varepsilon.$

(b) If there exist $\varepsilon_i \geq 0$, $i = 0, \dots, 3$, and $(\overline{z}^{0*}, \overline{z}^{1*}, \overline{\widetilde{z}}^{1*}, \overline{z}^{2*}) \in \mathcal{B}$ such that (i)-(v) *are fulfilled for some $\overline{x} \in X$, then $\overline{x}$ is an ε-optimal solution of the problem (P^F).*

Remark 4.1. *Like in the previous section, we are also able to give in this case a closedness type regularity condition, which provides a new stable strong duality statement and ε-optimality conditions where h_i, l_i and g are separated for $i = 1, \dots, n$. This can be achieved from the main results, for example, by introducing the following functions*

- $f : \overline{\mathbb{R}^n} \to \overline{\mathbb{R}}$ *defined by* $f(y^0) := \Phi(y^0)$, $y^0 = (y_1^0, \dots, y_n^0)^T \in \mathbb{R}^n$,
- $F^1 : \overline{\mathbb{R}^n \times \mathbb{R}^n} \to \overline{\mathbb{R}^n}$ *defined by*
$$F^1(y^1, \widetilde{y}^1) := \begin{cases} \left(-c_1 \frac{[y_1^1]^2}{\widetilde{y}_1^1}, \dots, -c_n \frac{[y_n^1]^2}{\widetilde{y}_n^1}\right)^T, & \text{if } y_i^1 \geq 0,\ \widetilde{y}_i^1 < 0\ \forall i = 1, \dots, n, \\ +\infty_{\mathbb{R}^n_+}, & \text{otherwise,} \end{cases}$$
- $F^2 : \overline{X^n \times X^n} \to \overline{\mathbb{R}^n \times \mathbb{R}^n}$ *defined by*
$$F^2(y^2, \widetilde{y}^2) = \begin{cases} (h_1(y_1^2), \dots, h_n(y_n^2), -l_1(\widetilde{y}_1^2), \dots, -l_n(\widetilde{y}_n^2)), & \text{if } y_i^2 \in \operatorname{dom} h_i,\ \widetilde{y}_i^2 \in \operatorname{dom}(-l_i),\ i = \overline{1, n} \\ +\infty_{\mathbb{R}^n_+ \times \mathbb{R}^n_+}, & \text{otherwise} \end{cases}$$

and

- $F^3 : X \to \overline{X^n \times X^n}$ *defined by* $F^3(x) = \begin{cases} (x, \dots, x), & \text{if } x \in \mathcal{A}, \\ +\infty_{K^n \times K^n}, & \text{otherwise.} \end{cases}$

4.2 Optimization problems with entropy-like objective functions

Other potential interesting applications of our main results from this paper are in entropy optimization. Inspired by our contribution [5], where we have presented duality investigations on optimization problems with entropy-like objective functions that encompassed as special cases the classical Kullback-Leibler, Shannon and Burg entropy functions, as well as other papers like [1, 13], we consider the following optimization problem

$$(P^E) \quad \inf_{x \in \mathcal{A}} \left\{ \sum_{i=1}^n h_i(x) \Phi_i\left(\frac{l_i(x)}{h_i(x)}\right) \right\},$$

where $\mathcal{A}$ is defined as in the beginning of this section, $\Phi_i : \mathbb{R}_+ \to \overline{\mathbb{R}}$ is a proper, convex, lower semicontinuous and increasing function fulfilling $\Phi_i(t) \geq 0$ for all $t \geq 0$, $h_i : X \to \overline{\mathbb{R}}$ is a proper, convex and lower semicontinuous function fulfilling $h_i(x) > 0$ for all $x \in X$ and $l_i : X \to \overline{\mathbb{R}}$ is a proper, concave and upper semicontinuous function fulfilling $l_i(x) \geq 0$ for all $x \in X$, $i = 1, \ldots, n$. As Φ_i is a convex function, $i = 1, \ldots, n$, we know by [1, Lemma 2.1] that the objective function of (P^E) is convex and hence, (P^E) is a convex optimization problem.
To consider it in the framework of the approach in the previous sections, take $X_0 = \mathbb{R}^n \times \mathbb{R}^n$ be partially ordered by $K_0 = \mathbb{R}^n_+ \times \mathbb{R}^n_+$, $X_1 = X^n \times X^n$ be partially ordered by $K_1 = K^n \times K^n$ and $X_2 = X$ be partially ordered by the closed, convex cone K as well as the following functions

- $f : \overline{\mathbb{R}^n \times \mathbb{R}^n} \to \overline{\mathbb{R}}$ defined by $f(y^0, \widetilde{y}^0) := \begin{cases} \sum\limits_{i=1}^n y_i^0 \Phi_i \left(-\frac{\widetilde{y}_i^0}{y_i^0} \right), & \text{if } (y^0, \widetilde{y}^0) \in \mathbb{R}^n_+ \times \left(-\mathbb{R}^n_+ \right), \\ +\infty, & \text{otherwise,} \end{cases}$

- $F^1 : \overline{X^n \times X^n} \to \overline{\mathbb{R}^n \times \mathbb{R}^n}$ defined by
$F^1(y^1, \widetilde{y}^1) = \begin{cases} (h_1(y_1^1), \ldots, h_n(y_n^1), -l_1(\widetilde{y}_1^1), \ldots, -l_n(\widetilde{y}_n^1)), & \text{if } (y^1, \widetilde{y}^1) \in \prod\limits_{i=1}^n \operatorname{dom} l_i \times \prod\limits_{i=1}^n \operatorname{dom}(-h_i), \\ +\infty_{\mathbb{R}^n_+ \times \mathbb{R}^n_+}, & \text{otherwise,} \end{cases}$

and

- $F^2 : X \to \overline{X^n \times X^n}$ defined by $F^2(x) = \begin{cases} (x, \ldots, x), & \text{if } x \in \mathcal{A}, \\ +\infty_{K^n \times K^n}, & \text{otherwise.} \end{cases}$

In order to keep the length of the paper reasonable, we leave the derivation of the corresponding duality, optimality and subdifferential statements to the interested reader.

Remark 4.2. *Other duality schemes may be employed for approaching the proposed applications, too, however, the separation of the conjugates of the involved functions in the corresponding dual problems may fail to happen. However, by introducing the function F^2 it is possible to separate the conjugates of the functions g, l_i and h_i, $i = 1, \ldots, n$, in the objective function of the conjugate dual problem of (P^E). This also underlines the benefit of the concept introduced in this paper.*

Remark 4.3. *For different hypotheses imposed on the involved functions, that can be written as multi-composed functions, too, by carefully choosing the corresponding functions and cones, the problem (P^E) turns out to encompass as special cases different important (entropy) optimization problems. In the following we present some of these situations, noting that as usual in entropy optimization we consider the convention $0 \ln 0 = 0$.*

1. *When Φ_i is decreasing, l_i is concave and h_i affine, for all $i = 1, \ldots, n$, one obtains a problem that, when $\Phi_i = -\ln$, $i = 1, \ldots, n$, collapses to the one treated in [5].*

2. *When Φ_i is increasing, l_i is convex and h_i affine, for all $i = 1, \ldots, n$, one obtains a problem that, when Φ_i is the identity function, $h_j(x) = 1$ for all $x \in X$ and $l_i(x) = k_i(x - y_i)$, where $k_i \in \mathbb{R}$ and $y_i \in X$, $i = 1, \ldots, n$, collapses to the Steiner-Fermat problem considered in [14].*

3. *When Φ_i is increasing and nonpositive on the set $\{l_i(x)/h_i(x) : x \in \mathcal{A}, i = 1, \ldots, n\}$, l_i is convex and h_i concave, for all $i = 1, \ldots, n$, one obtains a problem that, when $\Phi_i(x) = -1$ for all $x \in \mathbb{R}_+$ and $g_i(x) = \ln x_i$, where $x = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n$ collapses, for an adequate choice of the other involved functions and sets to the Burg entropy optimization problem treated in [12].*

4. *When Φ_i is decreasing and nonnegative on the set $\{l_i(x)/h_i(x) : x \in \mathcal{A}, i = 1, \ldots, n\}$, l_i is concave and h_i convex, for all $i = 1, \ldots, n$, one obtains a problem that, when $\Phi_i(x) := c_i(1/x)$ when $x > 0$ and for $c_i > 0$ $i = 1, \ldots, n$, turns out to be a special case of (P^F) (see also [11, 19]).*

Acknowledgements. Research partially supported by DFG (German Research Foundation), project WA 922/8-1.

References

[1] A. Ben-Tal, A. Ben-Israel, M. Teboulle: *Certainty Equivalents and Information Measures: Duality and Extremal Principles.* Journal of Mathematical Analysis and Applications 157, 211–236, 1991.

[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, G. Wanka: *A New Constraint Qualification for the Formula of the Subdifferential of Composed Convex Functions in Infinite Dimensional Spaces.* Mathematische Nachrichten 281(8), 1–20, 2008.

[4] R. I. Boţ, S.-M. Grad, G. Wanka: *Duality in Vector Optimization.* Springer-Verlag, Berlin Heidelberg, 2009.

[5] R.I. Boţ, S.-M. Grad, G. Wanka, *Duality for Optimization Problems with Entropy-Like Objective Functions.* Journal of Optimization and Information Sciences 26(2), 415–441, 2005.

[6] R. I. Boţ, S.-M. Grad, G. Wanka: *New Regularity Conditions for Lagrange and Fenchel-Lagrange Duality in Infinite Dimensional Spaces.* Mathematical Inequalities and Applications 12(1), 171–189, 2009.

[7] 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 and Applications 69(1), 323–336, 2008.

[8] R. I. Boţ, S.-M. Grad, G. Wanka: *A New Constraint Qualification and Conjugate Duality for Composed Convex Optimization Problems.* Journal of Optimization Theory and Applications 135(2), 241–255, 2007.

[9] R. I. Boţ, I. B. Hodrea, G. Wanka: *ε-Optimality Conditions for Composed Convex Optimization Problems.* Journal of Approximation Theory 153(1), 108–121, 2008.

[10] R. I. Boţ, G. Wanka: *Duality for Composed Convex Functions with Applications in Location Theory.* Multi-Criteria- und Fuzzy-Systeme in Theorie und Praxis, Deutscher Universitäts-Verlag, Wiesbaden, 1–18, 2003.

[11] R. I. Boţ, G. Wanka: *Multiobjective Duality for Convex Ratios.* Journal of Mathematical Analysis and Applications, 275(1), 354–368, 2002.

[12] Y. Censor, A.R. de Pierro, A.N. Iusem: *Optimization of Burg's Entropy over Linear Constraints.* Applied Numerical Mathematics 7(2), 151–165, 1991.

[13] I. Csiszár: *On the Dimension and Entropy of Order α of the Mixture of Probability Distributions.* Acta Mathematica Hungarica 13, 246–255, 1962.

[14] C.R. Glassey: *Explicit Duality for Convex Homogeneous Programs.* Mathematical Programming 10, 176–191, 1976.

[15] S.-M. Grad, H. V. Boncea: *Characterizations of ε-Duality Gap Statements for Composed Optimization Problems.* Nonlinear Analysis: Theory, Methods and Applications 92(1), 96–107, 2013.

[16] S.-M. Grad: *Vector Optimization and Monotone Operators via Convex Duality*, Springer-Verlag, Cham, 2015.

[17] D. T. Luc: *Theory of Vector Optimization.* Springer-Verlag, Berlin Heidelberg, 1989.

[18] R.T. Rockafellar: *Duality Theorems for Convex Functions.* Bulletin of the American Mathematical Society 70, 189–192, 1964.

[19] C. H. Scott, T. R. Jefferson: *Duality for a Sum of Convex Ratios.* Optimization: A Journal of Mathematical Programming and Operations Research 40(4), 303–312, 1997.

[20] G. Wanka, O. Wilfer: *A Lagrange Duality Approach for Multi-Composed Optimization Problems.* Preprint, 2015.