Conjugate Duality and the Control of Linear Discrete Systems

Radu Ioan Boţ *   Ernö Robert Csetnek †

July 2, 2013

Abstract. In this paper we deal with the minimization of a convex function over the solution set of a range inclusion problem determined by a multivalued operator with convex graph. We attach a dual problem to it, provide regularity conditions guaranteeing strong duality and derive for the resulting primal-dual pair necessary and sufficient optimality conditions. We also discuss the existence of optimal solutions for the primal and dual problems by using duality arguments. The theoretical results are applied in the context of the control of linear discrete systems.

Key Words. convex optimization, conjugate duality, control of linear systems

AMS subject classification. 90C25, 49N05, 49N15, 93C05

1 Introduction

In many optimal control problems the constraints set, whether this is described by linear discrete systems or by differential inclusions, is the set of zeros of a multivalued operator, see [1-11].

*Corresponding author. University of Vienna, Faculty of Mathematics, Nordbergstrasse 15, A-1090 Vienna, Austria, e-mail: radu.bot@univie.ac.at. Research partially supported by DFG (German Research Foundation), project BO 2516/4-1.

†Chemnitz University of Technology, Department of Mathematics, D-09107 Chemnitz, Germany, e-mail: robert.csetnek@mathematik.tu-chemnitz.de. Research supported by DFG (German Research Foundation), project BO 2516/4-1.

Having this in mind, we deal in this paper from the point of view of the convex conjugate duality theory with a convex optimization problem having such a constraint set and to which we attach a dual optimization problem by making use of the so-called lower support function. For the primal-dual pair of optimization problems we provide a number of sufficient regularity conditions guaranteeing strong duality, which is the situation when the optimal objective values of the two problems coincide and the dual has an optimal solution. To this end we make use of techniques relying on the perturbation theory in the conjugate duality. By relying on the apparatus of convex subdifferential calculus we are also able to formulate necessary and sufficient optimality conditions for the primal problem under investigation. By using duality arguments we also discuss the so-called reverse strong duality, which is the situation when the optimal objective values of the two problems coincide and the primal has an optimal solution.

The structure of the paper is as follows. In the next section we give a short overview of the notions and results used in the paper. In Section 3 we introduce the primal problem under investigation and its conjugate dual, address the issue of guaranteeing strong duality, formulate necessary and sufficient optimality conditions and study the existence of optimal solutions for the primal. In Section 4 we apply the general results in the context of the control of linear discrete systems and rediscover some results from the literature as particular cases.

2 Elements of Convex Analysis

For the notions and results which we recall in this section we refer the reader to [9, 12-17]. Consider X a real separated locally convex space and X^* its topological dual space. For $x^* \in X^*$ and $x \in X$ we denote by $\langle x^*, x\rangle$ the value of the linear and continuous functional x^* at x. The *interior* of a set $A \subseteq X$ is denoted by $\operatorname{int} A$, while, if A is convex, then $\operatorname{sqri} A$ denotes the set of those elements $x \in A$ with the property that $\cup_{\lambda>0}\lambda(A-x)$ is a closed linear subspace. Note that we always have $\operatorname{int} A \subseteq \operatorname{sqri} A$. If $X = \mathbb{R}^n$, then $\operatorname{sqri} A = \operatorname{ri} A$, where $\operatorname{ri} A$ denotes the *relative interior* of A, that is the interior of A with respect to the affine hull of A. If $K \subseteq X$ is a cone, then we denote by

$K^* = \{x^* \in X^* : \langle x^*, x\rangle \geq 0 \ \forall x \in K\}$ its *positive dual cone.*

For a given function $f : X \to \overline{\mathbb{R}} := \{\pm\infty\}$ we consider its *domain* defined by $\operatorname{dom} f = \{x \in X : f(x) < +\infty\}$ and say that f is *proper* if $\operatorname{dom} f \neq \emptyset$ and $f(x) > -\infty$ for all $x \in X$. By $f^* : X^* \to \overline{\mathbb{R}}$, defined by $f^*(x^*) = \sup_{x\in X}\{\langle x^*, x\rangle - f(x)\}$ for all $x^* \in X^*$, we denote the *Fenchel conjugate* of f. For all $(x, x^*) \in X \times X^*$ the *Young-Fenchel inequality* $f(x) + f^*(x^*) \geq \langle x^*, x\rangle$ holds. The notation $\delta_A : X \to \mathbb{R} \cup \{+\infty\}$ is used for the *indicator function* of a set $A \subseteq X$, which is the function that takes the value 0 on A and $+\infty$ on $X \setminus A$. Notice that the conjugate function of δ_A is nothing else than the *support function* of A, $\sigma_A : X^* \to \overline{\mathbb{R}}, \sigma_A(x^*) = \sup_{x\in A}\langle x^*, x\rangle$. We denote by $\operatorname{bar} A$ the *barrier cone* of A, which is defined as the domain of the support function of A, that is $\operatorname{bar} A = \operatorname{dom} \sigma_A$.

The *(convex) subdifferential* of f at a point $x \in X$ such that $f(x) \in \mathbb{R}$ is defined by $\partial f(x) = \{x^* \in X^* : f(y) - f(x) \geq \langle x^*, y - x\rangle \ \forall y \in X\}$. In case $f(x) \in \{\pm\infty\}$ we set $\partial f(x) := \emptyset$. This (global) notion is a generalization to the nonsmooth case of the gradient. Indeed, if $f : X \to \overline{\mathbb{R}}$ is proper, convex and Gâteaux differentiable at $x \in \operatorname{dom} f$, then $\partial f(x) = \{\nabla f(x)\}$ (cf. [17, Theorem 2.4.4(i)]). We notice the following characterization of the subdifferential via conjugate functions:

$$x^* \in \partial f(x) \text{ if and only if } f(x) + f^*(x^*) = \langle x^*, x\rangle.$$

We denote by $N_A := \partial\delta_A$ the *normal cone* of the set $A \subseteq X$. One can easily check that $N_A(x) = \{x^* \in X^* : \langle x^*, y - x\rangle \leq 0 \ \forall y \in A\}$, if $x \in A$, and $N_A(x) = \emptyset$, if $x \notin A$. For $x \in A$ we have $x^* \in N_A(x)$ if and only if $\sigma_A(x^*) = \langle x^*, x\rangle$.

For Y another real separated locally convex space we consider a multivalued operator $F : X \rightrightarrows Y$ and denote by $\operatorname{Gr} F = \{(x, y) \in X \times Y : y \in F(x)\}$ its *graph.* We consider the *domain* of F defined by $\operatorname{Dom} F := \operatorname{pr}_X \operatorname{Gr} F = \{x \in X : F(x) \neq \emptyset\}$ and its *range* $\operatorname{Range} F := \operatorname{pr}_Y \operatorname{Gr} F = \cup_{x\in X} F(x)$. Here, $\operatorname{pr}_X : X \times Y \to X$, $\operatorname{pr}_X(x, y) = x$, denotes the *projection operator* on X, while pr_Y is defined analogously. The inverse operator $F^{-1} : Y \rightrightarrows X$ is defined by $(y, x) \in \operatorname{Gr} F^{-1}$ if and only if $(x, y) \in \operatorname{Gr} F$. We say that F is convex, if $\operatorname{Gr} F$ is a convex set, and that F is closed, if $\operatorname{Gr} F$ is a closed set.

The following function associated to F, called *lower support function*, will play an important role in our investigations:

$$s_F : X \times Y^* \to \overline{\mathbb{R}},\ s_F(x, y^*) = \inf_{y \in F(x)} \langle y^*, y \rangle.$$

Notice that for every $(x, y^*) \in X \times Y^*$ it holds $s_F(x, y^*) = -\sigma_{F(x)}(-y^*)$. The lower support function is well-studied in the literature, see [9, 18-21], and some of it most important properties are mentioned as follows. These properties are probably well-known, however we include some details of their proofs for reader's convenience.

Proposition 2.1 The following properties are true:

(i) $\operatorname{dom} s_F(\cdot, y^*) = \operatorname{Dom} F$ for all $y^* \in Y^*$;

(ii) if F is convex, then $s_F(\cdot, y^*)$ is a convex function on X for all $y^* \in Y^*$;

(iii) if $(x, y) \in \operatorname{Gr} F$, then $s_F(x, y^*) = \langle y^*, y \rangle$ if and only if $-y^* \in N_{F(x)}(y)$;

(iv) $\big(s_F(\cdot, y^*)\big)^*(x^*) = \sigma_{\operatorname{Gr} F}(x^*, -y^*)$ for all $(x^*, y^*) \in X^* \times Y^*$;

(v) $(x^*, -y^*) \in N_{\operatorname{Gr} F}(x, y)$ if and only if $x^* \in \partial s_F(\cdot, y^*)(x)$ and $-y^* \in N_{F(x)}(y)$.

Proof. (i), (ii) and (iii) follow easily from the definition of the lower support function.

(iv) For $(x^*, y^*) \in X^* \times Y^*$ we have $\big(s_F(\cdot, y^*)\big)^*(x^*) = \sup_{x \in X}\{\langle x^*, x \rangle - s_F(x, y^*)\}$

$= \sup_{x \in X}\{\langle x^*, x \rangle - \inf_{y \in F(x)}\langle y^*, y \rangle\} = \sup_{x \in X, y \in F(x)}\{\langle x^*, x \rangle - \langle y^*, y \rangle\} = \sigma_{\operatorname{Gr} F}(x^*, -y^*)$.

(v) Take first $(x^*, -y^*) \in N_{\operatorname{Gr} F}(x, y)$. Then $(x, y) \in \operatorname{Gr} F$ and by using (iv) we obtain $\big(s_F(\cdot, y^*)\big)^*(x^*) = \sigma_{\operatorname{Gr} F}(x^*, -y^*) = \langle x^*, x \rangle - \langle y^*, y \rangle$. Further,

$$\begin{aligned}
\langle x^*, x \rangle &\leq s_F(x, y^*) + \big(s_F(\cdot, y^*)\big)^*(x^*) \\
&= s_F(x, y^*) + \langle x^*, x \rangle - \langle y^*, y \rangle \\
&\leq \langle y^*, y \rangle + \langle x^*, x \rangle - \langle y^*, y \rangle \\
&= \langle x^*, x \rangle,
\end{aligned}$$

hence $\langle x^*, x \rangle = s_F(x, y^*) + \big(s_F(\cdot, y^*)\big)^*(x^*)$ and $s_F(x, y^*) = \langle y^*, y \rangle$, which proves that $x^* \in \partial s_F(\cdot, y^*)(x)$ and $-y^* \in N_{F(x)}(y)$ (cf. (iii)).

Conversely, suppose that $x^* \in \partial s_F(\cdot, y^*)(x)$ and $-y^* \in N_{F(x)}(y)$. Then $(x, y) \in \operatorname{Gr} F$ and, by (iii) and (iv), $\sigma_{\mathrm{Gr}\,F}(x^*,-y^*) = \big(s_F(\cdot,y^*)\big)^*(x^*) = \langle x^*,x\rangle - s_F(x,y^*) = \langle x^*,x\rangle - \langle y^*,y\rangle$, thus $(x^*,-y^*) \in N_{\mathrm{Gr}\,F}(x,y)$. $\square$

3 Duality Results

Consider the primal optimization problem

$$(P) \qquad \inf_{0\in F(x)} f(x),$$

where X, Y are real separated locally convex spaces, $f : X \to \overline{\mathbb{R}}$ and $F : X \rightrightarrows Y$ are such that $\mathrm{dom}\, f \cap F^{-1}(0) \neq \emptyset$. To (P) we attach the following dual problem

$$(D) \qquad \sup_{y^*\in Y^*} \inf_{x\in X} [f(x) + s_F(x,y^*)].$$

In this section we investigate the primal-dual pair $(P) - (D)$ from the point of view of the existence of strong duality and some of its consequences. We denote by $v(P)$ and $v(D)$ the optimal objective values of the problems (P) and (D), respectively. *Weak duality*, namely $v(D) \leq v(P)$, holds, the proof of this inequality relying on the fact that $0 \in F(x)$ implies $s_F(x,y^*) \leq 0$ for all $y^* \in Y^*$. Of much more importance is the situation called *strong duality*, namely when $v(D) = v(P)$ and the dual problem has an optimal solution, for which we provide as follows several so-called regularity conditions. The proof of the next duality result relies on conjugate duality specific techniques.

Theorem 3.1 *Let $f : X \to \overline{\mathbb{R}}$ be a proper and convex function and $F : X \rightrightarrows Y$ a convex multivalued operator. If one of the following conditions is fulfilled:*

(i) $\exists x' \in \mathrm{dom}\, f$ such that $0 \in \mathrm{int}\, F(x')$;

(ii) X, Y are Fréchet spaces, f is lower semicontinuous, F is closed and

$$0 \in \mathrm{int}\, F(\mathrm{dom}\, f);$$

(iii) X, Y are Fréchet spaces, f is lower semicontinuous, F is closed and

$$0 \in \operatorname{sqri} F(\operatorname{dom} f).$$

(iv) Y is finite dimensional and

$$0 \in \operatorname{ri} F(\operatorname{dom} f),$$

then for (P) and (D) strong duality holds, namely, there exists $\overline{y}^* \in Y^*$ such that

$$\inf_{0 \in F(x)} f(x) = \sup_{y^* \in Y^*} \inf_{x \in X} [f(x) + s_F(x, y^*)] = \inf_{x \in X} [f(x) + s_F(x, \overline{y}^*)].$$

Proof. Consider the function $\Phi : X \times Y \to \overline{\mathbb{R}}$ defined for all $(x, y) \in X \times Y$ by

$$\Phi(x, y) = f(x) + \delta_{\operatorname{Gr} F}(x, -y) = f(x) + \delta_{-F(x)}(y).$$

According to the hypotheses, Φ is a convex function.

The condition $0 \in \operatorname{int} F(x')$ ensures that the function $\delta_{-F(\overline{x})}$ is continuous at 0, hence condition (i) reads as:

$$\exists x' \in X \text{ such that } (x', 0) \in \operatorname{dom} \Phi \text{ and } \Phi(x', \cdot) \text{ is continuous at } 0.$$

On the other hand, if f is lower semicontinuous and F is closed, then the function Φ is lower semicontinuous, too. Furthermore, it holds $F(\operatorname{dom} f) = -\operatorname{pr}_Y \operatorname{dom} \Phi$. Hence the conditions (ii) (respectively (iii)) ensure that Φ is a lower semicontinuous function and

$$0 \in \operatorname{int}(\operatorname{pr}_Y \operatorname{dom} \Phi) \text{ (respectively } 0 \in \operatorname{sqri}(\operatorname{pr}_Y \operatorname{dom} \Phi)).$$

Similarly, condition (iv) can be equivalently written as

$$0 \in \operatorname{ri}(\operatorname{pr}_Y \operatorname{dom} \Phi).$$

Now we can apply [14, Theorem 1.7] (see also [17, Theorem 2.7.1], [15, Proposition 2.3]) and conclude that there exists $\overline{y}^* \in Y^*$ such that

$$\inf_{x \in X} \Phi(x, 0) = \sup_{y^* \in Y^*} -\Phi^*(0, y^*) = -\Phi^*(0, \overline{y}^*). \tag{1}$$

It is immediate that $\inf_{x\in X}\Phi(x,0)=v(P)$. Let us compute the conjugate function of Φ. For $(x^*,y^*)\in X^*\times Y^*$ we have

$$\begin{aligned}
\Phi^*(x^*,y^*) &= \sup_{x\in X, y\in Y}\{\langle x^*,x\rangle+\langle y^*,y\rangle - f(x)-\delta_{F(x)}(-y)\}\\
&= \sup_{x\in X}\{\langle x^*,x\rangle - f(x)+\sup_{y\in Y}[\langle y^*,y\rangle-\delta_{F(x)}(-y)]\}\\
&= \sup_{x\in X}\{\langle x^*,x\rangle - f(x)-s_F(x,y^*)\}\\
&= (f+s_F(\cdot,y^*))^*(x^*).
\end{aligned}$$

Thus $-\Phi^*(0,y^*)=\inf_{x\in X}[f(x)+s_F(x,y^*)]$ and the conclusion follows now from (1). $\square$

Remark 3.1 If X is also finite dimensional, then the following more useful equivalent formulation of the condition (iv) in the above theorem, from the point of view of its verifiability, can be given:

$$0\in \operatorname{ri} F(\operatorname{dom} f) \Leftrightarrow \text{ there exists } x'\in \operatorname{ri}\operatorname{dom} f\cap \operatorname{ri}(\operatorname{Dom} F) \text{ and } 0\in \operatorname{ri} F(x'). \tag{2}$$

Indeed, for $C:=\operatorname{Gr} F\cap(\operatorname{dom} f\times Y)$, one has $F(\operatorname{dom} f)=\operatorname{pr}_Y(C)$, thus (see [16, Theorem 6.6])

$$0\in \operatorname{ri} F(\operatorname{dom} f)\Leftrightarrow 0\in \operatorname{pr}_Y(\operatorname{ri} C)\Leftrightarrow \text{there exists } x'\in X \text{ such that } (x',0)\in \operatorname{ri} C.$$

Finally, (2) follows by using the following characterization of the relative interior of C (see [16, Theorem 6.8])

$$(x,y)\in \operatorname{ri} C\Leftrightarrow x\in \operatorname{ri}\operatorname{dom} f\cap \operatorname{ri}(\operatorname{Dom} F) \text{ and } y\in \operatorname{ri} F(x). \tag{3}$$

The existence of strong duality gives rise to the following optimality conditions for the primal-dual pair $(P)-(D)$.

Theorem 3.2 *(a) Assume that the hypotheses of Theorem 3.1 are fulfilled and let $\overline{x}$ be an optimal solution to the primal problem (P). Then there exists $\overline{y}^*\in Y^*$, an optimal solution to (D), such that*

(i) $f(\overline{x})=\min_{x\in X}[f(x)+s_F(x,\overline{y}^*)]$;

(ii) $s_F(\overline{x},\overline{y}^*)=0$.

(b) Assume that $\overline{x}\in F^{-1}(0)$ and $\overline{y}^*\in Y^*$ satisfy the relations $(i)-(ii)$. Then $\overline{x}$ is an optimal solution to (P), $\overline{y}^*$ is an optimal solution to (D) and $v(P)=v(D)$.

Proof. The result is a direct consequence of Theorem 3.1. Alternatively, one can apply [15, Proposition 2.4] for the function Φ considered in the proof of Theorem 3.1. $\square$

Remark 3.2 (a) Notice that in Theorem 3.2(b) no regularity condition is needed for the validity of this result.

(b) The conditions (i)-(ii) in statement (a) of Theorem 3.2 ensure that $s_F(\cdot, \overline{y}^*)$ is a proper function, while relation (i) is nothing else than $0 \in \partial(f + s_F(\cdot, \overline{y}^*))(\overline{x})$. If, additionally to the hypotheses of Theorem 3.2(a), we suppose that f is finite and continuous at a point in $\operatorname{Dom} F$, then $0 \in \partial(f + s_F(\cdot, \overline{y}^*))(\overline{x})$ is the same as $0 \in \partial f(\overline{x}) + \partial s_F(\cdot, \overline{y}^*)(\overline{x})$ (see, for instance, [17, Theorem 2.8.7], [14, Theorem 2.2]). The latter reads as: there exists $x^* \in \partial f(\overline{x})$ such that $-x^* \in \partial s_F(\cdot, \overline{y}^*)(\overline{x})$. A direct consequence of (ii) and Proposition 2.1(iii) is that $-\overline{y}^* \in N_{F(\overline{x})}(0)$, hence $-(x^*, \overline{y}^*) \in N_{\operatorname{Gr} F}(\overline{x}, 0)$ (see Proposition 2.1(v)). In conclusion, the optimality conditions (i)-(ii) in Theorem 3.2(a) can be written as:

$$s_F(\overline{x}, \overline{y}^*) = 0 \text{ and there exists } x^* \in \partial f(\overline{x}) \text{ such that } -(x^*, \overline{y}^*) \in N_{\operatorname{Gr} F}(\overline{x}, 0).$$

Remark 3.3 Notice that if we consider the particular case $F(x) := g(x) + K$ for all $x \in X$, where $g : X \to Y$ is a given function and $K \subseteq Y$ is a convex cone, then (P) reduces to the optimization problem with cone constraints

$$(P_K) \quad \inf_{g(x) \in -K} f(x).$$

In this case the lower support function fulfills $s_F(x, y^*) = \langle y^*, g(x) \rangle$ for all $(x, y^*) \in X \times K^*$ and $s_F(x, y^*) = -\infty$ for all $(x, y^*) \in X \times (Y^* \setminus K^*)$, which means that the dual problem (D) is nothing else than the classical Lagrange dual to (P_K)

$$(D_K) \quad \sup_{y^* \in K^*} \inf_{x \in X} [f(x) + \langle y^*, g(x) \rangle].$$

The function g is K-convex, i.e. $\lambda g(u) + (1 - \lambda) g(v) - g(\lambda u + (1 - \lambda) v) \in K$ for all $u, v \in X$ and $\lambda \in [0, 1]$ if and only if F is a convex multivalued operator, while g is K-epi closed, i.e. its K-epigraph $\operatorname{epi}_K g := \{(x, y) \in X \times Y : y \in g(x) + K\}$ is a closed set (see [14]), if and only if F is a closed multivalued operator. Furthermore, condition (i) in Theorem 3.1 becomes the classical

Slater condition:

$$\exists x' \in \operatorname{dom} f \text{ such that } g(x') \in -\operatorname{int} K,$$

while, as $F(\operatorname{dom} f) = g(\operatorname{dom} f) + K$, the conditions (ii), (iii) and (iv) in the same result become the regularity conditions $(RC_{2''}^{C_L})$, $(RC_2^{C_L})$ and $(RC_3^{C_L})$, respectively, considered in the context of Lagrange duality in [14]. In this way we rediscover for f proper and convex the strong duality result Theorem 3.4 in [14] as a particular case of Theorem 3.1, while the optimality conditions in Theorem 3.2 give rise to the generalized Karush-Kuhn-Tucker conditions for the primal-dual pair $(P_K) - (D_K)$.

We close the section by addressing the so-called *reverse strong duality*, which is the situation when $v(P) = v(D)$ and the primal problem has an optimal solution.

Theorem 3.3 Let X and Y be reflexive Banach spaces, f a proper, convex and lower semicontinuous function and F a convex and closed multivalued operator such that $\operatorname{dom} f \cap F^{-1}(0) \neq \emptyset$ and

$$0 \in \operatorname{sqri}\Big(\bigcup_{y^* \in Y^*} \operatorname{dom}(f + s_F(\cdot, y^*))^* \Big). \tag{4}$$

Then $v(P) = v(D)$ and the primal problem has an optimal solution.

Proof. Consider again the function Φ defined in the proof of Theorem 3.1, which is in the present context proper, convex and lower semicontinuous. According to the formula of the conjugate function of Φ, we have that $\operatorname{pr}_{X^*} \operatorname{dom} \Phi^* = \bigcup_{y^* \in Y^*} \operatorname{dom}(f + s_F(\cdot, y^*))^*$, hence condition (4) reads $0 \in \operatorname{sqri}(\operatorname{pr}_{X^*} \operatorname{dom} \Phi^*)$. Let us introduce now the function $\Gamma : Y^* \times X^* \to \overline{\mathbb{R}}$, $\Gamma(y^*, x^*) = \Phi^*(x^*, y^*)$. The properties of the function Φ ensure that Γ is proper, convex, lower semicontinuous (with respect to the strong topology on $Y^* \times X^*$) and that it holds $0 \in \operatorname{sqri}(\operatorname{pr}_{X^*} \operatorname{dom} \Gamma)$. Taking into account that X, Y are reflexive Banach spaces, it follows by [14, Theorem 1.7] (see also [15, Proposition 2.3], [17, Theorem 2.7.1]) that there exists $\bar{x} \in X$ such that

$$\inf_{y^* \in Y^*} \Gamma(y^*, 0) = \sup_{x \in X} -\Gamma^*(0, x) = -\Gamma^*(0, \bar{x})$$

or, equivalently,

$$\sup_{y^* \in Y^*} -\Gamma(y^*, 0) = \inf_{x \in X} \Gamma^*(0, x) = \Gamma^*(0, \bar{x}). \tag{5}$$

The conclusion follows from (5) by taking into account that

$$-\Gamma(y^*,0) = -\Phi^*(0,y^*) = \inf_{x\in X}[f(x) + s_F(x,y^*)]\ \forall y^* \in Y^*$$

and

$$\Gamma^*(0,x) = \Phi(x,0) = f(x) + \delta_{-F(x)}(0)\ \forall x \in X.$$

$\square$

Remark 3.4 In the hypotheses of Theorem 3.3, a sufficient condition for (4) is given by

$$0 \in \operatorname{int}\Big(\operatorname{dom} f^* + \operatorname{pr}_{X^*}(\operatorname{bar}\operatorname{Gr} F)\Big). \tag{6}$$

This is a direct consequence of the fact that

$$\operatorname{dom} f^* + \operatorname{pr}_{X^*}(\operatorname{bar}\operatorname{Gr} F) \subseteq \bigcup_{y^*\in Y^*} \operatorname{dom}(f + s_F(\cdot,y^*))^*. \tag{7}$$

Indeed, take $u^* \in \operatorname{dom} f^*$ and $v^* \in \operatorname{pr}_{X^*}(\operatorname{bar}\operatorname{Gr} F)$. Then there exists $y^* \in Y^*$ such that $\sigma_{\operatorname{Gr} F}(v^*,y^*) < +\infty$, hence, due to Proposition 2.1(iv), $v^* \in \operatorname{dom}(s_F(\cdot,-y^*))^*$. Thus $u^* + v^* \in \operatorname{dom} f^* + \operatorname{dom}(s_F(\cdot,-y^*))^* \subseteq \operatorname{dom}(f + s_F(\cdot,-y^*))^*$ and (7) holds.

It is also worth mentioning, that, if the proper and convex function f is continuous at a point in $\operatorname{dom} F$, then (7) is fulfilled as equality (see [17, Theorem 2.8.7]).

Corollary 3.1 Let X and Y be reflexive Banach spaces, f a proper, convex, lower semicontinuous and coercive function (that is $\lim_{\|x\|\to+\infty} f(x) = +\infty$) and F a convex and closed multivalued operator such that $\operatorname{dom} f \cap F^{-1}(0) \neq \emptyset$. Then $v(P) = v(D)$ and the primal problem has an optimal solution.

Proof. The coercivity of the function f guarantees that $0 \in \operatorname{int}(\operatorname{dom} f^*)$ (see, for example, [17, Exercise 2.41]), thus the conclusion follows from Remark 3.4 by taking into account also that $0 \in \operatorname{pr}_{X^*}(\operatorname{bar}\operatorname{Gr} F)$. $\square$

Remark 3.5 In the hypotheses of Theorem 3.3, another sufficient condition for (6) reads

$$(-\operatorname{dom} f^*) \cap \operatorname{pr}_{X^*}(\operatorname{int}\operatorname{bar}\operatorname{Gr} F) \neq \emptyset.$$

Indeed, by using the open mapping principle, we have

$$\begin{aligned}
0 \in \operatorname{dom} f^* + \operatorname{pr}_{X^*}(\operatorname{int} \operatorname{bar} \operatorname{Gr} F) &\subseteq \operatorname{dom} f^* + \operatorname{int} \operatorname{pr}_{X^*}(\operatorname{bar} \operatorname{Gr} F) \\
&\subseteq \operatorname{int}\big(\operatorname{dom} f^* + \operatorname{pr}_{X^*}(\operatorname{bar} \operatorname{Gr} F)\big).
\end{aligned}$$

4 The Control of Linear Discrete Systems

In the following we apply the theoretical results obtained in the previous section in the context of the control of discrete linear systems. To this end we consider the optimization problem

$$\begin{aligned}
\inf \quad & f(x_0, u_0, x_1, u_1, ..., x_{k-1}, u_{k-1}, x_k), \\
\text{s.t.} \quad & x_0 \in S, u_i \in S_i, i = 0, ..., k-1 \\
& P_i x_{i+1} - L_i x_i - K_i u_i \in C_i, i = 0, ..., k-1
\end{aligned} \tag{8}$$

where $k \in \mathbb{N}$, $k \geq 2$, X_i, U_i, Y_i, $i = 0, ..., k-1$, and X_k are real Banach spaces, $S \subseteq X_0$ and $S_i \subseteq U_i, i = 0, ..., k-1$, are nonempty convex sets, $f : \prod_{i=0}^{k-1}(X_i \times U_i) \times X_k \to \overline{\mathbb{R}}$ is proper, convex and continuous at some point of its domain, $P_i : X_{i+1} \to Y_i$ are linear continuous and surjective operators, $L_i : X_i \to Y_i$ and $K_i : U_i \to Y_i$ are linear continuous operators and $C_i \subseteq Y_i$ are nonempty convex closed cones, $i = 0, ..., k-1$. The spaces $X_i, i = 0, ..., k$, are the so-called *state spaces*, while $U_i, i = 0, ..., k-1$, are the *control spaces*.

Our aim is to formulate optimality conditions for the problem (8) by making use of Theorem 3.2 and also to show in which circumstances this result is applicable. To this end we notice that (8) can be written in the form of (P) by taking

$$F : \prod_{i=0}^{k-1}(X_i \times U_i) \times X_k \rightrightarrows X_0 \times \prod_{i=0}^{k-1} U_i \times \prod_{i=0}^{k-1} X_{i+1},$$

$$F(x_0, u_0, ..., x_{k-1}, u_{k-1}, x_k) = (S - x_0) \times \prod_{i=0}^{k-1}(S_i - u_i) \times \prod_{i=0}^{k-1}\Big(P_i^{-1}(L_i x_i + K_i u_i + C_i) - x_{i+1}\Big).$$

The multivalued operator F is convex and, since P_i are surjective, $i = 0, ..., k-1$, it has full domain and full range.

Concerning the fulfillment of the regularity conditions, one can notice that the one formulated in Theorem (3.1)(i) holds, if there exists $(x_0', u_0', ..., x_{k-1}', u_{k-1}', x_k') \in \operatorname{dom} f$ such that

$$x_0' \in \operatorname{int} S, u_i' \in \operatorname{int} S_i \text{ and } P_i x_{i+1}' - L_i x_i' - K_i u_i' \in \operatorname{int} C_i, i = 0, ..., k-1.$$

If the sets S, S_i and $C_i, i = 0, ..., k-1$, have nonempty interiors and the function f takes only finite values (hence it is continuous on its whole domain of definition), then this regularity condition is automatically fulfilled and Theorem 3.2 can be applied.

On the other hand, it holds

$$F(\operatorname{dom} f) = \left\{ \begin{array}{l} (S - x_0) \times \prod_{i=0}^{k-1}(S_i - u_i) \times \prod_{i=0}^{k-1} \left(P_i^{-1}(L_i x_i + K_i u_i + C_i) - x_{i+1} \right) : \\ (x_0, u_0, ..., x_{k-1}, u_{k-1}, x_k) \in \operatorname{dom} f \end{array} \right\}.$$

If the sets S, S_i and $C_i, i = 0, ..., k-1$, are closed, then F is a closed multivalued operator. If the function f is lower semicontinuous and takes only finite values (hence it is continuous on its whole domain of definition), then the condition $0 \in \operatorname{sqri} F(\operatorname{dom} f)$ in Theorem 3.1(iii) is automatically fulfilled and Theorem 3.2 can be applied in this case, too.

Coming now to the formulation of the optimality conditions, one has that for two given elements $x = (x_0, u_0, ..., x_{k-1}, u_{k-1}, x_k)$ and $y^* = (y_0^*, u_0^*, u_1^*, ..., u_{k-1}^*, y_1^*, y_2^*, ..., y_k^*)$ it holds

$$\begin{aligned} s_F(x, y^*) &= \inf_{y_0 \in S - x_0} \langle y_0^*, y_0 \rangle + \sum_{i=0}^{k-1} \inf_{z_i \in S_i - u_i} \langle u_i^*, z_i \rangle \\ &\quad + \sum_{i=0}^{k-1} \inf_{y_{i+1} \in P_i^{-1}(L_i x_i + K_i u_i + C_i) - x_{i+1}} \langle y_{i+1}^*, y_{i+1} \rangle \\ &= -\langle y_0^*, x_0 \rangle - \sigma_S(-y_0^*) + \sum_{i=0}^{k-1} \Big(-\langle u_i^*, u_i \rangle - \sigma_{S_i}(-u_i^*) \Big) \\ &\quad + \sum_{i=0}^{k-1} \left(-\langle y_{i+1}^*, x_{i+1} \rangle + \inf_{v_{i+1} \in P_i^{-1}(L_i x_i + K_i u_i + C_i)} \langle y_{i+1}^*, v_{i+1} \rangle \right). \end{aligned}$$

Consider an optimal solution $\overline{x} = (\overline{x}_0, \overline{u}_0, ..., \overline{x}_{k-1}, \overline{u}_{k-1}, \overline{x}_k)$ to (8) and an optimal solution $\overline{y}^* = (\overline{y}_0^*, \overline{u}_0^*, \overline{u}_1^*, ..., \overline{u}_{k-1}^*, \overline{y}_1^*, \overline{y}_2^*, ..., \overline{y}_k^*)$ to its dual problem such that (i)-(ii) in Theorem 3.2 are fulfilled. According to Remark 3.2(b), we equivalently have $s_F(\overline{x}, \overline{y}^*) = 0$ and that there exists $x^* = (x_0^*, u_0^*, ..., x_{k-1}^*, u_{k-1}^*, x_k^*) \in \partial f(\overline{x})$ such that $-x^* \in \partial s_F(\cdot, \overline{y}^*)(\overline{x})$.

Since $\overline{x} \in S, \overline{u}_i \in S_i, P_i \overline{x}_{i+1} - L_i \overline{x}_i - K_i \overline{u}_i \in C_i, i = 0, ..., k-1$, we have

$$-\langle \overline{y}_0^*, \overline{x}_0 \rangle - \sigma_S(-\overline{y}_0^*) \leq 0,$$

$$-\langle \overline{u}_i^*, \overline{u}_i \rangle - \sigma_{S_i}(-\overline{u}_i^*) \leq 0, i = 0, ..., k-1,$$

and

$$-\langle \overline{y}_{i+1}^*, \overline{x}_{i+1} \rangle + \inf_{v_{i+1} \in P_i^{-1}(L_i \overline{x}_i + K_i \overline{u}_i + C_i)} \langle \overline{y}_{i+1}^*, v_{i+1} \rangle \leq 0, i = 0, ..., k-1.$$

As $s_F(\overline{x}, \overline{y}^*) = 0$, it follows that all above inequalities must be fulfilled as equalities, and from here we get that

$$-\overline{y}_0^* \in N_S(\overline{x}_0), \ -\overline{u}_i^* \in N_{S_i}(\overline{u}_i), i = 0, ..., k-1$$

and

$$\inf_{v_{i+1} \in P_i^{-1}(L_i \overline{x}_i + K_i \overline{u}_i + C_i)} \langle \overline{y}_{i+1}^*, v_{i+1} \rangle = \langle \overline{y}_{i+1}^*, \overline{x}_{i+1} \rangle, i = 0, ..., k-1.$$

For $i = 0, ..., k-1$, since P_i is surjective, by [14, Theorem 3.4] we obtain a Lagrange multiplier $\overline{\lambda}_i \in C_i^*$ such that

$$\begin{aligned}
\inf_{v_{i+1} \in P_i^{-1}(L_i \overline{x}_i + K_i \overline{u}_i + C_i)} \langle \overline{y}_{i+1}^*, v_{i+1} \rangle &= \inf_{\substack{v_{i+1} \in X_{i+1} \\ P_i v_{i+1} - L_i \overline{x}_i - K_i \overline{u}_i \in C_i}} \langle \overline{y}_{i+1}^*, v_{i+1} \rangle \\
&= \inf_{v_{i+1} \in X_{i+1}} \{ \langle \overline{y}_{i+1}^*, v_{i+1} \rangle + \langle \overline{\lambda}_i, -P_i v_{i+1} + L_i \overline{x}_i + K_i \overline{u}_i \rangle \} \\
&= \langle \overline{\lambda}_i, L_i \overline{x}_i + K_i \overline{u}_i \rangle - \delta_{\{P_i^* \overline{\lambda}_i\}}(\overline{y}_{i+1}^*).
\end{aligned}$$

Thus

$$\overline{y}_{i+1}^* = P_i^* \overline{\lambda}_i \text{ and } \langle \overline{y}_{i+1}^*, \overline{x}_{i+1} \rangle = \langle \overline{\lambda}_i, L_i \overline{x}_i + K_i \overline{u}_i \rangle, i = 0, ..., k-1.$$

This means that the optimality conditions (i)-(ii) in Theorem 3.2 read:

there exist $x^* = (x_0^*, u_0^*, ..., x_{k-1}^*, u_{k-1}^*, x_k^*) \in \partial f(\overline{x})$ and $\overline{\lambda}_i \in C_i^*$, $i = 0, ..., k-1$, such that

$$-\overline{y}_0^* \in N_S(\overline{x}_0),$$

$$-\overline{u}_i^* \in N_{S_i}(\overline{u}_i), \overline{y}_{i+1}^* = P_i^* \overline{\lambda}_i, \langle \overline{y}_{i+1}^*, \overline{x}_{i+1} \rangle = \langle \overline{\lambda}_i, L_i \overline{x}_i + K_i \overline{u}_i \rangle, i = 0, ..., k-1,$$

and

$$-x^* \in \partial s_F(\cdot, \overline{y}^*)(\overline{x}).$$

Hence, for every $x = (x_0, u_0, ..., x_{k-1}, u_{k-1}, x_k)$ it holds

$$\begin{aligned} s_F(x, \overline{y}^*) &= -\langle \overline{y}_0^*, x_0\rangle - \sigma_S(-\overline{y}_0^*) + \sum_{i=0}^{k-1}\Big(- \langle \overline{u}_i^*, u_i\rangle - \sigma_{S_i}(-\overline{u}_i^*)\Big) \\ &+ \sum_{i=0}^{k-1}\left(-\langle P_i^*\overline{\lambda}_i, x_{i+1}\rangle + \inf_{v_{i+1}\in P_i^{-1}(L_ix_i+K_iu_i+C_i)} \langle P_i^*\overline{\lambda}_i, v_{i+1}\rangle\right) \\ &= -\langle \overline{y}_0^*, x_0\rangle - \sigma_S(-\overline{y}_0^*) + \sum_{i=0}^{k-1}\Big(- \langle \overline{u}_i^*, u_i\rangle - \sigma_{S_i}(-\overline{u}_i^*)\Big) \\ &+ \sum_{i=0}^{k-1}\Big(- \langle \overline{y}_{i+1}^*, x_{i+1}\rangle + \langle L_i^*\overline{\lambda}_i, x_i\rangle + \langle K_i^*\overline{\lambda}_i, u_i\rangle\Big). \end{aligned}$$

Thus $-x^* \in \partial s_F(\cdot, \overline{y}^*)(\overline{x})$ if and only if $-x_i^* = -\overline{y}_i^* + L_i^*\overline{\lambda}_i$, $-u_i^* = -\overline{u}_i^* + K_i^*\overline{\lambda}_i, i = 0, ..., k-1$, and $-x_k^* = -\overline{y}_k^*$.

Altogether, the optimality conditions (i)-(ii) in Theorem 3.2 read:

there exist $\overline{\lambda}_i \in C_i^*$, $i = 0, ..., k-1$, such that

$$-\overline{y}_0^* \in N_S(\overline{x}_0), \tag{9}$$

$$-\overline{u}_i^* \in N_{S_i}(\overline{u}_i), \overline{y}_{i+1}^* = P_i^*\overline{\lambda}_i, \langle \overline{y}_{i+1}^*, \overline{x}_{i+1}\rangle = \langle \overline{\lambda}_i, L_i\overline{x}_i + K_i\overline{u}_i\rangle, i = 0, ..., k-1, \tag{10}$$

and

$$(\overline{y}_0^* - L_0^*\overline{\lambda}_0, \overline{u}_0^* - K_0^*\overline{\lambda}_0, ..., \overline{y}_{k-1}^* - L_{k-1}^*\overline{\lambda}_{k-1}, \overline{u}_{k-1}^* - K_{k-1}^*\overline{\lambda}_{k-1}, \overline{y}_k^*) \in \partial f(\overline{x}). \tag{11}$$

Remark 4.1 The following optimization problem has been investigated in [6] from the point of view of the formulation of optimality conditions

$$\begin{array}{ll} \inf & \sum\limits_{i=0}^{k-1} \left(\frac{1}{2}\langle x_i, Qx_i\rangle + \frac{1}{2}\langle u_i, Ru_i\rangle\right), \\ \text{s.t.} & x_{i+1} = \Phi x_i + Du_i, i = 0, ..., k-1 \end{array} \tag{12}$$

with H and U real Hilbert spaces, $D : U \to H$ and $\Phi : H \to H, i = 0, ..., k-1$, linear continuous operators, $Q : H \to H$ a linear continuous and self-adjoint positive semidefinite operator and $R : U \to U$ a linear continuous and self-adjoint positive definite operator.

Taking $S = X_i = Y_i = H$, $i = 0, ..., k-1$, $X_k = H$, $S_i = U_i = U$, $i = 0, ..., k-1$, $C_i = \{0\}$, P_i the identity operator on H, $L_i = \Phi$, $K_i = D$, $i = 0, ..., k-1$, and defining

$$f(x_0, u_0, ..., x_{k-1}, u_{k-1}, x_k) = \sum_{i=0}^{k-1}\left(\frac{1}{2}\langle x_i, Qx_i\rangle + \frac{1}{2}\langle u_i, Ru_i\rangle\right),$$

the problem under investigation becomes the optimization problem (8). Due to the fact that f is convex, continuous and with full domain, the regularity condition in Theorem 3.1(iii) is fulfilled.

Thus, if $\overline{x} = (\overline{x}_0, \overline{u}_0, ..., \overline{x}_{k-1}, \overline{u}_{k-1}, \overline{x}_k)$ is an optimal solution to (12), then there exist an optimal solution $\overline{y}^* = (\overline{y}_0^*, \overline{u}_0^*, \overline{u}_1^*, ..., \overline{u}_{k-1}^*, \overline{y}_1^*, \overline{y}_2^*, ..., \overline{y}_k^*)$ to its dual such that, according to (9)-(11),

$$\overline{y}_0^* = \overline{y}_k^* = 0, \overline{u}_i^* = 0, i = 0, ..., k-1,$$

$$\overline{u}_i = -R^{-1} D^* \overline{y}_{i+1}^* \text{ and } \overline{y}_i^* = \Phi^* \overline{y}_{i+1}^* + Q\overline{x}_i, i = 0, ..., k-1.$$

In this way we rediscover the optimality conditions given in [6, Theorem 3.2].

5 Conclusions

By means of conjugate duality techniques we investigate in this paper a convex optimization problem with the constraints set described by making use of a convex and closed multivalued operator. We attache a dual problem to it and study the relations between the primal-dual pair of optimization problems. Optimality conditions are delivered and the existence of optimal solutions is investigated, as well. The theoretical outcomes are applied to the control of linear discrete systems and in this way some results from the literature are rediscovered as special instances of the general approach. As further possible research directions, one can treat also other type of linear control systems, but also general continuous convex control problems.

References

1. Aubin, J.-P., Cellina, A.: Differential Inclusions. Set-valued Maps and Viability Theory, Grundlehren der Mathematischen Wissenschaften, Vol. 264, Springer-Verlag, Berlin, 1984

2. Cruceanu, S.: Duality and optimality for convex extremal problems described by discrete inclusions, Math. Operationsforsch. Statist. Ser. Optim. 11(1), 13–30, 1980

3. Fujita, M.: Duality and maximum principle in multi-period convex programming, J. Math. Econom. 1(3), 295–326, 1974

4. Kaczorek, T.: Two-dimensional Linear Systems, Lecture Notes in Control and Information Sciences, Vol. 68, Springer-Verlag, Berlin, 1985

5. Kaczorek, T.: Singular two-dimensional continuous-discrete linear systems, Dynam. Contin. Discrete Impuls. Systems 2(2), 193–204, 1996

6. Lee, K.Y., Chow, S., Barr, R.O.: On the control of discrete-time distributed parameter systems, SIAM J. Control 10, 361–376, 1972

7. Mahmudov, E.N.: Necessary and sufficient conditions for discrete and differential inclusions of elliptic type, J. Math. Anal. Appl. 323(2), 768–789, 2006

8. Mahmudov, E.N.: Locally adjoint mappings and optimization of the first boundary value problem for hyperbolic type discrete and differential inclusions, Nonlinear Anal. 67(10), 2966–2981, 2007

9. Mahmudov, E.N.: Approximation and Optimization of Discrete and Differential Inclusions, Elsevier, 2011

10. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation, I. Basic Theory and II. Applications, Series of Comprehensive Studies in Mathematics, Vol. 330, Springer-Verlag, Berlin Heidelberg, 2006

11. Zabczyk, J.: Remarks on the control of discrete-time distributed parameter systems, SIAM J. Control 12(4), 721–735, 1974

12. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, Springer, New York, 2011

13. Borwein, J.M., Vanderwerff, J.D.: Convex Functions: Constructions, Characterizations and Counterexamples, Encyclopedia of Mathematics and its Applications, Vol. 109, Cambridge University Press, Cambridge, 2010

14. Boţ, R.I.: Conjugate Duality in Convex Optimization, Lecture Notes in Economics and Mathematical Systems, Vol. 637, Springer-Verlag, Berlin Heidelberg, 2010

15. Ekeland, I., Temam, R.: Convex Analysis and Variational Problems, North-Holland Publishing Company, Amsterdam, 1976

16. Rockafellar, R.T.: Convex Analysis, Princeton University Press, Princeton, 1970

17. Zălinescu, C.: Convex Analysis in General Vector Spaces, World Scientific, Singapore, 2002

18. Aubin, J.-P., Ekeland, I.: Applied Nonlinear Analysis, Pure and Applied Mathematics, John Wiley & Sons, New York, 1984

19. Aubin, J.-P., Frankowska, H.: Set-Valued Analysis, Birkhäuser, Boston, 1990

20. Pshenichnyi, B.N.: Convex multi-valued mappings and their conjugates, in: Mathematical models in economics (eds.: J. Los and M.W. Los), pp. 333–349, North-Holland, Amsterdam, 1974

21. Rockafellar, R.T.: Monotone processes of convex and concave type, Memoirs of the American Mathematical Society, Vol. 77, American Mathematical Society, Providence, 1967