

## LAGRANGE THEORY OF DISCRETE-CONTINUOUS NONLINEAR OPTIMIZATION

JOHANNES JAHN\*, MARTIN KNOSSALLA

*Department Mathematik, Universität Erlangen-Nürnberg, Cauerstr. 11, 91058 Erlangen, Germany*

**Abstract.** This paper presents a new Lagrange theory of discrete-continuous conic optimization in an infinite dimensional setting. The following questions are answered for discrete-continuous optimization problems: how to define a Lagrange functional, how Karush-Kuhn-Tucker conditions look like, and which duality results can be obtained? This approach is based on new separation theorems for discrete sets, which are also given in this paper. The developed theory is finally applied to problems of discrete-continuous semidefinite and copositive optimization.

**Keywords.** Discrete-continuous optimization; Lagrange theory; Optimality conditions; Duality.

**2010 Mathematics Subject Classification.** 90C48, 90C11, 90C46.

### 1. INTRODUCTION

The well-known multiplier rule of Lagrange in continuous optimization is based on the fact that to each constraint *one* Lagrange multiplier is associated, which is then used for the formulation of the Lagrange functional. In mixed discrete-continuous optimization this point is completely different. In general, to each constraint of a discrete-continuous problem more than one Lagrange multiplier are associated. Recently, this important fact has been stated by Baes, Oertel and Weismantel [1]. Since much more Lagrange multipliers may arise in discrete-continuous optimization than in continuous optimization, the corresponding results are much more difficult to handle.

Duality results in mixed-integer optimization are known for a long time. Several approaches have been given (using linear programming (for instance, see [9, 13]), subadditive functions (compare [10, 14]) or Lagrangian relaxation (e.g., [8, 12])) during the last decades. Moreover, Murota [11] investigates discrete convexity. A more recent approach [1] is based on mixed-integer-free polyhedra and can also be used for the formulation of Karush-Kuhn-Tucker conditions.

The aim of this paper is to present new optimality conditions and duality results for discrete-continuous optimization problems in infinite dimensional spaces. This approach is completely different from the known theories. We investigate the separation of discrete sets where one set is convex and the other one is nonconvex. Such a separation can be done with piecewise affine linear functionals being non-convex in general. Based on these separation theorems we prove necessary and sufficient optimality conditions in infinite dimensional spaces. These variational inequalities can be simplified to a discrete-continuous version of the Karush-Kuhn-Tucker conditions. As a byproduct of this approach we are able

---

\*Corresponding author.

E-mail addresses: [johannes.jahn@fau.de](mailto:johannes.jahn@fau.de) (J. Jahn), [martin.knossalla@fau.de](mailto:martin.knossalla@fau.de) (M. Knossalla).

Received September 19, 2018; Accepted October 16, 2018.

to define a Lagrange functional for discrete-continuous optimization problems, which can be used for the proof of a saddle point result. This new approach is also the key for the formulation of duality results for nonlinear problems under certain convexity assumptions. The duality theory also incorporates linear discrete-continuous optimization problems. Since discrete-continuous semidefinite and copositive optimization problems are special problems of the problem class investigated in this paper, results on optimality conditions and duality can be directly applied to the semidefinite and copositive case.

This paper is organized as follows. In Section 2, we summarize the used notation and we discuss known concepts. Section 3 is devoted to theorems for the separation of a convex set and a certain type of a nonconvex set. Necessary and sufficient optimality conditions in discrete-continuous optimization are given in Section 4 and in the following section duality results are proven. An application of the theory of this paper to discrete-continuous problems in semidefinite and copositive optimization is presented in Section 6.

## 2. NOTATIONS AND BASICS

Throughout this paper we work with real linear spaces, which are sometimes specialized to real topological linear spaces, real locally convex spaces, real normed spaces, real Hilbert spaces or Euclidean spaces. For a real linear space  $E$  its *algebraical dual space* is denoted by  $E'$ . If  $E$  is a real linear topological space, its *topological dual space* is abbreviated by  $E^*$ . Discrete and continuous variables are indexed by a subscript  $d$  and  $c$ , respectively. For real linear spaces we also use these subscripts in a similar way.

The *convex hull* of a nonempty subset  $U$  of a real linear space is abbreviated by  $\text{co}(U)$ . If  $(E, \|\cdot\|_E)$  and  $(F, \|\cdot\|_F)$  are real normed spaces and  $f : E \rightarrow F$  is a given map, then the *Fréchet derivative* of  $f$  at some  $\bar{x} \in E$  is denoted by  $f'(\bar{x})$ .  $f'(\bar{x})(d)$  describes the Fréchet derivative of  $f$  at  $\bar{x}$  in the direction  $d \in E$ . The same notation is used for the *Gâteaux derivative*. It becomes clear from the text which type of derivative is actually used.

The *core* or the *algebraic interior* of a subset  $U$  of a real linear space  $E$  is defined by

$$\text{cor}(U) := \{ \bar{x} \in U \mid \forall x \in E \exists \bar{\lambda} > 0 : \bar{x} + \lambda x \in U \text{ for all } \lambda \in [0, \bar{\lambda}] \}.$$

If  $E$  is a real topological linear space and  $U$  is a nonempty subset of  $E$ , then  $\text{int}(U)$  denotes the *interior* of the set  $U$  and  $\text{cl}(U)$  describes the closure of the set  $U$ . For a nonempty convex subset  $U$  of a real topological linear space with  $\text{int}(U) \neq \emptyset$  it is well-known that  $\text{int}(U) = \text{cor}(U)$  and  $\text{cl}(U) = \text{cl}(\text{int}(U))$  (see [5, p. 59]). In this case we have  $U \subset \text{cl}(\text{int}(U))$ .

For the formulation of inequality constraints we use convex cones. A nonempty subset  $C$  of a real linear space is called a *cone*, if

$$x \in C, \lambda \geq 0 \Rightarrow \lambda x \in C.$$

Unfortunately, the definition of a cone is not consistent in the special literature. The definition used in this paper has the advantage that the partial ordering (sometimes also called preorder) induced by a convex cone is reflexive and transitive.

If  $C$  is a convex cone in a real linear space  $E$ , then

$$C' := \{ \ell \in E' \mid \ell(c) \geq 0 \text{ for all } c \in C \}$$

is called the *dual cone* for  $C$  and induces a partial ordering in the dual space  $E'$ . The notation  $C_{E'}$  is also in use. In the case of a real topological linear space the dual cone is denoted by  $C^*$ .

The separation theorem, which is fundamental for the investigations in the next section, is a well-known nontopological version of Eidelheit's separation theorem (see [3]): Let  $U$  and  $V$  be nonempty convex subsets of a real linear space  $E$  with  $\text{cor}(U) \neq \emptyset$ . Then  $\text{cor}(U) \cap V = \emptyset$  if and only if there exist a linear functional  $\ell \in E' \setminus \{0_{E'}\}$  and a real number  $\alpha$  with

$$\ell(u) \leq \alpha \leq \ell(v) \text{ for all } u \in U \text{ and all } v \in V$$

and

$$\ell(u) < \alpha \text{ for all } u \in \text{cor}(U).$$

### 3. SEPARATION THEOREMS FOR DISCRETE SETS

It is well-known that separation results are the fundamental tool for the theoretical investigation of continuous optimization problems. In discrete-continuous optimization one needs special separation results being the topic of this section.

In this paper we follow the definition of a discrete set often used in discrete optimization.

**Definition 3.1.** A nonempty subset of a real linear space is called *discrete*, if its number of elements is either finite or countably infinite.

This definition differs from that in a topological setting where for every vector of the considered set there is a neighborhood so that no other vector of this set belongs to this neighborhood. In a real topological linear space every vector of such a set is also called isolated.

**Example 3.1.** The number of elements of the set  $A := \{0\} \cup \{\frac{1}{n}\}_{n \in \mathbb{N}}$  is countably infinite and, therefore,  $A$  is a discrete set. But there is no neighborhood of 0 so that no nonzero element of  $A$  belongs to such a neighborhood, i.e., 0 is not an isolated point.

Based on Definition 3.1, we generally write a discrete subset  $A$  of a real linear space as  $A = \{a^i\}_{i \in N}$  with  $N := \{1, 2, \dots, n\}$  for some  $n \in \mathbb{N}$  or  $N := \mathbb{N}$ , where  $a^1, a^2, \dots$  are elements of the real linear space.

Next we extend Definition 3.1 to a more general class of sets.

**Definition 3.2.** A nonempty subset  $A$  of a real linear space  $E$  is called *extendedly discrete*, if  $A = \bigcup_{i \in N} A_i$  for  $N := \{1, 2, \dots, n\}$  (with  $n \in \mathbb{N}$ ) or  $N := \mathbb{N}$  and for nonempty sets  $A_i \subset E$  ( $i \in N$ ).

It is obvious that every discrete set is also extendedly discrete. Our first separation theorem can be formulated for extendedly discrete sets.

**Theorem 3.1.** Let  $E$  be a real linear space, let  $A$  be a nonempty subset of  $E$ , and let  $B$  be an extendedly discrete subset of  $E$  with the representation  $B = \bigcup_{i \in N} \text{cor}(\text{conv}(B_i))$  for  $N := \{1, 2, \dots, n\}$  (with  $n \in \mathbb{N}$ ) or  $N := \mathbb{N}$  where  $B_i \subset E$  and  $\text{cor}(\text{conv}(B_i)) \neq \emptyset$  for all  $i \in N$ . Then  $\text{conv}(A) \cap B = \emptyset$  if and only if for every  $i \in N$  there exist a linear functional  $\ell^i \in E' \setminus \{0_{E'}\}$  and a real number  $\alpha^i$  with

$$0 < \ell^i(y) + \alpha^i \text{ for all } y \in \text{cor}(\text{conv}(B_i)), \quad (3.1)$$

and the inequalities

$$\sup_{i \in N} \{\ell^i(x) + \alpha^i\} \leq 0 \leq \inf_{i \in N} \{\ell^i(y) + \alpha^i\} \text{ for all } x \in \text{conv}(A) \text{ and all } y \in \bigcup_{i \in N} \text{conv}(B_i) \quad (3.2)$$

are fulfilled.

*Proof.* (a) For all  $i \in N$  let there exist a linear functional  $\ell^i \in E' \setminus \{0_{E'}\}$  and a real number  $\alpha^i$  so that the inequalities (3.1) and in (3.2) are fulfilled, i.e., we have for all  $k \in N$

$$\sup_{i \in N} \{\ell^i(x) + \alpha^i\} \leq 0 < \ell^k(y) + \alpha^k \text{ for all } x \in \text{conv}(A) \text{ and all } y \in B.$$

Then it is obvious that  $\text{conv}(A) \cap B = \emptyset$ .

(b) Now assume that  $\text{conv}(A) \cap B = \emptyset$ , i.e.,  $\text{conv}(A) \cap (\bigcup_{i \in N} \text{cor}(\text{conv}(B_i))) = \emptyset$ . This implies

$$\text{conv}(A) \cap \text{cor}(\text{conv}(B_i)) = \emptyset \text{ for all } i \in N.$$

Let  $i \in N$  be arbitrarily chosen. By Eidelheit's separation theorem in the nontopological version (see [3] and compare [7, Theorem 3.14]) there exist a linear functional  $\ell^i \in E' \setminus \{0_{E'}\}$  and a real number  $-\alpha^i$  with

$$\ell^i(x) \leq -\alpha^i \leq \ell^i(y) \text{ for all } x \in \text{conv}(A) \text{ and all } y \in \text{conv}(B_i)$$

and

$$-\alpha^i < \ell^i(y) \text{ for all } y \in \text{cor}(\text{conv}(B_i)).$$

Then we get

$$\ell^i(x) + \alpha^i \leq 0 \leq \ell^i(y) + \alpha^i \text{ for all } x \in \text{conv}(A) \text{ and all } y \in \text{conv}(B_i) \quad (3.3)$$

and

$$0 < \ell^i(y) + \alpha^i \text{ for all } y \in \text{cor}(\text{conv}(B_i)).$$

So, inequality (3.1) is shown. Since the left inequality in (3.3) holds for all  $i \in N$ , we conclude

$$\sup_{i \in N} \{\ell^i(x) + \alpha^i\} \leq 0 \text{ for all } x \in \text{conv}(A)$$

and the left inequality in (3.2) is shown. The right inequality in (3.3) immediately implies

$$0 \leq \inf_{i \in N} \{\ell^i(y) + \alpha^i\} \text{ for all } y \in \bigcup_{i \in N} \text{conv}(B_i)$$

and the right inequality in (3.2) is shown. □

Figure 1 illustrates the separation result of Theorem 3.1. Since the set  $B$  in Theorem 3.1 may be

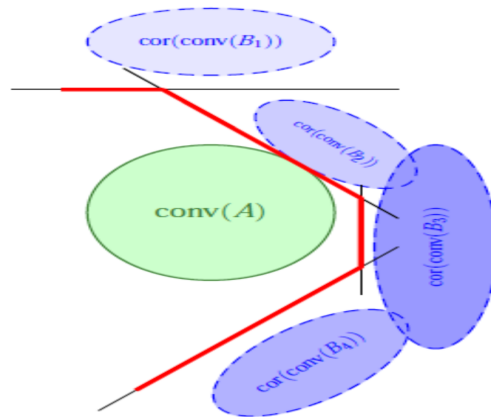


FIGURE 1. Separation of the sets  $\text{conv}(A)$  and  $B := \bigcup_{i \in \{1,2,3,4\}} \text{cor}(\text{conv}(B_i))$  by Theorem 3.1.

nonconvex, a separation is carried out with a nonlinear functional. From a geometrical point of view the separation of the sets  $\text{conv}(A)$  and  $B$  is done with a separating hyperface.

**Remark 3.1.** (a) The proof of Theorem 3.1 is based on the standard case of convex separation. But here we work with the supremum and infimum of affine linear functionals. If the index set  $N$  consists of finitely many indices, then the inf and sup terms reduce to a min and max term, respectively.

(b) Since the sets  $A, B_1, B_2, \dots$  in Theorem 3.1 may be discrete sets, the separation result can be applied to certain discrete problems.

(c) Theorem 3.1 can also be formulated in a topological space. We then replace the algebraic interior  $\text{cor}$  by the topological interior  $\text{int}$  and we apply the topological version of Eidelheit's separation theorem in the proof of Theorem 3.1 (compare [7, Thm. 3.16]). In this case the linear functionals  $\ell^i$  (with  $i \in N$ ) are even continuous, i.e.,  $\ell^i \in E^* \setminus \{0_{E^*}\}$  for all  $i \in N$ .

Next we specialize Theorem 3.1 to a singleton set  $A$ .

**Corollary 3.1.** *Let an element  $x$  of a real linear space  $E$  be given, and let  $B$  be an extendedly discrete subset of  $E$  with the representation  $B = \bigcup_{i \in N} \text{cor}(\text{conv}(B_i))$  for  $N := \{1, 2, \dots, n\}$  (with  $n \in \mathbb{N}$ ) or  $N := \mathbb{N}$  where  $B_i \subset E$  and  $\text{cor}(\text{conv}(B_i)) \neq \emptyset$  for all  $i \in N$ . Then  $x \notin B$  if and only if for every  $i \in N$  there exists a linear functional  $\ell^i \in E' \setminus \{0_{E'}\}$  with*

$$0 < \ell^i(y - x) \text{ for all } y \in \text{cor}(\text{conv}(B_i)). \quad (3.4)$$

In this case, we additionally have

$$0 \leq \inf_{i \in N} \{\ell^i(y - x)\} \text{ for all } y \in \bigcup_{i \in N} \text{conv}(B_i). \quad (3.5)$$

*Proof.* (a) First, for every  $i \in N$ , let  $\ell^i \in E' \setminus \{0_{E'}\}$  be a linear functional so that inequality (3.4) holds. If we assume that  $x \in B$ , then we get a contradiction from the inequality (3.4), if we set  $y = x$ .

(b) Next let  $x \notin B$ , which is equivalent to the condition  $0_E \notin B - \{x\}$ . This condition is again equivalent to  $\text{conv}(A) \cap (B - \{x\}) = \emptyset$  for  $A := \{0_E\}$ . By Theorem 3.1, for every  $i \in N$  there exist a linear functional  $\ell^i \in E' \setminus \{0_{E'}\}$  and a real number  $\alpha^i$  with

$$0 < \ell^i(y - x) + \alpha^i \text{ for all } y \in \text{cor}(\text{conv}(B_i)), \quad (3.6)$$

and the inequalities

$$\sup_{i \in N} \{\alpha^i\} \leq 0 \leq \inf_{i \in N} \{\ell^i(y - x) + \alpha^i\} \text{ for all } y \in \bigcup_{i \in N} \text{conv}(B_i) \quad (3.7)$$

are fulfilled. The left inequality in (3.7) implies

$$\alpha^i \leq 0 \text{ for all } i \in N \quad (3.8)$$

and the right inequality in (3.7) gives

$$0 \leq \inf_{i \in N} \{\ell^i(y - x) + \alpha^i\} \leq \inf_{i \in N} \{\ell^i(y - x)\} \text{ for all } y \in \bigcup_{i \in N} \text{conv}(B_i).$$

So, inequality (3.5) is shown. The proof of inequality (3.4) follows from inequality (3.6) together with inequality (3.8). This completes the proof.  $\square$

Another strict separation result can be shown in locally convex spaces.

**Theorem 3.2.** *Let an element  $x$  of a real locally convex space  $E$  be given, and let  $B$  be an extendedly discrete subset of  $E$  with the representation  $B = \bigcup_{i \in N} \text{cl}(\text{conv}(B_i))$  for  $N := \{1, 2, \dots, n\}$  (with  $n \in \mathbb{N}$ ) or  $N := \mathbb{N}$  where  $\text{conv}(B_i)$  are nonempty subsets of  $E$  for all  $i \in N$ . Then  $x \notin B$  if and only if for every  $i \in N$  there exists a continuous linear functional  $\ell^i \in E^* \setminus \{0_{E^*}\}$  with*

$$0 < \ell^i(y - x) \text{ for all } y \in \text{cl}(\text{conv}(B_i)). \quad (3.9)$$

*Proof.* (a) As in the part (a) of the proof of Corollary 3.1, we obtain that condition (3.9) implies  $x \notin B$ . (b) Now assume that  $x \notin B$ . Then for every  $i \in \mathbb{N}$ , the origin  $0_E$  does not belong to the set  $\text{cl}(\text{conv}(B_i) - \{x\})$ . By a strict separation theorem (e.g., compare [7, Thm. 3.18]), for every  $i \in \mathbb{N}$ , there are a continuous linear functional  $\ell^i \in E^* \setminus \{0_{E^*}\}$  and a real number  $\alpha^i$  with

$$0 = \ell^i(0_E) < \alpha^i \leq \ell^i(y) \text{ for all } y \in \text{cl}(\text{conv}(B_i) - \{x\})$$

implying

$$0 < \ell^i(y - x) \text{ for all } y \in \text{cl}(\text{conv}(B_i)),$$

which has to be shown.  $\square$

We now present a variant of the separation result of Theorem 3.1.

**Theorem 3.3.** *Let  $E$  be a real linear space, let  $A$  be a subset of  $E$  with  $\text{cor}(\text{conv}(A)) \neq \emptyset$ , and let  $B$  be an extendedly discrete subset of  $E$  with the representation  $B = \bigcup_{i \in N} \text{conv}(B_i)$  for  $N := \{1, 2, \dots, n\}$  (with  $n \in \mathbb{N}$ ) or  $N := \mathbb{N}$  where  $\emptyset \neq B_i \subset E$  for all  $i \in N$ . Then  $\text{cor}(\text{conv}(A)) \cap B = \emptyset$  if and only if for every  $i \in N$  there exist a linear functional  $\ell^i \in E' \setminus \{0_{E'}\}$  and a real number  $\alpha^i$  with*

$$0 < \ell^i(x) + \alpha^i \text{ for all } x \in \text{cor}(\text{conv}(A)), \quad (3.10)$$

and the inequalities

$$\sup_{i \in N} \{\ell^i(y) + \alpha^i\} \leq 0 \leq \inf_{i \in N} \{\ell^i(x) + \alpha^i\} \text{ for all } x \in \text{cor}(\text{conv}(A)) \text{ and all } y \in B \quad (3.11)$$

are fulfilled.

*Proof.* (a) First we assume that, for every  $i \in N$ , there exist a linear functional  $\ell^i \in E' \setminus \{0_{E'}\}$  (with  $i \in N$ ) and a real number  $\alpha^i$  (with  $i \in N$ ) so that the inequalities in (3.11) and (3.10) are fulfilled. Then we have for every  $k \in N$

$$\sup_{i \in N} \{\ell^i(y) + \alpha^i\} \leq 0 < \ell^k(x) + \alpha^k \text{ for all } x \in \text{cor}(\text{conv}(A)) \text{ and all } y \in B,$$

and this implies  $\text{cor}(\text{conv}(A)) \cap B = \emptyset$ .

(b) Next, we assume that

$$\emptyset = \text{cor}(\text{conv}(A)) \cap B = \text{cor}(\text{conv}(A)) \cap \left( \bigcup_{i \in N} \text{conv}(B_i) \right).$$

Then we have

$$\text{cor}(\text{conv}(A)) \cap \text{conv}(B_i) = \emptyset \text{ for all } i \in N.$$

For an arbitrarily chosen  $i \in N$  we now apply Eidelheit's separation theorem in the nontopological version (compare [3] and [7, Theorem 3.14]). Then there exist a linear functional  $-\ell^i \in E' \setminus \{0_{E'}\}$  and a real number  $\alpha^i$  with

$$-\ell^i(x) \leq \alpha^i \leq -\ell^i(y) \text{ for all } x \in \text{conv}(A) \text{ and all } y \in \text{conv}(B_i)$$

and

$$-\ell^i(x) < \alpha^i \text{ for all } x \in \text{cor}(\text{conv}(A)).$$

It follows that

$$\ell^i(y) + \alpha^i \leq 0 \leq \ell^i(x) + \alpha^i \text{ for all } x \in \text{conv}(A) \text{ and all } y \in \text{conv}(B_i)$$

and

$$0 < \ell^i(x) + \alpha^i \text{ for all } x \in \text{cor}(\text{conv}(A)).$$

Following the lines of the proof of Theorem 3.1, we conclude

$$\sup_{i \in N} \{\ell^i(y) + \alpha^i\} \leq 0 \leq \inf_{i \in N} \{\ell^i(x) + \alpha^i\} \text{ for all } x \in \text{conv}(A) \text{ and all } y \in B$$

and for every  $i \in N$ , we have

$$0 < \ell^i(x) + \alpha^i \text{ for all } x \in \text{cor}(\text{conv}(A)),$$

which completes the proof. □

In contrast to the first separation result (Theorem 3.1), Theorem 3.3 has the advantage that the sets  $B_i$  (with  $i \in N$ ) may be singletons, i.e.,  $B_i := \{b^i\}$  for some vectors  $b^i \in E$ . In this special case, the extended discrete set  $B = \bigcup_{i \in N} \text{conv}(B_i) = \{b^1, b^2, \dots\}$  is a discrete set. If  $A$  is a discrete set as well, i.e.,  $A := \{a^1, a^2, \dots\}$  for some  $a^1, a^2, \dots \in E$ , then we can separate  $\text{cor}(\text{conv}\{a^1, a^2, \dots\})$  and  $\{b^1, b^2, \dots\}$ . Figure 2 illustrates this special type of separation. The nonlinear separating functional makes use of only

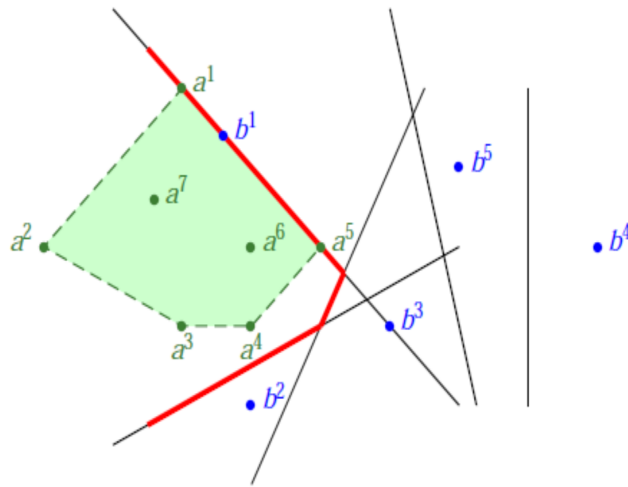


FIGURE 2. Separation of the sets  $\text{cor}(\text{conv}\{a^1, a^2, \dots, a^7\})$  and  $\{b^1, b^2, \dots, b^5\}$  by Theorem 3.3.

three affine linear functionals where one of them could be dropped. So, in Figure 2 one needs only two affine linear functionals for the nonlinear separation. This shows that the number of involved affine linear functionals may be less than expected.

It is well-known for a separation result in the finite dimensional space  $\mathbb{R}^m$  that the core of one of the two sets is not needed to be nonempty. This fact simplifies the two Theorems 3.1 and 3.3. This additional result is formulated in the following theorem.

**Theorem 3.4.** *Let  $A$  be a nonempty subset of  $\mathbb{R}^m$ , let  $B$  be an extendedly discrete subset of  $\mathbb{R}^m$  with the representation  $B = \bigcup_{i \in N} \text{conv}(B_i)$  for  $N := \{1, 2, \dots, n\}$  (with  $n \in \mathbb{N}$ ) or  $N := \mathbb{N}$  where  $\emptyset \neq B_i \subset \mathbb{R}^m$  for all  $i \in N$ . If  $\text{conv}(A) \cap B = \emptyset$ , then for every  $i \in N$  there exist a vector  $\ell^i \in \mathbb{R}^m \setminus \{0_{\mathbb{R}^m}\}$  and a real number  $\alpha^i$  with*

$$\sup_{i \in N} \{\ell^{iT} y + \alpha^i\} \leq 0 \leq \inf_{i \in N} \{\ell^{iT} x + \alpha^i\} \text{ for all } x \in \text{conv}(A) \text{ and all } y \in B.$$

*Proof.* The proof follows the lines of the proof of Theorem 3.1 or 3.3. Assume that the sets  $\text{conv}(A)$  and  $B$  are disjoint. This implies

$$\text{conv}(A) \cap \text{conv}(B_i) = \emptyset \text{ for all } i \in N.$$

For every  $i \in N$ , we then apply a finite dimensional separation theorem (see, e.g., [4, Satz 2.22]). Hence, for every  $i \in N$  there exist a vector  $\ell^i \in \mathbb{R}^m \setminus \{0_{\mathbb{R}^m}\}$  and a number  $\alpha^i \in \mathbb{R}$  so that

$$-\ell^{iT} x \leq \alpha^i \leq -\ell^{iT} y \text{ for all } x \in \text{conv}(A) \text{ and all } y \in \text{conv}(B_i).$$

As in the proof of Theorem 3.1, we get

$$\sup_{i \in N} \{\ell^{iT} y + \alpha^i\} \leq 0 \leq \inf_{i \in N} \{\ell^{iT} x + \alpha^i\} \text{ for all } x \in \text{conv}(A) \text{ and all } y \in B$$

and the proof is complete.  $\square$

#### 4. OPTIMALITY CONDITIONS

In this section, optimality conditions in discrete-continuous nonlinear optimization are given, which are based on the separation result of Corollary 3.1. First, we formulate the setting of the investigated optimization problems.

**Assumption 4.1.** Let  $X_d, X_c, Y$  and  $Z$  be real linear spaces, and let  $C \subset Y$  be a convex cone in  $Y$ . Let  $S_d$  be a discrete subset of  $X_d$ , i.e.  $S_d = \{x_d^i\}_{i \in N}$  for  $N := \{1, 2, \dots, n\}$  (with  $n \in \mathbb{N}$ ) or  $N := \mathbb{N}$  with elements  $x_d^1, x_d^2, \dots \in X_d$ . Let  $S_c$  be a nonempty subset of  $X_c$ . Let  $f : S_d \times S_c \rightarrow \mathbb{R}$ ,  $g : S_d \times S_c \rightarrow Y$  and  $h : S_d \times S_c \rightarrow Z$  be given maps and let the constraint set

$$S := \{(x_d, x_c) \in S_d \times S_c \mid g(x_d, x_c) \in -C, h(x_d, x_c) = 0_Z\}$$

be nonempty.

Under this assumption, we investigate the discrete-continuous optimization problem

$$\min_{(x_d, x_c) \in S} f(x_d, x_c). \quad (4.1)$$

Such a discrete-continuous optimization problem is much more complicated than standard problems of continuous optimization. For instance, the constraint set  $S$  of this problem can be written as

$$S = \bigcup_{i \in N} \{(x_d^i, x_c) \in X_d \times S_c \mid g(x_d^i, x_c) \in -C, h(x_d^i, x_c) = 0_Z\}$$

so that the constraint set is a union of layered subsets of  $X_d \times X_c$ . Figure 3 illustrates this multi-layered structure. It is obvious that these nonconvex constraint sets cannot be treated with standard approaches of continuous optimization.

The following result presents a first necessary optimality condition for the optimization problem (4.1).



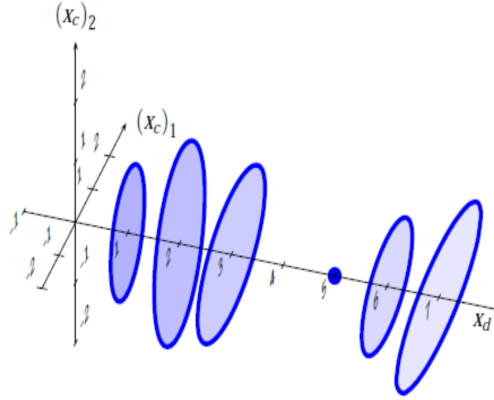


FIGURE 3. Illustration of a possible constraint set of problem (4.1) with  $S_d = \{1, 2, \dots, 7\}$  and  $S_c = \mathbb{R}^2$ .

**Theorem 4.1.** *Let Assumption 4.1 be satisfied and in addition, let  $X_c$ ,  $Y$  and  $Z$  be real topological linear spaces with  $\text{int}(S_c) \neq \emptyset$  and  $\text{int}(C) \neq \emptyset$ . Let  $\bar{x} = (x_d^j, \bar{x}_c)$  (for some  $j \in N$ ) be a minimal solution of the discrete-continuous optimization problem (4.1). For every  $i \in N$  let the set*

$$B_i := \left\{ \begin{pmatrix} f(x_d^i, x_c) - f(\bar{x}) + \alpha \\ g(x_d^i, x_c) + y \\ h(x_d^i, x_c) \end{pmatrix} \in \mathbb{R} \times Y \times Z \mid x_c \in \text{int}(S_c), \alpha > 0, y \in \text{int}(C) \right\}$$

*be convex and let  $h(x_d^i, \text{int}(S_c))$  be an open set. Then for every  $i \in N$  there exist a real number  $\mu^i \geq 0$  and continuous linear functionals  $\ell_g^i \in C^*$  and  $\ell_h^i \in Z^*$  with  $(\mu^i, \ell_g^i, \ell_h^i) \neq (0, 0_{Y^*}, 0_{Z^*})$ , and the inequality*

$$0 \leq \inf_{i \in N} \{ \mu^i (f(x_d^i, x_c) - f(\bar{x})) + \ell_g^i (g(x_d^i, x_c)) + \ell_h^i (h(x_d^i, x_c)) \} \text{ for all } x_c \in S_c \quad (4.2)$$

*and the equality*

$$\ell_g^j (g(\bar{x})) = 0 \quad (4.3)$$

*are fulfilled.*

*Proof.* Let  $\bar{x} = (x_d^j, \bar{x}_c)$  (for some  $j \in N$ ) be a minimal solution of the optimization problem (4.1). Choose an arbitrary  $i \in N$ . For the nonempty set  $B_i$  we show  $(0, 0_Y, 0_Z) \notin B_i$ . Assume that  $(0, 0_Y, 0_Z) \in B_i$ . Then there are some  $x_c \in \text{int}(S_c)$ , some  $\alpha > 0$  and some  $y \in \text{int}(C)$  with

$$f(x_d^i, x_c) - f(\bar{x}) = -\alpha \quad (4.4)$$

$$g(x_d^i, x_c) = -y \quad (4.5)$$

$$h(x_d^i, x_c) = 0_Z. \quad (4.6)$$

With  $x_c \in \text{int}(S_c)$  and  $y \in \text{int}(C)$ , (4.5) and (4.6) imply  $(x_d^i, x_c) \in S$ . Because of  $\alpha > 0$ , (4.4) contradicts the assumption that  $\bar{x}$  is a minimal solution of problem (4.1). Hence, we conclude

$$(0, 0_Y, 0_Z) \notin B_i \text{ for all } i \in N,$$

which implies

$$(0, 0_Y, 0_Z) \notin \bigcup_{i \in N} B_i.$$

Since  $B_i$  (with  $i \in N$ ) are convex and open, the separation result in Corollary 3.1 is applicable and for every  $i \in N$  there exist a real number  $\mu^i$  and continuous linear functionals  $\ell_g^i \in Y^*$  and  $\ell_h^i \in Z^*$  with  $(\mu^i, \ell_g^i, \ell_h^i) \neq (0, 0_{Y^*}, 0_{Z^*})$ , and the inequality

$$0 \leq \inf_{i \in N} \{ \mu^i (f(x_d^i, x_c) - f(\bar{x}) + \alpha) + \ell_g^i (g(x_d^i, x_c) + y) + \ell_h^i (h(x_d^i, x_c)) \} \\ \text{for all } x_c \in S_c, \alpha \geq 0 \text{ and } y \in C \quad (4.7)$$

is fulfilled (notice that  $S_c \subset \text{cl}(\text{int}(S_c))$  and  $C \subset \text{cl}(\text{int}(C))$ ). For an arbitrary  $k \in N$ , we get with  $x_c = \bar{x}_c$  and  $y = 0_Y$  from (4.7)

$$0 \leq \inf_{i \in N} \{ \mu^i (f(x_d^i, \bar{x}_c) - f(\bar{x}) + \alpha) + \ell_g^i (g(x_d^i, \bar{x}_c)) + \ell_h^i (h(x_d^i, \bar{x}_c)) \} \\ \leq \mu^k (f(x_d^k, \bar{x}_c) - f(\bar{x}) + \alpha) + \ell_g^k (g(x_d^k, \bar{x}_c)) + \ell_h^k (h(x_d^k, \bar{x}_c)) \text{ for all } \alpha \geq 0.$$

This implies for some  $\beta \in \mathbb{R}$

$$\mu^k \alpha \geq \beta \text{ for all } \alpha \geq 0$$

and  $\mu^k \geq 0$ . Since  $k \in N$  is arbitrarily chosen, we conclude

$$\mu^i \geq 0 \text{ for all } i \in N.$$

If we set  $x_c = \bar{x}_c$  and  $\alpha = 0$ , we obtain from (4.7), for an arbitrary  $k \in N$ , that

$$0 \leq \inf_{i \in N} \{ \mu^i (f(x_d^i, \bar{x}_c) - f(\bar{x})) + \ell_g^i (g(x_d^i, \bar{x}_c) + y) + \ell_h^i (h(x_d^i, \bar{x}_c)) \} \\ \leq \mu^k (f(x_d^k, \bar{x}_c) - f(\bar{x})) + \ell_g^k (g(x_d^k, \bar{x}_c) + y) + \ell_h^k (h(x_d^k, \bar{x}_c)) \text{ for all } y \in C.$$

We then conclude for some  $\gamma \in \mathbb{R}$

$$\ell_g^k(y) \geq \gamma \text{ for all } y \in C.$$

This inequality implies  $\ell_g^k \in C^*$ . Since  $k \in N$  is arbitrarily chosen, we have

$$\ell_g^i \in C^* \text{ for all } i \in N.$$

If we set  $\alpha = 0$  and  $y = 0_Y$  in (4.7), we immediately get (4.2). From (4.2) we obtain with  $x_c = \bar{x}_c$

$$0 \leq \inf_{i \in N} \{ \mu^i (f(x_d^i, \bar{x}_c) - f(\bar{x})) + \ell_g^i (g(x_d^i, \bar{x}_c)) + \ell_h^i (h(x_d^i, \bar{x}_c)) \} \\ \leq \underbrace{\mu^j (f(\bar{x}) - f(\bar{x}))}_{=0} + \ell_g^j (g(\bar{x})) + \underbrace{\ell_h^j (h(\bar{x}))}_{=0_Z} \\ = \ell_g^j (g(\bar{x}))$$

which implies  $\ell_g^j (g(\bar{x})) \geq 0$ . Since  $g(\bar{x}) \in -C$  and  $\ell_g^j \in C^*$ , we also get  $\ell_g^j (g(\bar{x})) \leq 0$ . Hence, we conclude  $\ell_g^j (g(\bar{x})) = 0$ , and the proof is complete.  $\square$

**Remark 4.1.** In Theorem 4.1, it is assumed that, for every  $i \in N$ , the set  $B_i$  is convex. It is well-known that  $B_i$  (for every  $i \in N$ ) is convex, if  $S_c$  is convex, the objective functional  $f$  is convex, the constraint map  $g$  is convex and the map  $h$  is affine linear. But these conditions can be weakened, if one extends the known notion of convex-likeness (see [7, Def. 2.9 and Ex. 2.10.(b)]) to the interior of a cone. Moreover, for every  $i \in N$  it is assumed that the set  $h(x_d^i, \text{int}(S_c))$  is open. For instance, this assumption is satisfied

by the open mapping theorem, if  $X_c$  and  $Z$  are real Banach spaces and the map  $h(x_d^i, \cdot)$  is continuous, linear and surjective.

**Corollary 4.1.** *Let the assumptions of Theorem 4.1 be satisfied. Again, let  $\bar{x} = (x_d^j, \bar{x}_c)$  (for some  $j \in N$  and some  $\bar{x}_c \in S_c$ ) denote a minimal solution of discrete-continuous optimization problem (4.1). In addition, we assume*

$$\forall i \in N \exists x_c \in S_c : 0 > \ell_g^i(g(x_d^i, x_c)) + \ell_h^i(h(x_d^i, x_c)). \quad (4.8)$$

Then, for every  $i \in N$ , there exist continuous linear functionals  $\bar{\ell}_g^i \in C^*$  and  $\bar{\ell}_h^i \in Z^*$ , and the inequality

$$f(\bar{x}) \leq \inf_{i \in N} \{f(x_d^i, x_c) + \bar{\ell}_g^i(g(x_d^i, x_c)) + \bar{\ell}_h^i(h(x_d^i, x_c))\} \text{ for all } x_c \in S_c \quad (4.9)$$

and the equality

$$\bar{\ell}_g^j(g(\bar{x})) = 0 \quad (4.10)$$

are fulfilled.

*Proof.* By Theorem 4.1, for every  $i \in N$ , there exist a real number  $\mu^i \geq 0$  and continuous linear functionals  $\ell_g^i \in C^*$  and  $\ell_h^i \in Z^*$  with  $(\mu^i, \ell_g^i, \ell_h^i) \neq (0, 0_{Y^*}, 0_{Z^*})$ , and the inequality (4.2) and the equality (4.3) are fulfilled. We first show that

$$\mu^i > 0 \text{ for all } i \in N. \quad (4.11)$$

Assume that there exists some  $k \in N$  with  $\mu^k = 0$ . Then we conclude with the additional assumption (4.8) for some  $x_c \in S_c$

$$\begin{aligned} 0 &> \ell_g^k(g(x_d^k, x_c)) + \ell_h^k(h(x_d^k, x_c)) \\ &= \mu^k(f(x_d^k, x_c) - f(\bar{x})) + \ell_g^k(g(x_d^k, x_c)) + \ell_h^k(h(x_d^k, x_c)) \\ &\geq \inf_{i \in N} \{ \mu^i (f(x_d^i, x_c) - f(\bar{x})) + \ell_g^i(g(x_d^i, x_c)) + \ell_h^i(h(x_d^i, x_c)) \}, \end{aligned}$$

which contradicts (4.2). So, (4.11) is proven.

For every  $i \in N$  we now set  $\bar{\ell}_g^i := \frac{1}{\mu^i} \ell_g^i \in C^*$  and  $\bar{\ell}_h^i := \frac{1}{\mu^i} \ell_h^i \in Z^*$ . From the inequality (4.2) it then follows

$$0 \leq \inf_{i \in N} \{f(x_d^i, x_c) - f(\bar{x}) + \bar{\ell}_g^i(g(x_d^i, x_c)) + \bar{\ell}_h^i(h(x_d^i, x_c))\} \text{ for all } x_c \in S_c,$$

which implies

$$f(\bar{x}) \leq \inf_{i \in N} \{f(x_d^i, x_c) + \bar{\ell}_g^i(g(x_d^i, x_c)) + \bar{\ell}_h^i(h(x_d^i, x_c))\} \text{ for all } x_c \in S_c.$$

Hence, the inequality (4.9) is shown. The equality (4.10) immediately follows from the equality (4.3).  $\square$

Condition (4.8) ensures that the multipliers  $\mu^i$  (for all  $i \in N$ ) are nonzero, i.e., it is a constraint qualification (CQ). This CQ extends the well-known Slater condition to discrete-continuous optimization problems.

The necessary optimality condition (4.9) can also be given in a different form.

**Corollary 4.2.** *Let the assumptions of Corollary 4.1 be satisfied, and let  $\bar{x} = (x_d^j, \bar{x}_c)$  (for some  $j \in N$  and some  $\bar{x}_c \in S_c$ ) denote a minimal solution of the discrete-continuous optimization problem (4.1). Then for every  $i \in N$  there exist continuous linear functionals  $\bar{\ell}_g^i \in C^*$  and  $\bar{\ell}_h^i \in Z^*$ , and the inequality*

$$\begin{aligned} & \inf_{i \in N} \{f(x_d^i, \bar{x}_c) + \bar{\ell}_g^i(g(x_d^i, \bar{x}_c)) + \bar{\ell}_h^i(h(x_d^i, \bar{x}_c))\} \\ & \leq \inf_{i \in N} \{f(x_d^i, x_c) + \bar{\ell}_g^i(g(x_d^i, x_c)) + \bar{\ell}_h^i(h(x_d^i, x_c))\} \text{ for all } x_c \in S_c \end{aligned} \quad (4.12)$$

and the equality

$$\bar{\ell}_g^j(g(\bar{x})) = 0$$

are fulfilled.

*Proof.* This corollary is a direct consequence of Corollary 4.1, if we prove (4.12). With (4.9) we obtain with  $\bar{x} = (x_d^j, \bar{x}_c)$

$$\begin{aligned} & \inf_{i \in N} \{f(x_d^i, \bar{x}_c) + \bar{\ell}_g^i(g(x_d^i, \bar{x}_c)) + \bar{\ell}_h^i(h(x_d^i, \bar{x}_c))\} \\ & \leq f(x_d^j, \bar{x}_c) + \bar{\ell}_g^j(g(x_d^j, \bar{x}_c)) + \bar{\ell}_h^j(h(x_d^j, \bar{x}_c)) \\ & = f(\bar{x}) + \underbrace{\bar{\ell}_g^j(g(\bar{x}))}_{=0} + \underbrace{\bar{\ell}_h^j(h(\bar{x}))}_{=0} \\ & = f(\bar{x}) \\ & \leq \inf_{i \in N} \{f(x_d^i, x_c) + \bar{\ell}_g^i(g(x_d^i, x_c)) + \bar{\ell}_h^i(h(x_d^i, x_c))\} \text{ for all } x_c \in S_c, \end{aligned}$$

which has to be shown. □

Corollary 4.2 motivates the following definition of a Lagrange functional in discrete-continuous optimization.

**Definition 4.1.** Let Assumption 4.1 be satisfied. The functional  $L : S_c \times \prod_{i \in N} C' \times \prod_{i \in N} Z' \rightarrow \mathbb{R}$  with

$$\begin{aligned} L(x_c, (\ell_g^i)_{i \in N}, (\ell_h^i)_{i \in N}) & = \inf_{i \in N} \{f(x_d^i, x_c) + \ell_g^i(g(x_d^i, x_c)) + \ell_h^i(h(x_d^i, x_c))\} \\ & \text{for all } x_c \in S_c, \ell_g^1, \ell_g^2, \dots \in C', \ell_h^1, \ell_h^2, \dots \in Z' \end{aligned}$$

is called *Lagrange functional* of the discrete-continuous optimization problem (4.1).

**Remark 4.2.** Using the Lagrange functional the result of Corollary 4.2 can be simplified. Under the assumptions of Corollary 4.2 for every minimal solution of the discrete-continuous optimization problem (4.1) there are continuous linear functionals  $\bar{\ell}_g^i \in C^*$  (for all  $i \in N$ ) and  $\bar{\ell}_h^i \in Z^*$  (for all  $i \in N$ ) with

$$L(\bar{x}_c, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N}) = \min_{x_c \in S_c} L(x_c, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N}).$$

**Corollary 4.3.** *Let the assumptions of Corollary 4.2 be satisfied, and let  $\bar{x} = (x_d^j, \bar{x}_c)$  (for some  $j \in N$  and some  $\bar{x}_c \in S_c$ ) denote a minimal solution of the discrete-continuous optimization problem (4.1). Then for every  $i \in N$  there exist continuous linear functionals  $\bar{\ell}_g^i \in C^*$  and  $\bar{\ell}_h^i \in Z^*$ , and the Lagrange functional  $L$  fulfills the inequalities*

$$\begin{aligned} L(\bar{x}_c, (\ell_g^i)_{i \in N}, (\ell_h^i)_{i \in N}) & \leq L(\bar{x}_c, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N}) \leq L(x_c, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N}) \\ & \text{for all } x_c \in S_c, \ell_g^1, \ell_g^2, \dots \in C^*, \ell_h^1, \ell_h^2, \dots \in Z^*. \end{aligned} \quad (4.13)$$

*Proof.* The right inequality in (4.13) is shown by Corollary 4.2. For the proof of the left inequality in (4.13) choose for every  $i \in N$  arbitrary linear functionals  $\ell_g^i \in C^*$  and  $\ell_h^i \in Z^*$ . Then we get with Corollary 4.1

$$\begin{aligned} & L(\bar{x}_c, (\ell_g^i)_{i \in N}, (\ell_h^i)_{i \in N}) \\ &= \inf_{i \in N} \{ f(x_d^i, \bar{x}_c) + \ell_g^i(g(x_d^i, \bar{x}_c)) + \ell_h^i(h(x_d^i, \bar{x}_c)) \} \\ &\leq f(\bar{x}) + \underbrace{\ell_g^j(g(\bar{x}))}_{\leq 0} + \underbrace{\ell_h^j(h(\bar{x}))}_{=0} \\ &\leq f(\bar{x}) \\ &\leq L(\bar{x}_c, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N}), \end{aligned}$$

which completes the proof.  $\square$

The inequalities in (4.13) mean that  $(\bar{x}_c, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N})$  is a saddle point of the Lagrange functional  $L$ . Corollary 4.3 implies a necessary optimality condition using the subdifferential of the Lagrangian.

**Corollary 4.4.** *Let the assumptions of Corollary 4.2 be satisfied, and in addition, let  $S_c = X_c$ . Let  $\bar{x} = (x_d^j, \bar{x}_c)$  (for some  $j \in N$  and some  $\bar{x}_c \in X_c$ ) denote a minimal solution of the discrete-continuous optimization problem (4.1). Then for every  $i \in N$  there exist continuous linear functionals  $\bar{\ell}_g^i \in C^*$  and  $\bar{\ell}_h^i \in Z^*$  with*

$$0_{X_c^*} \in \partial_{x_c} L(\bar{x}_c, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N}) \quad (4.14)$$

and

$$\bar{\ell}_g^j(g(\bar{x})) = 0, \quad (4.15)$$

where

$$\begin{aligned} & \partial_{x_c} L(\bar{x}_c, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N}) \\ &:= \{ \ell \in X_c^* \mid L(x_c, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N}) \geq L(\bar{x}_c, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N}) + \ell(x_c - \bar{x}_c) \\ & \quad \text{for all } x_c \in X_c \} \end{aligned}$$

denotes the classical subdifferential of convex analysis (w.r.t.  $x_c$ ) of the Lagrange functional  $L(\cdot, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N})$  at  $\bar{x}_c$ .

*Proof.* By Corollary 4.3, for every  $i \in N$ , there exist continuous linear functionals  $\bar{\ell}_g^i \in C^*$  and  $\bar{\ell}_h^i \in Z^*$  with

$$\begin{aligned} L(x_c, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N}) &\geq L(\bar{x}_c, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N}) + 0_{X_c^*}(x_c - \bar{x}_c) \\ &\text{for all } x_c \in X_c, \end{aligned}$$

and condition (4.14) is proven. Condition (4.15) is already shown in Corollary 4.1.  $\square$

Using the Lagrange functional  $L$ , (4.9) as part of a necessary optimality condition in Corollary 4.1 can be written as

$$f(\bar{x}) \leq L(x_c, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N}) \text{ for all } x_c \in S_c.$$

The next theorem says that this inequality (in a nontopological setting) is even a sufficient optimality condition.

**Theorem 4.2.** *Let Assumption 4.1 be satisfied, and let  $\bar{x} := (x_d^j, \bar{x}_c)$  be a feasible vector of discrete-continuous optimization problem (4.1). If, for every  $i \in N$ , there exist linear functionals  $\bar{\ell}_g^i \in C^j$  and  $\bar{\ell}_h^i \in Z^j$  and if*

$$f(\bar{x}) \leq L(x_c, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N}) \text{ for all } x_c \in S_c \quad (4.16)$$

*is satisfied, then  $\bar{x}$  is a minimal solution of problem (4.1).*

*Proof.* Let  $(x_d^k, x_c) \in S$  (with  $k \in N$ ) be arbitrarily chosen. It follows from (4.16) that

$$\begin{aligned} f(\bar{x}) &\leq L(x_c, (\bar{\ell}_g^i)_{i \in N}, (\bar{\ell}_h^i)_{i \in N}) \\ &= \inf_{i \in N} \{ f(x_d^i, x_c) + \bar{\ell}_g^i(g(x_d^i, x_c)) + \bar{\ell}_h^i(h(x_d^i, x_c)) \} \\ &\leq f(x_d^k, x_c) + \underbrace{\bar{\ell}_g^k(g(x_d^k, x_c))}_{\leq 0} + \underbrace{\bar{\ell}_h^k(h(x_d^k, x_c))}_{=0} \\ &\leq f(x_d^k, x_c). \end{aligned}$$

This implies that  $\bar{x}$  is a minimal solution of problem (4.1).  $\square$

**4.1. Specialization to discrete sets with finite cardinality.** The necessary optimality conditions are now specialized to discrete-continuous optimization problems with finitely many discrete points, i.e. we have  $S_d = \{x_d^1, \dots, x_d^n\}$  for some  $n \in \mathbb{N}$ .

**Assumption 4.2.** Let Assumption 4.1 be satisfied with  $N := \{1, \dots, n\}$  for some  $n \in \mathbb{N}$ .

**Theorem 4.3.** *Let Assumption 4.2 and the assumptions of Corollary 4.1 be satisfied, let  $(X_c, \|\cdot\|)$  be a real normed space, and let  $\bar{x} = (x_d^j, \bar{x}_c)$  (for some  $j \in \{1, \dots, n\}$  and some  $\bar{x}_c \in S_c$ ) denote a minimal solution of discrete-continuous optimization problem (4.1). For every  $i \in \{1, \dots, n\}$ , let the functional  $f(x_d^i, \cdot)$  and the maps  $g(x_d^i, \cdot)$  and  $h(x_d^i, \cdot)$  be Fréchet differentiable at  $\bar{x}_c$ . Then there exist continuous linear functionals  $\bar{\ell}_g^1, \dots, \bar{\ell}_g^n \in C^*$  and  $\bar{\ell}_h^1, \dots, \bar{\ell}_h^n \in Z^*$ , and the inequality*

$$\min_{i \in I(\bar{x}_c)} \{ (f'(x_d^i, \bar{x}_c) + \bar{\ell}_g^i(g'(x_d^i, \bar{x}_c)) + \bar{\ell}_h^i(h'(x_d^i, \bar{x}_c))) (x_c - \bar{x}_c) \} \geq 0 \text{ for all } x_c \in S_c \quad (4.17)$$

*and the equality*

$$\bar{\ell}_g^j(g(\bar{x})) = 0 \quad (4.18)$$

*are fulfilled. Here we use the abbreviation*

$$\begin{aligned} I(\bar{x}_c) &:= \left\{ i \in \{1, \dots, n\} \mid f(x_d^i, \bar{x}_c) + \bar{\ell}_g^i(g(x_d^i, \bar{x}_c)) + \bar{\ell}_h^i(h(x_d^i, \bar{x}_c)) \right. \\ &\quad \left. = \min_{1 \leq k \leq n} \{ f(x_d^k, \bar{x}_c) + \bar{\ell}_g^k(g(x_d^k, \bar{x}_c)) + \bar{\ell}_h^k(h(x_d^k, \bar{x}_c)) \} \right\}. \end{aligned} \quad (4.19)$$

*Proof.* Let  $\bar{x} = (x_d^j, \bar{x}_c)$  (for some  $j \in \{1, \dots, n\}$  and some  $\bar{x}_c \in S_c$ ) be a minimal solution of problem (4.1). By Remark 4.2, there exist continuous linear functionals  $\bar{\ell}_g^1, \dots, \bar{\ell}_g^n \in C^*$  and  $\bar{\ell}_h^1, \dots, \bar{\ell}_h^n \in Z^*$  with

$$L(\bar{x}_c, (\bar{\ell}_g^1, \dots, \bar{\ell}_g^n), (\bar{\ell}_h^1, \dots, \bar{\ell}_h^n)) = \min_{x_c \in S_c} L(x_c, (\bar{\ell}_g^1, \dots, \bar{\ell}_g^n), (\bar{\ell}_h^1, \dots, \bar{\ell}_h^n)),$$

where  $L$  denotes the Lagrange functional, i.e.,

$$\begin{aligned} L(x_c, (\bar{\ell}_g^1, \dots, \bar{\ell}_g^n), (\bar{\ell}_h^1, \dots, \bar{\ell}_h^n)) \\ &= \inf_{1 \leq i \leq n} \{f(x_d^i, x_c) + \bar{\ell}_g^i(g(x_d^i, x_c)) + \bar{\ell}_h^i(h(x_d^i, x_c))\} \\ &= \min_{1 \leq i \leq n} \{f(x_d^i, x_c) + \bar{\ell}_g^i(g(x_d^i, x_c)) + \bar{\ell}_h^i(h(x_d^i, x_c))\} \text{ for all } x_c \in S_c. \end{aligned}$$

Since (4.18) is already shown in Corollary 4.1, it remains to prove (4.17).  $\bar{x}_c$  is a minimal point of the Lagrange functional  $L(\cdot, (\bar{\ell}_g^1, \dots, \bar{\ell}_g^n), (\bar{\ell}_h^1, \dots, \bar{\ell}_h^n))$  on  $S_c$  and, therefore, we obtain the well-known necessary optimality condition

$$L'(\bar{x}_c, (\bar{\ell}_g^1, \dots, \bar{\ell}_g^n), (\bar{\ell}_h^1, \dots, \bar{\ell}_h^n))(x_c - \bar{x}_c) \geq 0 \text{ for all } x_c \in S_c, \quad (4.20)$$

where  $L'$  denotes the Gâteaux derivative (compare [6, Thm. 3.8,(a)] for directional derivatives). For simplification, for every  $i \in \{1, \dots, n\}$ , we use the functional  $\varphi_i : S_c \rightarrow \mathbb{R}$  with

$$\varphi_i(x_c) = f(x_d^i, x_c) + \bar{\ell}_g^i(g(x_d^i, x_c)) + \bar{\ell}_h^i(h(x_d^i, x_c)) \text{ for all } x_c \in S_c.$$

Since  $f(x_d^i, \cdot)$ ,  $\bar{\ell}_g^i(g(x_d^i, \cdot))$  and  $\bar{\ell}_h^i(h(x_d^i, \cdot))$  are Fréchet differentiable at  $\bar{x}_c$ , we obtain the Fréchet derivative  $\varphi_i'(\bar{x}_c)(\cdot)$  with

$$\begin{aligned} \varphi_i'(\bar{x}_c)(d) \\ &= f'(x_d^i, \cdot)(\bar{x}_c)(d) + (\bar{\ell}_g^i(g(x_d^i, \cdot)))'(\bar{x}_c)(d) + (\bar{\ell}_h^i(h(x_d^i, \cdot)))'(\bar{x}_c)(d) \\ &= (f'(x_d^i, \bar{x}_c) + \bar{\ell}_g^i(g'(x_d^i, \bar{x}_c)) + \bar{\ell}_h^i(h'(x_d^i, \bar{x}_c)))(d) \text{ for all } d \in X_c. \end{aligned} \quad (4.21)$$

The Gâteaux derivative of the Lagrange functional at  $\bar{x}_c$  can be written as

$$L'(\bar{x}_c, (\bar{\ell}_g^1, \dots, \bar{\ell}_g^n), (\bar{\ell}_h^1, \dots, \bar{\ell}_h^n))(d) = \left( \min_{1 \leq i \leq n} \{\varphi_i(\cdot)\} \right)'(\bar{x}_c)(d) \text{ for all } d \in X_c. \quad (4.22)$$

If we notice that with

$$I(\bar{x}_c) = \left\{ i \in \{1, \dots, n\} \mid \varphi_i(\bar{x}_c) = \min_{1 \leq k \leq n} \{\varphi_k(\bar{x}_c)\} \right\},$$

we have

$$\min_{1 \leq k \leq n} \{\varphi_k(\bar{x}_c)\} = \varphi_i(\bar{x}_c) \text{ for all } i \in I(\bar{x}_c)$$

and because of the continuity of the functionals  $\varphi_1, \dots, \varphi_n$  at  $\bar{x}_c$

$$\begin{aligned} \min_{1 \leq i \leq n} \{\varphi_i(\bar{x}_c + \lambda d)\} &= \min_{i \in I(\bar{x}_c)} \{\varphi_i(\bar{x}_c + \lambda d)\} \\ &\text{for all } d \in X_c \text{ and all } \lambda \text{ sufficiently close to } 0, \end{aligned}$$

we then get the Gâteaux derivative

$$\begin{aligned} \left( \min_{1 \leq i \leq n} \{\varphi_i(\cdot)\} \right)'(\bar{x}_c)(d) &= \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \left( \min_{1 \leq i \leq n} \{\varphi_i(\bar{x}_c + \lambda d)\} - \min_{1 \leq i \leq n} \{\varphi_i(\bar{x}_c)\} \right) \\ &= \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \left( \min_{i \in I(\bar{x}_c)} \{\varphi_i(\bar{x}_c + \lambda d)\} - \min_{1 \leq i \leq n} \{\varphi_i(\bar{x}_c)\} \right) \\ &= \lim_{\lambda \rightarrow 0} \min_{i \in I(\bar{x}_c)} \left\{ \frac{1}{\lambda} \left( \varphi_i(\bar{x}_c + \lambda d) - \varphi_i(\bar{x}_c) \right) \right\} \\ &= \min_{i \in I(\bar{x}_c)} \{\varphi_i'(\bar{x}_c)(d)\} \text{ for all } d \in X_c \end{aligned}$$

(compare [4, Satz 6.22] for the case of directional derivatives) where  $\varphi'_i$  (with  $i \in I(\bar{x}_c)$ ) denotes the Fréchet derivative, which equals the Gâteaux derivative. With (4.22), we then conclude

$$L'(\bar{x}_c, (\bar{\ell}_g^1, \dots, \bar{\ell}_g^n), (\bar{\ell}_h^1, \dots, \bar{\ell}_h^n))(d) = \min_{i \in I(\bar{x}_c)} \{\varphi'_i(\bar{x}_c)(d)\} \text{ for all } d \in X_c.$$

From (4.20) and (4.21), we obtain

$$\min_{i \in I(\bar{x}_c)} \{(f'(x_d^i, \bar{x}_c) + \bar{\ell}_g^i(g'(x_d^i, \bar{x}_c)) + \bar{\ell}_h^i(h'(x_d^i, \bar{x}_c)))(x_c - \bar{x}_c)\} \geq 0 \text{ for all } x_c \in S_c,$$

which has to be shown.  $\square$

If  $S_c$  equals  $X_c$ , we can simplify the result of Theorem 4.3.

**Corollary 4.5.** *Let the assumptions of Theorem 4.3 be satisfied and, in addition, let  $S_c = X_c$ . Again, let  $\bar{x} = (x_d^j, \bar{x}_c)$  (for some  $j \in \{1, \dots, n\}$  and some  $\bar{x}_c \in S_c$ ) denote a minimal solution of discrete-continuous optimization problem (4.1). Then there exist continuous linear functionals  $\bar{\ell}_g^1, \dots, \bar{\ell}_g^n \in C^*$  and  $\bar{\ell}_h^1, \dots, \bar{\ell}_h^n \in Z^*$ , and the inequality*

$$f'(x_d^j, \bar{x}_c) + \bar{\ell}_g^j(g'(x_d^j, \bar{x}_c)) + \bar{\ell}_h^j(h'(x_d^j, \bar{x}_c)) = 0_{X_c^*} \text{ for all } i \in I(\bar{x}_c) \quad (4.23)$$

(where  $I(\bar{x}_c)$  is defined in (4.19)) and the equality

$$\bar{\ell}_g^j(g(\bar{x})) = 0 \quad (4.24)$$

are fulfilled.

*Proof.* We write (4.17) as

$$(f'(x_d^i, \bar{x}_c) + \bar{\ell}_g^i(g'(x_d^i, \bar{x}_c)) + \bar{\ell}_h^i(h'(x_d^i, \bar{x}_c)))(x_c - \bar{x}_c) \geq 0$$

$$\text{for all } i \in I(\bar{x}_c) \text{ and } x_c \in S_c.$$

With the linearity of the Fréchet derivative, we then get

$$f'(x_d^i, \bar{x}_c) + \bar{\ell}_g^i(g'(x_d^i, \bar{x}_c)) + \bar{\ell}_h^i(h'(x_d^i, \bar{x}_c)) = 0_{X_c^*} \text{ for all } i \in I(\bar{x}_c)$$

and the condition (4.23) is shown.  $\square$

The optimality conditions (4.23), (4.24), which are valid in infinite dimensional spaces are already very close to the Karush-Kuhn-Tucker (KKT) conditions in finite dimensional spaces. KKT conditions are formulated in the next corollary, which is a direct consequence of Corollary 4.5.

**Corollary 4.6.** *Let the assumptions of Theorem 4.3 be satisfied and, in addition, let  $X_d = \mathbb{R}^{n_d}$ ,  $X_c = S_c = \mathbb{R}^{n_c}$ ,  $Y = \mathbb{R}^m$ ,  $Z = \mathbb{R}^p$  and  $C = \mathbb{R}_+^{n_c}$ . Again, let  $\bar{x} = (x_d^j, \bar{x}_c)$  (for some  $j \in \{1, \dots, n\}$  and some  $\bar{x}_c \in S_c$ ) denote a minimal solution of the discrete-continuous optimization problem (4.1). For every  $i \in \{1, \dots, n\}$  let  $f(x_d^i, \cdot)$ ,  $g_1(x_d^i, \cdot), \dots, g_m(x_d^i, \cdot)$  and  $h_1(x_d^i, \cdot), \dots, h_p(x_d^i, \cdot)$  be differentiable at  $\bar{x}_c$ . Then there exist vectors  $u^1, \dots, u^n \in \mathbb{R}_+^{n_c}$  and  $v^1, \dots, v^n \in \mathbb{R}^{n_c}$  with*

$$\nabla_{x_c} f(x_d^i, \bar{x}_c) + \sum_{k=1}^m u_k^i \nabla_{x_c} g_k(x_d^i, \bar{x}_c) + \sum_{k=1}^p v_k^i \nabla_{x_c} h_k(x_d^i, \bar{x}_c) = 0_{\mathbb{R}^{n_c}} \text{ for all } i \in I(\bar{x}_c)$$

(where  $I(\bar{x}_c)$  is defined in (4.19)) and

$$u_k^j g_k(x_d^j, \bar{x}_c) = 0 \text{ for all } k \in \{1, \dots, m\}.$$



**Example 4.1.** In Assumption 4.2 we set  $X_d = X_c = Y = \mathbb{R}$ ,  $S_c = \mathbb{R}$  and  $S_d = \{-10, -8, -6, \dots, 10\}$  (i.e. we have  $n = 11$ ). Let  $\psi : S_d \rightarrow \mathbb{R}$  be an arbitrary function with

$$\psi(x_d^i) > 0 \text{ for all } i \in \{1, \dots, n\},$$

and let  $x_d^j$  for some  $j \in \{1, \dots, 11\}$  be a unique minimal point of  $\psi$ , i.e.,

$$\psi(x_d^j) < \psi(x_d^i) \text{ for all } i \in \{1, \dots, 11\} \setminus \{j\}.$$

Then we consider the discrete-continuous optimization problem

$$\begin{aligned} & \min \psi(x_d) \cdot x_c^2 \\ & \text{subject to} \\ & x_c \geq 1 \\ & x_d \in \{-10, -8, \dots, 10\}, \quad x_c \in \mathbb{R}. \end{aligned}$$

It is evident that  $\bar{x} := (x_d^j, 1)$  is the unique minimal solution of this optimization problem. Next, we define the objective function  $f : S_d \times \mathbb{R} \rightarrow \mathbb{R}$  with

$$f(x_d, x_c) = \psi(x_d) \cdot x_c^2 \text{ for all } x_d \in S_d \text{ and all } x_c \in \mathbb{R}$$

and the constraint function  $g : S_d \times \mathbb{R} \rightarrow \mathbb{R}$  with

$$g(x_d, x_c) = 1 - x_c \text{ for all } x_d \in S_d \text{ and all } x_c \in \mathbb{R}.$$

Notice that there is no equality constraint. Then the Lagrangian  $L$  can be written as

$$\begin{aligned} L(x_c, (u^1, \dots, u^{11})) &= \inf_{1 \leq i \leq 11} \{f(x_d^i, x_c) + u^i g(x_d^i, x_c)\} \\ &= \min_{1 \leq i \leq 11} \{\psi(x_d^i) \cdot x_c^2 + u^i (1 - x_c)\} \\ &\text{for all } x_c \in \mathbb{R} \text{ and all } u^1, \dots, u^{11} \in \mathbb{R}_+. \end{aligned}$$

For the determination of the index set  $I(1)$ , we notice that

$$\begin{aligned} f(x_d^j, 1) + u^j g(x_d^j, 1) &= \psi(x_d^j) \\ &= \min_{1 \leq i \leq 11} \psi(x_d^i) \\ &= \min_{1 \leq i \leq 11} \{f(x_d^i, 1) + u^i g(x_d^i, 1)\}, \end{aligned}$$

i.e.,  $j \in I(1)$ . Since  $x_d^j$  is a unique minimal point of  $\psi$ , we obtain, for every  $i \in \{1, \dots, 11\} \setminus \{j\}$ ,

$$\begin{aligned} f(x_d^i, 1) + u^i g(x_d^i, 1) &= \psi(x_d^i) \\ &> \min_{1 \leq k \leq 11} \psi(x_d^k) \\ &= \min_{1 \leq k \leq 11} \{f(x_d^k, 1) + u^k g(x_d^k, 1)\}, \end{aligned}$$

i.e.,  $i \notin I(1)$ . Hence, we conclude  $I(1) = \{j\}$ . Then we get the KKT condition

$$\begin{aligned} 0 &= \nabla_{x_c} f(x_d^j, 1) + u^j \nabla_{x_c} g(x_d^j, 1) \\ &= 2\psi(x_d^j) - u^j, \end{aligned}$$

which gives  $u^j = 2\psi(x_d^j) > 0$ . So, the minimal solution  $\bar{x}$  fulfills the necessary optimality conditions in Corollary 4.6.

Finally, we show under which assumptions the optimality conditions in Theorem 4.3 are sufficient optimality conditions, at least in a local sense.

**Theorem 4.4.** *Let Assumption 4.2 be satisfied, and let  $\bar{x} = (x_d^j, \bar{x}_c) \in S$  (for some  $j \in \{1, \dots, n\}$ ) and some  $\bar{x}_c \in S_c$ ) be a feasible vector of discrete-continuous optimization problem (4.1). For every  $i \in \{1, \dots, n\}$  let the functional  $f(x_d^i, \cdot)$  and the maps  $g(x_d^i, \cdot)$  and  $h(x_d^i, \cdot)$  be Fréchet differentiable at  $\bar{x}_c$ . Let there exist continuous linear functionals  $\bar{\ell}_g^1, \dots, \bar{\ell}_g^n \in C^*$  and  $\bar{\ell}_h^1, \dots, \bar{\ell}_h^n \in Z^*$ , and let the inequality (4.17) and the equality (4.18) be satisfied. Moreover, let there exist some  $\varepsilon > 0$  so that*

$$j \in I(x_c) \text{ for all } x_c \in S_c \cap B(\bar{x}_c, \varepsilon)$$

(where  $I(x_c)$  is defined in (4.19) and  $B(\bar{x}_c, \varepsilon)$  denotes the ball with center  $\bar{x}_c$  and radius  $\varepsilon$ ) and the functional  $f(x_d^j, \cdot) + \bar{\ell}_g^j(g(x_d^j, \cdot)) + \bar{\ell}_h^j(h(x_d^j, \cdot))$  is convex on  $S_c \cap B(\bar{x}_c, \varepsilon)$ . Then  $\bar{x}$  is a locally minimal solution of the discrete-continuous optimization problem.

*Proof.* From the proof of Theorem 4.3, it is evident that the inequality (4.17) is equivalent to (4.20). Since by assumption the Lagrange functional  $L(\cdot, (\bar{\ell}_g^1, \dots, \bar{\ell}_g^n), (\bar{\ell}_h^1, \dots, \bar{\ell}_h^n))$  with

$$\begin{aligned} & L(x_c, (\bar{\ell}_g^1, \dots, \bar{\ell}_g^n), (\bar{\ell}_h^1, \dots, \bar{\ell}_h^n)) \\ &= \min_{1 \leq i \leq n} \{f(x_d^i, x_c) + \bar{\ell}_g^i(g(x_d^i, x_c)) + \bar{\ell}_h^i(h(x_d^i, x_c))\} \\ &= f(x_d^j, x_c) + \bar{\ell}_g^j(g(x_d^j, x_c)) + \bar{\ell}_h^j(h(x_d^j, x_c)) \text{ for all } x_c \in S_c \cap B(\bar{x}_c, \varepsilon) \end{aligned}$$

is convex on  $S_c \cap B(\bar{x}_c, \varepsilon)$ , we conclude with [6, Theorem 3.16] and (4.20)

$$\begin{aligned} & L(x_c, (\bar{\ell}_g^1, \dots, \bar{\ell}_g^n), (\bar{\ell}_h^1, \dots, \bar{\ell}_h^n)) \\ & \geq L(\bar{x}_c, (\bar{\ell}_g^1, \dots, \bar{\ell}_g^n), (\bar{\ell}_h^1, \dots, \bar{\ell}_h^n)) + \underbrace{L'(\bar{x}_c, (\bar{\ell}_g^1, \dots, \bar{\ell}_g^n), (\bar{\ell}_h^1, \dots, \bar{\ell}_h^n))}_{\geq 0} (x_c - \bar{x}_c) \\ & \geq L(\bar{x}_c, (\bar{\ell}_g^1, \dots, \bar{\ell}_g^n), (\bar{\ell}_h^1, \dots, \bar{\ell}_h^n)) \\ &= f(x_d^j, \bar{x}_c) + \bar{\ell}_g^j(g(x_d^j, \bar{x}_c)) + \bar{\ell}_h^j(h(x_d^j, \bar{x}_c)) \\ &= \underbrace{f(\bar{x})}_{=0} + \bar{\ell}_g^j(g(\bar{x})) + \bar{\ell}_h^j(h(\bar{x})) \\ &= f(\bar{x}) \text{ for all } x_c \in S_c \cap B(\bar{x}_c, \varepsilon). \end{aligned}$$

If we apply Theorem 4.2 with  $S_c \cap B(\bar{x}_c, \varepsilon)$  instead of  $S_c$ ,  $\bar{x}$  is a locally minimal solution of the discrete-continuous optimization problem.  $\square$

**Remark 4.3.** Since in general, the constraint set of a discrete-continuous optimization problem is non-convex (compare Figure 3), a locally minimal solution does not need to be a (globally) minimal solution, if the objective functional is convex (compare [6, Thm. 2.16]).

**Example 4.2.** We investigate the discrete-continuous optimization problem in Example 4.1. Since  $x_d^j$  (for some  $j \in \{1, \dots, 11\}$ ) is a unique minimal point of  $\psi$ , we have

$$\psi(x_d^j) < \min_{\substack{1 \leq i \leq 11 \\ i \neq j}} \{\psi(x_d^i)\}.$$

It is already discussed in Example 4.1 that  $\bar{x} = (x_d^j, 1)$  fulfills the KKT conditions. Notice that the Lagrange multipliers  $u^1, \dots, u^{11} \geq 0$  are given as  $u^i = 2\psi(x_d^i)$  for all  $i \in \{1, \dots, 11\}$ . For a sufficiently small  $\varepsilon > 0$  we then conclude

$$\begin{aligned} L(x_c, (u^1, \dots, u^{11})) &= \min_{1 \leq i \leq 11} \{ \psi(x_d^i) \cdot x_c^2 + u^i \cdot (1 - x_c) \} \\ &= \min_{1 \leq i \leq 11} \{ \psi(x_d^i) \cdot x_c^2 + 2\psi(x_d^i) \cdot (1 - x_c) \} \\ &= \min_{1 \leq i \leq 11} \{ \psi(x_d^i) + \psi(x_d^i) \cdot (x_c - 1)^2 \} \\ &= \psi(x_d^j) + \psi(x_d^j) \cdot (x_c - 1)^2 \text{ for all } x_c \in [1 - \varepsilon, 1 + \varepsilon]. \end{aligned}$$

This means that

$$j \in I(x_c) \text{ for all } x_c \in [1 - \varepsilon, 1 + \varepsilon].$$

The functional  $f(x_d^j, \cdot) + u^j \cdot g(x_d^j, \cdot)$  with

$$f(x_d^j, x_c) + u^j \cdot g(x_d^j, x_c) = \psi(x_d^j) \cdot x_c^2 + u^j \cdot (1 - x_c) \text{ for all } x_c \in \mathbb{R}$$

is convex so that all assumptions of Theorem 4.4 are fulfilled. Hence,  $\bar{x} = (x_d^j, 1)$  is a locally minimal solution of the considered discrete-continuous optimization problem.

## 5. DUALITY

Under Assumption 4.1, we continue to investigate the discrete-continuous optimization problem (4.1)

$$\begin{aligned} &\min f(x_d, x_c) \\ &\text{subject to} \\ &g(x_d, x_c) \in -C \\ &h(x_d, x_c) = 0_Z \\ &(x_d, x_c) \in S_d \times S_c, \end{aligned} \tag{P}$$

which is now called *primal problem (P)*. To this problem, we associate a so-called *dual problem (D)* given by

$$\begin{aligned} &\max \inf_{\substack{i \in N \\ x_c \in S_c}} \{ f(x_d^i, x_c) + \ell_g^i(g(x_d^i, x_c)) + \ell_h^i(h(x_d^i, x_c)) \} \\ &\text{subject to} \\ &\ell_g^i \in C' \text{ for all } i \in N \\ &\ell_h^i \in Z' \text{ for all } i \in N. \end{aligned} \tag{D}$$

Notice that the variables of the dual problem (D) are elements of the dual spaces  $Y'$  and  $Z'$ , respectively.

A first relationship between the primal and dual problem is given by the following *weak duality theorem*, which is simple to prove.

**Theorem 5.1.** *Let Assumption 4.1 be satisfied. For every  $\hat{x} \in S$  (feasible element of (P)) and for all  $\hat{\ell}_g^1, \hat{\ell}_g^2, \dots \in C'$  and  $\hat{\ell}_h^1, \hat{\ell}_h^2, \dots \in Z'$  (feasible elements of (D)) the following inequality holds*

$$\inf_{\substack{i \in N \\ x_c \in S_c}} \{ f(x_d^i, x_c) + \ell_g^i(g(x_d^i, x_c)) + \ell_h^i(h(x_d^i, x_c)) \} \leq f(\hat{x}).$$

*Proof.* For an arbitrary feasible vector  $\hat{x} = (x_d^j, \hat{x}_c)$  (for some  $j \in N$  and some  $\hat{x}_c \in S_c$ ) and arbitrary linear functionals  $\hat{\ell}_g^1, \hat{\ell}_g^2, \dots \in C'$  and  $\hat{\ell}_h^1, \hat{\ell}_h^2, \dots \in Z'$  we conclude

$$\begin{aligned} \inf_{\substack{i \in N \\ x_c \in S_c}} \{f(x_d^i, x_c) + \hat{\ell}_g^i(g(x_d^i, x_c)) + \hat{\ell}_h^i(h(x_d^i, x_c))\} &\leq f(\hat{x}) + \underbrace{\hat{\ell}_g^j(g(\hat{x}))}_{\leq 0} + \underbrace{\hat{\ell}_h^j(h(\hat{x}))}_{=0} \\ &\leq f(\hat{x}), \end{aligned}$$

which has to be shown.  $\square$

The next *strong duality theorem* answers the question under which assumptions the dual problem is solvable.

**Theorem 5.2.** *Let the assumptions of Corollary 4.1 be satisfied and let  $\bar{x} = (x_d^j, \bar{x}_c) \in S$  (for some  $j \in N$  and some  $\bar{x}_c \in S_c$ ) denote a minimal solution of primal problem (P). Then dual problem (D) is solvable and the extremal values of both problems are equal.*

*Proof.* By Corollary 4.1, we get for some continuous linear functionals  $\bar{\ell}_g^i \in C^*$  (with  $i \in N$ ) and  $\bar{\ell}_h^i \in Z^*$  (with  $i \in N$ )

$$f(\bar{x}) \leq \inf_{i \in N} \{f(x_d^i, x_c) + \bar{\ell}_g^i(g(x_d^i, x_c)) + \bar{\ell}_h^i(h(x_d^i, x_c))\} \text{ for all } x_c \in S_c,$$

which can also be written as

$$f(\bar{x}) \leq \inf_{\substack{i \in N \\ x_c \in S_c}} \{f(x_d^i, x_c) + \bar{\ell}_g^i(g(x_d^i, x_c)) + \bar{\ell}_h^i(h(x_d^i, x_c))\}.$$

By the weak duality theorem, we obtain

$$f(\bar{x}) = \inf_{\substack{i \in N \\ x_c \in S_c}} \{f(x_d^i, x_c) + \bar{\ell}_g^i(g(x_d^i, x_c)) + \bar{\ell}_h^i(h(x_d^i, x_c))\}.$$

This means that dual problem (D) is solvable and the extremal values of both problems are equal.  $\square$

**5.1. Specialization to extendedly linear problems.** Since the dual problem (D) seems to be quite general, we now investigate linear discrete-continuous optimization problems in an extended form. In this case the dual problem (D) has a simpler structure.

**Assumption 5.1.** Let  $X_d, X_c, Y$  and  $Z$  be real linear spaces, and let  $C_Y \subset Y$  and  $C_{X_c} \subset X_c$  be convex cones in  $Y$  and  $X_c$ , respectively. Let  $S_d$  be a discrete subset of  $X_d$ , i.e.  $S_d = \{x_d^i\}_{i \in N}$  for  $N := \{1, 2, \dots, n\}$  (with  $n \in \mathbb{N}$ ) or  $N := \mathbb{N}$  with elements  $x_d^1, x_d^2, \dots \in X_d$ . Let  $c \in X_c'$  be a linear functional, let  $A : X_c \rightarrow Y$  be a linear map, and let  $b \in Y$  be a given element. Let  $f_d : S_d \rightarrow \mathbb{R}$  and  $g_d : S_d \rightarrow Y$  be given maps and let the constraint set

$$S := \{(x_d, x_c) \in S_d \times C_{X_c} \mid g_d(x_d) + A(x_c) - b \in C_Y\}$$

be nonempty.

Notice that  $f_d$  and  $g_d$  do not need to be linear or convex. Under Assumption 5.1, we now consider the discrete-continuous optimization problem

$$\begin{aligned} & \min f_d(x_d) + c(x_c) \\ & \text{subject to} \\ & g_d(x_d) + A(x_c) - b \in C_Y \\ & x_d \in S_d, \quad x_c \in C_{X_c}. \end{aligned} \tag{ELP}$$

This is an *extendedly linear primal problem*, which is more general than standard linear optimization problems. According to our general formulation of the dual problem (D) we associate the *extendedly linear dual problem*

$$\begin{aligned} & \max \inf_{\substack{i \in N \\ x_c \in C_{X_c}}} \{ f_d(x_d^i) + c(x_c) + \ell^i(-g_d(x_d^i) - A(x_c) + b) \} \\ & \text{subject to} \\ & \ell^i \in C_{Y'} \text{ for all } i \in N \end{aligned} \tag{ELD}$$

to the original problem (ELP). Then the following lemma simplifies this problem in a way that it is more suitable in applications.

**Lemma 5.1.** *Under Assumption 5.1, the extendedly linear dual problem (ELD) is equivalent to*

$$\begin{aligned} & \max \inf_{i \in N} \{ f_d(x_d^i) - (\ell^i \circ g_d)(x_d^i) + \ell^i(b) \} \\ & \text{subject to} \\ & c - \ell^i \circ A \in C_{X_c'} \text{ for all } i \in N \\ & \ell^i \in C_{Y'} \text{ for all } i \in N. \end{aligned} \tag{5.1}$$

*Proof.* (a) Problem (ELD) is equivalent to

$$\begin{aligned} & \max \lambda \\ & \text{subject to} \\ & \lambda \leq \inf_{\substack{i \in N \\ x_c \in C_{X_c}}} \{ f_d(x_d^i) + c(x_c) + \ell^i(-g_d(x_d^i) - A(x_c) + b) \} \\ & \lambda \in \mathbb{R}, \quad \ell^i \in C_{Y'} \text{ for all } i \in N. \end{aligned} \tag{5.2}$$

The inequality constraint in problem (5.2) can also be written as

$$\lambda \leq f_d(x_d^i) + c(x_c) + \ell^i(-g_d(x_d^i) - A(x_c) + b) \text{ for all } i \in N \text{ and all } x_c \in C_{X_c}$$

or

$$\begin{aligned} & c(x_c) - (\ell^i \circ A)(x_c) \\ & \geq \lambda - f_d(x_d^i) + (\ell^i \circ g_d)(x_d^i) - \ell^i(b) \text{ for all } i \in N \text{ and all } x_c \in C_{X_c}. \end{aligned} \tag{5.3}$$

Since the right hand side of the inequality (5.3) does not depend on  $x_c$ , we conclude

$$c(x_c) - (\ell^i \circ A)(x_c) \geq 0 \text{ for all } x_c \in C_{X_c},$$

i.e.,

$$c - \ell^i \circ A \in C_{X_c'},$$

and the constraint in problem (5.1) is shown. If we set  $x_c = 0_{X_c}$  in (5.3), we obtain

$$\lambda \leq f_d(x_d^i) - (\ell^i \circ g_d)(x_d^i) + \ell^i(b) \text{ for all } i \in N$$

or

$$\lambda \leq \inf_{i \in N} \{f_d(x_d^i) - (\ell^i \circ g_d)(x_d^i) + \ell^i(b)\}.$$

By maximization of  $\lambda$ , we then get the formulation in (5.1).

(b) In a similar way we rewrite problem (5.1) in an equivalent form and we get the constraints

$$\begin{aligned} c(x_c) - (\ell^i \circ A)(x_c) \\ &\geq 0 \\ &\geq \lambda - f_d(x_d^i) + (\ell^i \circ g_d)(x_d^i) - \ell^i(b) \text{ for all } i \in N \text{ and all } x_c \in C_{X_c}. \end{aligned}$$

This leads to problem (ELD). □

**Remark 5.1.** In the finite dimensional case, i.e.,  $X_d = \mathbb{R}^{n_d}$ ,  $X_c = \mathbb{R}^{n_c}$ ,  $Y = \mathbb{R}^m$ ,  $A \in \mathbb{R}^{(m, n_c)}$ ,  $c \in \mathbb{R}^{n_c}$  and  $b \in \mathbb{R}^m$ , the equivalent dual problem (5.1) can be written as

$$\begin{aligned} \max \inf_{i \in N} \{f_d(x_d^i) + \ell^{iT}(b - g_d(x_d^i))\} \\ \text{subject to} \\ c - A^T \ell^i \in C_{\mathbb{R}^{n_c}}^* \text{ for all } i \in N \\ \ell^i \in C_{\mathbb{R}^m}^* \text{ for all } i \in N. \end{aligned}$$

**Example 5.1.** As a very simple example, we investigate a linear problem with only discrete variables:

$$\begin{aligned} \min -2(x_d)_1 + (x_d)_2 \\ \text{subject to} \\ (x_d)_1 + (x_d)_2 - 3 \leq 0 \\ x_d \in S_d := \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}. \end{aligned} \tag{5.4}$$

Here we have  $X_d = \mathbb{R}^2$ ,  $Y = \mathbb{R}$ ,  $C_Y = \mathbb{R}_+$ ,  $f_d(\cdot) = (-2, 1)(\cdot)$ ,  $g_d(\cdot) = (-1, -1)(\cdot)$  and  $b = -3$ . By Lemma 5.1, the equivalent dual problem reads

$$\max_{\ell^1, \dots, \ell^6 \geq 0} \min_{1 \leq i \leq 6} \{(-2, 1)x_d^i + \ell^i(1, 1)x_d^i - 3\ell^i\}. \tag{5.5}$$

The objective function of this problem can be written as

$$\begin{aligned} \min_{1 \leq i \leq 6} \{(-2, 1)x_d^i + \ell^i(1, 1)x_d^i - 3\ell^i\} \\ = \min \{-3\ell^1, 4 + \ell^2, -4 + 5\ell^3, -8 + \ell^4, 0, -3\}, \end{aligned}$$

and problem (5.5) is equivalent to

$$\begin{aligned}
& \max \lambda \\
& \text{subject to} \\
& \lambda \leq -3\ell^1 \\
& \lambda \leq 4 + \ell^2 \\
& \lambda \leq -4 + 5\ell^3 \\
& \lambda \leq -8 + \ell^4 \\
& \lambda \leq 0 \\
& \lambda \leq -3 \\
& \lambda \in \mathbb{R}, \ell^1, \dots, \ell^4 \geq 0.
\end{aligned}$$

Notice that the first, second and fifth inequality constraint are redundant so that the variables  $\ell^1$  and  $\ell^2$  can also be dropped. Then we get the equivalent dual problem

$$\begin{aligned}
& \max \lambda \\
& \text{subject to} \\
& \lambda \leq -4 + 5\ell^3 \\
& \lambda \leq -8 + \ell^4 \\
& \lambda \leq -3 \\
& \lambda \in \mathbb{R}, \ell^3, \ell^4 \geq 0.
\end{aligned}$$

It is evident that the maximal value of this problem equals  $-3$ . Notice that the primal problem (5.4) has only the three feasible points  $(0,0)$ ,  $(1,2)$  and  $(2,1)$  with the minimal value  $-3$ . Hence, the extremal values of the primal and dual problem coincide. In contrast to the classical Lagrange theory in continuous optimization there is no duality gap because we work with more than one Lagrange multiplier.

## 6. APPLICATION TO DISCRETE-CONTINUOUS SEMIDEFINITE AND COPOSITIVE OPTIMIZATION

It is well-known that semidefinite and copositive optimization have important applications in practice (e.g., compare [2, 6]). Now we extend these problems in such a way that the variables may be discrete and continuous. Since these problems are optimization problems in a finite dimensional Hilbert space, the theory of this paper can be used for this problem class.

**Assumption 6.1.** Let  $\mathcal{S}^k$  (for some  $k \in \mathbb{N}$ ) denote the real linear space of symmetric  $(k,k)$ -matrices. Let  $C \subset \mathcal{S}^k$  be either the Löwner cone, i.e.,

$$C := \left\{ M \in \mathcal{S}^k \mid M \text{ is positive semidefinite} \right\} =: \mathcal{S}_+^k, \quad (6.1)$$

or the copositive cone, i.e.,

$$C := \left\{ M \in \mathcal{S}^k \mid y^T M y \geq 0 \text{ for all } y \in \mathbb{R}_+^k \right\}. \quad (6.2)$$

Let  $m_d, m_c \in \mathbb{N}$  be given integers, and let  $S_d := \{x_d^1, x_d^2, \dots, x_d^n\} \subset \mathbb{R}^{m_d}$  be a discrete set with  $n \in \mathbb{N}$ . Let  $f : S_d \times \mathbb{R}^{m_c} \rightarrow \mathbb{R}$  and  $G : S_d \times \mathbb{R}^{m_c} \rightarrow \mathcal{S}^k$  be given maps, and let the constraint set

$$S := \{(x_d, x_c) \in S_d \times \mathbb{R}^{m_c} \mid G(x_d, x_c) \in -C\}$$

be nonempty.

It is well-known that the space  $\mathcal{S}^k$  is a finite dimensional Hilbert space with the scalar product  $\langle \cdot, \cdot \rangle$  defined by

$$\langle A, B \rangle = \text{trace}(A \cdot B) \text{ for all } A, B \in \mathcal{S}^k.$$

Under Assumption 6.1 we investigate the semidefinite/copositive optimization problem

$$\begin{aligned} & \min f(x_d, x_c) \\ & \text{subject to} \\ & G(x_d, x_c) \in -C \\ & (x_d, x_c) \in \mathcal{S}_d \times \mathbb{R}^{m_c}. \end{aligned} \tag{6.3}$$

If the cone  $C$  is given by (6.1), then problem (6.3) is a discrete-continuous semidefinite optimization problem. If  $C$  is given by (6.2), then problem (6.3) is a discrete-continuous copositive optimization problem.

We restrict our investigation to the formulation of an extended KKT condition for problem (6.3).

**Theorem 6.1.** *Let Assumption 6.1 be satisfied and let  $\bar{x} = (x_d^j, \bar{x}_c)$  (for some  $j \in \{1, \dots, n\}$ ) be a minimal solution of discrete-continuous semidefinite/copositive optimization problem (6.3). For every  $i \in \{1, \dots, n\}$ , let the set*

$$B_i := \left\{ \begin{pmatrix} f(x_d^i, x_c) - f(\bar{x}) + \alpha \\ G(x_d^i, x_c) + Y \end{pmatrix} \in \mathbb{R} \times \mathcal{S}^k \mid x_c \in \mathbb{R}^{m_c}, \alpha > 0, Y \in \text{int}(C) \right\}$$

be convex. For every  $i \in \{1, \dots, n\}$ , let the function  $f(x_d^i, \cdot)$  be differentiable at  $\bar{x}_c$  and let the map  $G(x_d^i, \cdot)$  be Fréchet differentiable at  $\bar{x}_c$ . Then there exist real numbers  $\mu^1, \dots, \mu^n \geq 0$  and matrices  $L^1, \dots, L^n \in \mathbb{C}^*$  with  $(\mu^i, L^i) \neq (0, 0_{\mathcal{S}^k})$  for all  $i \in \{1, \dots, n\}$ , and the equality

$$\mu^i \nabla_{x_c} f(x_d^i, \bar{x}_c) + \begin{pmatrix} \langle L^i, G_{(x_c)_1}(\bar{x}) \rangle \\ \vdots \\ \langle L^i, G_{(x_c)_{m_c}}(\bar{x}) \rangle \end{pmatrix} = 0_{\mathbb{R}^{m_c}} \text{ for all } i \in I(\bar{x}_c)$$

with

$$G_{(x_c)_i} := \begin{pmatrix} \frac{\partial}{\partial (x_c)_i} G_{11} & \cdots & \frac{\partial}{\partial (x_c)_i} G_{1k} \\ \vdots & & \vdots \\ \frac{\partial}{\partial (x_c)_i} G_{k1} & \cdots & \frac{\partial}{\partial (x_c)_i} G_{kk} \end{pmatrix} \text{ for all } i \in \{1, \dots, m_c\}.$$

and

$$\begin{aligned} I(\bar{x}_c) & := \left\{ i' \in \{1, \dots, n\} \mid f(x_d^{i'}, \bar{x}_c) + \langle L^{i'}, G(x_d^{i'}, \bar{x}_c) \rangle \right. \\ & \quad \left. = \min_{1 \leq i \leq n} \{ f(x_d^i, \bar{x}_c) + \langle L^i, G(x_d^i, \bar{x}_c) \rangle \} \right\} \end{aligned}$$

is fulfilled together with the equality

$$\langle L^j, G(\bar{x}) \rangle = 0.$$

If in addition to the above assumptions the condition

$$\forall i \in \{1, \dots, n\} \exists x_c \in \mathbb{R}^{m_c} : 0 > \langle L^i, G(x_d^i, x_c) \rangle$$

is satisfied, then it follows

$$\mu^i > 0 \text{ for all } i \in \{1, \dots, n\}.$$



*Proof.* To the minimal solution  $\bar{x} = (x_d^j, \bar{x}_c)$  we apply Theorem 4.1 and follow the proof of Theorem 4.3 and Corollary 4.5. Using the known formula for the Fréchet derivative of  $G(x_d^j, \cdot)$  for every  $i \in \{1, \dots, n\}$  (compare [6, Lemma 7.7]) we then get the first part of the assertion. The second part is a consequence of the CQ given in Corollary 4.1.  $\square$

Theorem 6.1 holds for the semidefinite variant of problem (6.3) and for the copositive case. The only difference is the dual cone  $C^*$ . Since the Löwner cone is self-dual, the condition  $L^*, \dots, L^n \in C^*$  means that the matrices  $L^1, \dots, L^n$  are positive semidefinite. If  $C$  equals the copositive cone, then its dual cone is given by

$$C^* = \text{co} \left\{ yy^T \mid y \in \mathbb{R}_+^k \right\}$$

and elements of  $C^*$  are called *completely positive matrices* (compare [6, p. 198–201]).

Finally, we turn our attention to the dual problem of the semidefinite/copositive problem (6.3). This dual problem is given by

$$\begin{aligned} \max \quad & \inf_{\substack{1 \leq i \leq n \\ x_c \in \mathbb{R}^{m_c}}} \left\{ f(x_d^i, x_c) + \langle L^i, G(x_d^i, x_c) \rangle \right\} \\ & \text{subject to} \\ & L^1, \dots, L^n \in C^*. \end{aligned} \tag{6.4}$$

So, the  $n$  variables of the dual semidefinite problem are positive semidefinite symmetric  $(k, k)$ -matrices and in the copositive case the dual variables are completely positive matrices. The weak duality theorem (Theorem 5.1) and the strong duality theorem (Theorem 5.2) are directly applicable to the primal problem (6.3) and its dual formulation (6.4).

## CONCLUSION

The main difference between a Lagrange theory in continuous optimization and such a theory in discrete-continuous optimization is the number of Lagrange multipliers being associated to a constraint. In general, in the discrete-continuous case one has several Lagrange multipliers per constraint. Therefore, optimality conditions in discrete-continuous optimization are much more difficult to handle than those in continuous optimization. It would be interesting to know how these conditions can be used for numerical methods. Since the minimum number of Lagrange multipliers of a concrete discrete-continuous optimization problem is unknown from the beginning, one possible strategy could be to start with only one Lagrange multiplier per constraint and then to increase the number of multipliers step by step. This expensive approach should be a topic for further research.

## REFERENCES

- [1] M. Baes, T. Oertel, R. Weismantel, Duality for mixed-integer convex minimization, *Math. Program. Ser. A* 158 (2016) 547–564.
- [2] G. Eichfelder, J. Jahn, Foundations of Set-Semidefinite Optimization, in: P.M. Pardalos, T.M. Rassias, A.A. Khan, (ed.), *Nonlinear Analysis and Variational Problems*, Chapter 18, pp. 259–284, Springer, New York, 2010.
- [3] M. Eidelheit, Zur Theorie der konvexen Mengen in linearen normierten Räumen, *Studia Math.* 6 (1936), 104–111.
- [4] C. Geiger, C. Kanzow, *Theorie und Numerik restringierter Optimierungsaufgaben*, Springer, Berlin, 2002.
- [5] R.B. Holmes, *Geometric functional analysis and its applications*, Springer, New York, 1975.
- [6] J. Jahn, *Introduction to the Theory of Nonlinear Optimization*, Springer, Berlin, 2007.
- [7] J. Jahn, *Vector Optimization - Theory, Applications, and Extensions*, Springer, Heidelberg, 2011.

- [8] J.B. Lasserre, A Lagrangian relaxation view of linear and semidefinite hierarchies, *SIAM J. Optim.* 23 (2013), 1742–1756.
- [9] M. Laurent, A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relations for 0-1 programming, *Math. Oper. Res.* 28 (2003), 470–496.
- [10] D.A. Morán, S.S. Dey, J.P. Vielma, A strong dual for conic mixed-integer programs, *SIAM J. Optim.* 22 (2012), 1136–1150.
- [11] K. Murota, *Discrete Convex Analysis*, SIAM, Philadelphia, 2003.
- [12] G.L. Nemhauser, L.A. Wolsey, *Integer and Combinatorial Optimization*, Wiley, New York, 1988.
- [13] H.D. Sherali, W.P. Adams, *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*, Kluwer, Norwell, 1999.
- [14] L.A. Wolsey, Integer programming: Price functions and sensitivity analysis, *Math. Program.* 20 (1981), 173–195.