

## Federation University ResearchOnline

**<https://researchonline.federation.edu.au>**

Copyright Notice

This is the post-peer-review, pre-copyedit version of an article published in 4OR: A Quarterly Journal of Operations Research. The final authenticated version is available online at:

<https://doi.org/10.1007/s10288-017-0350-6>

Copyright © 2017, Springer-Verlag GmbH Germany

See this record in Federation ResearchOnline at:

<http://researchonline.federation.edu.au/vital/access/HandleResolver/1959.17/159980>

# Recent contributions to linear semi-infinite optimization<sup>\*</sup>

M.A. Goberna<sup>\*\*</sup>, M.A. López

<sup>1</sup> Department of Mathematics, University of Alicante, Alicante, Spain, e-mail: [mgoberna@ua.es](mailto:mgoberna@ua.es)

<sup>2</sup> Department of Mathematics, University of Alicante, Alicante, Spain and Federation University of Australia, Ballarat, Australia, e-mail: [marco.antonio@ua.es](mailto:marco.antonio@ua.es)

The date of receipt and acceptance will be inserted by the editor

**Abstract** This paper reviews the state-of-the-art in the theory of deterministic and uncertain linear semi-infinite optimization, presents some numerical approaches to this type of problems, and describes a selection of recent applications in a variety of fields. Extensions to related optimization areas, as convex semi-infinite optimization, linear infinite optimization, and multi-objective linear semi-infinite optimization, are also commented.

**Key words** Linear semi-infinite optimization – Theory – Methods – Applications

## 1 Introduction

*Linear semi-infinite optimization* (LSIO in short) deals with linear optimization problems in which either the dimension of the decision space or the number of constraints (but not both) is infinite. We say that a linear optimization problem is *ordinary* (respectively, *infinite*) when the dimension of the decision space and the number of constraints are both finite (respectively, infinite).

The first three known contributions to LSIO are due to A. Haar (1924), E. Remez (1934), and G. Dantzig (1939), but they were basically ignored until the 1960s due to either the low diffusion, inside the mathematical community, of the journals where Haar and Remez published their discoveries,

---

<sup>\*</sup> This work was supported by the MINECO of Spain and ERDF of EU, Grant MTM2014-59179-C2-1-P, and by the Australian Research Council, Project DP160100854.

<sup>\*\*</sup> Corresponding author.

and the languages used (German and French, respectively), or to the temporary leave by Dantzig of his incipient academic career when he joined the Pentagon; in fact, he only wrote on his findings on LSIO within his professional memories, published many years later [52]. More in detail, Haar's paper [108] was focussed on the extension of the homogeneous Farkas lemma for linear systems from  $\mathbb{R}^n$  to an Euclidean space equipped with a scalar product  $\langle \cdot, \cdot \rangle$  (actually, the space  $\mathcal{C}([0, 1], \mathbb{R})$  of real-valued continuous functions on  $[0, 1]$  equipped with the scalar product  $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$ ); these systems are semi-infinite because they involve finitely many linear inequalities while the variable ranges on an infinite dimensional space. Remez's paper [161], in turn, proposed an exchange numerical method for a class of LSIO problems arising in polynomial approximation, e.g., computing the best uniform approximation to  $f \in \mathcal{C}([0, 1], \mathbb{R})$  by means of real polynomials of degree less than  $n - 1$ , i.e.,

$$\inf_{x \in \mathbb{R}^n} \max_{t \in [0, 1]} \left| f(t) - \sum_{i=1}^{n-1} x_i t^{i-1} \right|,$$

which is equivalent to the LSIO problem

$$\begin{aligned} P_A : \quad & \inf_{x \in \mathbb{R}^n} x_n \\ \text{s.t.} \quad & -x_n \leq f(t) - \sum_{i=1}^{n-1} t^{i-1} x_i \leq x_n, \quad t \in [0, 1]. \end{aligned} \quad (1)$$

Finally, Dantzig reformulated a Neyman-Pearson-type problem on statistical inference (posed by the same J. Neyman in a doctoral course attended by Dantzig) as a linear optimization problem with finitely many constraints and an infinite number of variables; Dantzig observed that the feasible set of this LSIO problem was the convex hull of its extreme points and conceived a geometry of columns allowing to jump from a given extreme point to a better adjacent one, which is a clear antecedent of the celebrated simplex method for linear optimization problems he proposed in 1947.

The next contributions to LSIO came in the 1960s, and are due to A. Charnes, W. Cooper and their doctoral student K. Kortanek; they conceived LSIO as a natural extension of ordinary linear optimization (also called linear programming, LP in short) and coined the term "semi-infinite" in [41] (for more details, see the description by the third author of the inception of LSIO in [126]). In [40] and [41] the LSIO problems with finitely many variables were called *primal*. These problems can be written as

$$\begin{aligned} P : \quad & \inf_{x \in \mathbb{R}^n} c'x \\ \text{s.t.} \quad & a'_t x \geq b_t, \quad t \in T, \end{aligned} \quad (2)$$

where  $c'$  represents the transpose of  $c \in \mathbb{R}^n$ ,  $a_t = (a_1(t), \dots, a_n(t))' \in \mathbb{R}^n$ , and  $b_t = b(t) \in \mathbb{R}$  for all  $t \in T$ . As in any field of optimization, the first theoretical results on LSIO dealt with optimality conditions and duality, and showed that LSIO is closer to ordinary convex optimization than to LP as the finiteness of the optimal value does not imply the existence of

an optimal solution and a positive duality gap can occur for the so-called *Haar's dual problem* of  $P$  (term also introduced by Charnes, Cooper and Kortanek),

$$\begin{aligned} D : \quad & \sup_{\lambda \in \mathbb{R}_+^{(T)}} \sum_{t \in T} \lambda_t b_t \\ \text{s.t.} \quad & \sum_{t \in T} \lambda_t a_t = c, \end{aligned} \quad (3)$$

where  $\mathbb{R}_+^{(T)}$  is the positive cone in the linear space of *generalized finite sequences*  $\mathbb{R}^{(T)}$ , whose elements are the functions  $\lambda \in \mathbb{R}^T$  that vanish everywhere on  $T$  except on a finite subset of  $T$  called *supporting set* of  $\lambda$ , and that we represent by  $\text{supp } \lambda$ . Observing that

$$\inf_{x \in \mathbb{R}^n} \left\{ L(x, \lambda) := c'x + \sum_{t \in T} \lambda_t (b_t - a_t'x) \right\} = \sum_{t \in T} \lambda_t b_t \quad (4)$$

if  $\lambda \in \mathbb{R}_+^{(T)}$  is a feasible solution of  $D$  and  $-\infty$  otherwise, one concludes that  $D$  is nothing else than a simplified form of the classical *Lagrange dual* of  $P$ :

$$D_L : \quad \sup_{\lambda \in \mathbb{R}_+^{(T)}} \inf_{x \in \mathbb{R}^n} L(x, \lambda). \quad (5)$$

As in convex optimization, the first characterizations of the optimal solutions of  $P$  and the duality theorems required some constraint qualification (CQ in brief) to be fulfilled. The problem  $P$  in (2) is said to be *continuous* whenever  $T$  is a compact topological space and the function  $t \mapsto (a_t, b_t)$  is continuous on  $T$ . The approximation problem  $P_A$  in (1) can easily be reformulated as a continuous LSIO problem by taking  $T = [0, 1] \times \{1, 2\} \subset \mathbb{R}^2$ . The *continuous dual problem* of a continuous LSIO problem  $P$  is

$$\begin{aligned} D_C : \quad & \sup_{\mu \in C'_+(T)} \int_T b_t d\mu(t) \\ \text{s.t.} \quad & \int_T a_t d\mu(t) = c, \end{aligned}$$

where  $C'_+(T)$  represents the cone of non-negative regular Borel measures on  $T$ . Since the elements of  $\mathbb{R}_+^{(T)}$  can be identified with the non-negative atomic measures, the optimal value of  $D_C$  is greater or equal than the optimal value of  $D$ .

The first numerical approach to sufficiently smooth LSIO problems, proposed by S. Gustafson and K. Kortanek in the early 1970s, consisted in the reduction of  $P$  to a nonlinear system of equations to be solved by means of a quasi-Newton method ([106], [107]). This approach was improved in [73] by aggregating a first phase, based in discretization by grids, in order to get a starting point for the quasi-Newton second phase. Simplex-like methods for particular classes of LSIO problems were proposed in [6], under the assumption that  $T$  is an interval and the  $n + 1$  functions  $a_1(\cdot), \dots, a_n(\cdot), b(\cdot)$  are analytic functions on  $T$ , and in [5], under the assumption that the feasible set of  $P$  is *quasipolyhedral* (meaning that its nonempty intersections with

polytopes are polytopes). An interior cutting plane method for continuous LSIO problems, inspired by the Elzinga-Moore method of ordinary convex optimization, was proposed in [90] and improved by an accelerated version in [16].

In many applications of optimization to real word problems the data defining the nominal problem are uncertain due to measurement errors, prediction errors, and round-off errors, so that the user must choose a suitable uncertainty model.

*Parametric models* are based on embedding the nominal problem into some topological space of admissible perturbed problems, the so-called *space of parameters*. *Qualitative stability analysis* provides conditions under which sufficiently small perturbations of the nominal problem provoke small changes in the optimal value, the optimal set and the feasible set; to be more precise, conditions ensuring the lower and upper semicontinuity of the mappings associating to each parameter the corresponding optimal value, optimal set and feasible set. These desirable stability properties are *generic* in certain set of parameters when they hold for most (in some sense) elements of that set. The first works on qualitative stability analysis of LSIO continuous problems were published in the 1980s by the group of Goethe University Frankfurt, formed by B. Brosowski and his collaborators T. Fischer and S. Helbig, together with their frequent visitor M.I. Todorov, who provided conditions for the semicontinuity of the above mappings, for continuous LSIO problems, under a variety of perturbations including those affecting the right-hand-side function  $b(\cdot)$  (see, e.g., [23], [24], [70], [178]). The extension of these results to (non-necessarily continuous) LSIO problems, and to their corresponding Haar's dual problems, was carried out by Todorov and the authors of this review in the second half of the 1990s ([89], [92], [93], [94]). In [129] several results about the stability of the boundary of the feasible set in LSIO are given.

*Quantitative stability analysis*, in turn, yields exact and approximate distances, in the space of parameters, from the nominal problem  $\bar{P}$  to important families of problems (e.g., from either a given consistent problem  $\bar{P}$  to the inconsistent ones or from a given bounded problem  $\bar{P}$  to the class of solvable problems), error bounds, and moduli of different Lipschitz-type properties which are related to the complexity analysis of the numerical methods; see, e.g., the works published during the last 15 years by the second author of this review together with the group of University Miguel Hernández of Elche, Spain (M.J. Cánovas, J. Parra, F. Toledo) and their collaborator A. Dontchev ([25], [33], [34], [35], [36]). *Sensitivity analysis* provides estimations of the impact of a given perturbation of the nominal problem on the optimal value, so that it can be seen as the quantitative stability analysis specialized to the optimal value. Results of this type for LSIO problems can be found in some of the above mentioned works and in the specific ones of the first author, the group of Puebla, Mexico (S. Gómez, F. Guerra-Vázquez, M.I. Todorov) and their collaborator T. Terlaky ([76], [100]).

Linear semi-infinite optimization has attracted during the last decades the attention, on the one hand, of optimization theorists as it is a simple, but non-trivial, extension of LP, and, on the other hand, of the linear optimization community, typically oriented towards numerical issues, as a primal LSIO problem can be seen as a huge LP problem. In fact, during the 1990s, authors like M. Powell, M. Todd, L. Tunçel or R. Vanderbei explained the numerical difficulties encountered by the interior-point methods on huge LP problems by analyzing the convergence of their adaptations to simple LSIO problems ([156], [177], [181], [183]). Finally, from the modeling perspective, LSIO has been systematically applied in those fields where uncertain LP problems arise in a natural way (as it happens in engineering and finance), specially when the user appeals to the robust approach. For all these reasons, several survey papers and extended reviews on LSIO have been published in the past, the last ones dated in 2005 ([74], [75]), and 2014 ([91], exclusively focused on uncertain LSIO). Linear semi-infinite optimization was also considered in two other surveys, the first one on semi-infinite optimization, published in 2007 ([139]), and the second one on the stability analysis of (ordinary, semi-infinite and infinite) linear optimization problems, published in 2012 ([138]). It is worth mentioning another survey on non-linear semi-infinite optimization [170] which explicitly precludes LSIO. Coherently with these antecedents, this review is intended to cover the period 2007-2016, for deterministic LSIO, and 2014-2016, for uncertain LSIO.

## 2 Deterministic linear semi-infinite optimization (2007-2016)

Let us introduce the necessary notation and basic concepts. Given a real linear space  $X$ , we denote by  $0_X$  the zero vector of  $X$ , except in the particular cases that  $X = \mathbb{R}^n$  or  $X = \mathbb{R}^T$ , whose null-vector are represented by  $0_n$  and  $0_T$ , respectively. Given a nonempty set  $Y \subset X$ , by  $\text{span } Y$ , and  $\text{conv } Y$  we denote the linear hull and the convex hull of  $Y$ , respectively, while  $\text{cone } Y$  denotes the conical convex hull of  $Y \cup \{0_X\}$ . We also denote by  $X'$  the algebraic dual of  $X$ , and by  $\langle \cdot, \cdot \rangle$  the duality product (i.e.,  $\langle \psi, x \rangle = \psi(x)$  for all  $\psi \in X'$  and  $x \in X$ ). Obviously,  $(\mathbb{R}^n)' = \mathbb{R}^n$  whereas  $(\mathbb{R}^T)' \not\subseteq \mathbb{R}^{(T)} = \{\lambda \in \mathbb{R}^T : \text{supp } \lambda \text{ is finite}\}$ . Indeed,  $(\mathbb{R}^{\mathbb{N}})'$  is of uncountable dimension while the dimension of  $\mathbb{R}^{(\mathbb{N})}$  is countable (see, e.g., [2] and [7]).

Given a topological space  $X$  and a set  $Y \subset X$ ,  $\text{int } Y$ ,  $\text{cl } Y$ , and  $\text{bd } Y$  represent the interior, the closure, and the boundary of  $Y$ , respectively. When  $X$  is a locally convex Hausdorff topological space (lcs in short) and  $Y \subset X$ ,  $\text{rint } Y$  denotes the relative interior of  $Y$ . We denote by  $X^*$  the topological dual of  $X$ , i.e.,  $X^* = X' \cap \mathcal{C}(X, \mathbb{R})$ . It is known (e.g., [6]) that  $(\mathbb{R}^n)^* = \mathbb{R}^n$  whereas  $(\mathbb{R}^T)^* = \mathbb{R}^{(T)}$  when one considers  $\mathbb{R}^T$  equipped with the product topology.

The Euclidean (respectively,  $l_\infty$ ) norm and the associated distance in  $\mathbb{R}^n$  are denoted by  $\|\cdot\|$  and  $d$  (respectively,  $\|\cdot\|_\infty$  and  $d_\infty$ ). We denote

by  $\mathbb{B}_n$  the closed unit ball of  $\mathbb{R}^n$  for the Euclidean norm. If  $Y$  is a non-empty convex subset of  $\mathbb{R}^n$ ,  $\dim Y$  and  $\text{extr } Y$  denote the dimension (i.e., the dimension of the affine hull of  $Y$ ) and the set of extreme points of  $Y$ , respectively. If  $Y$  is a convex cone in  $\mathbb{R}^n$ , its (negative) *polar cone* is  $Y^\circ = \{x \in \mathbb{R}^n : \langle y, x \rangle \leq 0 \ \forall y \in Y\}$ . Moreover, the *cone of feasible directions* (respectively, *normal cone*) of  $Y$  at  $x \in Y$  is

$$D(X; x) := \{d \in \mathbb{R}^n : \exists \theta > 0, x + \theta d \in X\}$$

(respectively,  $D(X; x)^\circ$ ). Since  $Y$  is convex,  $D(Y; x)$  is convex too (while  $D(Y; x)^\circ$  is, additionally, closed).

We denote by  $F$ ,  $S$  and  $v(P)$  the feasible set, the optimal set and the optimal value of the problem  $P$  in (2), and by  $F^D$ ,  $S^D$  and  $v(D)$  the feasible set, the optimal set and the optimal value of the problem  $D$  in (3).

A basic result on the feasibility of linear systems establishes ([44], [90]) that

$$F \neq \emptyset \Leftrightarrow (0_n, 1) \notin \text{cl cone} \{(a_t, b_t) : t \in T\}. \quad (6)$$

For the sake of simplicity, we only consider in this survey three constraint qualifications for a problem  $P$  such that  $F \neq \emptyset$ . The first one is an interior-type condition which was already used in the 1950s:

- $P$  satisfies the *Slater constraint qualification* (SCQ in short) if there exists  $\hat{x} \in \mathbb{R}^n$  (called *Slater point*) such that  $a'_t \hat{x} > b_t$  for all  $t \in T$ .

When  $P$  is continuous,  $\text{int } F \neq \emptyset$  implies SCQ and the converse holds whenever  $(a_t, b_t) \neq 0_{n+1}$  for all  $t \in T$  (in this case,  $\text{int } F$  coincides with the set of Slater points [90, Corollary 5.9.1]).

The second CQ is a closed cone condition introduced by Charnes, Cooper and Kortanek in [42]:

- $P$  satisfies the *Farkas-Minkowsky constraint qualification* (FMCQ in brief) if

$$\begin{aligned} \{x \in \mathbb{R}^n : a'_t x \geq b_t, t \in T\} &\subset \{x \in \mathbb{R}^n : a'x \geq b\} \\ &\Downarrow \\ \exists \tilde{T} \subset T, |\tilde{T}| < \infty : \{x \in \mathbb{R}^n : a'_t x \geq b_t, t \in \tilde{T}\} &\subset \{x \in \mathbb{R}^n : a'x \geq b\} \end{aligned}$$

or, equivalently, if the so-called *characteristic cone*

$$\text{cone} \{(a_t, b_t) : t \in T\} + \text{cone} \{(0_n, -1)\}$$

is closed. The closedness of the *second moment cone*

$$\text{cone} \{(a_t, b_t) : t \in T\}$$

implies FMCQ. If  $P$  is continuous, SCQ implies that the second moment cone is closed, entailing that the characteristic cone is also closed, and so  $P$  satisfies FMCQ [90, Theorem 5.3(ii)].

The third CQ is local, and involves the convex cone generated by the gradients of the active constraints at a given point. Given  $x \in F$ , the *active cone* of  $P$  at  $x$  is

$$A(x) := \text{cone} \{a_t : t \in T(x)\},$$

where  $T(x) = \{t \in T : a'_t x = b_t\}$  is the *set of active indices* at  $x$ . This CQ was introduced by R. Puente and V. Vera de Serio in [160], and extended to convex SIO in [90, §7.5] (see also [62], [63], [64]).

- $P$  satisfies the *local Farkas-Minkowsky constraint qualification* (with abbreviated acronym LFMQC) at  $x \in F$  if  $D(F; x)^\circ = A(x)$ .

It is easy to prove that FMCQ implies that LFMQC holds at any feasible point [90, §5.2].

### 2.1 Optimality

The classical LSIO optimality theorem establishes that the Karush-Kuhn-Tucker (KKT in short) condition at  $\bar{x} \in F$ , which is given by  $c \in A(\bar{x})$ , guarantees the optimality and the converse holds whenever LFMQC holds at  $\bar{x}$ . More precisely:

**Theorem 1 (Optimality in LSIO under LFMQC)** ([90, Theorem 7.1], [160]) *Each one of the following statements is sufficient for the (primal) optimality of  $\bar{x} \in F$ , and they are also necessary when LFMQC holds at  $\bar{x}$ :*

- (i)  $c \in \text{cl } A(\bar{x})$ .
- (ii)  $c \in A(\bar{x})$  (KKT condition).
- (iii) *There exists a feasible solution  $\bar{\lambda}$  of  $D$  such that  $\bar{\lambda}_t(a'_t \bar{x} - b_t) = 0$  for all  $t \in T$  (complementarity condition).*
- (iv) *There exists  $\bar{\lambda} \in \mathbb{R}_+^{(T)}$  such that  $L(\bar{x}, \lambda) \leq L(\bar{x}, \bar{\lambda}) \leq L(x, \bar{\lambda})$  for all  $x \in \mathbb{R}^n$ , with  $L(\cdot, \cdot)$  defined as in (4), and for all  $\lambda \in \mathbb{R}_+^{(T)}$  (Lagrange saddle point condition).*

Consequently, the optimality is characterized by the KKT condition under FMCQ and also under SCQ when  $P$  is continuous. A data qualification involving  $c$  which is weaker than LFMQC has been used in [135]. The above optimality theorem under CQs has been extended to infinite optimization in different ways ([48], [54], [65], [66], [133]), to multi-objective LSIO in [78, Theorem 23] and to multi-objective convex semi-infinite optimization in [79] and [98].

The problem of checking the optimality of a given feasible solution of  $P$  without CQs was first handled in [87] and [95] via the concept of extended active constraints at  $x^*$ , approach that has been recently extended to convex semi-infinite optimization in [145]. The next theorem gathers some asymptotic characterizations (through conditions involving limits) of the optimal solutions which have been obtained by using a novel geometric approach

(based in the so-called ladder regions introduced by Y. Liu and M. Ding in LSIO [136]).

We associate with each finite set  $T_0 \subset T$  and vector  $c^0 \in \mathbb{R}^n$ , the linear optimization problem

$$P(T_0, c^0) : \begin{array}{ll} \inf_{x \in \mathbb{R}^n} & \langle c^0, x \rangle \\ \text{s.t.} & a'_t x \geq b_t, \quad t \in T_0, \end{array}$$

whose feasible set and solution set we denote by  $F_0$  (called *ladder*) and  $S_0$ , respectively. Moreover, given  $x^0 \in F_0$ , we denote by  $T_0(x^0)$  and by  $A_0(x^0)$  the set of active indices and the active cone of  $x^0$  with respect to the problem  $P(T_0, c^0)$ , respectively, i.e.,  $T_0(x^0) = \{t \in T_0 : a'_t x = b_t\}$  and  $A_0(x^0) = \text{cone}\{a_t : t \in T_0(x^0)\}$ .

**Theorem 2 (Optimality in LSIO without CQs)** [137] *Let  $\bar{x} \in F$ . The following statements are equivalent:*

- (i)  $\bar{x}$  is an optimal solution of  $P$ .
- (ii) There exist sequences  $\{T_r\}_{r \in \mathbb{N}}$  of finite subsets of  $T$ ,  $\{c^r\}_{r \in \mathbb{N}} \subset \mathbb{R}^n$  and  $\{x^r\}_{r \in \mathbb{N}} \subset \mathbb{R}^n$  such that

$$c^r \rightarrow c, \quad (7)$$

$$x^r \rightarrow \bar{x}, \quad (8)$$

$$x^r \in F_r, \quad r = 1, 2, \dots, \quad (9)$$

and

$$c^r \in A_r(x^r), \quad r = 1, 2, \dots \quad (10)$$

(Asymptotic KKT condition.)

- (iii) There exist sequences  $\{T_r\}_{r \in \mathbb{N}}$  of finite subsets of  $T$ ,  $\{c^r\}_{r \in \mathbb{N}} \subset \mathbb{R}^n$ ,  $\{x^r\}_{r \in \mathbb{N}} \subset \mathbb{R}^n$  and  $\{\lambda^r\}_{r \in \mathbb{N}} \subset \mathbb{R}^{(T)}$  such that (7), (8) and (9) hold as well as

$$c^r = \sum_{t \in T} \lambda_t^r a_t, \quad r = 1, 2, \dots, \quad (11)$$

$$\text{supp } \lambda^r \subset T_r(x^r), \quad r = 1, 2, \dots, \quad (12)$$

$$\{a_t : t \in \text{supp } \lambda^r\} \text{ is linearly independent, } r = 1, 2, \dots, \quad (13)$$

$$a'_t x = b_t, \quad \forall t \in \text{supp } \lambda^r, \quad r = 1, 2, \dots \quad (14)$$

(Asymptotic complementarity condition.)

- (iv) There exist sequences  $\{T_r\}_{r \in \mathbb{N}}$  of finite subsets of  $T$ ,  $\{c^r\}_{r \in \mathbb{N}} \subset \mathbb{R}^n$ ,  $\{x^r\}_{r \in \mathbb{N}} \subset \mathbb{R}^n$  and  $\{\lambda^r\}_{r \in \mathbb{N}} \subset \mathbb{R}^{(T)}$  such that (7), (8), (9) and (12) hold, and

$$L(x^r, \lambda) \leq L(x^r, \lambda^r) \leq L(x, \lambda^r), \quad \forall x \in F_r, \forall \lambda \in \mathbb{R}_+^{(T)} \text{ with } \text{supp } \lambda \subset T_r(x^r). \quad (15)$$

(Asymptotic Lagrange saddle point condition.)

(v) *There exists a sequence  $\{(\lambda^r, \varepsilon_r)\}_{r \in \mathbb{N}} \subset \mathbb{R}_+^{(T)} \times \mathbb{R}_+$  such that*

$$\sum_{t \in T} \lambda_t^r b_t \geq c' \bar{x} - \varepsilon_r, \quad r = 1, 2, \dots, \quad (16)$$

and

$$\left( \sum_{t \in T} \lambda_t^r a_t, \varepsilon_r \right) \rightarrow (c, 0). \quad (17)$$

The asymptotic characterizations of optimality in (ii)-(iv) appeared by the first time in [137, Theorem 4.4] and their equivalence with (v) is shown in [137, Corollary 4.5]. The latter characterization is a particular case of optimality conditions in convex infinite optimization (see [55, Corollary 5], [174], and [22]). Observe the existence of many relationships among the conditions involved in Theorem 2, e.g., (9) and (10) imply that  $\{x^r\}_{r \in \mathbb{N}}$  is a sequence of optimal solution of the LP problems  $P(T_r, c^r)$ , which in turn shows the existence of a sequence  $\{\lambda^r\}_{r \in \mathbb{N}}$  such that (12), (11), and (15) hold.

## 2.2 Duality

The main classical LSIO duality theorems give conditions guaranteeing the identity  $v(P) = v(D)$  (i.e., the zero duality gap) together with the attainment of either the dual or the primal optimal value (i.e., solvability of either  $D$  or  $P$ ). Recall that we denote by  $F$  and  $S$  the primal feasible and optimal sets and by  $F^D$  and  $S^D$  their (Haar) dual counterparts.

**Theorem 3 (Zero Haar-duality gap with solvability)** *Let  $(a, b, c) \in (\mathbb{R}^n)^T \times \mathbb{R}^T \times \mathbb{R}^n$  be such that  $F \neq \emptyset \neq F^D$ .*

- (i) *If  $P$  satisfies FMCQ, then  $v(P) = v(D) \in \mathbb{R}$  and  $S^D \neq \emptyset$  (strong or infmax duality).*
- (ii) *If  $c \in \text{rint cone}\{(a_t, b_t) : t \in T\}$ , the following statements are true:*
  - (a)  *$v(P) = v(D) \in \mathbb{R}$  and  $S \neq \emptyset$  (converse strong or minsup duality).*
  - (b)  *$S$  is the sum of a non-empty compact convex set with a linear subspace.*
  - (c) *There exists a sequence of finite subproblems of  $P$ ,  $\{P(T_r, c)\}_{r \in \mathbb{N}}$ , such that*

$$v(P) = \lim_{r \rightarrow \infty} v(P(T_r, c))$$

(discretizability).

- (d) *If  $\{P(T_r, c)\}_{r \in \mathbb{N}}$  is a sequence of finite subproblems of  $P$  such that the sequence  $\{T_r\}_{r \in \mathbb{N}}$  is expansive (i.e.,  $T_r \subset T_{r+1}$  for all  $r \in \mathbb{N}$ ) and satisfies  $F = \bigcap_{r \in \mathbb{N}} F_r$ , then*

$$S \cap \limsup_{r \in \mathbb{N}} S_r \neq \emptyset;$$

*in other words, there exist a subsequence of optimal solutions of the problems  $P(T_r, c)$  converging to an optimal solution of  $P$ .*

Statement (i) was proved in [43]. Concerning (ii), (a) is in [165]; (b) in [88] (in [96], for convex infinite-dimensional optimization); (c) in [90, §8.3], although the study of the relationship between discretizability and absence of duality gap started with A. Charnes, W. Cooper and K. Kortanek ([41], [42], [43]), R.J. Duffin and L.A. Karlovitz [58], among others; and (d) is Theorem 8.6 in [90].

A new non-homogeneous Farkas lemma for linear semi-infinite inequality systems composed by  $m$  blocks of constraints with affine coefficient functions  $a_1(\cdot), \dots, a_n(\cdot), b(\cdot)$  defined on *spectrahedral sets* (i.e., solutions sets of semi-definite systems)  $T_j, j = 1, \dots, m$ , has been proved in [46]. When the constraint system of an LSIO problem belongs to this class of systems, the Haar dual problem can be formulated as a semi-definite optimization problem. A duality theorem [46, Theorem 2.1] is provided assuming the FMCQ. Sufficient conditions for this constraint qualification to be held are SCQ and that  $T_j$  is a polytope,  $j = 1, \dots, m$ .

Some recent contributions to LSIO duality theory are related with the semi-infinite versions of the Fourier (also called Fourier-Motzkin) elimination method, whose classical version provides linear representations of the projections of polyhedral convex sets on the coordinate hyperplanes. To the best of our knowledge, the first semi-infinite version of that method, which provides linear representations of the projections of closed convex sets on the coordinate hyperplanes, was introduced in [77] to characterize the so-called *Motzkin decomposable sets* (i.e., those sets which can be expressed as sums of polyhedral convex sets with closed convex cones, as the optimal set  $S$  of  $P$  when  $c \in \text{rint } M$ ), see also [81] and [77]. The second and third semi-infinite versions of the Fourier elimination method are due to A. Basu, K. Martin, and C. Ryan ([12], [13], [14]) and to K. Kortanek and Q. Zhang [127], respectively, these four papers dealing with LSIO duality theory.

The primal problem  $P$  is reformulated in [13] by aggregating an additional variable  $z$  representing upper bounds for  $\langle c, \cdot \rangle$  to be minimized on the feasible set, contained in  $\mathbb{R}^{n+1}$ , in order to get known and new duality results for the pair  $P - D$ . This formulation of  $P$ , called *standard* by the authors, has been reformulated again in [127] as a conic linear program from which the so-called classification duality theorems for semi-infinite linear programs [125] have been recovered.

A duality scheme for LSIO inspired by [7] is used in [12] (where the index set  $T$  is countable), [13], and [14]. Denoting by  $Y$  a linear subspace of  $\mathbb{R}^T$  (called *constraint space*) such that

$$U := \text{span} \{a_1(\cdot), \dots, a_n(\cdot), b(\cdot)\} \subset Y, \quad (18)$$

by  $Y'$  its algebraic dual, and by  $Y'_+ = \{\psi \in Y' : \psi(y) \geq 0, \forall y \in Y \cap \mathbb{R}_+^T\}$  the *positive cone* in  $Y'$ , the *algebraic dual problem* associated with  $Y$  is defined as

$$\begin{aligned} D_Y : \quad & \sup_{\psi \in Y'_+} \psi(b) \\ \text{s.t.} \quad & \psi(a_k) = c_k, k = 1, \dots, n. \end{aligned}$$

In [12] it is shown that  $D_{\mathbb{R}^T}$  is equivalent to  $D$  when  $T = \mathbb{N}$ , while the equivalence may fail when  $T$  is uncountable. The main merit of [14] is having posed the challenging problem of identifying vector spaces  $Y$  satisfying (18) such that the pair  $P - D_Y$  enjoys desirable properties. Indeed, the authors pay attention to the classical strong duality property and to dual pricing, the property (related with sensitivity analysis) which consists of the existence, for any perturbation  $d$  of  $b$ , of an optimal dual solution  $\lambda^d$  such that the rate of growth of the dual optimal value along the ray  $\mathbb{R}_+d$  is  $\sum_{t \in T} \lambda_t^d b_t$ .

**Theorem 4 (Strong duality and dual pricing I)** [14] *Assume that  $P$  is bounded. Then,  $P - D_U$  satisfies strong duality and dual pricing. Moreover, the dual optimal set  $S^{D_U}$  is a singleton.*

The proof that  $P - D_U$  satisfies strong duality (respectively, dual pricing) can be found in [14, Theorem 4.1] (respectively, [14, Theorem 4.3]).

**Theorem 5 (Strong duality and dual pricing II)** [14, Proposition 5.2] *If  $P - D_Y$  satisfies strong duality (respectively, dual pricing), then  $P - D_Q$  also satisfies strong duality (respectively, dual pricing) for any subspace  $Q$  such that  $U \subset Q \subset Y$ .*

Theorems 5.7 and 5.12 in [14] provide vector spaces  $Y$ , with  $U \subsetneq Y$ , such that  $D_Y$  satisfies strong duality and dual pricing, respectively.

The three papers [12], [13], [14] use almost exclusively arguments based on the Fourier elimination method, even to prove statements that can be easily obtained via convex analysis. Doing that, the authors try to defend the thought-provoking thesis, refuted by K. Kortanek and Q. Zhang [127], that LSIO is a subdiscipline of algebra. In our modest opinion, this claim has no practical implication and the consequence of the exclusive use of this methodology, less attractive than the elegant tools of convex analysis, is to require the readers an extra effort that not everybody is ready to pay. Besides this aesthetic drawback, the Fourier elimination method has the inconvenient, in comparison with the conventional tools of convex analysis, that it is hardly extensible to infinite dimensions, so that it cannot be used to repair the failure of strong duality and dual pricing for the Lagrange dual problems in infinite dimensional convex optimization, despite the reasonable conjecture that the above interesting results admit extensions to more general settings in the same way the classical Lagrange duality theorems have been already extended (see, e.g., [97] and references therein).

### 2.3 Numerical methods

A variety of approaches have been proposed for solving LSIO problems which satisfy different assumptions (discretization, cutting plane, reduction, penalty, interior-point, simplex-like, smoothing, etc.). Next, we describe briefly the most popular approaches.

### – Grid discretization Methods

Let  $P$  be an arbitrary LSIO problem. We associate with each nonempty finite set  $T_0 \subset T$  the corresponding discretization of  $P$  :

$$P(T_0, c) : \quad \inf_{x \in \mathbb{R}^n} c'x \\ \text{s.t.} \quad a'_t x \geq b_t, \quad t \in T_0.$$

Obviously,  $P(T_0, c)$  is a finite program when  $T_0$  is finite. Discretization methods generate sequences of points in  $\mathbb{R}^n$  converging to an optimal solution of  $P$  by solving a sequence of problems of the form  $P(T_k, c)$ , where  $T_k$  is a nonempty finite subset of  $T$  for  $k = 1, 2, \dots$

**Grid discretization algorithmic scheme** Let  $\varepsilon > 0$  be a fixed small scalar (called *accuracy*).

Step  $k$  : Let  $T_k$  be given.

(i) Compute a solution  $x_k$  of  $P(T_k, c)$ .

(ii) Stop if  $x_k$  is feasible within the fixed accuracy  $\varepsilon$ , i.e.,  $a(t)'x_k - b(t) \geq -\varepsilon$  for all  $t \in T$ . Otherwise, replace  $T_k$  with a new grid  $T_{k+1}$ .

Obviously,  $x_k$  is infeasible before optimality. Grid discretization methods select a priori sequences of *grids*,  $T_1, T_2, \dots$ , (usually expansive, i.e. satisfying  $T_k \subset T_{k+1}$  for all  $k$ ). The alternative discretization approaches generate the sequence  $\{T_k\}_{k \in \mathbb{N}}$  inductively. For instance, the classical Kelley cutting plane approach consists of taking  $T_{k+1} = T_k \cup \{t_k\}$ , for some  $t_k \in T$ , or  $T_{k+1} = (T_k \cup \{t_k\}) \setminus \{t'_k\}$  for some  $t'_k \in T_k$  (if an elimination rule is included).

Convergence of discretization methods requires  $P$  to be continuous [90, Theorem 11.1], which is based on Theorem 3(ii-d). The main difficulties with these methods are undesirable jamming in the proximity of an optimal solution and the increasing size of the auxiliary problems  $P(T_k, c)$  (unless elimination rules are implemented). These methods are only efficient for low-dimensional index sets. For more details, see [139] and references therein. To the authors knowledge, no significant contribution to this approach has been published during the last decade.

### – Central cutting plane methods

Let  $P$  be an LSIO problem. Given  $\alpha \in \mathbb{R}$ , the set  $\{x \in F : c'x \leq \alpha\}$  is called the *sublevel set* of  $P$  for  $\alpha$ . We associate with any polytope (i.e., a bounded polyhedral set)  $Q$  a certain *center* (e.g., the centre of the greatest ball contained in  $Q$  or the analytic centre of  $Q$ ) and the ordinary LP problem

$$P(Q) : \quad \inf_{x \in Q} c'x.$$

Central cutting plane methods generate sequences of points in  $\mathbb{R}^n$  converging to an optimal solution of  $P$  by solving a sequence of problems of the form  $P(Q_k)$ , where  $Q_k$  is a polytope for  $k = 1, 2, \dots$

**Central cutting plane algorithmic scheme** Let  $\varepsilon > 0$  be a fixed accuracy.

Step  $k$  : Let  $Q_k$  be a polytope containing some sublevel set of  $P$ .

(i) Compute a center  $x_k$  of  $Q_k$ .

(ii) If  $x_k \notin F$ , set  $Q_{k+1} = \{x \in Q_k : a(t_0)'x - b(t_0) \geq 0\}$ , where  $t_0 \in T$  satisfies  $a(t_0)'x_k - b(t_0) < 0$ , and  $k = k + 1$ . Otherwise, continue.

(iii) If  $c'x_k \leq \min\{c'x : x \in Q_k\} + \varepsilon$ , stop. Otherwise, replace  $Q_k$  with a new set  $Q_{k+1} = \{x \in Q_k : c'x \leq c'x_k\}$ .

Obviously, (ii) aggregates to the current polytope  $Q_k$  a feasibility-cut (some constraint of  $P$  violated by  $x_k$ ) when  $x_k$  is unfeasible, whereas (iii) checks the  $\varepsilon$ -optimality of  $x_k$ , aggregating to  $Q_k$  an objective-cut when the result is negative.

The comments on the convergence and drawbacks of discretization methods also apply to cutting plane algorithms. Nevertheless, efficient implementations of the latter methods turn out to be computationally faster than the grid discretization counterparts and, moreover, they stop before optimality at a feasible solution. For more details see [16] and references therein.

Algorithm 2 in [53] is a (non central) cutting plane method generating at each step, from the current non-feasible point  $x_k$ , a feasible one which is the result of shifting  $x_k$  towards a fixed Slater point with step length easily computable.

The two cutting plane discretization algorithms for a continuous LSIO problem  $P$  whose feasible set is a convex body (a compact convex full dimensional set proposed in [151] are inspired by the logarithmic barrier decomposition method in [141]. Both algorithms update the current discretization  $P(T_k, c)$  of  $P$  by selecting a point in the proximity of the central path of  $P(T_k, c)$  and aggregating to the constraints of  $P(T_k, c)$  some violated constraints. Then the full dimension of the new feasible set is recovered and the central path is updated. This process continues until the barrier parameter is small enough. The novelties introduced by [151], in comparison with [141], are that the violated constraints are aggregated without changing their right-hand-side (RHS in short) coefficients (as it was proposed by [141] in order to maintain the full dimension of the feasible sets), that more than one violated constraints can be aggregated in each step, and that the barrier function is updated at step  $k$ . The method in [151], called "interior point constraint generation algorithm" even though it does not generate feasible points of  $P$ , converges to an  $\varepsilon$ -optimal solution after a finite number of constraints is generated. The authors provide complexity bounds on the number of Newton steps needed and on the total number of constraints that is required for the overall algorithm to converge. Computational experiments are reported.

#### – KKT reduction methods

The basic idea of the KKT reduction methods, which were already known from Chebyshev approximation, consists of replacing  $P$  with a non-

linear system of equations obtained from the KKT local optimality conditions for  $P$ .

In fact, if  $\bar{x}$  is a local minimizer of  $P$  satisfying LFMCQ, Theorem 1 guarantees the existence of indices  $\bar{t}_j \in T(\bar{x})$ ,  $j = 1, \dots, q(\bar{x})$ , with  $q(\bar{x})$  depending on  $\bar{x}$ , and nonnegative multipliers  $\bar{\lambda}_j$ ,  $j = 1, \dots, q(\bar{x})$ , such that

$$c = \sum_{j=1}^{q(\bar{x})} \bar{\lambda}_j a(\bar{t}_j).$$

We also assume the availability of a description of  $T \subset \mathbb{R}^p$  as

$$T = \{t \in \mathbb{R}^p : u_i(t) \geq 0, i = 1, \dots, m\},$$

where  $u_i$  is smooth for all  $i = 1, \dots, m$ . Observe that  $q(\bar{x})$  is the number of global minima of the so-called *lower level problem* at  $\bar{x}$ ,

$$Q(\bar{x}) : \inf \{a(t)' \bar{x} - b(t) : u_i(t) \geq 0, i = 1, \dots, m, t \in T\},$$

provided that the optimal value of this problem is zero, i.e.  $v(Q(\bar{x})) = 0$ . This is a global optimization problem which can be solved whenever  $T$  is a finite dimensional interval and  $a_1(\cdot), \dots, a_n(\cdot), b(\cdot) \in \mathbb{R}[t]$  (the ring of polynomials in the single variable  $t$  with real coefficients), in which case  $t \mapsto a(t)' \bar{x} - b(t)$  is polynomial too. Then, under some constraint qualification, for each  $\bar{t}_j$ ,  $j = 1, \dots, q(\bar{x})$ , the classical KKT theorem yields the existence of nonnegative multipliers  $\bar{\theta}_i^j$ ,  $i = 1, \dots, m$ , such that

$$\langle \nabla_t a(\bar{t}_j), \bar{x} \rangle - \nabla_t b(\bar{t}_j) = \sum_{i=1}^m \bar{\theta}_i^j \nabla_t u_i(\bar{t}_j)$$

and

$$\bar{\theta}_i^j u_i(\bar{t}_j) = 0, i = 1, \dots, m.$$

**KKT reduction algorithmic scheme** Step  $k$  : Start with a given  $x_k$  (not necessarily feasible).

(i) Estimate  $q(x_k)$ .

(ii) Apply  $N_k$  steps of a quasi-Newton method (for finite systems of equations) to

$$\left\{ \begin{array}{l} c = \sum_{j=1}^{q(x_k)} \lambda_j a(t_j) \\ a(t_j)' x = b(t_j), j = 1, \dots, q(x_k) \\ \langle \nabla_t a(t_j), x \rangle - \nabla_t b(t_j) = \sum_{i=1}^m \theta_i^j \nabla_t u_i(t_j), j = 1, \dots, q(x_k) \\ \theta_i^j u_i(t_j) = 0, i = 1, \dots, m, j = 1, \dots, q(x_k) \end{array} \right\},$$

- (with unknowns  $x, t_j, \lambda_j, \theta_i^j, i = 1, \dots, m, j = 1, \dots, q(x_k)$ ) leading to iterates  $x_{k,l}, l = 1, \dots, N_k$ .  
 (iii) Set  $x_{k+1} = x_{k,N_k}$  and  $k = k + 1$ .

The main advantage of the KKT reduction methods on the discretization and cutting plane methods is their fast local convergence (as they are adaptations of quasi-Newton methods) provided that they start sufficiently close to an optimal solution. The so-called *two-phase methods* combine a discretization or central cutting plane method providing a rough approximation of an optimal solution of  $P$  (1st phase) and a reduction method afterwards improving this approximation (2nd phase). No theoretical result supports the decision to switch from phase 1 to phase 2. For more details see [73], [139], and references therein. The last extension we know of the above reduction scheme to smooth non-linear semi-optimization appeared in [189].

Numerical methods which are not based on the three approaches above have been proposed for particular types of LSIO problems. On the one hand, an exchange method for continuous LSIO problems such that  $T$  is a compact interval in  $\mathbb{R}$  based on the geometric concept of ladder point has been proposed in [136]. On the other hand, [190] proposed an interior-point method for polynomial LSIO problems (i.e., assuming that  $T$  is a compact interval in  $\mathbb{R}$  and  $a_1(\cdot), \dots, a_n(\cdot), b(\cdot) \in \mathbb{R}[t]$ ; this method is inspired by Lasserre's method [130] for polynomial nonlinear semi-infinite optimization; error bounds for the used approximations are given; the efficiency of this approach has been illustrated by two classical test problems. The latter method, for polynomial LSIO problems with one-dimensional index set, has been extended in [104] to similar problems with semialgebraic index set  $T = \{t \in \mathbb{R}^m : g_j(t) \geq 0, j = 1, \dots, s\}$ , where  $g_j$  is a polynomial on the variables  $t_1, \dots, t_m$  for all  $j = 1, \dots, s$ . Since unconstrained minimization of a polynomial on  $\mathbb{R}^n$  is NP-hard when  $n > 1$ , the polynomial LSIO problem  $P$  is NP-hard too, so that it cannot be expected to be solved in polynomial time (unless  $P=NP$ ), [104] proposes to construct a sequence of polynomial time problems whose optimal solutions converge to some optimal solution of  $P$ . The author associates to the set of polynomial functions  $G := \{g_1, \dots, g_s\}$  its *quadratic module*  $Q(G)$ , which is formed by the sums  $\sum_{j=0}^s \sigma_j g$ , where

$g_0 = 1$  and the  $s + 1$  functions  $\sigma_0, \dots, \sigma_s$  are sums of squares of polynomials. Aggregating the condition that  $\deg(\sigma_j g_j) \leq 2k$  for  $k \in \{0\} \cup \mathbb{N}$ , one gets the so-called *k-th truncation*  $Q_k(G)$  of  $Q(G)$ , with  $Q(G)$  being the expansive union of the sequence  $\{Q_k(G)\}_{k \in \mathbb{N}}$ . Under certain assumptions,  $F := \{x \in \mathbb{R}^n : a(t)^T x + b(t) \in Q(G)\}$ . Thus,  $P$  is approached by the sequence of problems

$$(P_k) \quad \inf_{x \in \mathbb{R}^n} c^T x \\ \text{s.t.} \quad a(t)^T x + b(t) \in Q_k(G),$$

$k \in \mathbb{N}$ , with  $v(P_k) \searrow v(P)$  as  $k \rightarrow \infty$ . Finally, each problem  $(P_k)$  is reformulated as a semi-definite problem whose Lagrange dual, called *semi-definite relaxation* of  $(P)$ , can be solved in polynomial time.

Descent methods (respectively, simplex-like methods) for  $P$  start with a feasible solution (respectively, an extreme point of  $F$  obtained from an arbitrary feasible solution via purification). This initial feasible solution can be obtained with some semi-infinite version of the relaxation method for ordinary linear inequality systems proposed independently by Agmon [1] and by Motzkin and Schoenberg [149] in 1954. The semi-infinite fixed step relaxation algorithm can be described as follows:

**Relaxation algorithmic scheme** Let  $\varepsilon > 0$  be a fixed accuracy.

Step  $k$ : Let  $x_k \notin F$  be given.

(i) Compute an approximate value  $\mu_k$  of  $d(x_k, H_k)$ , where

$$H_k := \{x \in \mathbb{R}^n : \langle a(t_k), x \rangle = b(t_k)\}$$

is the hyperplane determined by some constraint  $\langle a(t_k), x \rangle \geq b(t_k)$ ,  $t_k \in T$ , violated by  $x_k$ . Take  $\lambda_k \in ]0, 2]$  according to a given rule and compute

$$x_{k+1} := x_k + \lambda_k \mu_k \frac{a(t_k)}{\|a(t_k)\|}.$$

(ii) Stop if  $x_{k+1}$  is feasible within the fixed accuracy  $\varepsilon$ , i.e.,  $a(t)'x_{k+1} - b(t) \geq -\varepsilon$  for all  $t \in T$ . Otherwise, replace  $x_k$  with  $x_{k+1}$ .

This relaxation scheme admits different implementations, e.g., if  $\lambda_k = 1$  for all  $k$ , then  $x_{k+1}$  is an approximate projection of  $x_k$  onto  $H_k$ , and, if  $\lambda_k = 2$  for all  $k$ , then  $x_{k+1}$  is an approximate symmetric of  $x_k$  with respect to  $H_k$ . The relaxation parameter  $\lambda_k$  is maintained fixed in [113] and [118] (with  $\lambda_k = 1$  for all  $k$ ), as well as in [102] and [103] (with  $\lambda_k = \lambda$  for all  $k$ , for some  $\lambda \in (0, 2]$ ), whereas  $\lambda_k$  is taken at random in some subinterval of  $(0, 2]$  in [68], where all this variants of the relaxation algorithm are compared from the computational efficiency point of view.

It is worth mentioning that LSIO problems can also be solved by means of methods which have been conceived for more general problems, as infinite linear optimization problems ([72]), convex semi-infinite optimization problems ([9], [10], [67], [194]), or non-linear semi-infinite optimization problems. The NEOS Server allows to solve the latter type of problems by the program NSIPS, coded in AMPL, see [184] and [185]: <http://www.neos-server.org/neos/solvers/sio:nsips/AMPL.html>

Finally, let us observe that a unique numerical method has been proposed so far for integer LSIO problems, more precisely, for a class of problems which arises in solid waste management [105].

### 3 Uncertain linear semi-infinite optimization (2014-2016)

Let

$$\begin{aligned} \bar{P}: \quad & \inf_{x \in \mathbb{R}^n} \bar{c}'x \\ \text{s.t.} \quad & \bar{a}_t'x \geq \bar{b}_t, \quad t \in T, \end{aligned} \quad (19)$$

be a given LSIO primal problem, called *nominal*, whose data, gathered in the triple  $(\bar{a}, \bar{b}, \bar{c}) \in (\mathbb{R}^n)^T \times \mathbb{R}^T \times \mathbb{R}^n$ , are uncertain for some reason. We consider the linear space  $(\mathbb{R}^n)^T \times \mathbb{R}^T \times \mathbb{R}^n$  equipped with the topology induced by the *box* (or *Chebyshev*) *pseudo-norm*

$$\|(a, b, c)\|_\infty := \max \{ \|c\|_\infty, \sup_{t \in T} \|(a_t, b_t)\|_\infty \}, \quad \forall (a, b, c) \in (\mathbb{R}^n)^T \times \mathbb{R}^T \times \mathbb{R}^n,$$

which is a proper norm whenever  $T$  is finite. In this way,  $(\mathbb{R}^n)^T \times \mathbb{R}^T \times \mathbb{R}^n$  is equipped with the pseudometric  $d_\infty$  defined as

$$d_\infty((a^1, b^1, c^1), (a^2, b^2, c^2)) := \|(a^1 - a^2, b^1 - b^2, c^1 - c^2)\|_\infty,$$

for any pair  $(a^1, b^1, c^1), (a^2, b^2, c^2) \in (\mathbb{R}^n)^T \times \mathbb{R}^T \times \mathbb{R}^n$ . Obviously,  $d_\infty$  describes the topology of the uniform convergence on  $(\mathbb{R}^n)^T \times \mathbb{R}^T \times \mathbb{R}^n$ .

Any perturbation of the triple  $(\bar{a}, \bar{b}, \bar{c})$  which preserves the decision space  $\mathbb{R}^n$  and the index set  $T$  of  $\bar{P}$  provides a perturbed problem

$$\begin{aligned} P: \quad & \inf_{x \in \mathbb{R}^n} c'x \\ \text{s.t.} \quad & a_t'x \geq b_t, \quad t \in T, \end{aligned}$$

with associated data triple  $(a, b, c) \in (\mathbb{R}^n)^T \times \mathbb{R}^T \times \mathbb{R}^n$ . The *parameter space*  $\Theta \subset (\mathbb{R}^n)^T \times \mathbb{R}^T \times \mathbb{R}^n$  is formed by the admissible perturbations of the *nominal data*  $(\bar{a}, \bar{b}, \bar{c})$ . For instance, if the RHS of  $\bar{P}$  cannot be perturbed, the parameter space is formed by the triples  $(a, \bar{b}, c)$ ; and abusing of the language, we can identify  $(a, \bar{b}, c)$  with the pair  $\theta = (a, c)$ , so that we have, in that case,  $\Theta = (\mathbb{R}^n)^T \times \mathbb{R}^n$ . Then,  $d_\infty(\theta, \bar{\theta}) = d_\infty((a, c), (\bar{a}, \bar{c})) \in [0, +\infty]$  measures the size of the perturbation.

In the next two subsections we consider the seven parameter spaces which appear in Table 1.

Scenario	Perturbable data	Parameter space $\Theta$
1	$(a, b, c)$	$(\mathbb{R}^n)^T \times \mathbb{R}^T \times \mathbb{R}^n$
2	$(a, b)$	$(\mathbb{R}^n)^T \times \mathbb{R}^T$
3	$(a, c)$	$(\mathbb{R}^n)^T \times \mathbb{R}^n$
4	$(b, c)$	$\mathbb{R}^T \times \mathbb{R}^n$
5	$a$	$(\mathbb{R}^n)^T$
6	$b$	$\mathbb{R}^T$
7	$c$	$\mathbb{R}^n$

Table 1

Scenarios 1, 4 and 6 correspond to the so-called *full*, *canonical* and *RHS* perturbations, respectively. It is easy to prove that the set of parameters  $\Theta_{con}$  providing primal and dual consistent problems is a convex cone in Scenarios 4, 6 and 7. The stability behavior of Scenario 5 is the most difficult to be analyzed ([51]).

### 3.1 Qualitative stability

This subsection provides conditions, preferably expressed in terms of the data, under which sufficiently small perturbations of the nominal problem provoke small changes in the following four mappings (two of them set-valued, the remaining two extended real-valued):

- The (primal) feasible set mapping  $\mathcal{F} : \Theta \rightrightarrows \mathbb{R}^n$  associating with each  $\theta \in \Theta$  the corresponding primal feasible set.
- The (primal) optimal set mapping  $\mathcal{S} : \Theta \rightrightarrows \mathbb{R}^n$  associating with each  $\theta \in \Theta$  the corresponding primal optimal set.
- The (primal) optimal value function  $\vartheta : \Theta \rightarrow \overline{\mathbb{R}}$  assigning to  $\theta \in \Theta$  the corresponding primal optimal value.
- The gap function  $g : \Theta \rightarrow \mathbb{R} \cup \{+\infty\}$  such that

$$g(\theta) := \begin{cases} \vartheta(\theta) - \vartheta^D(\theta), & \text{if } \vartheta(\theta) - \vartheta^D(\theta) \in \mathbb{R}, \\ +\infty, & \text{otherwise,} \end{cases}$$

where  $\vartheta^D(\theta)$  denotes the corresponding dual optimal value. It can easily be shown that  $g$  is an homogeneous function such that  $g^{-1}(\mathbb{R}) = \Theta_{con}$ . Moreover,  $g$  is the difference of two convex functions (i.e., a DC function) in Scenarios 6 and 7, so that  $g(\cdot, c)$  and  $g(b, \cdot)$  are DC functions for all  $c \in \mathbb{R}^n$  and for all  $b \in \mathbb{R}^T$ , respectively, in Scenario 4.

The stability properties of  $g$  at  $\bar{\theta} \in \Theta_{con}$  are related with the preservation of the primal-dual consistency under sufficiently small perturbations, i.e., with the condition that  $\bar{\theta} \in \text{int } \Theta_{con}$ .

**Theorem 6 (The interior of  $\Theta_{con}$ )** [99, Theorem 1] *The membership of an element of  $\Theta_{con}$  to its interior  $\text{int } \Theta_{con}$  is characterized in Table 2, where the involved conditions (20) and (21) are*

$$0_{n+1} \notin \text{clconv} \{(a_t, b_t) : t \in T\} \quad (20)$$

and

$$c \in \text{int cone} \{a_t : t \in T\}. \quad (21)$$

Scenario	Parameter	Characterization of $\theta \in \text{int } \Theta_{con}$
1	$(a, b, c)$	(20) and (21)
2	$(a, b)$	(20), if $c = 0_n$ (20) and (21), else
3	$(a, c)$	(21), if $\sup_{t \in T} b_t \leq 0$ (20) and (21), else
4	$(b, c)$	(20) and (21)
5	$a$	(20) and (21), if $c \neq 0_n$ and $\sup_{t \in T} b_t > 0$ (21), if $c \neq 0_n$ and $\sup_{t \in T} b_t \leq 0$ (20), if $c = 0_n$ and $\sup_{t \in T} b_t > 0$ It holds for sure, if $c = 0_n$ and $\sup_{t \in T} b_t \leq 0$
6	$b$	(20)
7	$c$	(21)

Table 2

**Theorem 7 (Stability of the duality gap)** [99, Theorem 2] *Let  $\bar{\theta} \in \Theta_{con}$ , with  $\Theta$  being the parameter space for any scenario  $k \in \{1, 3, 4, 5, 7\}$ , or  $k = 2$  with  $c \neq 0_n$ . Then, the following statements are equivalent:*

- (i)  $g$  is identically zero in some neighborhood of  $\bar{\theta}$  (0–stability).
  - (ii)  $g$  is continuous at  $\bar{\theta}$ .
  - (iii)  $g$  is upper semicontinuous at  $\bar{\theta}$ .
  - (iv)  $\bar{\theta} \in \text{int } \Theta_{con}$ .
- Moreover, if  $|T| < \infty$ , then the four statements (i) - (iv) are equivalent for any  $k \in \{1, \dots, 7\}$ .

A partial extension of this result to linear infinite optimization, only for Scenarios 4, 6 and 7, and for a type of dual problem which coincides with  $D_C$  when  $P$  is a continuous LSIO problem, can be found in [187].

The adjective *generic* applied to certain desirable property, as the 0–stability, may have different meanings. Topological genericity on  $\Theta_{con}$  means that the property holds in a dense open subset. This kind of genericity is called *weak* in [59, Remark 1], as it does not imply that the set of parameters of  $\Theta_{con}$  where that property fails has Lebesgue measure zero.

**Theorem 8 (Topological genericity of the 0–stability)** [99, Theorem 3] *Table 3 below provides sufficient conditions guaranteeing the weak genericity of the 0–stability of  $g$  in  $\Theta_{con}$  for the seven scenarios (assuming  $|T| < \infty$  in Scenario 6).*

Scenario	Parameter	Suff. cond. for $\text{int } \Theta_{con}$ to be dense in $\Theta_{con}$
1	$(a, b, c)$	$ T  \geq n$
2	$(a, b)$	$ T  \geq n$ or $c = \mathbf{0}_n$
3	$(a, c)$	$ T  \geq n$ and $\sup_{t \in T} b_t \leq 0$
4	$(b, c)$	$\dim \text{cone}(a) = n$
5	$a$	$ T  \geq n$ and $\sup_{t \in T} b_t \leq 0$
6	$b$	No condition is needed
7	$c$	$\dim \text{cone}(a) = n$

Table 3

### 3.2 Quantitative stability

We assume in this subsection that the nominal problem  $\bar{P}$  is continuous and only perturbations  $P$  of  $\bar{P}$  providing continuous problems are allowed, i.e., that the corresponding triples  $(a, b, c)$  belong to  $\mathcal{C}(T, \mathbb{R}^n) \times \mathcal{C}(T, \mathbb{R}) \times \mathbb{R}^n$ . In other words, we consider in this subsection the restriction of  $\mathcal{F}$  and  $\mathcal{S}$  to the linear subspace of *continuous* parameters  $\Theta_0 := \mathcal{C}(T, \mathbb{R}^n) \times \mathcal{C}(T, \mathbb{R}) \times \mathbb{R}^n \subset \Theta$ .

Recall that the *graph* of a mapping  $\mathcal{M} : \Theta_0 \rightrightarrows \mathbb{R}^n$  is

$$\text{gph } \mathcal{M} := \{(\theta, x) \in \Theta_0 : x \in \mathcal{M}(\theta)\}.$$

- A set-valued mapping  $\mathcal{M} : \Theta_0 \rightrightarrows \mathbb{R}^n$  is *pseudo-Lipschitz* (also called *Aubin continuous* or *Lipschitz-like*) at  $(\bar{\theta}, \bar{x}) \in \text{gph } \mathcal{M}$  if there exist neighborhoods  $V_{\bar{\theta}}$  of  $\bar{\theta}$  and  $U_{\bar{x}}$  of  $\bar{x}$ , and a scalar  $\kappa \geq 0$  such that

$$d(x, \mathcal{M}(\theta)) \leq \kappa d(\theta, \theta'), \quad \forall \theta, \theta' \in V_{\bar{\theta}}, \forall x \in \mathcal{M}(\theta') \cap U_{\bar{x}}. \quad (22)$$

The infimum of such  $\kappa$  for all the triples  $(V_{\bar{\theta}}, U_{\bar{x}}, \kappa)$  verifying (22) is called *Lipschitz modulus* of  $\mathcal{M}$  at  $(\bar{\theta}, \bar{x})$  and it is denoted by  $\text{lip } \mathcal{M}(\bar{\theta}, \bar{x})$ .

The reader will find in the monographs [57], [123], [148], [163], etc., a comprehensive study of this property. Formulas involving  $\text{lip } \mathcal{F}(\bar{\theta}, \bar{x})$  and  $\text{lip } \mathcal{S}(\bar{\theta}, \bar{x})$  can be found in [91] and references therein.

For many authors, the pseudo-Lipschitz property is too demanding, so that they prefer the *weaker* notion of calmness. We devote the rest of this section to the study of the calmness property of  $\mathcal{F}$  and  $\mathcal{S}$  in different scenarios, as well as to propose either exact expressions or estimates of the corresponding moduli. Let us start by defining this property.

- A set-valued mapping  $\mathcal{M} : \Theta_0 \rightrightarrows \mathbb{R}^n$  is *calm* at  $(\bar{\theta}, \bar{x}) \in \text{gph } \mathcal{M}$  if there exist neighborhoods  $V_{\bar{\theta}}$  of  $\bar{\theta}$  and  $U_{\bar{x}}$  of  $\bar{x}$ , and a scalar  $\kappa \geq 0$  such that

$$d(x, \mathcal{M}(\bar{\theta})) \leq \kappa d(\theta, \bar{\theta}), \quad \forall \theta \in V_{\bar{\theta}}, \forall x \in \mathcal{M}(\theta) \cap U_{\bar{x}}. \quad (23)$$

The infimum of such  $\kappa$  for all the triples  $(V_{\bar{\theta}}, U_{\bar{x}}, \kappa)$  verifying (23) is called *calmness modulus* of  $\mathcal{M}$  at  $(\bar{\theta}, \bar{x})$  and it is denoted by  $\text{clm } \mathcal{M}(\bar{\theta}, \bar{x})$ .

Calmness was introduced in 1976 by Clarke [47] in the context of functions and in connection with constraint qualification conditions, but the term calmness was coined in [163]. It plays a key role in many issues of mathematical programming like optimality conditions, error bounds and stability of solutions, among others, and this fact motivated an increasing interest among researchers in the last twenty years. Calmness property is known to be equivalent to the *metric subregularity* of the inverse mapping, property introduced by Ioffe [116] under a different name.

In Scenario 6 we shall consider the feasible set mapping depending only on the right-hand-side (RHS); specifically, let us consider a fixed  $\bar{a} \equiv (\bar{a}_t)_{t \in T} \in \mathcal{C}(T, \mathbb{R}^n)$  and the corresponding feasible set mapping  $\mathcal{F}_{\bar{a}} : \mathcal{C}(T, \mathbb{R}) \rightrightarrows \mathbb{R}^n$  given by

$$\mathcal{F}_{\bar{a}}(b) := \mathcal{F}(\bar{a}, b) = \{x \in \mathbb{R}^n : \bar{a}'_t x \geq b_t \text{ for all } t \in T\}. \quad (24)$$

In this framework,  $\mathcal{C}(T, \mathbb{R})$  is our parameter space and the parameter is  $\theta \equiv b \in \mathcal{C}(T, \mathbb{R})$ ; the parameter space is endowed with the supremum norm  $\|b\|_{\infty} := \max_{t \in T} |b_t|$ . It makes sense to deal first with the calmness property of  $\mathcal{F}_{\bar{a}}$  at  $(\bar{\theta}, \bar{x}) \equiv (\bar{b}, \bar{x}) \in \text{gph} \mathcal{F}_{\bar{a}}$  and, in a second stage, to study the calmness of  $\mathcal{F}$ . Let us announce that calmness of  $\mathcal{F}_{\bar{a}}$  at  $(\bar{b}, \bar{x})$  is equivalent to calmness of  $\mathcal{F}$  at  $((\bar{a}, \bar{b}), \bar{x}) \in \text{gph} \mathcal{F}$ , and the calmness modulus of  $\mathcal{F}_{\bar{a}}$  at  $(\bar{b}, \bar{x})$  is closely related to the calmness modulus of  $\mathcal{F}$  at  $((\bar{a}, \bar{b}), \bar{x})$  (see (32) below).

In the case of finite linear systems (i.e., when  $T$  is finite), it is well-known that  $\mathcal{F}_{\bar{a}}$  is always calm at any point of its (polyhedral) graph as a consequence of a classical result by Robinson [162]. For  $T$  infinite, it is known that the Slater constraint qualification (SCQ) is a sufficient condition for the calmness of  $\mathcal{F}$  and of  $\mathcal{F}_{\bar{a}}$ , as far as SCQ characterizes, in both cases, the Aubin continuity. In fact,  $\mathcal{F}_{\bar{a}}$  has a closed and convex graph and the equivalence between the Aubin property and SCQ comes from the *Robinson-Ursescu Theorem* (as SCQ is equivalent to  $\bar{b} \in \text{int dom } \mathcal{F}_{\bar{a}}$ ). However, this classical result does not apply for multifunction  $\mathcal{F}$ , whose graph is closed but not convex. In this case, the equivalence between the Aubin continuity and SCQ is established via specific arguments of linear systems (see again [88, Theorem 6.9]).

For continuous LSIO problems, the maximum function  $\bar{s} : \mathbb{R}^n \rightarrow \mathbb{R}$ , given by

$$\bar{s}(x) := \max \{\bar{b}_t - \bar{a}'_t x, t \in T\}, \quad (25)$$

is a key tool in the calmness analysis of  $\mathcal{F}$ . Obviously

$$\mathcal{F}_{\bar{a}}(\bar{b}) = \{x \in \mathbb{R}^n : \bar{s}(x) \leq 0\},$$

and

$$d(\bar{b}, \mathcal{F}_{\bar{a}}^{-1}(x)) = [\bar{s}(x)]_+.$$

By the Ioffe-Tikhomirov theorem (see, for instance, [193, Theorem 2.4.18]), we have, the convex subdifferential of  $\bar{s}$  at any  $x \in \mathbb{R}^n$  is

$$\partial \bar{s}(x) = \text{conv} \{-\bar{a}_t : \bar{b}_t - \bar{a}'_t x = \bar{s}(x), t \in T\}.$$

Observe that when the inequality system does not contain the trivial inequality  $0'_n x \geq 0$ ,  $\bar{x} \in \text{int}\mathcal{F}_{\bar{a}}(\bar{b})$  if and only if  $\bar{x}$  is a Slater point, and then  $\mathcal{F}_{\bar{a}}$  is clearly calm at  $(\bar{b}, \bar{x})$ . So, we will focus exclusively on the case when  $\bar{x} \in \text{bd}\mathcal{F}_{\bar{a}}(\bar{b})$ , in which case  $\bar{s}(\bar{x}) = 0$ , and so

$$\partial\bar{s}(\bar{x}) = \text{conv} \left\{ -\bar{a}_t : t \in T_{(\bar{a}, \bar{b})}(\bar{x}) \right\},$$

with  $T_{(\bar{a}, \bar{b})}(\bar{x}) := \{t_0 \in T : \bar{b}_{t_0} - \bar{a}'_{t_0}x = 0\}$  denoting the set of active indices at  $\bar{x}$ . The associated active cone will be

$$A_{(\bar{a}, \bar{b})}(\bar{x}) = \text{cone}\{\bar{a}_t : t \in T_{(\bar{a}, \bar{b})}(\bar{x})\}.$$

We can establish the following list of equivalences:

**Theorem 9 (Calmness of the feasible set mapping)** (a) Let  $\bar{x} \in \text{bd}\mathcal{F}_{\bar{a}}(\bar{b})$ . The following are equivalent:

- (i)  $\mathcal{F}_{\bar{a}}$  is calm at  $(\bar{b}, \bar{x}) \in \text{gph}\mathcal{F}$ ;
- (ii)  $\bar{s}$  has a local error bound at  $\bar{x}$ ; i.e., there exist a constant  $\kappa \geq 0$  and a neighborhood  $U_{\bar{x}}$  of  $\bar{x}$  such that

$$d(x, [\bar{s} \leq 0]) \leq \kappa [\bar{s}(x)]_+, \forall x \in U_{\bar{x}};$$

$$(iii) \liminf_{\substack{x \rightarrow \bar{x}, \\ \bar{s}(x) > 0}} d(0_n, \partial\bar{s}(x)) > 0;$$

- (iv) the strong basic constraint qualification at  $\bar{x}$  holds; i.e., there exist  $\tau > 0$  and neighborhood of  $\bar{x}$ ,  $U_{\bar{x}}$ , such that

$$N(\mathcal{F}_{\bar{a}}(\bar{b}), x) \cap \mathbb{B}_n \subset [0, \tau]\partial\bar{s}(x), \forall x \in U_{\bar{x}} \cap \text{bd}\mathcal{F}_{\bar{a}}(\bar{b}), \quad (26)$$

where  $N(\mathcal{F}_{\bar{a}}(\bar{b}), x)$  is the normal cone to  $\mathcal{F}_{\bar{a}}(\bar{b})$  at  $x$ .

$$(v) \liminf_{\substack{x \rightarrow \bar{x}, \\ \bar{s}(x) > 0}} \sup_{u \neq x} \frac{[\bar{s}(x) - [\bar{s}(u)]_+]_+}{\|x - u\|} > 0;$$

- (vi) there exist  $\lambda_0 > 0$  and a neighborhood  $U_{\bar{x}}$  of  $\bar{x}$  such that, for each  $x \in U_{\bar{x}}$  with  $\bar{s}(x) > 0$  we can find  $u_x \in \mathbb{R}^n$ , with  $\|u_x\| = 1$ , and  $\mu_x > 0$  satisfying

$$\bar{s}(x + \mu_x u_x) \leq \bar{s}(x) - \mu_x \lambda_0. \quad (27)$$

- (vii) there exists a neighborhood  $U_{\bar{x}}$  of  $\bar{x}$  such that

$$N(\mathcal{F}_{\bar{a}}(\bar{b}), x) = \text{cl}A_{(\bar{a}, \bar{b})}(x), \forall x \in \mathcal{F}_{\bar{a}}(\bar{b}) \cap U_{\bar{x}}, \quad (28)$$

and, additionally, there is  $\tau > 0$  such that

$$A_{(\bar{a}, \bar{b})}(x) \cap \mathbb{B}_n \subset [0, \tau](\partial\bar{s}(x)), \forall x \in \text{bd}\mathcal{F}_{\bar{a}}(\bar{b}) \cap U_{\bar{x}}. \quad (29)$$

Moreover, in relation to (iii) and (v) we have

$$\text{clm}\mathcal{F}_{\bar{a}}(\bar{b}, \bar{x}) = \left( \liminf_{\substack{x \rightarrow \bar{x}, \\ \bar{s}(x) > 0}} d(0_n, \partial \bar{s}(x)) \right)^{-1} \quad (30)$$

$$= \left( \liminf_{\substack{x \rightarrow \bar{x}, \\ \bar{s}(x) > 0}} \sup_{u \neq x} \frac{[\bar{s}(x) - [\bar{s}(u)]_+]_+}{\|x - u\|} \right)^{-1}. \quad (31)$$

Concerning (vi), the infimum of constants  $\lambda_0$  in (27) (for some associated  $U_{\bar{x}}$ ) coincides with  $1/\text{clm}\mathcal{F}_{\bar{a}}(\bar{b}, \bar{x})$ . Finally, with respect to (vii), the infimum of constants  $\tau$  in (29) (for some associated  $U_{\bar{x}}$ ) also coincides with the calmness modulus  $\text{clm}\mathcal{F}_{\bar{a}}(\bar{b}, \bar{x})$ .

(b) One has

$$\text{clm}\mathcal{F}((\bar{a}, \bar{b}), \bar{x}) = (\|\bar{x}\| + 1) \text{clm}\mathcal{F}_{\bar{a}}(\bar{b}, \bar{x}), \quad (32)$$

and therefore,  $\mathcal{F}$  is calm at  $((\bar{a}, \bar{b}), \bar{x}) \in \text{gph}\mathcal{F}$  if and only if  $\mathcal{F}_{\bar{a}}$  is calm at  $(\bar{b}, \bar{x}) \in \text{gph}\mathcal{F}_{\bar{a}}$ .

In part (a), the equivalence (i)  $\Leftrightarrow$  (ii) comes from Azé and Corvellec [11, Proposition 2.1 and Theorem 5.1]; (i)  $\Leftrightarrow$  (v) follows from Fabian *et al.* [61, Theorem 2(ii)]; (v)  $\Leftrightarrow$  (vi) is a consequence of the convexity of  $\bar{s}$ ; (i)  $\Leftrightarrow$  (vi) is a linear semi-infinite version of [124, Theorem 3]; (i)  $\Leftrightarrow$  (iv) follows from Zheng and Ng [195, Theorem 2.2]; (i)  $\Leftrightarrow$  (vii) is Theorem 3 in Cánovas *et al.* [37], (30) can be found in Kruger *et al.* [128, Th. 1]; and (31) follows from [61, Theorem 1(ii)]. Condition (28) in (vii) is called in [37] *Abadie constraint qualification around  $\bar{x}$* , whereas (29) constitutes a kind of uniform boundedness of the scalars involved in certain conic combinations. In part (b), (32) is established in Theorem 5 in [37].

When  $T$  is finite (i.e. for finite systems of inequalities), Theorem 3.1 in [30]) gives the following characterization of  $1/\text{clm}\mathcal{F}_{\bar{a}}(\bar{b}, \bar{x})$ :

$$\limsup_{x \rightarrow \bar{x}, \bar{s}(x) > \bar{s}(\bar{x})} \partial \bar{s}(x) = \bigcup_{D \in \mathcal{D}(\bar{x})} \text{conv}\{-\bar{a}_i, i \in D\}, \quad (33)$$

where

$$\mathcal{D}(\bar{x}) := \left\{ D \subset T_{(\bar{a}, \bar{b})}(\bar{x}) \left| \begin{array}{l} \text{there exists } d \text{ verifying:} \\ \left\{ \begin{array}{ll} \bar{a}'_i d = 1, & i \in D, \\ \bar{a}'_i d < 1, & i \in T_{(\bar{a}, \bar{b})}(\bar{x}) \setminus D \end{array} \right\} \end{array} \right. \right\},$$

and (30) gives rise to the following result [37, Theorem 4]:

$$\text{clm}\mathcal{F}_{\bar{a}}(\bar{b}, \bar{x}) = \max_{D \in \mathcal{D}(\bar{x})} (d(0_n, \text{conv}\{\bar{a}_i, i \in D\}))^{-1}. \quad (34)$$

By introducing two new families of indices sets, namely

$$\mathcal{D}_{AI}(\bar{x}) := \{D \in \mathcal{D}(\bar{x}) : \text{such that } \{\bar{a}_i, i \in D\} \text{ is affinely independent}\}, \quad (35)$$

and

$$\mathcal{D}^0(\bar{x}) := \left\{ D \subset T_{(\bar{a}, \bar{b})}(\bar{x}) \left| \begin{array}{l} \text{there exists } d \neq 0_n \text{ verifying:} \\ \left\{ \begin{array}{ll} \bar{a}'_i d = 0, & i \in D, \\ \bar{a}'_i d < 0, & i \in T_{(\bar{a}, \bar{b})}(\bar{x}) \setminus D \end{array} \right\} \end{array} \right. \right\},$$

we get the following relations (see [30, Theorem 3.2] for the inclusions, and (37) below for the equality):

$$\begin{aligned} \bigcup_{D \in \mathcal{D}_{AI}(\bar{x})} \text{conv} \{-\bar{a}_i, i \in D\} &\subset \limsup_{x \rightarrow \bar{x}, \bar{s}(x) > \bar{s}(\bar{x})} \partial \bar{s}(x) \\ &\subset \bigcup_{D \in \mathcal{D}(\bar{x}) \cup \mathcal{D}^0(\bar{x})} \text{conv} \{-\bar{a}_i, i \in D\} \subset \text{bd } \partial \bar{s}(\bar{x}) \\ &= \limsup_{x \rightarrow \bar{x}} \partial \bar{s}(x). \end{aligned} \quad (36)$$

Under the *Linear Independence Constraint Qualification*, i.e. when the vectors  $\{\bar{a}_i, i \in T_{(\bar{a}, \bar{b})}(\bar{x})\}$  are linearly independent, all the inclusions above become equalities since the system  $\{\bar{a}'_i d = 1; i \in I(\bar{x})\}$  obviously has a solution; in other words,

$$T_{(\bar{a}, \bar{b})}(\bar{x}) \in \mathcal{D}_{AI}(\bar{x}),$$

and so

$$\begin{aligned} \bigcup_{D \in \mathcal{D}_{AI}(\bar{x})} \text{conv} \{-\bar{a}_i, i \in D\} &= \text{conv} \{-\bar{a}_i, i \in T_{(\bar{a}, \bar{b})}(\bar{x})\} \\ &= \text{bdconv} \{-\bar{a}_i, i \in T_{(\bar{a}, \bar{b})}(\bar{x})\} = \text{bd } \partial \bar{s}(\bar{x}). \end{aligned}$$

Moreover, the well-known *Mangasarian-Fromovitz Constraint Qualification* is not sufficient to guarantee equality in any of the inclusions in (36) as the example in [30, Remark 3.3(ii)] shows. Let us note that (36) is established in [30, Remark 3.3(ii)] for finite inequality systems involving continuously differentiable functions.

In [60] the authors generalize the outer subdifferential construction (33), given by Cánovas *et al.* in [30, Theorem 3.2], from supremum functions to pointwise minima of regular Lipschitz functions. They provide an improvement of Theorem 3.2 in [30] by replacing the family  $\mathcal{D}_{AI}(\bar{x})$  in (36) with  $\mathcal{D}(\bar{x})$ . Also in [132] a refinement of this theorem in [30] is given.

In [26] specific formulas for the calmness modulus of feasible set mappings associated with partially perturbed linear inequality systems are given, taking [37, Section 4] as starting point. The paper analyzes the relationship between the calmness modulus of a particular feasible set mapping (related to the KKT conditions for the original LP problem) and the linear rate of convergence for the distance between the central path of a linear program and its primal-dual optimal set.

Another couple of applications of the modulus of calmness for fully perturbed linear programs (Scenario 1) can be found in [28, §5]. The first one

to a descent method in LP, and the second to a regularization method for linear programs with complementarity constraints.

Concerning the optimal set mapping  $\mathcal{S}$  in LSIO, Theorem 3.1 in [27] establishes that, in Scenario 4 and under SCQ, perturbations of  $\bar{c}$  are negligible when characterizing the calmness of  $\mathcal{S}_{\bar{a}}$ , and that this property is equivalent to the calmness of the level set mapping

$$(\alpha, b) \mapsto \{x \in \mathbb{R}^n : \bar{c}'x \leq \alpha; \bar{a}_t'x \geq b_t \quad \forall t \in T\}$$

at  $((\bar{c}'\bar{x}, \bar{b}), \bar{x})$ . If  $T$  is finite, the SCQ continues to be a key property but the finiteness of  $T$  is no longer a sufficient condition for the calmness of  $\mathcal{S}$ , in the fully-perturbed setting or Scenario 1, as the following result shows:

**Theorem 10 (Calmness of the optimal set mapping)** [29, Theorem 4.1] *Assume that  $T$  is finite ( $T = \{1, 2, \dots, m\}$ ) and  $\mathcal{S}(\bar{c}, \bar{a}, \bar{b}) = \{\bar{x}\}$ . The following are equivalent:*

- (i)  $\mathcal{S}$  is calm at  $((\bar{c}, \bar{a}, \bar{b}), \bar{x})$ ;
- (ii) either SCQ holds at  $(\bar{a}, \bar{b})$  or  $\mathcal{F}(\bar{a}, \bar{b}) = \{\bar{x}\}$ ;
- (iii)  $0_n \notin \text{bd conv} \left\{ \bar{a}_i, i \in T_{(\bar{a}, \bar{b})}(\bar{x}) \right\}$ .

Additionally, [29] provides, for  $T$  finite, a formula for the calmness modulus of  $\mathcal{S}$  which is exact in Scenario 4 and an upper bound in Scenario 1, respectively.

In [32] a lower bound on  $\text{clm}\mathcal{S}_{\bar{a}}((\bar{c}, \bar{b}), \bar{x})$  for an LSIO with a unique optimal solution, under SCQ and in Scenario 4, is given in Theorem 6. It turns out that this lower bound equals the exact modulus when  $T$  is finite without requiring either SCQ or the uniqueness of  $\bar{x}$  as optimal solution of  $P$ . Also when  $T$  is finite and the optimal set is a singleton, a new upper bound for  $\text{clm}\mathcal{S}_{\bar{a}}((\bar{c}, \bar{b}), \bar{x})$  is proposed in Theorem 13. This upper bound is easily computable, as it is formulated exclusively in terms of the nominal data  $\bar{c}$ ,  $\bar{b}$ , and  $\bar{x}$ , but examples show that it may not be attained.

In [28], the main result (Theorem 3.1) gives the following characterization of the boundary of the subdifferential set of a convex function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  at any point  $\bar{x} \in \mathbb{R}^n$ :

$$\text{bd } \partial f(\bar{x}) = \limsup_{x \neq \bar{x}, x \rightarrow \bar{x}} \partial f(x). \quad (37)$$

Using (37), the authors provide in Theorem 4.1 an upper bound of  $\text{clm}\mathcal{S}((\bar{c}, \bar{a}, \bar{b}), \bar{x})$  for a continuous LSIO such that  $\mathcal{S}(\bar{c}, \bar{a}, \bar{b}) = \{\bar{x}\}$  and SCQ holds.

When  $T$  is finite, in order to provide estimates of  $\text{clm}\mathcal{S}((\bar{c}, \bar{a}, \bar{b}), \bar{x})$ , with  $((\bar{c}, \bar{a}, \bar{b}), \bar{x}) \in \text{gph}\mathcal{S}$ , and taking KKT optimality conditions into account, we define:

$$\mathcal{M}(\bar{x}) := \left\{ \begin{array}{l} D \subset T_{(\bar{a}, \bar{b})}(\bar{x}) : -\bar{c} \in \text{cone} \{\bar{a}_i, i \in D\} \\ \text{and } D \text{ is minimal for the inclusion order} \end{array} \right\}.$$

If  $D \in \mathcal{M}(\bar{x})$ , Carathéodory's Theorem yields  $|D| \leq n$ .

Associated with each  $D \in \mathcal{M}(\bar{x})$  we consider the mapping  $\mathcal{L}_D : \mathbb{R}^m \times \mathbb{R}^D \rightrightarrows \mathbb{R}^n$

$$\mathcal{L}_D(b, d) := \{x \in \mathbb{R}^n : \bar{a}'_i x \geq b_i, i = 1, \dots, m; -\bar{a}'_i x \geq d_i, i \in D\}.$$

Corollary 4.1 in [30, Corollary 4.1] establishes that

$$\text{clm}\mathcal{S}_{\bar{a}}((\bar{c}, \bar{b}), \bar{x}) = \sup_{D \in \mathcal{M}(\bar{x})} \text{clm}\mathcal{L}_D((\bar{b}, -\bar{b}_D), \bar{x}),$$

meanwhile  $\sup_{D \in \mathcal{M}(\bar{x})} \text{clm}\mathcal{L}_D((\bar{b}, -\bar{b}_D), \bar{x})$  is only a lower bound of  $\text{clm}\mathcal{S}_{\bar{a}}((\bar{c}, \bar{b}), \bar{x})$

for the semi-infinite optimization problem [30, Theorem 4.2]. Moreover, under SCQ,  $\mathcal{S}(\bar{c}, \bar{a}, \bar{b}) = \{\bar{x}\}$ , and for small  $\|\bar{c}\|$ , one has [31, Section 5]

$$\text{clm}\mathcal{S}((\bar{c}, \bar{a}, \bar{b}), \bar{x}) = (\|\bar{x}\| + 1) \text{clm}\mathcal{S}_{\bar{a}}((\bar{c}, \bar{b}), \bar{x}).$$

The Aubin property of  $S$  for linear programs under full perturbations is characterized in the first part of [38], whereas in the second part of that paper a formula for computing the exact Lipschitz modulus is given.

Table 4 summarizes the recent literature on calmness in LSIO that we have commented above. Column 1 enumerates the references in chronological order, columns 2 and 3 inform on the type of LSIO problem each reference deals with (either a finite one or a continuous one, or both), column 4 indicates the considered scenarios, column 5 whether the corresponding paper provides the exact value of  $\text{clm}\mathcal{F}(\bar{\theta}, \bar{x})$  or not, while columns 6 and 7 provide a similar information regarding  $\text{clm}\mathcal{S}(\bar{\theta}, \bar{x})$  through estimation or exact formulas, respectively.

Paper	Finite	Cont.	Scenarios	$\mathcal{F}$ exact	$\mathcal{S}$ est	$\mathcal{S}$ exact
[27]	✓	✓	4		✓	
[32]	✓	✓	4		✓	✓
[37]	✓	✓	2, 6	✓		
[28]		✓	1		✓	
[29]	✓		1, 4		✓	✓
[30]	✓	✓	4			✓
[31]	✓		1			
[38]	✓		1, 4			
[26]	✓		1, 4	✓		

Table 4

### 3.3 Robust linear semi-infinite optimization

Suppose that we are given an uncertain LSIO problem  $\bar{P}$  as in (19), where  $\bar{c} \in \mathbb{R}^n$  and  $(\bar{a}, \bar{b}) \in (\mathbb{R}^n)^T \times \mathbb{R}^T$  are uncertain. The robust approach transforms  $\bar{P}$  into deterministic problems involving the so-called *uncertainty sets*.

We assume, for the sake of simplicity, that the uncertainty sets for the vectors  $(\bar{a}_t, \bar{b}_t)$  are Euclidean closed balls centered at  $(\bar{a}_t, \bar{b}_t)$  of the same radius, say  $\alpha \geq 0$ , i.e., the sets

$$U_t^\alpha := (\bar{a}_t, \bar{b}_t) + \alpha \mathbb{B}_{n+1}, \quad t \in T.$$

We also assume that the uncertainty sets for the cost vector  $\bar{c}$  is an Euclidean closed ball centered at  $\bar{c}$  of radius  $\beta \geq 0$ , i.e., the set

$$V^\beta := \bar{c} + \beta \mathbb{B}_n.$$

The user must decide the value of the parameters  $\alpha$  and  $\beta$  in such a way that the uncertainty sets are sufficiently large to contain any conceivable value of the real data but sufficiently small to guarantee the existence of robust solutions.

Pessimistic decision makers only accept solutions which are immunized against infeasibility. Once some  $\alpha \geq 0$  has been chosen, the solutions of the linear semi-infinite system

$$\{a'x \geq b, (a, b) \in U_t^\alpha, t \in T\} \quad (38)$$

are called *robust feasible solutions*, which form the so-called *robust feasible set*  $F_R^\alpha$ .

The following question arises in a natural way: how large can we take  $\alpha$  in order to guarantee that  $F_R^\alpha \neq \emptyset$ ? Since  $\{\alpha \geq 0 : F_R^\alpha \neq \emptyset\}$  is a bounded interval containing 0, it is enough to compute its supremum, the so-called *radius of robust feasibility* of  $\bar{P}$ :

$$\rho_M(\bar{P}) := \sup \{\alpha \geq 0 : F_R^\alpha \neq \emptyset\}, \quad (39)$$

where  $\sup \emptyset = -\infty$ .

The following characterization of  $\rho_M(\bar{P})$ , proved in [84, Theorem 2.5 and Proposition 2.3], is based on the existence theorem (6) and arguments similar to those used in [33] to compute the distance to ill-posedness in the consistency sense in the parametric setting (Scenarios 1 and 2). The formula involves the set

$$H(\bar{a}, \bar{b}) := \text{conv} \{(\bar{a}_t, \bar{b}_t), t \in T\} + \mathbb{R}_+ \{(0_n, -1)\},$$

which is called *hypographical set* of  $\bar{P}$  in [33].

**Theorem 11 (The radius of robust feasibility)** *The following equation holds:*

$$\rho_M(\bar{P}) = d(0_{n+1}, H(\bar{a}, \bar{b})). \quad (40)$$

*Moreover, if the recession cone of the feasible set of  $\bar{P}$  is a linear subspace, then the supremum in (39) is attained.*

In contrast with the existing consensus on the type of acceptable solutions by the robust optimization community, there is a controversy on the type of optimal solutions to be selected once  $\alpha \geq 0$  such that  $F_R^\alpha \neq \emptyset$  and  $\beta \geq 0$  have been chosen. These are the most popular concepts of robust optimal solutions:

- A vector  $\bar{x} \in F_R^\alpha$  is a *minmax (or worst-case) optimal solution* when it is an optimal solution of the so-called *robust (or pessimistic) counterpart* of  $\bar{P}$  (a deterministic convex semi-infinite optimization problem):

$$\begin{aligned} & \inf_{x \in \mathbb{R}^n} \max_{c \in V^\beta} c'x \\ & \text{s.t.} \quad a'x \geq b, \quad (a, b) \in U_t^\alpha, \quad t \in T. \end{aligned} \quad (41)$$

We denote by  $S_M^{\alpha, \beta}$  the set of minmax optimal solutions.

- A vector  $\bar{x} \in F_R^\alpha$  is a *highly robust optimal solution* when it is an optimal solution of

$$\begin{aligned} & \inf_{x \in \mathbb{R}^n} c'x \\ & \text{s.t.} \quad a'x \geq b, \quad (a, b) \in U_t^\alpha, \quad t \in T, \end{aligned}$$

for all  $c \in V^\beta$ . We denote by  $S_H^{\alpha, \beta}$  the set of highly robust optimal solutions.

Once  $\alpha \geq 0$  has been chosen, it is also natural to ask how large the parameter  $\beta$  can be taken in order to guarantee the existence of minmax optimal solutions and/or highly robust optimal solutions.

We define the *radius of minmax robust efficiency* of  $\bar{P}$  for a given  $\alpha \geq 0$  such that  $F_R^\alpha \neq \emptyset$  as

$$\rho_M^\alpha := \sup \left\{ \beta \geq 0 : S_M^{\alpha, \beta} \neq \emptyset \right\}.$$

Analogously, the *radius of highly robust efficiency* of  $\bar{P}$  for a given  $\alpha$  such that  $F_R^\alpha \neq \emptyset$  as

$$\rho_H^\alpha := \sup \left\{ \beta \geq 0 : S_H^{\alpha, \beta} \neq \emptyset \right\}.$$

The value of  $\rho_H^\alpha$  is related to the existence of strongly unique solution to the problem

$$\begin{aligned} & \inf_{x \in \mathbb{R}^n} \bar{c}'x \\ & \text{s.t.} \quad a'x \geq b, \quad (a, b) \in U_t^\alpha, \quad t \in T. \end{aligned} \quad (42)$$

Recall that a vector  $\bar{x} \in F_R^\alpha$  is said to be *isolated (or strongly unique) solution* for the problem in (42) if there exists  $\kappa > 0$  such that

$$\bar{c}'x \geq \bar{c}'\bar{x} + \kappa \|x - \bar{x}\| \quad \text{for all } x \in F_R^\alpha. \quad (43)$$

The infimum of such  $\kappa$  verifying (43) is called *isolation modulus* of  $\bar{x}$ , that we denote by  $\text{iso}(\alpha; \bar{x})$ . It is known [88, Theorem 10.5] that  $\bar{x} \in F_R^\alpha$  is an isolated solution for the problem in (42) if and only if  $\bar{c} \in \text{int } D(F_R^\alpha; \bar{x})^\circ$ . Recall that  $D(F_R^\alpha; \bar{x})^\circ$  can be recovered from  $(\bar{a}, \bar{b})$  when the system (38) satisfies LFMCQ at  $\bar{x}$ .

The next theorem is the specialization to scalar optimization of results in the recent work [86] (on robust multi-objective convex optimization).

**Theorem 12 (The radii of minmax and highly robustness )** *Let  $\alpha \geq 0$  be such that  $F_R^\alpha \neq \emptyset$ . Then, the following statements hold:*

- (i)  $\rho_M^\alpha = +\infty$ .
- (ii)  $\rho_H^\alpha > 0$  if and only if the problem in (42) has an isolated solution  $\bar{x}$ , in which case  $\rho_H^\alpha = \text{iso}(\alpha; \bar{x})$ .

Let us observe that

$$\inf \{ \alpha \geq 0 : F_R^\alpha \neq \emptyset \} = \inf \{ \beta \geq 0 : S_H^{\alpha, \beta} \neq \emptyset \} = 0,$$

while

$$\text{dist} \left( -\bar{c}, \text{conv} \left\{ \bigcup_{x \in F_R^\alpha} D(F_R^\alpha; x)^\circ \right\} \right) \leq \inf \{ \beta \geq 0 : S_M^{\alpha, \beta} \neq \emptyset \} \leq \|\bar{c}\|.$$

Optimality conditions for robust multi-objective (and so for scalar) LSIO problems with arbitrary uncertainty sets can be found in [84, Theorem 3.3] whereas duality theorems for robust scalar (respectively, multi-objective) LSIO problems involving the optimistic counterparts are given in [83] (respectively, [84]).

Concerning the numerical treatment of the robust counterpart problem in (41), let us observe that, taken into account the special structure of the index set  $\bigcup_{t \in \mathcal{T}} U_t^\alpha \subset \mathbb{R}^{n+1}$ , only discretization methods can be applied, with the serious inconvenient that grid discretization methods are only efficient provided the dimension  $n + 1$  of the index set is sufficiently small.

The stability of robust optimization problems with respect to perturbations in their uncertainty sets is studied in [39]. In particular, the paper focus on robust LSIO problems, and considers uncertainty in both the cost function and constraints. The authors prove Lipschitz continuity properties of the optimal value and the  $\varepsilon$ -optimal solution set with respect to the Hausdorff distance between uncertainty sets.

#### 4 Selected applications 2007-2016

Chapter 1 of [90] describes the way the primal LSIO problem  $P$  arises in functional approximation, pattern recognition, environmental decision making, generalized Neyman-Pearson problem for grouped data, optimal experimental design in regression, maximum likelihood estimation, semi-definite programming, and geometric programming. In turn, Chapter 2 describes applications of the dual problem  $D$  to data envelopment analysis, location problems, and robustness in Bayesian statistics. We now describe three new applications of linear semi-infinite optimization and provide comments on other applications spread on the recent literature.

#### 4.1 Linear conic optimization

Consider the conic optimization problem

$$\begin{aligned} \bar{P}_K : \quad & \inf_{x \in \mathbb{R}^n} \bar{c}'x \\ \text{s.t.} \quad & \begin{bmatrix} \bar{a}'_1 x + \bar{b}_1 \\ \vdots \\ \bar{a}'_m x + \bar{b}_m \end{bmatrix} \in -K, \end{aligned}$$

where  $\bar{c} \in \mathbb{R}^n$ ,  $(\bar{a}_i, \bar{b}_i) \in \mathbb{R}^{n+1}$ ,  $1 \leq i \leq m$ , and  $K$  is a closed pointed convex cone in  $\mathbb{R}^m$  such that  $\text{int } K \neq \emptyset$ . Since  $K^{\circ\circ} = K$ , given  $x \in \mathbb{R}^n$ ,

$$\begin{bmatrix} \bar{a}'_1 x + \bar{b}_1 \\ \vdots \\ \bar{a}'_m x + \bar{b}_m \end{bmatrix} \in -K \iff \sum_{i=1}^m \lambda_i \bar{a}'_i x \geq - \sum_{i=1}^m \lambda_i \bar{b}_i, \forall \lambda = (\lambda_1, \dots, \lambda_m) \in K^\circ,$$

so that  $\bar{P}_K$  can be reformulated as an LSIO problem. Since the convergence of most numerical approaches described in Subsection 2.3 requires the given LSIO problem to be continuous, it is convenient to reformulate again  $\bar{P}_K$  as

$$\begin{aligned} \bar{P}_K : \quad & \inf_{x \in \mathbb{R}^n} \bar{c}'x \\ \text{s.t.} \quad & \sum_{i=1}^m \lambda_i \bar{a}'_i x \geq - \sum_{i=1}^m \lambda_i \bar{b}_i, \lambda \in \mathcal{B}, \end{aligned} \tag{44}$$

where  $\mathcal{B}$  is any *compact base* of  $K^\circ$ ; i.e., a compact convex subset of  $K^\circ$  such that  $0_m \notin \mathcal{B}$  and  $\text{cone } \mathcal{B} = K^\circ$  (called *compact sole* in [7, p. 57]). The condition that  $0_m \notin \mathcal{B}$  plays an important role in theoretical and numerical aspects of  $\bar{P}_K$  as  $0_m \in \mathcal{B}$  is incompatible with the desirable SCQ.

The cone  $K \subset \mathbb{R}^m$ , with  $m = q + \sum_{j=1}^q n_j$ , in [109] is the cartesian product of  $q$  second order cones of the form  $\{x \in \mathbb{R}^{n_j+1} : x_{n_j+1} \geq \|(x_1, \dots, x_{n_j})\|\}$ ,  $j = 1, \dots, q$ ; so, the equivalent continuous LSIO problem can be solved by means of some LSIO method; the authors preferred a dual simplex method whose convergence is not guaranteed (continuity is not enough). Since the dimension of  $\mathcal{B}$  is  $m$ ,  $\bar{P}_K$  could be solved using discretization methods when  $m$  is small.

The LSIO theory has also been used in this setting. Indeed, some results in [59], on genericity of strong duality and weak genericity of uniqueness in linear conic optimization (specially those in Subsection 4.1), have been obtained appealing to the LSIO reformulation of  $\bar{P}_K$ . However, the known genericity results on LSIO (see [101], [150] and references therein) cannot be directly transferred to linear conic optimization as the class of LSIO reformulations represents a small subset of  $\Theta_0$  (see [59, Subsection 4.4]).

#### 4.2 Robust linear and convex programming

The robust counterpart,  $P_R^\alpha$  in (41), with  $\alpha > 0$ , of any linear optimization problem  $\bar{P}$  is a continuous LSIO problem even though  $T$  is finite, so that  $P_R^\alpha$  could be solved by means of one of the numerical approaches described in Subsection 2.3. Due to the difficulties encountered by the numerical LSIO methods to solve  $P_R^\alpha$  (recall the argument in the last paragraph of Subsection 3.3), many authors impose strong conditions on  $\bar{P}$  allowing to reformulate  $P_R^\alpha$  as a tractable finite convex problem (see [15] and references therein). However, when these assumptions fail, LSIO methods can be useful. Indeed, there is a stream of works dealing with waste management under uncertainty where the uncertain LP problem is solved through its reformulation as an LSIO problem. This approach is compared with the interval linear optimization one in [105], [110], [111], [114], [134], [147] (a paper on water resources management), whose authors conclude that the LSIO approach has the following advantages:

- (i) It reflects better the association of the total system revenue with gas and power prices;
- (ii) it generates more reliable solutions with a lower risk of system failure due to the possible constraints violation; and
- (iii) it provides a more flexible management strategy since the capital availability can be adjusted with the variations in gas prices.

An integer LSIO model has been proposed in [196] for planning municipal energy systems.

However, LSIO theory, in particular the existence theorem (6), plays a crucial role in getting formulas for the radius of robust feasibility for finite linear and convex problems with uncertain constraints (in the latter case, together with the calculus rules for conjugate functions). A direct simple proof of (40) for robust linear programs has been given in [85]. The Euclidean balls are replaced with the much more general class of spectrahedral sets (including ellipsoids, balls, polytopes and boxes) in [45], where the formulas for the radius of robust feasibility involves the Minkowski function (or gauge) of the corresponding spectrahedral set.

Formulas for the radius of robust feasibility have been provided in [82] for convex programs with uncertain polynomial constraints. More in detail, this paper considers a nominal convex problem of the form

$$\begin{aligned} \bar{P} : \quad & \inf_{x \in \mathbb{R}^n} \bar{f}(x) \\ & \text{s.t. } \bar{g}_j(x) \leq 0, \quad j = 1, \dots, q, \end{aligned}$$

where  $\bar{f} : \mathbb{R}^n \rightarrow \mathbb{R}$  is a convex function and  $\bar{g}_j : \mathbb{R}^n \rightarrow \mathbb{R}$  is a convex polynomial for all  $j$ . The robust counterpart of  $\bar{P}$ , depending on a parameter  $\alpha \geq 0$ , is the problem

$$\begin{aligned} P_R^\alpha : \quad & \inf_{x \in \mathbb{R}^n} \bar{f}(x) \\ & \text{s.t. } \bar{g}_j(x) + \sum_{l=1}^p v_j^l \bar{g}_j^l(x) + a_j'x + b_j \leq 0, \\ & (v_j, a_j, b_j) \in \alpha(M \times \mathbb{B}_{n+1}), \quad j = 1, \dots, q, \end{aligned}$$

where  $\bar{g}_j^l$  are given convex polynomials on  $\mathbb{R}^n$ ,  $l = 1, \dots, p$ ,  $v_j = (v_j^1, \dots, v_j^p) \in \mathbb{R}^p$ ,  $M \subset \mathbb{R}_+^p$  is a given convex compact set with  $0_p \in M$  and  $\bar{\alpha}_j \geq 0$ ,  $j \in J$ . An upper bound for the radius of robust feasibility is given in [82, Theorem 2.1] and an exact formula in [82, Theorem 3.1] under the assumption that  $\bar{g}_j$  is a *sum-of-squares* (i.e., it is the finite sum of squares of polynomial functions) for all  $j$ . This formula expresses  $\rho_M(\bar{P})$  as the square root of the optimal value of some semi-definite optimization problem.

#### 4.3 Machine Learning

Pattern classification is to classify some unseen object into one of the given classes (or categories). This is done by means of mathematical models, called *classifiers*, which can be of different nature: statistical models (which select a discriminant function on the basis of a priori information on the probabilistic distributions of the classes), neural networks, fuzzy systems, deterministic systems or optimization models. A classifier is implemented with the help of a given (training) data set of examples that have been previously classified (or labelled), say  $\{x^i, i = 1, \dots, N\} \subset \mathbb{R}^n$ . An optimization model for classification is called a *machine learning* or *support vector machine* (SVM in brief), and such a model becomes an optimization problem introducing the training data (the input); the optimal solution provides *decision functions* (the output) whose values on unseen objects determine their classification.

In the simplest case of two classes contained in  $\mathbb{R}^n$ , a unique decision function  $g : \mathbb{R}^n \mapsto \mathbb{R}$  belonging to a predetermined family classifies an object  $x \in \mathbb{R}^n$  in the 1st (respectively, 2nd) class if  $g(x) > 0$  (respectively,  $g(x) < 0$ ). The decision function  $g$  is computed by solving an optimization problem with one constraint associated with each example. A training example  $x^i$  is called *support vector* if the corresponding dual variable is non-zero, i.e., if a small perturbation of  $x^i$  affects the output  $g$ . For instance, if the sets of examples of both classes can be strictly separated by means of hyperplanes, it is sensible to select an affine function  $g(x) = w'x - b$ , with  $w \in \mathbb{R}^n \setminus \{0_n\}$  (called *weight vector*) and  $b \in \mathbb{R}$  (called *bias*), such that the corresponding hyperplane  $H := \{w \in \mathbb{R}^n \mid w'x = b\}$  separates the examples in both classes maximizing the distance to the training data set. In that case, the training examples closest to  $H$  are the support vectors. Most SVM algorithms exploit the empirical fact that the cardinality of the set of support vectors is small in comparison with the whole training data set. In some applications, together with the training data set there is either a (possibly infinite) *knowledge set* necessarily contained in one of the two classes or *logical constraints* to be respected. The ordinary (respectively, semi-infinite) version of the non-homogeneous Farkas lemma allows to formulate the inclusion constraint when the contained set is formed by the solution of a given ordinary (respectively, semi-infinite) system. Concerning the introduction of a logical constraint of the form  $t \in T \Rightarrow (Ex - a)'t \geq b$ , where  $T \subset \mathbb{R}^p$  is an infinite set,  $E \in \mathbb{R}^{p \times n}$ ,  $a \in \mathbb{R}^p$  and  $b \in \mathbb{R}$ , as pointed out

by O. Mangasarian [142], it is equivalent to the linear semi-infinite system  $\{(Ex - a)'t \geq b, t \in T\}$ . Prior knowledge over arbitrary general sets has been incorporated in [143] and [144] into nonlinear kernel approximation problems in the form of an LSIO problem by means of a theorem of the alternative for convex functions. The resulting LSIO problem is then solved by means of a discretization procedure. The mentioned papers include numerical examples and an important lymph node metastasis prediction application.

When the two sets cannot be separated by means of hyperplanes, different strategies are possible, for instance, separating the two sets by means of a certain family of nonlinear surfaces (e.g., quadrics), minimizing a certain unconstrained function penalizing (for each given hyperplane in  $\mathbb{R}^n$ ) the misclassified examples, or separating through hyperplanes the images of the examples in a certain space  $\mathbb{R}^l$  (called *featured space*) by means of a suitable mapping  $\phi : \mathbb{R}^n \mapsto \mathbb{R}^l$ . These two strategies involve LSIO problems. First, if the training sets are separated by means of ellipsoids, the corresponding linear semi-definite optimization problem can be reformulated and solved as an LSIO problem. Second, if the examples are mapped on the featured space, the decision function has the form  $f(x) = w'\phi(x) + b$ , and  $x$  is classified into the 1st (respectively, 2nd) class if  $f(x) > 0$  (respectively,  $f(x) < 0$ ). As a consequence of the KKT condition, if  $(w, b)$  is an optimal solution,

there exist multipliers  $\alpha_i$ ,  $i = 1, \dots, N$ , such that  $w = \sum_{i=1}^N \alpha_i y_i \phi(x^i)$ , where

$y_i$  denotes the label  $y_i = +1$  (respectively,  $y_i = -1$ ) if example  $x^i$  belongs to the 1st class (respectively, 2nd class). Replacing in the expression of  $f$ , we get

$$f(x) = \sum_{i=1}^N \alpha_i y_i \phi(x^i)' \phi(x) + b = \sum_{i=1}^N \alpha_i y_i \kappa(x^i, x) + b, \quad (45)$$

where the function  $\kappa : \mathbb{R}^{2n} \mapsto \mathbb{R}$  such that  $\kappa(x, z) = \phi(x)' \phi(z)$ , called *kernel*, is symmetric and can be interpreted as a measure of the similarity between  $x$  and  $z$ . Among the kernels commonly used to classify let us mention the *polynomial kernel*  $\kappa(x, z) = (x'z + 1)^m$ , with  $m \in \mathbb{N}$ , the *Gaussian* (or *radial basis function*) *kernel*  $\kappa(x, z) = \exp(-\gamma \|x - z\|^2)$ , with  $\gamma > 0$ , and the *three-layer neural network kernel*  $\kappa(x, z) = \frac{1}{1 + \exp(\nu x'z - \mu)}$ , with  $\nu, \mu \in \mathbb{R}$ .

Since typical learning problems involve multiple, heterogeneous data sources, it is usually convenient to use a convex combination of kernels instead of just one. Linear semi-infinite optimization is also used in *multiple kernel learning* where, according to [169], an optimal decision function for the given training data is obtained by solving a problem of the form

$$\left. \begin{array}{l} \sup_{\theta \in \mathbb{R}, \beta \in \mathbb{R}^K} \theta \\ \text{s.t.} \quad \theta \leq \sum_{k=1}^K \beta_k S_k(\alpha), \alpha \in A, \\ \sum_{k=1}^K \beta_k = 1, \beta_k \geq 0, k = 1, \dots, K, \end{array} \right\}$$

where the functions  $S_k(\alpha)$  are continuous and

$$A := \left\{ \alpha \in \mathbb{R}^N : \sum_{i=1}^N y_i \alpha_i = 0, 0 \leq \alpha_i \leq C, i = 1, \dots, N \right\}$$

is a polytope. This continuous LSIO problem is solved in [169] by means of a simplex dual algorithm.

Finally, consider the case in which there is an infinite family of kernels, say  $\{\kappa(t), t \in T\}$ , where  $T$  is a one dimensional compact interval (the model could be generalized by replacing the restrictive assumption that  $T$  is an interval by the weaker one that  $T$  is a compact topological space). One of the ways to handle infinitely many kernels consists of considering nonnegative convex combinations of finite subsets, a limitation of the feasible decisions which does not always allow to represent the similarity or dissimilarity of data points, specifically highly nonlinearly distributed and large-scaled ones. Thus, [152], [153], and [154] consider weighted combinations of all the kernels by means of probability measures on  $T$ . To each  $\beta \in \mathcal{C}'_+(T)$  is associated the weighted combination of elements of  $\{\kappa(t), t \in T\}$ , i.e.,  $\kappa_\beta(x, z) := \int_T \kappa(x, z, t) d\beta(t)$ . For instance, if the family is formed by Gaussian kernels, the form of the weighted combination would be  $\kappa_\beta(x, z) := \int_T \exp(-t \|x - z\|^2) d\beta(t)$ . This way, after the change of variable  $\eta = -\theta$ , the LSIO problem becomes

$$\left. \begin{array}{l} \inf_{\eta \in \mathbb{R}, \beta \in \mathcal{C}'_+(T)} \eta \\ \text{s.t.} \quad \eta + \int_T S(t, \alpha) d\beta(t) \geq 0, \alpha \in A, \\ \int_T d\beta(t) = 1, \end{array} \right\} \quad (46)$$

where

$$S(t, \alpha) := \frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j y_i y_j \kappa_\beta(x^i, x^j) - \sum_{i=1}^N \alpha_i.$$

The connections between machine learning, LSIO and other optimization fields have been discussed in [176].

#### 4.4 Other applications

##### – Finance, economy and games

The uncertain portfolio problem is another source of applications of LSIO. First, the general uncertain portfolio problem has been formulated in [186] as an LSIO problem, where the reader can find examples solved via feasible direction methods. Second, the problem consisting of minimizing the portfolio cost subject to constraints guaranteeing that the portfolio payout is always greater or equal to the payoff of the barrier option has been modeled in [131] and [146] as an LSIO problem whose corresponding dual is solved via discretization and its optimal solution is interpreted in economic terms; this model is studied in detail in the monograph [146], which provides numerical examples showing that this approach is superior to those developed previously in the literature. Third, basket options (a popular way to hedge portfolio risk as they offer the unique characteristic of a strike price based on the weighted value of the basket of assets) has been modelled in [155] as a stochastic optimization problem involving  $m$  assets, each asset with its corresponding strike  $K_i$ , spot price  $C_{0,i}$ , and weight,  $w_i$ ,  $i = 1, \dots, m$ ; the horizon is  $\tau$ , and there is a riskless zero bound with initial price 1 and future payoff  $e^{r\tau}$ . The price of asset  $i$  at the horizon is a random variable  $S_i$ . The payoff of the call option  $i$  at maturity  $\tau$  is  $(S_i - K_i)_+ := \max\{S_i - K_i, 0\}$ . An important problem in stochastic finance consists of computing the interval of variation of the expected value of the price of the basket option at maturity, i.e., the random variable  $e^{r\tau} (w^T S - K_B)_+$ , where  $K_B$  denotes the strike. Under mild assumptions, [155] reformulates the dual problems of maximizing and minimizing the mentioned expected value as LSIO problems. In particular, the LSIO problem providing the upper bound of the expected value of  $e^{r\tau} (w^T S - K_B)_+$  turns out to be solvable by means of a closed formula. This problem has been studied in detail in [53] from the numerical LSIO perspective.

Complementary references: [8], [166], [167], [171], [172], [175], [182].

##### – Probability and Statistics

The most important application of LSIO in statistics is Bayesian robustness. Two central problems in this field consist of analyzing the robustness of Bayesian procedures with respect to the prior and calculating minimax decision rules under generalized moment conditions.

The univariate moment problem has been reformulated as a dual LSIO problem by assuming that the prior belongs to a class defined in terms of the so-called generalized moment conditions. These problems have been solved by the cutting-plane discretization method proposed by Betrò in [16] ([17], [18], [157], [158]).

Concerning the second problem, the corresponding decision rules are obtained by minimizing the maximum of the integrals of the risk function with respect to a given family of distributions on a certain space of parameters; [69] considers systems whose components have exponential life

distributions. The corresponding optimal component testing problem is formulated as an LSIO problem which is solved by means of an algorithmic procedure that computes optimal test times based on the column generation technique (dual simplex method).

The deterministic semi-Markov decision processes in Borel spaces are usually formulated as infinite LP problems. These problems have been approximated by means of primal/dual LSIO problems in [122], where the authors show that strong duality holds.

Complementary references: [3], [50], [188], [191].

– **Stochastic programming**

In [112] the authors deal with the optimal scenario generation problem and show that, for linear two-stage stochastic models, this problem can be formulated as a generalized semi-infinite program, which is convex in some cases and enjoys stability properties. Additionally, the problem of optimal scenario reduction for two-stage models admits a new formulation (based on the minimal information distance) as a mixed-integer linear semi-infinite program. The latter decomposes into solving binary and linear semi-infinite programs recursively. Also in this paper, a mixed-integer linear semi-infinite program is proposed for optimal scenario generation in chance constrained programming. The approach is illustrated with an application to scenario generation for the classical newsvendor problem.

Distributionally robust two-stage stochastic linear optimization problems with higher-order are reformulated as LSIO problems in [71]. A numerical experiment is reported, where the LSIO problem is solved with the reduction method of [189].

– **Best approximate solutions of inconsistent systems**

Best approximate solutions of inconsistent linear systems of the form  $\{Ax = b\}$ , with  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ ,  $m \geq n$ , are usually computed, following Legendre, by minimizing the Euclidean norm of the *residual vector*  $r(x) := Ax - b \in \mathbb{R}^m$ , whose optimal set has the closed form  $\{x \in \mathbb{R}^n : (A^\top A)x = A^\top b\}$  (a unique optimal solution  $(A^\top A)^{-1}A^\top b$  whenever the columns of  $A$  are linearly independent). When the inconsistent system is formed by inequalities,  $\{a_j^\top x \leq b_j, j = 1, \dots, m\}$ , the minimization of the Euclidean norm of the residual vector

$$\begin{aligned} r_+(x) &:= (\max\{a_1^\top x - b_1, 0\}, \dots, \max\{a_m^\top x - b_m, 0\}) \\ &:= ((a_1^\top x - b_1)_+, \dots, (a_m^\top x - b_m)_+) \end{aligned}$$

on  $\mathbb{R}^n$  requires an iterative method (as the least-squares method proposed by Han in 1980). When the finite index set  $\{1, \dots, m\}$  is replaced by an infinite set  $T$ , the residual vector becomes a real-valued function  $T \ni t \mapsto (a_t^\top x - b_t)_+$ , whose size can certainly be measured with the uniform norm, but not always with the Euclidean one (except under continuity assumptions). Computing the best uniform solution can so

be formulated as the LSIO problem

$$\begin{aligned} P : \inf_{(x, x_{n+1}) \in \mathbb{R}^{n+1}} & x_{n+1} \\ \text{s.t.} & a_t^\top x - x_{n+1} \leq b_t, \quad t \in T, \end{aligned}$$

which allows to give conditions guaranteeing the existence and uniqueness of best uniform solutions via LSIO theory and compute them with LSIO methods [80].

– **Big Data Analytics**

As claimed by different authors in [121], mathematical optimization is an important tool used in big data analytics, in particular in artificial intelligence and machine learning. We can mention, among the most used optimization tools, dynamic programming, heuristic and metaheuristic methods, multi-objective and multi-modal methods, such as Pareto optimization and evolutionary algorithms. It is evident that identifying and selecting the most important features from a ultrahigh dimensional big dataset play a major role in processing large volume of data to take instant decision in short period of time. Hence, ranking the features based on their relevance and selecting the most relevant features can vastly improve the generalization performance. Feature selection is also considered very important for big data analytics due to its characteristics of semi-infinite optimization problem. To address an associated convex SIO problem, Tan et al. [173] propose an efficient feature selection algorithm that works iteratively and selects a subset of features, and solves a sequence of multiple kernel learning subproblems. The authors assert that the proposed method converges globally under mild condition and yields low biasness on feature selection.

– **Engineering and Chemistry**

In [56] the model reduction of high order linear-in-parameters discrete-time systems is considered. The coefficients of the original system model are assumed to be known only within given intervals, and the coefficients of the derived reduced order model are also obtained in intervals, such that the complex value sets of the uncertain original and reduced models will be optimally close to each other on the unit circle. The authors apply a novel approach based on the minimization of the infinity norm of “distance” between two polygons representing the original and the reduced uncertain systems. Thanks to a special definition of this distance, the problem admits an LSIO reformulation which reduces significantly the computation time.

In [180] the authors addressed the problem of optimizing the power flow in power systems with transient stability constraints. They translate this problem into a SIO problem and, based on the KKT-system of this reformulated SIO, a smoothing quasi-Newton algorithm is presented in which numerical integration is used. The convergence of the algorithm is also established.

– **Other references**

Selected applications of LSIO to different fields can be found in [3], [4], [19], [20], [21], [49], [56], [69], [115], [117], [119], [120], [134], [140], [159], [164], [168], [180], [179], [191], [192] and [196].

**Acknowledgment** Thanks are due to the Editors-in-Chief of 4OR for their kind invitation to submit a review on linear semi-infinite optimization and, in particular, to Michel Grabisch for his valuable comments and suggestions on the preliminary versions.

## References

1. Agmon, S (1954) The relaxation method for linear inequalities. *Canadian J Math* 6:382-392
2. Aliprantis C, Border K (2005) *Infinite dimensional analysis: a Hitchhiker's Guide* (3rd Edition). Springer, Berlin
3. Altinel IK, Çekyay BÇ, Feyzioğlu O, Keskin M.E., Özekici S (2011) Mission-based component testing for series systems. *Ann Oper Res* 186:1-22
4. Altinel IK, Çekyay BÇ, Feyzioğlu O, Keskin M.E., Özekici S (2013) The design of mission-based component test plans for series connection of sub-systems. *IIE Transactions* 45:1202-1220
5. Anderson EJ., Goberna MA, López MA (2001) Simplex-like trajectories on quasi-polyhedral convex sets. *Math Oper Res* 26:147-162
6. Anderson EJ, Lewis AS (1989) An extension of the simplex algorithm for semi-infinite linear programming. *Math Programming* 44A:247-269
7. Anderson EJ, Nash P (1987) *Linear programming in infinite-dimensional spaces: theory and applications*. Wiley, Chichester
8. Audy J-F, D'Amours S, Rönnqvist M (2012) An empirical study on coalition formation and cost/savings allocation. *Int J Prod Econ* 136:13-27
9. Auslender A, Ferrer A, Goberna MA, López MA (2015) Comparative study of RPSALG algorithms for convex semi-infinite programming. *Comput Optim Appl* 60:59-87
10. Auslender A, Goberna MA, López MA (2009) Penalty and smoothing methods for convex semi-infinite programming. *Math Oper Res* 34:303-319
11. Azé D, Corvellec J-N (2004) Characterizations of error bounds for lower semicontinuous functions on metric spaces. *ESAIM Control Optim Calc Var* 10:409-425
12. Basu A, Martin K, Ryan CT (2014) On the sufficiency of finite support duals in semi-infinite linear programming. *Oper Res Lett* 42:16-20
13. Basu A, Martin K, Ryan CT (2015) A Unified approach to semi-infinite linear programs and duality in convex programming. *Math Oper Res* 40:146-170
14. Basu A, Martin K, Ryan CT (2017) Strong duality and sensitivity analysis in semi-infinite linear programming. *Math Programming* 161A:451-485
15. Ben-Tal A, Ghaoui LE, Nemirovski A (2009) *Robust optimization*. Princeton, New Jersey
16. Betró B (2004) An accelerated central cutting plane algorithm for linear semi-infinite programming. *Math Programming* 101A:479-495
17. Betró, B. (2007) Bayesian robustness: theory and computation. In Ruggeri F, Kenett RS and Faltin F (eds.) *Encyclopedia of statistics in quality and reliability*, pp 203-207, Wiley, Chichester

18. Betrò, B (2009) Numerical treatment of Bayesian robustness problems. *Int J Appr Reasoning* 50:279-288
19. Bisbos CD, Ampatzis AT (2008) Shakedown analysis of spatial frames with parameterized load domain. *Eng Structures* 303:119-3128
20. Blado D, Hu W, Toriello A (2016) Semi-infinite relaxations for the dynamic knapsack problem with stochastic item sizes. *SIAM J Optim* 26:1625-1648
21. Bodirsky M, Jonsson P, Oertzen T (2012) Essential convexity and complexity of semi-algebraic constraints. *Log. Methods Comput. Sci.* 8:4-25
22. Boţ RI, Csetnek ER, Wanka G. (2008) Sequential optimality conditions in convex programming via perturbation approach. *J. Convex Anal* 15:149-164
23. Brosowski B (1982) Parametric semi-infinite optimization. Peter Lang, Frankfurt am Main
24. Brosowski B (1984) Parametric semi-infinite linear programming I. Continuity of the feasible set and the optimal value. *Math Programming Study* 21:18-42
25. Cánovas MJ, Dontchev AL, López MA, Parra J (2009) Isolated calmness of solution mappings in convex semi-infinite optimization. *J Math Anal Appl* 350:892-837
26. Cánovas MJ, Hall JAJ, López MA, Parra J. Calmness of partially perturbed linear systems with an application to interior point methods. Submitted
27. Cánovas MJ, Hantoute A, Parra J, Toledo FJ (2014) Calmness of the argmin mapping in linear semi-infinite optimization. *J Optim Theory Appl* 160:111-126
28. Cánovas MJ, Hantoute A, Parra J, Toledo FJ (2015) Boundary of subdifferentials and calmness moduli in linear semi-infinite optimization. *Optim Lett* 9:513-521
29. Cánovas MJ, Hantoute A, Parra J, Toledo FJ (2016) Calmness modulus of fully perturbed linear programs. *Math Programming* 158A: 267-290
30. Cánovas MJ, Henrion R, López MA, Parra J (2016) Outer limit of subdifferentials and calmness moduli in linear and nonlinear programming. *J Optim Theory Appl* 169: 925-952
31. Cánovas MJ, Henrion R, Parra J, Toledo FJ. Critical objective size and calmness modulus in linear programming, *Set-Valued Var. Anal.*, to appear
32. Cánovas MJ, Kruger AY, López MA, Parra J, Théra MA (2014) Calmness modulus of linear semi-infinite programs. *SIAM J Optim* 24:29-48
33. Cánovas MJ, López MA, Parra J, Toledo FJ (2005) Distance to ill-posedness and the consistency value of linear semi-infinite inequality systems. *Math Programming* 103A:95-126
34. Cánovas MJ, López MA, Parra J, Toledo FJ (2006) Distance to solvability / unsolvability in linear optimization. *SIAM J Optim* 16:629-649
35. Cánovas MJ, López MA, Parra J, Toledo FJ (2006) Ill-posedness with respect to the solvability in linear optimization. *Linear Algebra Appl* 416:520-540
36. Cánovas MJ, López MA, Parra J, Toledo FJ (2007) Error bounds for the inverse feasible set mapping in linear semi-infinite optimization via a sensitivity dual approach. *Optimization* 56:547-563
37. Cánovas MJ, López MA, Parra J, Toledo FJ (2014) Calmness of the feasible set mapping for linear inequality systems. *Set-Valued Var Anal* 22:375-389
38. Cánovas MJ, López MA, Parra J, Toledo FJ. Lipschitz modulus of fully perturbed linear programs. *Pacific J Optim*, to appear.
39. Chan TCY, Mar PhA (2017) Stability and Continuity in Robust Optimization. *SIAM J Optim* 27:817-841

40. Charnes A, Cooper WW, Kortanek KO (1962) Duality, Haar programs, and finite sequence spaces. *Proc Natl Acad Sci USA* 48:783-786
41. Charnes A, Cooper WW, Kortanek KO (1963) Duality in semi-infinite programs and some works of Haar and Carathéodory. *Manag Sci* 9:209-228
42. Charnes A, Cooper WW, Kortanek KO (1965) On representations of semi-infinite programs which have no duality gaps. *Manag Sci* 12:113-121
43. Charnes A, Cooper WW, Kortanek KO (1969) On the theory of semi-infinite programming and a generalization of the Kuhn-Tucker saddle point theorem for arbitrary convex functions. *Nav. Res. Log Quart* 16:41-51
44. Chu YC (1966) Generalization of some fundamental theorems on linear inequalities. *Acta Math Sinica* 16:25-40
45. Chuong TD, Jeyakumar V (2017) An exact formula for radius of robust feasibility of uncertain linear programs. *J Optim Theory Appl* 173: 203-226
46. Chuong TD, Jeyakumar V (2017) A generalized Farkas lemma with a numerical certificate and linear semi-infinite programs with SDP duals. *Linear Algebra Appl* 515:38-52
47. Clarke FH (1976) A new approach to Lagrange multipliers. *Math Oper Res* 1:165-174
48. Correa R, Hantoute A, López MA (2016) Weaker conditions for subdifferential calculus of convex functions. *J Funct Anal* 271:1177-1212
49. Cozad A, Sahinidis NV, Miller DC (2015) A combined first-principles and data-driven approach to model building. *Comput Chem Eng* 73:116-127.
50. Cozman FG, Polpo de Campos C (2014) Kuznetsov independence for interval-valued expectations and sets of probability distributions: properties and algorithms. *Int J Appr Reasoning* 55:666-682
51. Daniilidis A, Goberna MA, López MA, Lucchetti R (2015) Stability in linear optimization under perturbations of the left-hand side coefficients. *Set-Valued Var Anal* 23:737-758
52. Dantzig GB (1991) Linear programming. In: Lenstra JK et al (eds) *History of mathematical programming: a collection of personal reminiscences*, pp 19-31, North-Holland, Amsterdam
53. Daum S, Werner R (2011) A novel feasible discretization method for linear semi-infinite programming applied to basket options pricing. *Optimization* 60:1379-1398
54. Dinh N, Goberna MA, López MA, Song TQ (2007) New Farkas-type constraint qualifications in convex infinite programming. *ESAIM: Control Optim Calc Var* 13:580-597
55. Dinh N, Goberna MA, López MA, Volle M (2010) Convex inequalities without constraint qualification nor closedness condition, and their applications in optimization. *Set-valued Var Anal* 18:423-445
56. Dolgin Y, Zeheb E (2005) Model reduction of uncertain systems retaining the uncertainty structure. *Systems & Control Letters* 54:771-779
57. Dontchev AL, Rockafellar RT (2009) *Implicit Functions and Solution Mappings: A View from Variational Analysis*. Springer, New York
58. Duffin R, Karlovitz LA (1965) An infinite linear program with a duality gap. *Management Sci* 12:122-134
59. Dür M., Jargalsaikhan B, Still G (2016) Genericity results in linear conic programming -a tour d'horizon. *Math Oper Res* 42:77-94
60. Eberhard A, Roshchina V, Sang T (2017) Outer limits of subdifferentials for min-max type functions. Manuscript (arXiv:1701.02852v1[math.OC])

61. Fabian M, Henrion R, Kruger AY, Outrata J (2010) Error bounds: Necessary and sufficient conditions. *Set-Valued Anal* 18:121-149
62. Fajardo MD, López MA (1999) Locally Farkas-Minkowski systems in convex semi-infinite programming. *J Optim Theory Appl* 103:313-335
63. Fajardo MD, López MA (2006) Some results about the facial geometry of convex semi-infinite systems. *Optimization* 55:661-684
64. Fajardo MD, López MA, Puente R (2008) Linear representations and quasisipolyhedrality of a finite-valued convex function. *Optimization* 57:215-237
65. Fang DH, Li C, Ng KF (2009) Constraint qualifications for extended Farkas's lemmas and Lagrangian dualities in convex infinite programming. *SIAM J Optim* 20:1311-1332
66. Fang DH, Li C, Ng KF (2010) Constraint qualifications for optimality conditions and total Lagrange dualities in convex infinite programming. *Nonlinear Anal* 73:1143-1159
67. Faybusovich L, Mouktonglang T, Tsuchiya, T (2008) Numerical experiments with universal barrier functions for cones of Chebyshev systems. *Comput Optim Appl* 41:205-223
68. Ferrer A, Goberna MA, González-Gutiérrez E, Todorov MI, A comparative study of relaxation algorithms for the linear semi-infinite feasibility problem, *Ann Oper Res*, to appear. DOI: 10.1007/s10479-016-2135-2
69. Feyzioglu O, Altinel IK, Ozekici S (2008) Optimum component test plans for phased-mission systems. *European J Oper Res* 185:255-265
70. Fischer, T (1983) Contributions to semi-infinite linear optimization. *Meth Verf Math Phys* 27:175-199
71. Gao SY, Sun J, Wu S-Y. A semi-infinite programming approach to two-stage stochastic linear programs with high-order moment constraints. *Optim Lett*, to appear. DOI: 10.1007/s11590-016-1095-4
72. Ghaté A, Sharma D, Smith RL (2010) A shadow simplex method for infinite linear programs. *Oper Res* 58:865-877
73. Glashoff K, Gustafson SA (1983) *Linear optimization and approximation*. Springer, Berlin
74. Goberna MA (2005) Linear semi-infinite optimization: recent advances. In: Rubinov A., Jeyakumar V (eds) *Continuous Optimization: Current Trends and Modern Applications*, pp 3-22, Springer, New York
75. Goberna MA (2005) *Linear semi-infinite programming: A Guided Tour*. IMCA Monographs, Lima
76. Goberna MA, Gómez S, Guerra-Vázquez F, Todorov MI (2007) Sensitivity analysis in linear semi-infinite programming: perturbing cost and right-hand-side coefficients. *European J Oper Res* 181:1069-1085
77. Goberna MA, González E, Martínez-Legaz JE, Todorov MI (2010) Motzkin decomposition of closed convex sets. *J Math Anal Appl* 364:209-221
78. Goberna MA, Guerra-Vázquez F, Todorov MI (2013) Constraint qualifications in linear vector semi-infinite optimization. *European J Oper Res* 227:12-21
79. Goberna MA, Guerra-Vázquez F, Todorov MI (2016) Constraint qualifications in convex vector semi-infinite optimization. *European J Oper Res* 249:32-40
80. Goberna MA, Hiriart-Urruty J.-B., López MA, Best approximate solutions of inconsistent linear inequality systems, submitted

81. Goberna MA, Iusem A, Martínez-Legaz JE, Todorov MI (2013) Motzkin decomposition of closed convex sets via truncation. *J Math Anal Appl* 400:35-47
82. Goberna MA, Jeyakumar V, Li G, Linh N (2016) Radius of robust feasibility formulas for classes of convex programs with uncertain polynomial constraints. *OR Letters* 44:67-73
83. Goberna MA, Jeyakumar V, Li G, López MA (2013) Robust linear semi-infinite programming duality under uncertainty. *Math Programming* 139B:185-203
84. Goberna MA, Jeyakumar V, Li G, Vicente-Pérez J (2014) Robust solutions of multi-objective linear semi-infinite programs under constraint data uncertainty. *SIAM J Optim* 24:1402-1419
85. Goberna MA, Jeyakumar V, Li G, Vicente-Pérez J (2015) Robust solutions to multi-objective linear programs with uncertain data. *European J Oper Res* 242:730-743
86. Goberna MA, Jeyakumar V, Li G, Vicente-Pérez J. Radii of robust efficiency in robust multi-objective convex optimization. Submitted.
87. Goberna MA, Lancho A, Todorov MI, Vera de Serio VN (2011) On implicit active constraints in linear semi-infinite programs with unbounded coefficients. *Appl Math Optim* 63:239-256
88. Goberna MA, López MA (1988) Optimal value function in semi-infinite programming. *J Optim Theory Appl* 59:261-279
89. Goberna MA, López MA (1998) Topological stability of linear semi-infinite inequality systems. *J Optim Theory Appl* 89:227-236
90. Goberna MA, López MA (1998) Linear semi-infinite optimization. Wiley, Chichester
91. Goberna MA, López MA (2014) Post-optimal analysis in linear semi-infinite optimization. Springer Briefs, Springer New York
92. Goberna MA, López MA, Todorov MI (1996) Stability theory for linear inequality systems. *SIAM J Matrix Anal Appl* 17:730-743
93. Goberna MA, López MA, Todorov MI (1997) Stability theory for linear inequality systems II: upper semicontinuity of the solution set mapping. *SIAM J Optim* 7:1138-1151
94. Goberna MA, López MA, Todorov MI (2001) On the stability of the feasible set in linear optimization. *Set-Valued Anal* 9:75-99
95. Goberna MA, López MA, Todorov MI (2003) Extended active constraints in linear optimization with applications. *SIAM J Optim* 14:608-619
96. Goberna MA, López MA, Volle, M (2014) Primal attainment in convex infinite optimization duality. *J Convex Anal* 21:1043-1064
97. Goberna MA, López MA, Volle, M (2014) Modified Lagrangian duality for the supremum of convex functions. *Pacific J Optim*, to appear
98. Goberna MA, Kanzi N. Optimality conditions in convex multi-objective SIP. *Math Programming (Series A)*, to appear. DOI: 10.1007/s10107-016-1081-8
99. Goberna MA, Ridolfi A, Vera de Serio VN, Stability of the duality gap in linear optimization. *Set-Valued Var Anal*, to appear. DOI: 10.1007/s11228-017-0405-z
100. Goberna MA, Terlaky T, Todorov MI (2010) Sensitivity analysis in linear semi-infinite programming via partitions. *Math Oper Res* 35:14-25
101. Goberna MA, Todorov MI (2009) Primal-dual stability in continuous linear optimization. *Math Programming* 116B:129-146

102. González-Gutiérrez E, Rebollar LA, Todorov MI (2012) Relaxation methods for solving linear inequality systems: converging results. *Top* 20: 426-436.
103. González-Gutiérrez E, Todorov MI (2012) A relaxation method for solving systems with infinitely many linear inequalities. *Optim Letters* 6:291-298
104. Guo F (2015) Semidefinite programming relaxations for linear semi-infinite polynomial programming. *OALib Journal*, online (<http://www.oalib.com/paper/3855833#.WSRhOdztbX4>)
105. Guo P, Huang GH, He L (2008) ISMISIP: an inexact stochastic mixed integer linear semi-infinite programming approach for solid waste management and planning under uncertainty. *Stoch Env Res Risk Assess* 22:759-775
106. Gustafson SA (1970) On the computational solution of a class of generalized moment problems. *SIAM J Numer Anal* 7:343-357
107. Gustafson SA, Kortanek KO (1973) Numerical treatment of a class of semi-infinite programming problems. *Nav Res Logist Quart* 20:477-504
108. Haar A (1924) Über lineare ungleichungen (in German). *Acta Math Szeged* 2:1-14
109. Hayashi S, Okuno T, Ito Y (2016) Simplex-type algorithm for second-order cone programmes via semi-infinite programming reformulation. *Optim Met Soft* 31:1272-1297
110. He L., Huang GH (2008) Optimization of regional waste management systems based on inexact semi-infinite programming. *Can J Civil Eng* 35:987-998
111. He L., Huang GH, Lu H (2011) Bivariate interval semi-infinite programming with an application to environmental decision-making analysis. *European J Oper Res* 211:452-465
112. Henrion R, Roemisch W (2017) Optimal scenario generation and reduction in stochastic programming. Preprint (<https://www.researchgate.net/publication/316038597>)
113. Hu H (1994) A Projection method for solving infinite systems of linear inequalities. In: Du D-Z, Sun J (eds) *Advances in optimization and approximation*, pp 186-194, Kluwer, Dordrecht
114. Huang GH, He L, Zeng GM, Lu HW (2008) Identification of optimal urban solid waste flow schemes under impacts of energy prices. *Env Eng Sci* 25: 685-695
115. Huynh DBP, Rozza G, Sen S, Patera AT (2007) A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants. *C R Math Acad Sci Paris* 345:473-478.
116. Ioffe AD (1979) Necessary and sufficient conditions for a local minimum. I. A Reduction theorem and first order conditions. *SIAM J Control Optim* 17: 245-250
117. Jeyakumar V, Li GY, Lee GM (2011) A robust von Neumann minimax theorem for zero-sum games under bounded payoff uncertainty. *Oper. Res. Lett.* 39:109-114
118. Jeroslow RG (1979) Some relaxation methods for linear inequalities. *Cahiers du Cero* 21:43-53
119. Jinglai S (2012) Positive Invariance of Constrained Affine Dynamics and Its Applications to Hybrid Systems and Safety Verification. *IEEE Trans Automatic Control* 57:3-18
120. Karimi, A., Galdos, G.: Fixed-order  $H_\infty$  controller design for nonparametric models by convex optimization. *Automatica* 46, 1388-1394 (2010)

121. Kashyap H, Ahmed HA, Hoque N, Roy S, Bhattacharyya DK (2016) Big data analytics in bioinformatics: architectures, techniques, tools and issues. *Netw Model Anal Health Inform Bioinf* 5:28
122. Klabjan D, Adelman D (2007) An infinite-dimensional linear programming algorithm for deterministic semi-Markov decision processes on Borel spaces. *Math Oper Res* 32:528-550
123. Klatte D, Kummer B (2002) Nonsmooth equations in optimization: regularity, calculus, methods and applications. Kluwer, Dordrecht
124. Klatte D, Kummer B (2009) Optimization methods and stability of inclusions in Banach spaces. *Math. Programming* 117B:305-330
125. Kortanek KO (1974) Classifying convex extremum problems over linear topologies having separation properties. *J Math Anal Appl* 46:725-755
126. Kortanek KO (2001) On the 1962-1972 decade of semi-infinite programming: a subjective view. In Goberna MA, López MA (eds) *Semi-infinite programming: recent advances*, pp 3-34, Kluwer, Dordrecht
127. Kortanek KO, Zhang Q (2016) Extending the mixed algebraic-analysis Fourier-Motzkin elimination method for classifying linear semi-infinite programmes. *Optimization* 65:707-727
128. Kruger AY, Van Ngai H, Théra M (2010) Stability of error bounds for convex constraint systems in Banach spaces. *SIAM J Optim* 20:3280-3296
129. Larriqueta M, Vera de Serio, VN (2014) On metric regularity and the boundary of the feasible set in linear optimization. *Set-Valued Var Anal* 22:1-17.
130. Lasserre JB (2012) An algorithm for semi-infinite polynomial optimization. *Top* 20:119-129
131. Leibfritz F, Maruhn JH (2009) A successive SDP-NSDP approach to a robust optimization problem in finance. *Comput Optim Appl* 44:443-466
132. Li MH, Meng KW, Yang XQ (2016) On error bound moduli for locally Lipschitz and regular functions. Manuscript (arXiv: 1608.03360v1[math.OC])
133. Li C, Ng KF, Pong TK (2008) Constraint qualifications for convex inequality systems with applications in constrained optimization. *SIAM J Optim* 19:163-187
134. Li X, Lu H, He L, Shi B (2014) An inexact stochastic optimization model for agricultural irrigation management with a case study in China. *Stoch Env Res Risk* 28A:281-295
135. Liu Y (2016) New constraint qualification and optimality for linear semi-infinite programming. *Pac J Optim* 12: 223-232
136. Liu Y, Ding MF (2014) A ladder method for semi-infinite programming. *J Ind Manag Optim* 10:397-412
137. Liu Y, Goberna MA (2016) Asymptotic optimality conditions for linear semi-infinite programming. *Optimization* 65:387-414
138. López MA (2012) Stability in linear optimization and related topics. A personal tour. *Top* 20:217-244
139. López MA, Still G (2007) Semi-infinite programming. *European J Oper Res* 180:491-518
140. Lou Y, Yin Y, Lawphongpanich S (2010) Robust congestion pricing under boundedly rational user equilibrium. *Transportation Research Part B: Methodological* 44:15-28
141. Luo Z-Q, Roos C, Terlaky T (1999) Complexity analysis of a logarithmic barrier decomposition method for semi-infinite linear programming. *Appl Numer Math* 29:379-394

142. Mangasarian OL (2004) Knowledge-based linear programming. *SIAM J Optim* 12:375-382
143. Mangasarian OL, Wild EW (2007) Nonlinear knowledge in kernel approximation. *IEEE Trans Neural Netw* 18:300-306
144. Mangasarian OL, Wild EW (2008) Nonlinear knowledge-based classification. *IEEE Trans. Neural Netw* 19:1826-1832
145. Martínez-Legaz JE, Todorov MI, Zetina C (2014)  $\gamma$ -Active constraints in convex semi-infinite programming. *Numer Funct Anal Appl* 35:1078-1094
146. Maruhn JH (2009) Robust static super-replication of barrier options. De Gruyter, Berlin
147. Miao DY, Li YP, Huang GH, Yang ZF (2014) Optimization model for planning regional water resource systems under uncertainty. *J Water Resources Plan Manag* 140:238-249
148. Mordukhovich BS (2006) *Variational Analysis and Generalized Differentiation, I: Basic Theory*. Springer, Berlin
149. Motzkin TS, Schoenberg IJ (1954) The relaxation method for linear inequalities. *Canadian J. Math.* 6:393-404
150. Ochoa PD, Vera de Serio VN (2012) Stability of the primal-dual partition in linear semi-infinite programming. *Optimization* 61:1449-1465
151. Oskoorouchi MR, Ghaffari HR, Terlaky T, Aleman DM (2011) An interior point constraint generation algorithm for semi-infinite optimization with health-care application. *Oper Res* 59:1184-1197
152. Ozogur S, Ustunkar G, Weber G-W (2016) Adapted infinite kernel learning by multi-local algorithm, *Int. J. Patt. Recogn. Artif. Intell.* 30:1651004
153. Ozogur S, Weber G-W (2010) On numerical optimization theory of infinite kernel learning. *J Global Optim* 48:215-239
154. Ozogur S, Weber G-W (2010) Infinite kernel learning via infinite and semi-infinite programming. *Optim Meth Soft* 25:937-970
155. Peña J, Vera JC, Zuluaga LF (2010) Static-arbitrage lower bounds on the prices of basket options via linear programming. *Quant Finance* 10:819-827
156. Powell MJD (1990) Karmarkar's algorithm: a view from nonlinear programming. *Bull Inst Math Appl* 26: 165-181
157. Prékopa A (2009) Inequalities for discrete higher order convex functions. *J Math Inequalities* 3:485-498
158. Prékopa A, Ninh A, Alexe G (2016) On the relationship between the discrete and continuous bounding moment problems and their numerical solutions. *Ann Oper Res* 238:521-575
159. Priyadarsini PI, Devarakonda N, Babu IR (2013) A chock-full survey on support vector machines. *Int J Adv Res Comput Sci Soft Eng* 3:780-799
160. Puente R, Vera de Serio VN (1999) Locally Farkas-Minkowski linear semi-infinite systems. *Top* 7: 103-121.
161. Remez E. (1934) Sur la détermination des polynômes d'approximation de degré donné (in French). *Commun Soc Math Kharkoff and Inst Sci Math et Mecan* 10:41-63
162. Robinson SM (1981) Some continuity properties of polyhedral multifunctions. *Mathematical programming at Oberwolfach (Proc. Conf., Math. Forschungsinstitut, Oberwolfach, 1979)*. *Math. Programming Stud* 14:206-214
163. Rockafellar RT, Wets RJB (1998) *Variational analysis*. Springer, Berlin

164. Rozza G, Huynh DBP, Patera AT (2008) Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: application to transport and continuum mechanics. *Arch Comput Meth Eng* 15:229-275
165. Rubinstein GS (1981) A comment on Voigt's paper "a duality theorem for linear semi-infinite programming" (in Russian). *Optimization* 12:31-32
166. Shani B, Solan E (2014) Strong approachability. *J Dyn Games* 71:507-535
167. Singh C, Sarkar S, Aram A, Kumar A (2012) Cooperative profit sharing in coalition-based resource allocation in wireless networks. *IEEE/ACM Trans Netw* 20B:69-83.
168. Sommer B, Dingersen T, Gamroth C, Schneider SE, Rubert S, Krüger J, Dietz KJ (2009) CELLmicrocosmos 2.2 MembraneEditor: A modular interactive shape-based software approach to solve heterogenous membrane packing problems, *J Chem Inf Model* 5:1165-1182
169. Sonnenburg S, Rätsch G, Schäfer C, Schölkopf, B (2006) Large scale multiple kernel learning. *J Machine Learning Res* 7:1531-1565
170. Stein O (2012) How to solve a semi-infinite optimization problem. *European J Oper Res* 223:312-320
171. Suakkaphong N, Dror M (2011) Managing decentralized inventory and transshipment. *Top* 19:480-506
172. Summerfield NS, Dror M (2012) Stochastic programming for decentralized newsvendor with transshipment. *Int J Prod Econ* 137:292-303
173. Tan M, Tsang IW, Wang L (2014) Towards ultrahigh dimensional feature selection for big data. *J Machine Learning Res* 15:1371-1429
174. Thibault L (1997) Sequential convex subdifferential calculus and sequential Lagrange multipliers. *SIAM J Control Optim* 35:1434-1444
175. Tian Y (2013) Strategy-proof and efficient offline interval scheduling and cake. In: Chen Y, Immorlica N (eds) *Web and internet economics*, pp 436-437, Springer, New York
176. Tian Y, Shi Y, Liu X (2012) Recent advances on support vector machines research. *Tech Econ Develop Econ* 18:5-33
177. Todd MJ (1994) Interior-point algorithms for semi-infinite programming. *Math Programing* 65A:217-245
178. Todorov MI (1985/86) Generic existence and uniqueness of the solution set to linear semi-infinite optimization problems. *Numer Funct Anal Optim* 8:27-39
179. Tong X, Wu S-Yi, Zhou R (2010) New approach for the nonlinear programming with transient stability constraints arising from power systems. *Comput. Optim. Appl.* 45:495-520
180. Tong X, Ling Ch, Qi L (2008) A semi-infinite programming algorithm for solving optimal power flow with transient stability constraints. *J Comput Appl Math* 217: 432-447
181. Tunçel L, Todd MJ (1996) Asymptotic behavior of interior-point methods: a view from semi-infinite programming. *Math Oper Res* 21:354-381
182. Uhan NA (2015) Stochastic linear programming games with concave preferences. *European J Oper Res* 243:637-646
183. Vanderbei RJ (1995) Affine-scaling trajectories associated with a semi-infinite linear program. *Math Oper Res* 20:163-174
184. Vaz A, Fernandes E, Gomes M (2003) A sequential quadratic programming with a dual parametrization approach to nonlinear semiinfinite programming. *Top* 11:109-130

185. Vaz A, Fernandes E, Gomes M (2004) SIPAMPL: Semi-infinite programming with AMPL. *ACM Trans. Math Soft* 30:47-61
186. Vercher E (2008) Portfolios with fuzzy returns: selection strategies based on semi-infinite programming. *J Comput Appl Math* 217:381-393
187. Vinh NT, Kim DS, Tam NN, Yen ND (2016) Duality gap function in infinite dimensional linear programming. *J Math Anal Appl* 437:1-15
188. Wang Y, Ni H (2012) Multivariate convex support vector regression with semi-definite programming. *Knowledge-Based Systems* 30:87-94
189. Wu S-Y, Li DH, Qi LQ, Zhou GL (2005) An iterative method for solving KKT system of the semi-infinite programming. *Opt Meth Soft* 20:629-643
190. Xu Y, Sun W, Qi LQ (2015) On solving a class of linear semi-infinite programming by SDP method. *Optimization* 64:603-616
191. Yamangil E, Altinel IK, Çekyay B, Feyzioğlu O, Özekici S (2011) Design of optimum component test plans in the demonstration of diverse system performance measures. *IIE Trans* 43:535-546
192. Yiu KFC, Gao MJ, Shiu TJ, Wu SY, Tran T, Claesson I (2013) A fast algorithm for the optimal design of high accuracy windows in signal processing. *Optim. Meth Softw.* 28:900-916
193. Zălinescu C (2002) *Convex analysis in general vector spaces*. World Scientific, Singapore
194. Zhang L, Wu S-Y, López MA (2010) A new exchange method for convex semi-infinite programming. *SIAM J Optim* 20:2959-2977
195. Zheng XY, Ng KF (2003) Metric regularity and constraint qualifications for convex inequalities on Banach spaces. *SIAM J. Optim.* 14:757-772
196. Zhu Y, Huang GH, Li YP, He L, Zhang XX (2011) An interval full-infinite mixed-integer programming method for planning municipal energy systems: a case study of Beijing. *Appl Energy* 88:2846-2862