

## SOLVING SEMI-INFINITE OPTIMIZATION PROBLEMS WITH INTERIOR POINT TECHNIQUES\*

OLIVER STEIN<sup>†</sup> AND GEORG STILL<sup>‡</sup>

**Abstract.** We introduce a new numerical solution method for semi-infinite optimization problems with convex lower level problems. The method is based on a reformulation of the semi-infinite problem as a Stackelberg game and the use of regularized nonlinear complementarity problem functions. This approach leads to central path conditions for the lower level problems, where for a given path parameter a smooth nonlinear finite optimization problem has to be solved. The solution of the semi-infinite optimization problem then amounts to driving the path parameter to zero.

We show convergence properties of the method and give a number of numerical examples from design centering and from robust optimization, where actually so-called generalized semi-infinite optimization problems are solved. The presented method is easy to implement, and in our examples it works well for dimensions of the semi-infinite index set at least up to 150.

**Key words.** generalized semi-infinite optimization, convexity, Stackelberg game, nonlinear complementarity problem function, smoothing, optimality condition

**AMS subject classifications.** 90C34, 90C25, 49M37, 65K05

**DOI.** 10.1137/S0363012901398393

**1. Introduction.** In this article we introduce a bi-level solution method for so-called *generalized semi-infinite optimization problems*. These problems have the form

$$GSIP : \quad \text{minimize } f(x) \quad \text{subject to } x \in M$$

with

$$M = \{x \in \mathbb{R}^n \mid g_j(x, y) \leq 0, y \in Y(x), j \in J\}$$

and

$$Y(x) = \{y \in \mathbb{R}^m \mid v_\ell(x, y) \leq 0, \ell \in L\}.$$

All defining functions  $f, g_j, j \in J = \{1, \dots, p\}$ ,  $v_\ell, \ell \in L = \{1, \dots, s\}$ , are assumed to be real-valued and  $d$  times continuously differentiable on their respective domains with  $d \geq 2$ .

As opposed to a standard semi-infinite optimization problem *SIP*, the possibly infinite index set  $Y(x)$  of inequality constraints is  $x$ -dependent in a *GSIP*. For surveys about *standard* semi-infinite optimization we refer to [10, 18, 37].

Engineering applications that give rise to generalized semi-infinite optimization problems include robot design [12, 19], reverse Chebyshev approximation [29], time-optimal control [31], and design centering [35, 36]. In finite optimization with uncertainty about parameters  $y$  from a fixed set  $Y$ , the robust (i.e., worst-case) formulation of inequality constraints gives rise to a standard semi-infinite problem [1]. If the set

---

\*Received by the editors November 19, 2001; accepted for publication (in revised form) January 22, 2003; published electronically June 18, 2003.

<http://www.siam.org/journals/sicon/42-3/39839.html>

<sup>†</sup>Department of Mathematics – C, Aachen University, Aachen, Germany (stein@mathC.rwth-aachen.de).

<sup>‡</sup>Faculty of Mathematical Sciences, University of Twente, Enschede, The Netherlands (g.still@math.utwente.nl).

of uncertain parameters is state-dependent, then the worst-case formulation takes the form of *GSIP*. Furthermore, min-max problems can be reformulated as either standard or generalized semi-infinite programs, depending on whether the feasible set of the maximization (inner) problem depends on the minimization (outer) variable.

The growing interest in *GSIP* over recent years has resulted in various contributions on the structure of the feasible set  $M$  [25, 41, 48, 49, 50, 57] and on first and second order optimality conditions [19, 25, 29, 40, 51, 53, 57]. The articles [55] and [57] investigate how the known methods from *SIP* have to be modified in order to cover the more general situation of *GSIP*. The Newton-SQP approach, which works well in standard semi-infinite programming (see, e.g., [13]), can be transferred to *GSIP* if the so-called reduction ansatz holds. In [16] such a Newton-type method is applied to the terminal variational problems from [29]. Since the reduction ansatz is of local nature, also generalizations of the discretization and exchange methods from *SIP* are desired. In [55] it is shown that discretization methods converge if the  $x$ -dependent grid points are chosen such that they depend continuously on  $x$ . Moreover, [56] studies how for a discretization method the rate of convergence depends on a consistent treatment of the boundary points of  $Y(x)$ .

As these generalized discretization methods are not easy to implement, in the present article we concentrate on the case of convex lower level problems. Based upon the observation that, under natural assumptions, *GSIP* can be reformulated as a special Stackelberg game (cf. also [54]), we design a numerical solution method, which exploits the lower level convexity. As opposed to the exchange and discretization methods presented in [55], this approach is not a generalization of known methods from standard semi-infinite programming, but it provides a new and different way of numerical treatment. Moreover, as standard semi-infinite programming is a special case of *GSIP*, as a by-product we obtain a new solution method for standard semi-infinite optimization problems, too. Our point of view also was implicitly taken in [33], where a branch and bound method is developed for generalized semi-infinite optimization problems with linear-quadratic lower level problems and additional convexity in the upper level. An approach using exact penalization to transform *GSIP* to *SIP* is given in [32].

We remark that the inclusion of equality constraints in the definitions of  $M$  and  $Y(x)$ , as well as a  $j$ -dependence of the index set  $Y(x)$ , is straightforward and will not be considered here for ease of presentation.

The article is organized as follows. In section 2 we recall the so-called reduction ansatz and the resulting first order optimality condition for *GSIP*. Furthermore, we give analogous results under a convexity assumption and recall the concept of nonlinear complementarity problem (NCP) functions. Section 3 presents the numerical method, and in section 4 we study its convergence properties. A number of numerical results conclude the article in section 5.

## 2. Preliminaries.

**2.1. The reduction ansatz.** In this section we briefly recall the reduction ansatz and explain the concept of Fritz John points for *GSIP*. Since optimality conditions are well known for points from the topological interior of  $M$ , in the following we focus our attention on a given feasible boundary point of  $M$ , i.e., a point  $\bar{x} \in M \cap \partial M$ , where  $\partial M$  denotes the topological boundary of  $M$ .

Recall that the set-valued mapping  $Y$  is called locally bounded around  $\bar{x}$  if there exists a neighborhood  $U$  of  $\bar{x}$  such that the set  $\cup_{x \in U} Y(x)$  is bounded.

*Assumption 1* (local boundedness). The set-valued mapping  $Y$  is locally bounded around  $\bar{x}$ .

Let Assumption 1 hold throughout this article, and fix  $U$  to be some corresponding bounded open neighborhood of  $\bar{x}$ .

The  $n$ -parametric so-called lower level problems of *GSIP* are given by

$$Q^j(x) : \quad \text{maximize } g_j(x, y) \quad \text{subject to } y \in Y(x)$$

with  $j \in J$ . Associated with  $Q^j(x)$  are its optimal value function

$$\varphi_j(x) = \begin{cases} \max_{y \in Y(x)} g_j(x, y) & \text{if } Y(x) \neq \emptyset, \\ -\infty & \text{else} \end{cases}$$

and, in case of solvability, its solution set

$$Y_\star^j(x) = \{y \in Y(x) \mid g_j(x, y) = \varphi_j(x)\}.$$

It is easily seen that  $M$  and the set  $\{x \in \mathbb{R}^n \mid \varphi_j(x) \leq 0, j \in J\}$  coincide.

Since the defining functions of  $Y(x)$  are continuous, the set-valued mapping  $Y$  is closed. Together with Assumption 1 this means that  $Y$  is upper semicontinuous in the sense of Berge [2] around  $\bar{x}$ . As a consequence (cf., e.g., [21]), the  $\varphi_j, j \in J$ , are upper semicontinuous on  $U$ . From this it is not hard to derive that the sets  $M \cap U$  and  $\{x \in U \mid \varphi_j(x) \leq 0, j \in J_0(\bar{x})\}$  coincide (possibly after shrinking  $U$ ), where  $J_0(\bar{x}) = \{j \in J \mid \varphi_j(\bar{x}) = 0\}$  denotes the set of active indices at  $\bar{x}$ . Since for  $j \in J_0(\bar{x})$  the problem  $Q^j(\bar{x})$  has vanishing optimal value, the set of its solution points can be described as

$$Y_\star^j(\bar{x}) = Y_0^j(\bar{x}) = \{y \in Y(\bar{x}) \mid g_j(\bar{x}, y) = 0\},$$

and we have  $j \in J_0(\bar{x})$  if and only if  $Y_0^j(\bar{x}) \neq \emptyset$ .

Next, we give a local description of  $M$  by finitely many *smooth* constraints for the case when certain regularity assumptions hold in the lower level problems.

The *linear independence constraint qualification (LICQ)* is said to hold at a point  $\bar{y} \in Y(\bar{x})$  if the family of vectors  $D_y v_\ell(\bar{x}, \bar{y}), \ell \in L_0(\bar{x}, \bar{y})$ , is linearly independent, and the weaker *Mangasarian–Fromovitz constraint qualification (MFCQ)* holds if there exists some vector  $\eta \in \mathbb{R}^m$  such that  $D_y v_\ell(\bar{x}, \bar{y}) \eta < 0, \ell \in L_0(\bar{x}, \bar{y})$ . Here  $L_0(\bar{x}, \bar{y}) = \{\ell \in L \mid v_\ell(\bar{x}, \bar{y}) = 0\}$  denotes the set of lower level active indices, and  $D_y v_\ell$  stands for the row vector of partial derivatives of  $v_\ell$  with respect to  $y$ .

In what follows, let  $v$  be the column vector of the functions  $v_\ell, \ell \in L = \{1, \dots, s\}$ , let  $\text{diag}(\gamma)$  stand for the  $(s, s)$ -diagonal matrix with diagonal vector  $\gamma \in \mathbb{R}^s$ , and let  $j \in J_0(\bar{x})$ . Since each  $\bar{y} \in Y_0^j(\bar{x})$  is a solution of  $Q^j(\bar{x})$ , upon definition of the lower level Lagrange function

$$\mathcal{L}_j(x, y, \gamma) = g_j(x, y) - \gamma^\top v(x, y)$$

the Karush–Kuhn–Tucker theorem states that the following system of equalities and inequalities has a solution  $\gamma$  if the MFCQ holds at such a  $\bar{y}$ :

- (2.1)  $D_y^\top \mathcal{L}_j(\bar{x}, \bar{y}, \gamma) = 0,$
- (2.2)  $-\text{diag}(\gamma) v(\bar{x}, \bar{y}) = 0,$
- (2.3)  $\gamma \geq 0,$
- (2.4)  $-v(\bar{x}, \bar{y}) \geq 0.$

Note that  $\gamma$  is uniquely determined under the LICQ. We denote the set of Kuhn–Tucker multipliers corresponding to  $\bar{x}$  and  $\bar{y} \in Y_0^j(\bar{x})$  by

$$KT^j(\bar{x}, \bar{y}) = \{\gamma \in \mathbb{R}^s \mid \gamma \text{ satisfies (2.1)–(2.4)}\}.$$

The point  $\bar{y}$  is said to satisfy the *strict complementary slackness (SCS)* condition if  $\gamma_\ell > 0$ ,  $\ell \in L_0(\bar{x}, \bar{y})$ .

Under the LICQ the tangent space to  $Y(\bar{x})$  at  $\bar{y}$  can be described as  $T_{\bar{y}}Y(\bar{x}) = \{\eta \in \mathbb{R}^m \mid D_y v_\ell(\bar{x}, \bar{y})\eta = 0, \ell \in L_0(\bar{x}, \bar{y})\}$ . Let  $\bar{y} \in Y_0^j(\bar{x})$  and let  $\bar{\gamma}$  be the corresponding solution of (2.1)–(2.4). The point  $\bar{y}$  is said to satisfy the *second order sufficiency condition (SOSC)* if the matrix  $D_y^2 \mathcal{L}_j(\bar{x}, \bar{y}, \bar{\gamma})|_{T_{\bar{y}}Y(\bar{x})}$  possesses only negative eigenvalues. Here,  $D_y^2 \mathcal{L}_j = D_y D_y^\top \mathcal{L}_j$  denotes the Hessian matrix of  $\mathcal{L}_j$  with respect to  $y$ , and  $D_y^2 \mathcal{L}_j(\bar{x}, \bar{y}, \bar{\gamma})|_{T_{\bar{y}}Y(\bar{x})} = V^\top D_y^2 \mathcal{L}_j(\bar{x}, \bar{y}, \bar{\gamma})V$  for any matrix  $V$  of  $m$ -vectors which form a basis of the tangent space  $T_{\bar{y}}Y(\bar{x})$ .

**DEFINITION 2.1.** *Let  $\bar{x} \in \partial M \cap M$  and  $j \in J_0(\bar{x})$ . A point  $\bar{y} \in Y_0^j(\bar{x})$  is called the nondegenerate global maximizer of  $Q^j(\bar{x})$  if the LICQ holds at  $\bar{y}$  and if SCS and the SOSC are valid with the vector  $\gamma$  from (2.1)–(2.4).*

*Assumption 2 (reduction ansatz).* For each  $j \in J_0(\bar{x})$  all global maximizers of  $Q^j(\bar{x})$  are nondegenerate.

The reduction ansatz was originally formulated for standard semi-infinite optimization problems in [58, 17] under weaker regularity assumptions and was transferred to generalized semi-infinite optimization problems in [19]. For standard semi-infinite optimization problems the reduction ansatz is a natural assumption in the sense that for problems with defining functions in general position *it holds at each local minimizer* (cf. [59, 47]). For *GSIP* this result could be transferred to local minimizers  $\bar{x}$  with  $\sum_{j \in J_0(\bar{x})} |Y_0^j(\bar{x})| \geq n$  in [49]. Moreover, in [54] it is shown that it holds in the “completely linear” case, i.e., when all defining functions  $f, g_j, j \in J, v_\ell, \ell \in L$ , of *GSIP* are affine linear on their respective domains.

As under Assumption 2 the global maximizers of  $Q^j(\bar{x})$  are isolated points in  $Y(\bar{x})$ , and the latter set is compact, there are only finitely many global maximizers, say

$$Y_0^j(\bar{x}) = \{\bar{y}^{j,k}, k \in J_0^j(\bar{x})\}$$

with  $|J_0^j(\bar{x})| < \infty$ . An application of the implicit function theorem (cf. [7]) shows that for each  $\bar{y}^{j,k}$  with  $k \in J_0^j(\bar{x})$  and corresponding multiplier vector  $\bar{\gamma}^{j,k}$  there are locally defined  $C^{d-1}$ -functions  $y^{j,k}$  and  $\gamma^{j,k}$  with  $y^{j,k}(\bar{x}) = \bar{y}^{j,k}$  and  $\gamma^{j,k}(\bar{x}) = \bar{\gamma}^{j,k}$  such that  $y^{j,k}(x)$  is the locally unique local maximizer of  $Q^j(x)$  with multiplier  $\gamma^{j,k}(x)$ . Hence, we may introduce the locally defined optimal value functions

$$\varphi_{j,k}(x) = g_j(x, y^{j,k}(x)), \quad k \in J_0^j(\bar{x}), j \in J_0(\bar{x}).$$

**LEMMA 2.2** (cf., e.g., [24]). *The functions  $\varphi_{j,k}$  are of differentiability class  $C^d$ , and their gradients satisfy*

$$D\varphi_{j,k}(\bar{x}) = D_x \mathcal{L}_j(\bar{x}, \bar{y}^{j,k}, \bar{\gamma}^{j,k}).$$

The next result follows from a more general reduction lemma that we shall prove in section 4 (cf. Lemma 4.2).

**THEOREM 2.3** (reduction lemma; cf. [19, 47]). *Let Assumption 2 be satisfied at  $\bar{x}$ . Then the sets  $M$  and*

$$M_{\bar{x}} = \{x \in U \mid \varphi_{j,k}(x) \leq 0, k \in J_0^j(\bar{x}), j \in J_0(\bar{x})\}$$

coincide locally around  $\bar{x}$ .

Theorem 2.3 shows that under the reduction ansatz the original problem *GSIP* is locally equivalent to the reduced problem  $\min f|_{M_{\bar{x}}}$ . Hence, local optimality conditions from finite optimization may be applied to yield results for the semi-infinite case. In particular, we obtain a Fritz John-type first order necessary optimality condition (cf. [23]).

**THEOREM 2.4.** *Let  $\bar{x}$  be a local minimizer of *GSIP*, and let Assumption 2 hold. Then there exist multipliers  $\kappa \geq 0$ ,  $\lambda_{j,k} \geq 0$ ,  $k \in J_0^j(\bar{x})$ ,  $j \in J_0(\bar{x})$ , not all vanishing, such that*

$$(2.5) \quad \kappa Df(\bar{x}) + \sum_{j \in J_0(\bar{x})} \sum_{k \in J_0^j(\bar{x})} \lambda_{j,k} D_x \mathcal{L}_j(\bar{x}, \bar{y}^{j,k}, \bar{\gamma}^{j,k}) = 0.$$

Note that in Theorem 2.4 we do not need to impose a complementarity condition since all appearing constraints are active by definition. Moreover, recall that due to Carathéodory’s theorem at most  $n + 1$  nonvanishing multipliers  $\lambda_{j,k}$  are required in (2.5). In what follows we will call each point  $\bar{x}$  that satisfies the reduction ansatz and the necessary optimality condition from Theorem 2.4 a *Fritz John point* for *GSIP*.

**2.2. Convex lower level problems.** We call a problem  $Q^j(x)$ ,  $j \in J$ , convex if the functions  $-g_j(x, \cdot)$ ,  $v_\ell(x, \cdot)$ ,  $\ell \in L$ , are convex on  $\mathbb{R}^m$ . The main assumption of the present article is the following.

*Assumption 3.* The lower level problems  $Q^j(x)$ ,  $j \in J$ , are convex for all  $x \in \mathbb{R}^n$ .

Under Assumption 3 a set  $Y(x)$  with  $x \in \mathbb{R}^n$  is said to satisfy the Slater condition if there exists  $y^*$  such that  $v_\ell(x, y^*) < 0$  for all  $\ell \in L$ .

*Assumption 4.* The sets  $Y(x)$  are bounded and satisfy the Slater condition for all  $x \in \mathbb{R}^n$ .

Under Assumptions 3 and 4 the sets  $Y_x^j(x)$  are nonempty and locally bounded around each  $\bar{x} \in \mathbb{R}^n$  (cf. [22, Lemma 2]), so that the optimal value functions  $\varphi_j(x) = \max_{y \in Y(x)} g_j(x, y)$ ,  $j \in J$ , are well defined and continuous on  $\mathbb{R}^n$  [22]. Hence, the feasible set  $M$  is closed, and as in section 2.1 we have that  $M$  and the set  $\{x \in \mathbb{R}^n \mid \varphi_j(x) \leq 0, j \in J_0(\bar{x})\}$  coincide around each boundary point  $\bar{x}$ .

For the next theorem, which is a slight generalization of [42, Theorem 4.2], recall that the sets  $KT^j(\bar{x})$ ,  $j \in J_0(\bar{x})$ , do not depend on the variable  $y$  in the convex case (cf., e.g., [11]). Without loss of generality we set  $J_0(\bar{x}) = \{1, \dots, p_0\}$ .

**THEOREM 2.5.** *Let  $\bar{x}$  be a local minimizer of *GSIP* and let Assumptions 3 and 4 be satisfied. Then for each selection  $(\bar{\gamma}^1, \dots, \bar{\gamma}^{p_0}) \in KT^1(\bar{x}) \times \dots \times KT^{p_0}(\bar{x})$  there exist  $y^{j,k} \in Y_0^j(\bar{x})$ ,  $k = 1, \dots, p_j$ ,  $j \in J_0(\bar{x})$ ,  $\sum_{j \in J_0(\bar{x})} p_j \leq n + 1$ , and multipliers  $\kappa \geq 0$ ,  $\lambda_{j,k} \geq 0$ , not all equal to zero, such that*

$$\kappa Df(\bar{x}) + \sum_{j \in J_0(\bar{x})} \sum_{k=1}^{p_j} \lambda_{j,k} D_x \mathcal{L}^j(\bar{x}, y^{j,k}, \bar{\gamma}^j) = 0.$$

If, in addition to Assumptions 3 and 4, the reduction ansatz (Assumption 2) also holds, then the sets  $KT^j(\bar{x}) = \{\bar{\gamma}^j\}$  and  $Y_0^j(\bar{x}) = \{\bar{y}^j\}$ ,  $j \in J_0(\bar{x})$ , are singletons. In this case, Theorems 2.4 and 2.5 obviously simplify to the following result.

**COROLLARY 2.6.** *Let  $\bar{x}$  be a local minimizer of *GSIP* and let Assumptions 2, 3, and 4 be satisfied. Then there exist multipliers  $\kappa \geq 0$ ,  $\lambda_j \geq 0$ ,  $j \in J_0(\bar{x})$ , not all equal to zero, such that*

$$\kappa Df(\bar{x}) + \sum_{j \in J_0(\bar{x})} \lambda_j D_x \mathcal{L}^j(\bar{x}, \bar{y}^j, \bar{\gamma}^j) = 0.$$

The following lemma is well known. A short proof can be found in [52].

LEMMA 2.7. *Let Assumptions 3 and 4 be satisfied. Then a point  $\bar{y}^j$  is a nondegenerate global maximizer of  $Q^j(\bar{x})$  with corresponding multiplier vector  $\bar{\gamma}^j$  if and only if (2.1)–(2.4) hold and if the Jacobian of (2.1), (2.2) with respect to  $(y^j, \gamma^j)$ ,*

$$A^j = A^j(\bar{x}, \bar{y}^j, \bar{\gamma}^j) = \begin{pmatrix} D_y^2 \mathcal{L}_j(\bar{x}, \bar{y}^j, \bar{\gamma}^j) & -D_y^\top v(\bar{x}, \bar{y}^j) \\ -\text{diag}(\bar{\gamma}^j) D_y v(\bar{x}, \bar{y}^j) & -\text{diag}(v(\bar{x}, \bar{y}^j)) \end{pmatrix},$$

is nonsingular.

**2.3. NCP functions.** A function  $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$  with

$$\psi(a, b) = 0 \quad \text{if and only if} \quad a \geq 0, \quad b \geq 0, \quad ab = 0$$

is called an *NCP function*. Let us remark that the existence of a  $C^\infty$ -NCP function is clear from a theorem by Whitney [3]. However, as smooth NCP functions are degenerate at the origin, in the following we will work with the nonsmooth NCP functions

$$\psi^{NR}(a, b) = \frac{1}{2} \left( a + b - \sqrt{(a - b)^2} \right)$$

and

$$\psi^{FB}(a, b) = a + b - \sqrt{a^2 + b^2}.$$

The function  $\psi^{NR}$  is the so-called natural residual or min-function since it coincides with  $\min(a, b)$ , and  $\psi^{FB}$  is known as the Fischer–Burmeister function [8].

For numerical purposes one can regularize these nondifferentiable NCP functions. The so-called Chen–Harker–Kanzow–Smale function [4, 27, 46] is given by

$$\psi_\tau^{NR}(a, b) = \frac{1}{2} \left( a + b - \sqrt{(a - b)^2 + 4\tau^2} \right),$$

whereas the so-called smoothed Fischer–Burmeister function is

$$\psi_\tau^{FB}(a, b) = a + b - \sqrt{a^2 + b^2 + 2\tau^2}.$$

Obviously,  $\psi_\tau^{NR}$  and  $\psi_\tau^{FB}$  are continuously differentiable for all  $\tau \neq 0$ , and for  $\tau = 0$  they coincide with  $\psi^{NR}$  and  $\psi^{FB}$ , respectively. Moreover, both functions share the following important properties.

LEMMA 2.8. *Let  $\tau \neq 0$ , and let  $\psi_\tau$  denote one of the functions  $\psi_\tau^{NR}$  and  $\psi_\tau^{FB}$ . Then the following assertions hold:*

- (i) *We have  $\psi_\tau(a, b) = 0$  if and only if  $a > 0, b > 0, ab = \tau^2$ .*
- (ii) *For a zero  $(a, b)$  of  $\psi_\tau$  the gradient  $D\psi_\tau(a, b)$  does not explicitly depend on  $\tau$  and is given by  $(a + b)^{-1}(b, a)$ .*

*Proof.* Part (i) was observed in [27], and part (ii) is easily verified. □

In what follows we mainly need the results of Lemma 2.8, so we will not distinguish between  $\psi_\tau^{NR}$  and  $\psi_\tau^{FB}$  but simply write  $\psi_\tau$ .

**3. The numerical approach.** The aim of our numerical method is to replace *GSIP* by a sequence of finite nonlinear programming problems which are numerically tractable and whose solutions or stationary points converge to a solution or a stationary point of *GSIP*, respectively. Unlike other numerical methods for semi-infinite

programming, our approach does not discretize the index set  $Y(x)$ , but we take advantage of the fact that the solution set of a regular convex lower level problem is characterized by its first order optimality condition. Thus, let Assumptions 3 and 4 hold throughout this section.

In a first step we reformulate *GSIP* as a special Stackelberg game:

$$SG : \min_{x, y^1, \dots, y^p} f(x) \text{ subject to (s.t.) } g_j(x, y^j) \leq 0, \text{ and } y^j \text{ solves } Q^j(x), j \in J.$$

Note that the *SG* possesses two special features: its objective function  $f$  does not depend on the variables  $y^j, j \in J$ , and its upper level inequality constraint functions coincide with its lower level objective functions. In [54] it is shown that *GSIP* and *SG* are equivalent problems whenever the index set  $Y(x)$  is nonempty for all  $x \in \mathbb{R}^n$ . The latter is the case under Assumption 4. We point out that for  $Y(x) = \emptyset$  the point  $x$  would be feasible for *GSIP* but infeasible for *SG*.

Next, since the problems  $Q^j(x)$  are convex, we may replace the restrictions “ $y^j$  solves  $Q^j(x)$ ” in *SG* equivalently by their first order optimality conditions: for each  $j \in J$  there is a solution  $\gamma^j$  of (2.1)–(2.4). The latter statement is true under Assumption 4, since Slater’s condition guarantees the existence of Kuhn–Tucker multipliers. However, unlike in the case of the LICQ, these multipliers are not necessarily uniquely determined. By this reformulation, *SG* is equivalent to the following mathematical programming problem with equilibrium constraints:

$$MPEC : \min_{x, y^1, \gamma^1, \dots, y^p, \gamma^p} f(x) \text{ s.t.} \begin{aligned} g_j(x, y^j) &\leq 0, \\ D_y^\top \mathcal{L}_j(x, y^j, \gamma^j) &= 0, \\ -\text{diag}(\gamma^j) v(x, y^j) &= 0, \\ \gamma^j &\geq 0, \\ -v(x, y^j) &\geq 0, j \in J. \end{aligned}$$

At this point, *GSIP* has been replaced by an equivalent finite nonlinear programming problem. However, numerical standard software cannot be expected to solve this problem since due to the appearance of complementarity conditions the MFCQ is violated at all points of the feasible set of *MPEC* (cf. [43]). In [26, 38] it is shown that the MFCQ is a necessary condition for the stability of smooth nonlinear programs under data perturbations and thus for the stability of numerical methods in the presence of round-off errors.

Given an NCP function  $\psi$  and  $a, b \in \mathbb{R}^s$  we define the vectorization

$$\Psi(a, b) = (\psi(a_1, b_1), \dots, \psi(a_s, b_s))^\top$$

so that *MPEC* can be equivalently rewritten as

$$P : \min_{x, y^1, \gamma^1, \dots, y^p, \gamma^p} f(x) \text{ s.t.} \begin{aligned} g_j(x, y^j) &\leq 0, \\ D_y^\top \mathcal{L}_j(x, y^j, \gamma^j) &= 0, \\ \Psi(\gamma^j, -v(x, y^j)) &= 0, j \in J. \end{aligned}$$

We now apply an interior point approach to the lower level problems  $Q^j(x)$ . For  $j \in J$  we replace the Karush–Kuhn–Tucker system (2.1)–(2.4) at  $y^j$  and its corresponding multiplier vector  $\gamma^j$  by the perturbed system

$$(3.1) \quad D_y^\top \mathcal{L}_j(x, y^j, \gamma^j) = 0,$$





LEMMA 3.2. *Let  $j \in J$  and  $\tau \neq 0$ .*

- (i) *The point  $y^j$  is a solution of  $Q_\tau^j(x)$  if and only if  $(y^j, \gamma^j)$  with  $\gamma_\ell^j = -\tau^2/v_\ell(x, y^j)$ ,  $\ell \in L$ , is a solution of (3.1)–(3.4). Moreover, for the latter solutions  $D_y^2 b_\tau^j(x, y^j)$  is nonsingular if and only if  $A^j(x, y^j, \gamma^j)$  is nonsingular.*
- (ii) *If at least one of the matrices  $D_y^2 g_j(x, y^j)$ ,  $D_y^2 v_\ell(x, y^j)$ ,  $\ell \in L$ , is nonsingular, then  $A^j(x, y^j, \gamma^j)$  is nonsingular, too.*

*Proof.* The first part of (i) is evident from a comparison between the relations  $D_y \mathcal{L}_j(x, y^j, \gamma^j) = 0$  and  $D_y b_\tau^j(x, y^j) = 0$ . For the second part note that in view of Lemma 2.8(i) the matrix  $\text{diag}(v(x, y^j))$  is nonsingular so that  $A^j = A^j(x, y^j, \gamma^j)$  is nonsingular if and only if the Schur complement  $S^j$  of  $\text{diag}(v(x, y^j))$  in  $A^j$  is nonsingular. This Schur complement is just  $S^j = D_y^2 b_\tau^j(x, y^j)$ .

Due to our convexity assumptions (Assumption 3) it is easily seen that  $S^j$  is the sum of negative semidefinite matrices. Under the assumption of part (ii), at least one of the matrices  $D_y^2 g_j(x, y^j)$ ,  $-D_y^2 v_\ell(x, y^j)$ ,  $\ell \in L$ , is actually negative definite and, since all numbers  $\tau^2/v_\ell(x, y^j)$ ,  $\ell \in L$ , are negative,  $S^j$  is negative definite, too. Together with part (i) this shows the assertion of part (ii).  $\square$

A different proof for a weaker result related to part (ii) of the preceding lemma can be found in [28]. Let us point out that Lemma 3.2 provides sufficient conditions for the assumption of nonsingular matrices  $A^j(x, y^j, \gamma^j)$ ,  $j \in J$ , in Proposition 3.1. From Lemma 2.7 and a simple continuity argument it follows that  $A^j(x, y^j, \gamma^j)$  is also nonsingular if  $(x, y^j, \gamma^j)$  is sufficiently close to a point  $(\bar{x}, \bar{y}^j, \bar{\gamma}^j)$  such that  $\bar{y}^j$  is a nondegenerate solution of  $Q^j(\bar{x})$  with corresponding multiplier vector  $\bar{\gamma}^j$ .

The ideas presented so far lead to a simple continuation method for the numerical solution of *GSIP* which is easy to implement and can be given conceptually in the following form.

**Numerical method.**

*Step 1.* Choose a sequence  $\{\tau_\nu\}$  of nonzero reals with  $\lim_{\nu \rightarrow \infty} \tau_\nu = 0$  and a starting point  $x^0 \in \mathbb{R}^n$ .

*Step 2.* Compute a starting point  $(x^{0,0}, y^{1,0,0}, \gamma^{1,0,0}, \dots, y^{p,0,0}, \gamma^{p,0,0})$  of  $P_{\tau_0}$  and set  $\nu = 0$ .

*Step 3.* Find a solution  $(x^{\nu,*}, y^{1,\nu,*}, \dots, \gamma^{p,\nu,*})$  of  $P_{\tau_\nu}$ .

*Step 4.* Set  $(x^{\nu+1,0}, y^{1,\nu+1,0}, \dots, \gamma^{p,\nu+1,0}) = (x^{\nu,*}, y^{1,\nu,*}, \dots, \gamma^{p,\nu,*})$ ,  $\nu := \nu + 1$ , and go to Step 3.

In Step 2, we clearly choose  $x^{0,0} = x^0$ . In order to obtain the corresponding values  $(y^{1,0,0}, \dots, \gamma^{p,0,0})$  numerically, one might try to find a zero of

$$\begin{pmatrix} D_y^\top \mathcal{L}_j(x^0, y^j, \gamma^j) \\ \Psi_{\tau_0}(\gamma^j, -v(x^0, y^j)) \end{pmatrix}$$

for each  $j \in J$ . Another method will be given below.

Step 3 is a “black box” which stands for any standard solution method for nonlinear finite optimization problems. In view of Steps 2 and 4, a minimal requirement is that the method should be able to process infeasible starting points.

Conceptually, termination criteria might be the relative error of optimal points or of optimal values, as well as the error in the first order optimality condition for *GSIP* (cf. Corollary 2.6) and combinations thereof. We emphasize that the availability of an easily checkable first order optimality condition is crucial for the numerical performance of the method.

**4. Convergence results.** Recall that in section 3 we reformulated *GSIP* equivalently first as an *SG* and then as a finite optimization problem  $P$ . Then we embedded  $P$  in the parametric family  $P_\tau$  with  $\tau \in \mathbb{R}$ . Let us now clarify what the equivalent embedding for *GSIP* is. Throughout this section, let Assumptions 3 and 4 hold.

With the observations in section 3 the problem  $P_\tau$  with  $\tau \neq 0$  can equivalently be written in the form of an *SG*:

$$SG_\tau : \min_{x, y^1, \dots, y^p} f(x) \text{ s.t. } g_j(x, y^j) \leq 0, \text{ and } y^j \text{ solves } Q_\tau^j(x), j \in J.$$

Under the reduction ansatz, locally near a feasible point  $\bar{x}$  of *GSIP* (i.e., for  $\bar{\tau} = 0$ ) there exists a unique solution of (3.1)–(3.4) with  $\tau \neq 0$ .

**PROPOSITION 4.1.** *Let  $\bar{x} \in M$  and  $j \in J$  be given. Assume that  $\bar{y}^j$  is a solution of  $Q^j(\bar{x})$  with corresponding multiplier vector  $\bar{\gamma}^j$  such that  $A^j(\bar{x}, \bar{y}^j, \bar{\gamma}^j)$  is nonsingular. Then there exist neighborhoods  $U$  of  $\bar{x}$  and  $T$  of  $\bar{\tau} = 0$  as well as  $C^{d-1}$ -functions  $y^j : V \rightarrow \mathbb{R}^m, \gamma^j : V \rightarrow \mathbb{R}^s$  ( $V := U \times T$ ) such that  $y^j(\bar{x}, 0) = \bar{y}^j, \gamma^j(\bar{x}, 0) = \bar{\gamma}^j$  and such that for all  $(x, \tau) \in V$   $(y^j(x, \tau), \gamma^j(x, \tau))$  is the unique solution of (3.1)–(3.4).*

*Proof.* The proof follows directly by applying the implicit function theorem to the system of equations (3.1), (3.2), and using that  $A^j(\bar{x}, \bar{y}^j, \bar{\gamma}^j)$ , its Jacobian with respect to  $(y, \gamma)$  at  $(\bar{x}, \bar{y}^j, \bar{\gamma}^j, \bar{\tau})$ , is nonsingular. Note that (3.3) and (3.4) hold thanks to continuity arguments.  $\square$

Under the reduction ansatz (Assumption 2) the assumptions of Proposition 4.1 are satisfied for  $j \in J_0(\bar{x})$  due to Lemma 2.7. Because of Lemma 3.2(i) for all  $x \in U, \tau \in T$  the condition

$$g_j(x, y^j) \leq 0 \text{ and } y^j \text{ solves } Q_\tau^j(x)$$

is equivalent with

$$g_j(x, y^j(x, \tau)) \leq 0.$$

Note that under Assumptions 3 and 4 for  $j \in J \setminus J_0(\bar{x})$  the value function  $\tilde{\varphi}_j(x, \tau)$  of  $Q_\tau^j(x)$  is continuous in the neighborhood  $U \times T$  of  $(\bar{x}, 0)$ . So for all  $\tau \in T$  the problem  $P_\tau$ , locally restricted to  $x \in U$ , is equivalent to the reduced problem

$$GSIP_\tau(\bar{x}) : \min_{x \in U} f(x) \text{ s.t. } g_j(x, y^j(x, \tau)) \leq 0 \text{ for all } j \in J_0(\bar{x}).$$

Summarizing, we have obtained the following reduction lemma, which provides the basis for our convergence analysis.

**LEMMA 4.2** (parametric reduction lemma). *Let the reduction ansatz hold at a point  $\bar{x} \in M$ . Then locally in a neighborhood  $U \times T$  of  $(\bar{x}, 0)$  (cf. Proposition 4.1) the problems  $P_\tau$  and  $GSIP_\tau(\bar{x})$  are equivalent in the sense that for all  $\tau \in T$  the vector  $x_\tau \in U$  is a solution of  $GSIP_\tau(\bar{x})$  if and only if  $(x_\tau, y^1, \dots, y^p)$  with  $x_\tau \in U$  solves  $SG_\tau$ . In particular, problem  $GSIP_0(\bar{x})$  is locally in  $U$  equivalent with  $SG_0 = SG$  and, hence, with *GSIP*.*

The above lemma yields local reductions for all problems  $P_\tau$  when  $(x, \tau)$  is sufficiently close to  $(\bar{x}, 0)$ , which shows that the parametric reduction lemma implies the reduction lemma (Theorem 2.3).

The following theorem is related to a result by Shimizu and Aiyoshi [45] about the convergence behavior of the solutions of  $SG_\tau$  for  $\tau \rightarrow 0$ . However, our proof relies only on the parametric reduction lemma and some well-known results from parametric optimization.

**THEOREM 4.3.** *Let  $(\tau_\nu)_{\nu \in \mathbb{N}}$  be a sequence with  $\lim_{\nu \rightarrow \infty} \tau_\nu = 0$ , and let  $(x^\nu, y^{1,\nu}, \gamma^{1,\nu}, \dots, y^{p,\nu}, \gamma^{p,\nu})_{\nu \in \mathbb{N}}$  be a sequence of global solutions of  $P_{\tau_\nu}$ ,  $\nu \in \mathbb{N}$  (cf. (3.5)). If  $x^*$  is an accumulation point of the sequence  $(x^\nu)_{\nu \in \mathbb{N}}$  such that Assumption 2 holds at  $x^*$  and such that the MFCQ holds at some solution of  $GSIP_0(x^*)$ , then  $x^*$  is a global solution of  $GSIP$ .*

*Proof.* Without loss of generality, let  $(x^\nu, \tau_\nu)$  converge to  $(x^*, 0)$  for  $\nu \rightarrow \infty$ . By Proposition 4.1 for the solutions  $y^{j,\nu}$ ,  $j \in J_0(x^*)$ , we have (for sufficiently large  $\nu$ )  $y^{j,\nu} = y(x^\nu, \tau_\nu)$ , and  $x^\nu$  is a solution of  $GSIP_{\tau_\nu}(x^*)$ . By continuity  $x^*$  is feasible for  $GSIP_0(x^*) = GSIP$  (locally in  $U$ ). As the MFCQ holds at some solution of  $GSIP_0(x^*)$ , by a result due to Gauvin and Dubeau [9] the value function  $\omega(\tau)$  of  $GSIP_\tau(x^*)$  is continuous in  $\tau$ . Consequently  $\omega(\tau_\nu) = f(x^\nu) \rightarrow \omega(0)$  and  $x^*$  is a solution of  $GSIP$ .  $\square$

Theorem 4.3 is primarily of theoretical interest, as numerical standard software may usually not find global solution points of the problems  $P_{\tau_\nu}$ ,  $\nu \in \mathbb{N}$ . One can at most expect a point which satisfies a first order optimality condition like the one of Fritz John. Consequently, a numerical solution method for  $GSIP$  can also only be expected to find Fritz John points in the sense of section 2.

In the following we study how Fritz John points of the finite problems  $P_\tau$  ( $\tau \neq 0$ ) are related to the Fritz John points of  $GSIP_\tau(\bar{x})$  and  $GSIP$ .

**LEMMA 4.4.** *Let the reduction ansatz hold at a point  $\bar{x} \in M$ , and let  $(x, y^1, \gamma^1, \dots, y^p, \gamma^p)$  be a Fritz John point of  $P_\tau$  (cf. (3.5)) with  $(x, \tau)$  sufficiently close to  $(\bar{x}, 0)$ . Moreover, let the matrices*

$$A^j = \begin{pmatrix} D_y^2 \mathcal{L}_j(x, y^j, \gamma^j) & -D_y^\top v(x, y^j) \\ -\text{diag}(\gamma^j) D_y v(x, y^j) & -\text{diag}(v(x, y^j)) \end{pmatrix}, \quad j \in J \setminus J_0(\bar{x}),$$

*be nonsingular. Then  $x$  is a Fritz John point of  $GSIP_\tau(\bar{x})$ .*

*Proof.* It is not hard to see that  $x$  is feasible for  $GSIP_\tau(\bar{x})$ . The feasibility of  $(x, y^1, \gamma^1, \dots, y^p, \gamma^p)$  for  $P_\tau$  implies particularly that  $y^j$  solves  $Q_\tau^j(x)$ . Since the matrices  $A^j$  are nonsingular for all  $j \in J$ , for  $(x, \tau)$  sufficiently close to  $(\bar{x}, 0)$  the point  $y^j$  coincides with the unique solution  $y^j(x, \tau)$  of  $Q_\tau^j(x)$  (cf. Proposition 4.1). These observations, together with some simple continuity arguments, yield the reduction of the Fritz John condition of  $P_\tau$  to the one of  $GSIP_\tau(\bar{x})$ .  $\square$

Now we can prove the main result of this section.

**THEOREM 4.5.** *Let  $(\tau_\nu)_{\nu \in \mathbb{N}}$  be a sequence with  $\lim_{\nu \rightarrow \infty} \tau_\nu = 0$ , and let  $(x^\nu, y^{1,\nu}, \gamma^{1,\nu}, \dots, y^{p,\nu}, \gamma^{p,\nu})$  be Fritz John points of  $P_{\tau_\nu}$ ,  $\nu \in \mathbb{N}$  (cf. (3.5)), with an accumulation point  $(x^*, y^{1,*}, \gamma^{1,*}, \dots, y^{p,*}, \gamma^{p,*})$ . Let the reduction ansatz (Assumption 2) hold at  $x^*$ , and let the matrices*

$$A^j = \begin{pmatrix} D_y^2 \mathcal{L}_j(x^*, y^{j,*}, \gamma^{j,*}) & -D_y^\top v(x^*, y^{j,*}) \\ -\text{diag}(\gamma^{j,*}) D_y v(x^*, y^{j,*}) & -\text{diag}(v(x^*, y^{j,*})) \end{pmatrix}, \quad j \in J \setminus J_0(x^*),$$

*be nonsingular. Then  $x^*$  is a Fritz John point of  $GSIP$ .*

*Proof.* For sufficiently large  $\nu \in \mathbb{N}$  all assumptions of Lemma 4.4 are satisfied so that  $x^\nu$  is a Fritz John point of  $GSIP_{\tau_\nu}(x^*)$ . By a continuity argument and Lemma 2.2  $x^*$  is thus a Fritz John point for  $GSIP$  in the sense of Corollary 2.6.  $\square$

The following proposition gives a criterion for the existence of an accumulation point in the assumption of Theorem 4.5.

**PROPOSITION 4.6.** *Let  $(\tau_\nu)_{\nu \in \mathbb{N}}$  be a sequence with  $\lim_{\nu \rightarrow \infty} \tau_\nu = 0$ , and let  $(x^\nu, y^{1,\nu}, \gamma^{1,\nu}, \dots, y^{p,\nu}, \gamma^{p,\nu})$  be feasible points of  $P_{\tau_\nu}$ ,  $\nu \in \mathbb{N}$  (cf. (3.5)). Moreover, let  $x^*$  be an accumulation point of the sequence  $(x^\nu)_{\nu \in \mathbb{N}}$  such that the LICQ holds*

everywhere in  $Y(x^*)$ . Then the sequence  $(x^\nu, y^{1,\nu}, \gamma^{1,\nu}, \dots, y^{p,\nu}, \gamma^{p,\nu})_{\nu \in \mathbb{N}}$  possesses an accumulation point  $(x^*, y^{1,*}, \gamma^{1,*}, \dots, y^{p,*}, \gamma^{p,*})$ .

*Proof.* After taking a subsequence, let  $x^\nu \rightarrow x^*$  for  $\nu \rightarrow \infty$ . For  $\nu \in \mathbb{N}$  the feasibility of  $(x^\nu, y^{1,\nu}, \gamma^{1,\nu}, \dots, y^{p,\nu}, \gamma^{p,\nu})$  for  $P_{\tau_\nu}$  implies  $\psi_{\tau_\nu}(\gamma_\ell^{j,\nu}, -v_\ell(x^\nu, y^{j,\nu})) = 0$  and thus, by Lemma 2.8(i),  $v_\ell(x^\nu, y^{j,\nu}) < 0$  for all  $\ell \in L, j \in J$ . Hence,  $y^{j,\nu} \in Y(x^\nu), j \in J$ , and the upper semicontinuity of  $Y$  yields that an accumulation point  $y^{j,*}$  of  $y^{j,\nu}, \nu \in \mathbb{N}$ , exists and is contained in  $Y(x^*), j \in J$ .

Next assume that for some  $j \in J$  the sequence  $(\gamma^{j,\nu})_{\nu \in \mathbb{N}}$  is unbounded. Then for sufficiently large  $\nu$  we can consider the vectors  $\|\gamma^{j,\nu}\|^{-1}\gamma^{j,\nu}$ , which converge to a vector  $\eta^j$  with  $\|\eta^j\| = 1$ , possibly after taking a subsequence.

The feasibility of  $(x^\nu, y^{1,\nu}, \gamma^{1,\nu}, \dots, y^{p,\nu}, \gamma^{p,\nu})$  for  $P_{\tau_\nu}$  yields

$$(4.1) \quad D_y g_j(x^\nu, y^{j,\nu}) - \gamma^{j,\nu} D_y v(x^\nu, y^{j,\nu}) = 0$$

and, by Lemma 2.8(i),

$$(4.2) \quad -\gamma_\ell^{j,\nu} \cdot v_\ell(x^\nu, y^{j,\nu}) = \tau_\nu^2, \quad \ell \in L.$$

Division of (4.2) by  $\|\gamma^{j,\nu}\|$  and taking the limit for  $\nu \rightarrow \infty$  yields  $\eta_\ell^j = 0$  for  $\ell \in L \setminus L_0(x^*, y^{j,*})$ . However, in the same way (4.1) then implies that the LICQ is violated at  $y^{j,*}$  in  $Y(x^*)$ , contradicting our assumptions. Consequently, for each  $j \in J$  the sequence  $(\gamma^{j,\nu})_{\nu \in \mathbb{N}}$  is bounded and, thus, has an accumulation point.  $\square$

The existence of some accumulation point  $x^*$  in the assumption of Proposition 4.6 can of course be guaranteed if an additional restriction  $x \in X$  is incorporated into *GSIP*, with  $X$  nonempty and compact.

We end this section with a result on the rate of convergence of the method.

LEMMA 4.7. For  $\bar{x} \in M$  and  $j \in J$  let  $y^j$  be a nondegenerate solution of  $Q^j(\bar{x})$ , and let  $(y^j(x, \tau), \gamma^j(x, \tau))$  denote the locally unique solution of (3.1)–(3.4) around  $(\bar{x}, 0)$ . Then we have

$$\begin{pmatrix} D_\tau y^j(x, 0) \\ D_\tau \gamma^j(x, 0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

for  $x$  sufficiently close to  $\bar{x}$ .

In the following proposition we call  $\bar{x} \in M$  a nondegenerate solution of *GSIP* if the reduction ansatz holds at  $\bar{x}$  and if  $\bar{x}$  is a nondegenerate local minimizer for the locally reduced problem  $GSIP_0(\bar{x})$ .

PROPOSITION 4.8. Let the assumptions of Theorem 4.3 hold, and let the solution  $x^*$  of *GSIP* be nondegenerate. Then for each subsequence of  $(x^\nu)_{\nu \in \mathbb{N}}$  that converges to  $x^*$ , the optimal values of  $P_{\tau_\nu}$  satisfy

$$f(x^\nu) - f(x^*) = O(\tau_\nu^2).$$

*Proof.* Let  $\omega(\tau)$  denote the optimal value of  $P_\tau$ . In the proof of Theorem 4.3 we have seen that  $\omega(\tau)$  coincides with the optimal value of  $GSIP_\tau(x^*)$  if  $\tau$  is sufficiently close to zero. Now Lemma 2.2, Lemma 4.7, and a Taylor expansion of  $\omega$  around 0 yield the assertion.  $\square$

**5. Numerical examples.** For the numerical illustrations we implemented the method from section 3 in *Matlab* 5.3 and used the routine *fmincon* from its *Optimization Toolbox* 2.0, i.e., an SQP method with BFGS updates for the Hessian of the

Lagrangian, to replace the “black box” in Step 3 of the method. All examples were run on an 800 MHz Linux PC.

In Step 2 we do *not* solve the nonlinear problems

$$\begin{pmatrix} D_y^\top \mathcal{L}_j(x^0, y^j, \gamma^j) \\ \Psi_{\tau_0}(\gamma^j, -v(x^0, y^j)) \end{pmatrix} = 0, \quad j \in J,$$

since an appropriate starting point for an iteration procedure is not at hand. Instead we solve the unconstrained, concave problems

$$Q_{\tau_0}^j(x^0) : \quad \max_y g_j(x^0, y) + \tau_0^2 \sum_{\ell \in L} \ln(-v_\ell(x^0, y)), \quad j \in J,$$

by the routine *fminunc* to obtain  $y^{j,0,0}$ ,  $j \in J$ , and put

$$\gamma_\ell^{j,0,0} = -\frac{\tau_0^2}{v_\ell(x^0, y^{j,0,0})}, \quad \ell \in L, \quad j \in J.$$

Here the problem of finding an appropriate starting point is solved easily by the determination of some Slater point of the set  $Y(x^0)$ . The latter task can be fulfilled by solving the convex problem

$$\min_{y, \eta} \eta \quad \text{s.t.} \quad v_\ell(x^0, y) - \eta \leq 0, \quad \ell \in L,$$

where the choice of a starting point is obvious.

As an a priori  $\tau$ -sequence we use  $\tau_\nu = 10 \cdot 100^{-\nu}$ ,  $\nu \in \mathbb{N}$ . The iteration terminates if either the relative error in the optimal point or in the optimal value is less than  $10^{-6}$ . If the method is provided with the gradients of  $f$ ,  $g$ , and  $v$  with respect to  $x$ , then also the first order condition for *GSIP* from Corollary 2.6 is checked (in the Euclidean norm).

**5.1. Design centering in two dimensions.** The general design centering problem (see also [14]) consists of maximizing some measure, e.g., the volume, of a parameterized body  $B(x)$  which is contained in a second body  $G$ :

$$\max_{x \in \mathbb{R}^n} \text{Vol}(B(x)) \quad \text{s.t.} \quad B(x) \subset G.$$

In the first examples we let  $G = \{y \in \mathbb{R}^2 \mid g(y) \leq 0\}$  with

$$g(y) = \begin{pmatrix} -y_1 - y_2^2 \\ y_1/4 + y_2 - 3/4 \\ -y_2 - 1 \end{pmatrix}.$$

The two-dimensional volume of the resulting body is easily calculated to be  $20/3$ . An equivalent formulation of the general design centering problem as *GSIP* is

$$\max_{x \in \mathbb{R}^n} \text{Vol}(B(x)) \quad \text{s.t.} \quad g(y) \leq 0 \quad \text{for all } y \in B(x).$$

*Problem 1.* We look for the largest ball with free center and radius that is contained in  $G$ . Thus, we have  $n = 3$  and

$$B(x) = \{y \in \mathbb{R}^2 \mid (y_1 - x_1)^2 + (y_2 - x_2)^2 - x_3^2 \leq 0\}, \quad \text{Vol}(B(x)) = \pi x_3^2.$$

TABLE 1  
Approximation of  $G$  with  $\psi^{NR}$ .

Problem	ov	$\varepsilon_{ov}$	$\varepsilon_{op}$	$\varepsilon_{FOC}$	CPU <sub>init</sub>	CPU <sub>iter</sub>	# <sub>iter</sub>
1	1.8606	2.5596e-06	*	*	2.94	1.04	3
2	3.4838	*	*	4.8781e-06	2.96	3.55	4
3	3.7234	*	*	n.a.	3.31	7.50	4
4	3.0792	3.9278e-06	*	2.1812e-06	3.76	1.85	3

As an initial point we use the infeasible point  $x^0 = (0, 0, 1)^\top$ .

*Problem 2.* We search the largest ellipsoid with free center and axis lengths that is contained in  $G$ . The axes are supposed to be parallel to the coordinate axes. We have  $n = 4$  and

$$B(x) = \left\{ y \in \mathbb{R}^2 \mid \frac{(y_1 - x_1)^2}{x_3^2} + \frac{(y_2 - x_2)^2}{x_4^2} - 1 \leq 0 \right\}, \text{Vol}(B(x)) = \pi x_3 x_4.$$

The initial point is  $x^0 = (0, 0, 1, 1)^\top$ .

*Problem 3.* Now the ellipsoid from Problem 2 is allowed to have axes in arbitrary position, and with  $n = 6$  we can set

$$B(x) = \left\{ y \in \mathbb{R}^2 \mid \left( y - \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right)^\top \left( \begin{pmatrix} x_3 & x_4 \\ x_5 & x_6 \end{pmatrix} \begin{pmatrix} x_3 & x_5 \\ x_4 & x_6 \end{pmatrix} \right)^{-1} \left( y - \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) - 1 \leq 0 \right\}.$$

Then we have

$$\text{Vol}(B(x)) = \pi \left| \det \begin{pmatrix} x_3 & x_4 \\ x_5 & x_6 \end{pmatrix} \right|,$$

and we choose the initial point  $x^0 = (0, 0, 1, 0, 0, 1)^\top$ .

*Problem 4.* In this problem we inscribe the largest box with sides parallel to the coordinate axes into  $G$ . For  $n = 4$  we have

$$B(x) = \{ y \in \mathbb{R}^2 \mid y_1 - x_1 \leq 0, \quad y_2 - x_2 \leq 0, \quad -y_1 + x_3 \leq 0, \quad -y_2 + x_4 \leq 0 \}$$

with

$$\text{Vol}(B(x)) = (x_1 - x_3) \cdot (x_2 - x_4),$$

and we choose the infeasible initial point  $x^0 = (1, 1, -1, -1)^\top$ .

The columns of the following tables are labeled as follows:  $ov$ , optimal value;  $\varepsilon_{ov}$ , relative error in optimal value;  $\varepsilon_{op}$ , relative error in optimal point;  $\varepsilon_{FOC}$ , error in first order optimality condition; CPU<sub>init</sub>, CPU time for initialization step in seconds; CPU<sub>iter</sub>, CPU time for iterations in seconds; #<sub>iter</sub>, number of outer iterations. If (relative) errors are below  $10^{-6}$ , we replace the actual number by the symbol “\*”. In Problem 3 the gradients entering the first order optimality condition are not available for the method, so this criterion is not checked.

For the results in Table 1 we used the natural residual function  $\psi^{NR}$  as the NCP function. In this example the performance of the method does not change significantly if the NCP function  $\psi^{NR}$  is replaced by  $\psi^{FB}$  (see [52] for more details).

Concerning the solution of Problem 4 it is worth mentioning, as expected, we have  $x_4^* = -1$ . This means that the method converges although  $Q^3(x^*)$  clearly is a *degenerate* problem. In fact, the computed “maximal box” in  $G$  is the set  $B(x^*) = [-0.024, 3.619] \times [-1, -0.155]$ , so that the solution set of  $Q^3(x^*)$  coincides with the facet  $[-0.024, 3.619] \times \{-1\}$  of  $B(x^*)$ . From all these optimal points the method chooses their “midpoint” as  $y^{3,*} = (1.7975, -1)$ . In fact, in our approach the lower level linear problems are solved by the central path method. It is well known in linear programming that under weak assumptions the interior point sequence converges to the so-called analytic center of the optimal facet (cf. [39]).

**5.2. Robust optimization.** Robustness questions arise when an optimization problem is subject to uncertain data. If an inequality constraint function  $g(x, y)$  depends on some uncertain parameter vector  $y \in Y \subset \mathbb{R}^m$ , then the “most cautious” way to deal with this constraint is to use its worst-case reformulation

$$g(x, y) \leq 0 \text{ for all } y \in Y,$$

which is clearly of semi-infinite type. When the uncertainty set  $Y$  also depends on the state variable  $x$ , we arrive at a generalized semi-infinite constraint.

The following robust optimization problem is studied in [1] for elliptic uncertainty sets. In the case of ellipticity the lower level optimal value functions can be computed explicitly in such a way that the semi-infinite problem is reduced to a nonsmooth finite problem which can be tackled by SDP methods (cf. [1] for details). We will show that our numerical method solves not only this specially structured problem but also two nontrivial generalizations.

Let 1 euro be invested in a portfolio comprised of  $N$  shares. At the end of a given period the return per 1 euro invested in share  $i$  is  $y_i > 0$ . The goal is to determine the amount  $x_i$  to be invested in share  $i$ ,  $i = 1, \dots, N$ , so as to maximize the end-of-period portfolio value  $y^\top x$ .

If the vector  $y$  was certain, the solution of this optimization problem would be evident. A more realistic assumption is that  $y$  varies in some nonempty compact set  $Y \subset \mathbb{R}^N$ . Upon moving the objective function to the constraint set we obtain the following standard semi-infinite optimization problem with  $n = N + 1$  and  $m = N$ :

$$\max_{x, x_{N+1}} x_{N+1} \quad \text{s.t.} \quad x_{N+1} - y^\top x \leq 0, \quad y \in Y, \quad \sum_{i=1}^N x_i = 1, \quad x \geq 0.$$

Apart from its special structure used in [1] for the case of an ellipsoidal set  $Y$ , this is also a linear semi-infinite optimization problem, meaning that the semi-infinite constraint function is linear in the variable  $(x, x_{N+1})$ . Solution methods for these types of problems are described, e.g., in [10, 20]. Note, however, that the index set of the semi-infinite constraint is  $N$ -dimensional, where  $N$  might be a large number.

*Problem 5.* In [1] the set  $Y$  has the form

$$Y = \left\{ y \in \mathbb{R}^N \mid \sum_{i=1}^N \frac{(y_i - \bar{y}_i)^2}{\sigma_i^2} \leq \theta^2 \right\},$$

where  $\bar{y}_i$  is some “nominal” value of  $y_i$ ,  $\sigma_i$  is a scaling parameter,  $i = 1, \dots, N$ , and  $\theta$  measures the risk aversion of the decision maker. With the particular choices

$$\bar{y}_i = 1.15 + i \cdot \frac{0.05}{N}, \quad i = 1, \dots, N,$$

TABLE 2  
*Optimal portfolio with ellipsoidal uncertainty and  $\psi^{NR}$ .*

$N$	ov	$\varepsilon_{ov}$	$\varepsilon_{op}$	$\varepsilon_{FOC}$	CPU <sub>init</sub>	CPU <sub>iter</sub>	#iter
10	1.15	★	1.3693e-03	3.0860e-05	1.55	3.71	3
50	1.15	★	5.4195e-05	6.6652e-05	7.24	27.23	4
100	1.15	★	3.3458e-05	7.5987e-05	66.27	241.6	4
150	1.15	★	1.9149e-05	4.1678e-05	272.94	884.22	4

TABLE 3  
*Optimal portfolio with ellipsoidal uncertainty and  $\psi^{FB}$ .*

$N$	ov	$\varepsilon_{ov}$	$\varepsilon_{op}$	$\varepsilon_{FOC}$	CPU <sub>init</sub>	CPU <sub>iter</sub>	#iter
10	1.15	★	7.7231e-04	2.8254e-05	1.53	3.68	4
50	1.15	★	5.4165e-05	5.5333e-05	7.21	34.52	4
100	1.15	★	4.7199e-05	1.5974e-04	65.17	333.8	5
150	1.15	★	4.7277e-05	3.7021e-04	271.49	912.72	4

$$\sigma_i = \frac{0.05}{3N} \sqrt{2N(N+1)i}, \quad i = 1, \dots, N,$$

$$\theta = 1.5,$$

one can show that the optimal policy is to invest equally in all shares, i.e.,  $x_i = 1/N$ ,  $i = 1, \dots, N$ , with optimal value 1.15. We use the starting point  $x^0 = (1, 0, \dots, 0)$  in  $\mathbb{R}^{N+1}$ .

Having the large dimensions of  $Y$  in mind, the method performs very well when  $\psi^{NR}$  is used (cf. Table 2). In the case of  $\psi^{FB}$  as the NCP function, the “black box” part of the method (i.e., the *Matlab* routine *fmincon*) does not converge for  $N = 100$  (cf. Table 3), so that we solved this particular problem with the a priori sequence  $\tau_\nu = 10^{-\nu}$ ,  $\nu \in \mathbb{N}$ .

*Problem 6.* A more general choice of  $Y$  is

$$Y_\delta = \{ y \in \mathbb{R}^N \mid \|\text{diag}(\sigma)^{-1}(y - \bar{y})\|_\delta \leq \theta \}$$

with  $\delta \in [1, \infty]$ . Whereas  $Y_2$  is the ellipsoid from Problem 5, the sets  $Y_1$  and  $Y_\infty$  are polytopes. For all other choices of  $\delta$  we still obtain a nonempty compact convex set  $Y_\delta$ . As polytopes can be considered as ellipsoidal sets in the sense of [1], let us use our method for a *nonellipsoidal* set like  $Y_{10}$ . Tables 4 and 5 show the results for the starting point  $x^0 = 1/N \cdot (1, \dots, 1, 0)$ . The method performs well for dimensions up to  $N = 150$ . In this example the initialization phase takes about as long as the main iterations. Note that for increasing dimensions the attainment of the first order condition becomes worse when the method terminates because of a small relative error in the optimal value.

*Problem 7.* Finally, since our method works for  $x$ -dependent sets  $Y$ , we can also consider the case in which the risk aversion of the decision maker depends on the point  $x$ . If for instance his risk aversion increases when the values  $x_i$  deviate from  $1/N$ ,  $i = 1, \dots, N$ , we can replace  $\theta$  by the expression

$$\Theta(x) = \theta \cdot \left( 1 + \sum_{i=1}^N \left( x_i - \frac{1}{N} \right)^2 \right)$$

and obtain the *generalized* semi-infinite optimization problem

$$\max_{x, x_{N+1}} x_{N+1} \quad \text{s.t.} \quad x_{N+1} - y^\top x \leq 0, \quad y \in Y(x), \quad \sum_{i=1}^N x_i = 1, \quad x \geq 0,$$



TABLE 4  
*Optimal portfolio with nonellipsoidal uncertainty and  $\psi^{NR}$ .*

$N$	ov	$\varepsilon_{ov}$	$\varepsilon_{op}$	$\varepsilon_{FOC}$	CPU <sub>init</sub>	CPU <sub>iter</sub>	#iter
10	1.1190	★	7.8075e-05	1.5139e-03	2.82	5.63	3
50	1.1155	★	1.5402e-05	6.1081e-02	21.36	41.14	3
100	1.1151	★	9.0467e-06	6.0789e-02	246.35	231.22	3
150	1.1150	★	2.4410e-05	2.0368e-01	809.94	708.99	3

TABLE 5  
*Optimal portfolio with nonellipsoidal uncertainty and  $\psi^{FB}$ .*

$N$	ov	$\varepsilon_{ov}$	$\varepsilon_{op}$	$\varepsilon_{FOC}$	CPU <sub>init</sub>	CPU <sub>iter</sub>	#iter
10	1.1190	★	5.4529e-04	1.1783e-03	2.81	5.58	3
50	1.1155	★	4.3946e-04	6.0965e-02	21.26	41.71	3
100	1.1151	★	1.5527e-05	5.7950e-02	241.95	228.89	3
150	1.1150	★	1.0349e-05	2.1809e-01	811.68	703.06	3

TABLE 6  
*Optimal portfolio with state-dependent uncertainty and  $\psi^{NR}$ .*

$N$	ov	$\varepsilon_{ov}$	$\varepsilon_{op}$	$\varepsilon_{FOC}$	CPU <sub>init</sub>	CPU <sub>iter</sub>	#iter
10	0.7033	★	★	6.9574e-06	2.81	1.45	4
50	0.9638	★	1.1912e-05	1.1828e-03	14.78	5.49	4
100	1.0259	★	2.5937e-05	5.5778e-03	132.54	37.05	3
150	1.0535	★	7.0987e-06	2.3336e-03	643.39	73.96	3

TABLE 7  
*Optimal portfolio with state-dependent uncertainty and  $\psi^{FB}$ .*

$N$	ov	$\varepsilon_{ov}$	$\varepsilon_{op}$	$\varepsilon_{FOC}$	CPU <sub>init</sub>	CPU <sub>iter</sub>	#iter
10	0.7033	★	1.3743e-05	1.8857e-04	2.77	1.53	4
50	0.9638	★	3.7616e-05	3.7318e-03	14.76	4.84	3
100	1.0259	★	4.5000e-05	9.6580e-03	133.00	20.48	3
150	1.0535	★	3.6030e-05	1.1934e-02	642.96	58.05	3

with

$$Y(x) = \left\{ y \in \mathbb{R}^N \mid \sum_{i=1}^N \frac{(y_i - \bar{y}_i)^2}{\sigma_i^2} \leq \Theta(x)^2 \right\}.$$

The choice  $\sigma_i = 1, i = 1, \dots, N$ , prevents the solution of the original Problem 5 to be  $x_i = 1/N, i = 1, \dots, N$ , so that the modified term  $\Theta(x)$  can take effect, and we observe a good performance of our method in Tables 6 and 7 up to dimension  $N = 150$ . Note, however, that in this example the initialization phase takes considerably longer than the main iterations.

**6. Final remarks.** The essential idea behind the numerical method presented in section 3 is to reformulate the generalized semi-infinite optimization problem as a Stackelberg game and to make use of the convexity in the lower level problems. Starting at this point, there are several possible routes to the design of a numerical method.

First, there are other ways to treat the convex lower level problems than the one used in this article, e.g., penalty, barrier, or cutting plane methods. Furthermore, if one decides to replace the convex optimization problems by their first order optimality conditions and obtain a mathematical program with equilibrium constraints, this

program does not necessarily have to be solved with regularized NCP functions. Instead, the exact penalization approaches from [34, 44] or nonsmooth Newton methods (see, e.g., [30]) could be promising alternatives. Finally, in the case when one uses NCP functions, there are a multitude of functions other than the natural residual and the Fischer–Burmeister function to choose from. For a survey see, e.g., [5].

Further questions concern the required accuracy for the solution of the auxiliary problems by the “black box” method, the implementation of an active set strategy, and the design of a pathfollowing method (cf. [15]) to solve the finite parametric optimization problems  $P_\tau$  for parameter values tending to  $\tau = 0$ . Let us point out, however, that the resulting numerical method will then not just solve a sequence of finite dimensional optimization problems that are easily constructed from the problem data, so that the implementation effort for the user increases drastically.

Moreover, generalizations of the convergence proofs from section 4 to cases such as convex lower level problems in which the reduction ansatz is not necessarily satisfied (recall Problem 4 in section 5.1) will be the subject of future research. We finally remark that a direct application of the presented ideas to semi-infinite optimization problems with *nonconvex* lower level problems results, in general, only in a *relaxation* of the original problem, since lower level optimality can then not be replaced equivalently with a first order optimality condition. However, our method can then still be used to obtain lower bounds for the optimal value of *GSIP*.

**Acknowledgment.** We express our thanks to the referee, whose precise and substantial remarks led to an improved version of the article.

#### REFERENCES

- [1] A. BEN-TAL AND A. NEMIROVSKI, *Robust solutions of uncertain linear programs*, Oper. Res. Lett., 25 (1999), pp. 1–13.
- [2] C. BERGE, *Topological Spaces*, Oliver and Boyd, Edinburgh, 1963.
- [3] T. BRÖCKER AND L. LANDER, *Differentiable Germs and Catastrophes*, London Math. Soc. Lecture Notes Ser. 17, Cambridge University Press, Cambridge, UK, 1975.
- [4] B. CHEN AND P. T. HARKER, *A non-interior-point continuation method for linear complementarity problems*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 1168–1190.
- [5] B. CHEN, X. CHEN, AND C. KANZOW, *A penalized Fischer–Burmeister NCP-function*, Math. Program., 88 (2000), pp. 211–216.
- [6] F. FACCHINEI, H. JIANG, AND L. QI, *A smoothing method for mathematical programs with equilibrium constraints*, Math. Program., 85 (1999), pp. 107–134.
- [7] A. V. FIACCO AND G. P. MCCORMICK, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley, New York, 1968.
- [8] A. FISCHER, *A special Newton-type optimization method*, Optimization, 24 (1992), pp. 269–284.
- [9] J. GAUVIN AND F. DUBEAU, *Differential properties of the marginal function in mathematical programming*, Math. Program. Study, 19 (1982), pp. 101–119.
- [10] M. A. GOBERNA AND M. A. LÓPEZ, *Linear Semi-Infinite Optimization*, John Wiley, Chichester, 1998.
- [11] E. G. GOL’STEIN, *Theory of Convex Programming*, Transl. Math. Monogr. 36, AMS, Providence, RI, 1972.
- [12] T. J. GRAETTINGER AND B. H. KROGH, *The acceleration radius: A global performance measure for robotic manipulators*, IEEE J. Robotics Automat., 4 (1988), pp. 60–69.
- [13] G. GRAMLICH, R. HETTICH, AND E. W. SACHS, *Local convergence of SQP methods in semi-infinite programming*, SIAM J. Optim., 5 (1995), pp. 641–658.
- [14] P. GRITZMANN AND V. KLEE, *On the complexity of some basic problems in computational convexity. I. Containment problems*, Discrete Math., 136 (1994), pp. 129–174.
- [15] J. GUDDAT, F. GUERRA VASQUEZ, AND H. TH. JONGEN, *Parametric Optimization: Singularities, Pathfollowing and Jumps*, John Wiley, Chichester, Teubner, Stuttgart, 1990.
- [16] M. GUGAT, *Semi-infinite terminal problems: A Newton-type method*, Optimization, 44 (1998), pp. 25–48.

- [17] R. HETTICH AND H. TH. JONGEN, *Semi-infinite programming: Conditions of optimality and applications*, in Optimization Techniques, Part 2, Lecture Notes in Control and Inform. Sci. 7, J. Stoer, ed., Springer-Verlag, Berlin, 1978, pp. 1–11.
- [18] R. HETTICH AND K. O. KORTANEK, *Semi-infinite programming: Theory, methods, and applications*, SIAM Rev., 35 (1993), pp. 380–429.
- [19] R. HETTICH AND G. STILL, *Second order optimality conditions for generalized semi-infinite programming problems*, Optimization, 34 (1995), pp. 195–211.
- [20] R. HETTICH AND P. ZENCKE, *Numerische Methoden der Approximation und semi-infiniten Optimierung*, Teubner Studienbücher, Stuttgart, 1982.
- [21] W. W. HOGAN, *Point-to-set maps in mathematical programming*, SIAM Rev., 15 (1973), pp. 591–603.
- [22] W. W. HOGAN, *Directional derivatives for extremal value functions with applications to the completely convex case*, Oper. Res., 21 (1973), pp. 188–209.
- [23] F. JOHN, *Extremum problems with inequalities as subsidiary conditions*, in Studies and Essays, R. Courant Anniversary Volume, Interscience Publishers, New York, 1948, pp. 187–204.
- [24] H. TH. JONGEN, P. JONKER, AND F. TWILT, *Nonlinear Optimization in Finite Dimensions*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000.
- [25] H. TH. JONGEN, J.-J. RÜCKMANN, AND O. STEIN, *Generalized semi-infinite optimization: A first order optimality condition and examples*, Math. Program., 83 (1998), pp. 145–158.
- [26] H. TH. JONGEN AND G. W. WEBER, *Nonlinear optimization: Characterization of structural stability*, J. Global Optim., 1 (1991), pp. 47–64.
- [27] C. KANZOW, *Some noninterior continuation methods for linear complementarity problems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 851–868.
- [28] C. KANZOW AND H. JIANG, *A continuation method for (strongly) monotone variational inequalities*, Math. Program., 81 (1998), pp. 103–125.
- [29] A. KAPLAN AND R. TICHATSCHKE, *On a class of terminal variational problems*, in Parametric Optimization and Related Topics IV, J. Guddat, H. Th. Jongen, F. Nožička, G. Still, F. Twilt, eds., Peter Lang, Frankfurt, Germany, 1997, pp. 185–199.
- [30] M. KOČVARA, J. OUTRATA, AND J. ZOWE, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.
- [31] W. KRABS, *On time-minimal heating or cooling of a ball*, in Numerical Methods of Approximation Theory, Vol. 8, Internat. Schriftenreihe. Numer. Math. 81, Birkhäuser, Basel, 1987, pp. 121–131.
- [32] E. LEVITIN, *Reduction of Generalized Semi-Infinite Programming Problems to Semi-Infinite or Piece-Wise Smooth Programming Problems*, Preprint no. 8-2001, University of Trier, Trier, Germany, 2000.
- [33] E. LEVITIN AND R. TICHATSCHKE, *A branch-and-bound approach for solving a class of generalized semi-infinite programming problems*, J. Global Optim., 13 (1998), pp. 299–315.
- [34] Z. LUO, J. PANG, AND D. RALPH, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, UK, 1996.
- [35] V. H. NGUYEN AND J. J. STRODIOT, *Computing a global optimal solution to a design centering problem*, Math. Program., 53 (1992), pp. 111–123.
- [36] E. POLAK, *An implementable algorithm for the optimal design centering, tolerancing and tuning problem*, J. Optim. Theory Appl., 37 (1982), pp. 45–67.
- [37] R. REEMTSEN AND J.-J. RÜCKMANN, EDs., *Semi-Infinite Programming*, Kluwer Academic Publishers, Boston, 1998.
- [38] S. M. ROBINSON, *Stability theory for systems of inequalities, part II: Differentiable nonlinear systems*, SIAM J. Numer. Anal., 13 (1976), pp. 497–513.
- [39] C. ROOS, T. TERLAKY, AND J.-P. VIAL, *Theory and Algorithms for Linear Optimization*, John Wiley, Chichester, UK, 1997.
- [40] J.-J. RÜCKMANN AND A. SHAPIRO, *First-order optimality conditions in generalized semi-infinite programming*, J. Optim. Theory Appl., 101 (1999), pp. 677–691.
- [41] J.-J. RÜCKMANN AND O. STEIN, *On linear and linearized generalized semi-infinite optimization problems*, Ann. Oper. Res., 101 (2001), pp. 191–208.
- [42] J.-J. RÜCKMANN AND O. STEIN, *On convex lower level problems in generalized semi-infinite optimization*, in Semi-Infinite Programming – Recent Advances, M. A. Goberna, M. A. Lopez, eds., Kluwer Academic Publishers, Dordrecht, The Netherlands, 2001, pp. 121–134.
- [43] H. SCHEEL AND S. SCHOLTES, *Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity*, Math. Oper. Res., 25 (2000), pp. 1–22.
- [44] S. SCHOLTES AND M. STÖHR, *Exact penalization of mathematical programs with equilibrium constraints*, SIAM J. Control Optim., 37 (1999), pp. 617–652.

- [45] K. SHIMIZU AND E. AIYOSHI, *A new computational method for Stackelberg and min-max problems by use of a penalty method*, IEEE Trans. Automat. Control, AC-26 (1981), pp. 460–466.
- [46] S. SMALE, *Algorithms for solving equations*, in Proceedings of the International Congress of Mathematicians, AMS, Providence, RI, 1987, pp. 172–195.
- [47] O. STEIN, *On Parametric Semi-Infinite Optimization*, Shaker-Verlag, Aachen, Germany, 1997.
- [48] O. STEIN, *The reduction ansatz in absence of lower semi-continuity*, in Parametric Optimization and Related Topics V, J. Guddat, R. Hirabayashi, H. T. Jongen, and F. Twilt, eds., Peter Lang, Frankfurt, Germany, 2000, pp. 165–178.
- [49] O. STEIN, *On level sets of marginal functions*, Optimization, 48 (2000), pp. 43–67.
- [50] O. STEIN, *The feasible set in generalized semi-infinite programming*, in Approximation, Optimization and Mathematical Economics, M. Lassonde, ed., Physica-Verlag, Heidelberg, 2001, pp. 313–331.
- [51] O. STEIN, *First order optimality conditions for degenerate index sets in generalized semi-infinite programming*, Math. Oper. Res., 26 (2001), pp. 565–582.
- [52] O. STEIN, *Bi-level Strategies in Semi-infinite Programming*, Kluwer Academic Publishers, Boston, 2003.
- [53] O. STEIN AND G. STILL, *On optimality conditions for generalized semi-infinite programming problems*, J. Optim. Theory Appl., 104 (2000), pp. 443–458.
- [54] O. STEIN AND G. STILL, *On generalized semi-infinite optimization and bilevel optimization*, European J. Oper. Res., 142 (2002), pp. 444–462.
- [55] G. STILL, *Generalized semi-infinite programming: Numerical aspects*, Optimization, 49 (2001), pp. 223–242.
- [56] G. STILL, *Discretization in semi-infinite programming: The rate of convergence*, Math. Program., 91 (2001), pp. 53–69.
- [57] G.-W. WEBER, *Generalized Semi-Infinite Optimization and Related Topics*, Habilitation Thesis, Darmstadt University of Technology, Darmstadt, Germany, 1999.
- [58] W. WETTERLING, *Definitheitsbedingungen für relative Extrema bei Optimierungs- und Approximationsaufgaben*, Numer. Math., 15 (1970), pp. 122–136.
- [59] G. ZWIER, *Structural Analysis in Semi-Infinite Programming*, Thesis, University of Twente, Enschede, The Netherlands, 1987.