J Optim Theory Appl (2010) 146: 419–443
DOI 10.1007/s10957-010-9674-5

Feasible Method for Generalized Semi-Infinite Programming

O. Stein · A. Winterfeld

Published online: 19 March 2010

Abstract In this paper, we analyze the outer approximation property of the algorithm for generalized semi-infinite programming from Stein and Still (SIAM J. Control Optim. 42:769–788, 2003). A simple bound on the regularization error is found and used to formulate a *feasible* numerical method for generalized semi-infinite programming with convex lower-level problems. That is, all iterates of the numerical method are feasible points of the original optimization problem. The new method has the same computational cost as the original algorithm from Stein and Still (SIAM J. Control Optim. 42:769–788, 2003). We also discuss the merits of this approach for the adaptive convexification algorithm, a feasible point method for standard semi-infinite programming from Floudas and Stein (SIAM J. Optim. 18:1187–1208, 2007).

Keywords Semi-infinite programming · Interior-point method · Mathematical program with equilibrium constraints · Bilevel programming · Design centering

1 Introduction

In this article, we study a numerical solution method for generalized semi-infinite optimization problems of the type

$$\begin{aligned} \text{(GSIP)} \qquad & \min \quad f(x), \\ & \text{s.t.} \quad x \in M, \end{aligned}$$

O. Stein (✉)
Institute of Operations Research, Karlsruhe Institute of Technology, Karlsruhe, Germany
e-mail: stein@kit.edu

A. Winterfeld
Fraunhofer Institut für Techno- und Wirtschaftsmathematik, Kaiserslautern, Germany
e-mail: winterfeld@itwm.fraunhofer.de

with

$$M = \{x \in \mathbb{R}^n \mid g_i(x, y) \leq 0, \text{ for all } y \in Y(x),\ i \in I\},$$
$$Y(x) = \{y \in \mathbb{R}^m \mid v_l(x, y) \leq 0,\ l \in L\}$$

and finite index sets $I = \{1, \ldots, r\}$, $L = \{1, \ldots, s\}$. All defining functions f, g_i, $i \in I$, v_l, $l \in L$, are assumed to be real-valued and at least twice continuously differentiable on their respective domains.

For an introduction to theory, applications, and numerical methods for GSIP we refer to the monography [3] and to the survey article [4]. A semi-infinite problem is called standard semi-infinite if the mapping $Y(\cdot)$ is constant.

Several real-life applications of generalized semi-infinite optimization problems are described in [3], for example robust optimization, design centering, minimax problems, Chebyshev approximation, and disjunctive programming. These problem classes are of major interest, among others, in chemical engineering, in economical and physical models, as well as in the geometrical layout of systems.

In particular, the theoretical and numerical treatment of design centering (DC) problems is essential for cutting stock and nesting problems, where given shapes are cut from a valuable material resource so that the amount of wasted material is minimized [5]. The nesting problem of cutting a gem of maximal volume with prescribed shape features from a raw gem is treated in the recent thesis [6] as a generalized semi-infinite program. Among other applications of design centering are the so-called maneuverability problem of a robot from [7] and the determination of 'innermost' points of a container set of uncertain quality parameters to find a point where a company can safely produce a good [8].

More formally, a design centering problem deals with the maximization of some measure f of a parameterized body Y, the *design*, which must be contained in a second body C, the *container*, where the dimension of the embedding space will be denoted by m. Both the container C and the design Y can depend on a parameter vector $x \in \mathbb{R}^n$,

$$Y(x) = \{y \in \mathbb{R}^m \mid v_l(x, y) \leq 0,\ l \in L\},$$
$$C(x) = \{y \in \mathbb{R}^m \mid g_i(x, y) \leq 0,\ i \in I\}.$$

Then, problem DC can be formulated as

$$\begin{array}{lll} \text{(DC)} & \min_{x \in \mathbb{R}^n} & f(x), \\ & \text{s.t.} & Y(x) \subseteq C(x), \end{array}$$

which, after writing down the inclusion constraint explicitly, is just the above problem GSIP.

The following assumptions are made throughout this paper.

Assumption 1 For all $x \in \mathbb{R}^n$, $Y(x)$ is a nonempty compact set.

Assumption 2 For all $x \in \mathbb{R}^n$, $Y(x)$ is *convex* and the functions $g_i(x,\cdot)$ are *concave* in y for $i \in I$.

Assumption 3 For all $x \in \mathbb{R}^n$, $Y(x)$ has a Slater point, i.e. a point $y(x)$ such that

$$v_l(x, y(x)) < 0, \quad l \in L. \tag{1}$$

The optimal solution of a generalized semi-infinite problem is closely related to its *lower-level problems*, which have the form

$$(Q^i(x)) \qquad \begin{aligned} &\max_y \ g_i(x,y), \\ &\text{s.t.} \quad y \in Y(x), \quad i \in I. \end{aligned}$$

Let $\varphi_i(x)$ denote the optimal value of $Q^i(x)$, $i \in I$. Then it is not hard to see that $x \in \mathbb{R}^n$ is feasible for GSIP if and only if $\varphi_i(x) \le 0$ is satisfied for all $i \in I$. The main computational problem in semi-infinite programming stems from the fact that the evaluation of $\varphi_i(x)$ involves the *global* maximization of $Q^i(x)$. Under Assumptions 2 and 3, however, the lower level problem $Q^i(x)$ is a regular convex problem for all $x \in \mathbb{R}^n$ and $i \in I$, making its global solution numerically tractable. Moreover, under Assumptions 1, 2 and 3 the optimal value functions $\varphi_i(x)$, $i \in I$, are well-defined and continuous on $\mathbb{R}^n$ [9], so that the feasible set M is closed.

In recent years some efforts were made to obtain feasible numerical methods for *standard* semi-infinite programming (SIP) *without* convex lower level problems. To our knowledge, [10] describes the first algorithm with feasible iterates for SIP. The main idea leading to the feasibility of the iterates is bounding the optimal values $\varphi_i(x)$, $i \in I$, of the lower level problems *from above*. The techniques used in [10] in order to compute such upper bounds are interval arithmetic and inclusion functions.

A reference dealing with the feasibility issue in more explicit form is [2]. To formulate a method for standard SIP with convex index set but non-concave upper level constraints, the lower level problems are convexified by the αBB method, a technique of global optimization. The fact that the convexified lower level problems are *relaxations* of the original lower level problems again yields upper bounds on $\varphi_i(x)$, and an adaptive convexification method with feasible iterates is formulated. We will return to this method in Sect. 6. We point out that, since lower level convexity is not assumed in these approaches, both algorithms have to employ heavy-weight methods of global optimization which leads to high computational costs.

On the other hand, for *generalized* semi-infinite programming *with* convex lower level problems a different numerical method was introduced in [1]. Although it promises to be much more efficient than approaches without convexity assumptions, the regularization approach of this algorithm leads to an outer approximation method and, thus, all its iterates have to be expected to be infeasible for the original feasible set. The aim of the present article is to show how a simple modification of the latter algorithm transforms it into a feasible point method, that is, all its iterates are feasible for the original feasible set.

This paper is organized as follows. In Sect. 2 we briefly recall the main ideas of the algorithm from [1]. In Sect. 3 we derive an upper bound of the lower level

optimal values and thus a quantification of the regularization error. It turns out that the regularization error is easy to compute and *a-priori known*. In Sect. 4 this observation is used to construct a parametric inner approximation of M. In Sect. 5, the inner approximation is used to formulate a GSIP method with the inner approximation property. By construction, the method has the same computational cost as the original algorithm from [1]. The problem of finding a feasible initial point for the method is also discussed in this section and leads to a two-phase approach. Section 6 shows how the adaptive convexification algorithm from [2] for standard semi-infinite programs benefits from the obtained results, before Sect. 7 ends the article with conclusions and some open questions.

2 Outer Approximation Method for Generalized Semi-infinite Programming

This section recalls the basic ideas behind the numerical method for GSIP from [1]. The approach bases on the observation from [11] that GSIP and the Stackelberg game [12]

$$\begin{array}{lll} \text{(SG)} & \min\limits_{x,y^1,\ldots,y^r} & f(x), \\ & \text{s.t.} & g_i(x,y^i) \le 0, \\ & & y^i \ \text{ solves } Q^i(x), \quad i \in I, \end{array}$$

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 3.

Given a Stackelberg game reformulation of GSIP, the special structure of the lower level problems can be exploited in order to replace GSIP by an equivalent one-level finite optimization problem. In fact, recall that under Assumptions 2 and 3 all lower level problems $Q^i(x)$ are regular convex problems. Thus the Karush-Kuhn-Tucker conditions of $Q^i(x)$ are equivalent to global optimality for all x and for all $i \in I$. Replacing 'y^i solves $Q^i(x)$' in the formulation of SG by the KKT system of $Q^i(x)$ yields the following equivalent reformulation of SG as a special *mathematical program with equilibrium constraints*. Here

$$\mathcal{L}_i(x,y,\gamma) = g_i(x,y) - \gamma^T v(x,y)$$

denotes the Lagrange function of $Q^i(x)$ with multiplier vector $\gamma \in \mathbb{R}^s$, and D_y stands for the partial gradient with respect to the variable vector y:

$$\begin{array}{lll} \text{(MPEC)} & \min\limits_{x,y^1,\gamma^1,\ldots,y^r,\gamma^r} & f(x), \\ & \text{s.t.} & g_i(x,y^i) \le 0, \quad i \in I, \qquad (2) \\ & & D_y \mathcal{L}_i(x,y^i,\gamma^i) = 0, \quad i \in I, \qquad (3) \\ & & -\operatorname{diag}(\gamma^i) v(x,y^i) = 0, \quad i \in I, \qquad (4) \\ & & \gamma^i \ge 0, \quad i \in I, \qquad (5) \\ & & -v(x,y^i) \ge 0, \quad i \in I. \qquad (6) \end{array}$$

Mathematical programs with equilibrium constraints frequently arise in the context of game theory and have been successively applied to various practical problems. See [13] for a large collection of engineering and economic applications of MPEC and [14] for an annotated bibliography.

Unfortunately, MPEC is an ill-posed problem even under very strong assumptions. In fact, it is known (cf. [15, 16]) that MFCQ is violated at any feasible point of MPEC due to the presence of degenerate complementary slackness constraints (4). Since MFCQ is known to be a necessary condition for stability in presence of round-off errors (cf. [17]), the direct application of standard non-linear solvers to MPEC may fail.

We note, however, that an SQP method has been found to be competitive compared to other approaches on a large set of MPECs in [18]. Since direct application of SQP to our instances of MPEC failed, we follow the proposal made in [3] and use an *explicit smoothing method*. For more information on smoothing methods, see, e.g., [19–21]. We note that there are many alternative ways to remove the ill-posedness of MPEC. Some examples are relaxation schemes [22], penalty methods [23], extensions of the standard SQP methods [15, 24], and a recent lifting approach [25]. An extensive overview of methods for MPEC and further references can be found in [26].

The main idea behind the explicit regularization method is to replace the malignant constraint (4) by

$$-\operatorname{diag}(\gamma^i)v(x,y^i)=\tau^2 e,\quad i\in I, \tag{7}$$

where $\tau>0$ is a perturbation parameter and $e=(1,\ldots,1)^T\in\mathbb{R}^s$. So-called *smoothed NCP-functions* can now be used to replace (5)–(7). Here a function $\psi:\mathbb{R}^2\to\mathbb{R}$ satisfying

$$\psi(a,b)=0\quad\Leftrightarrow\quad a,b\geq 0,\quad ab=0 \tag{8}$$

is called NCP function, the terminology owing to the fact that for each $i\in I$ the constraints (4)–(6) constitute a non-linear complementary slackness problem. There is a large number of NCP functions available (see, e.g., [27]). Since the zero set of any NCP function has to exhibit a kink at the origin, those functions are either nonsmooth or degenerate. Nonsmooth ones may be smoothed in such a way that their perturbed zero set models, for example, (7).

In the sequel, we will use the *Chen-Harker-Kanzow-Smale* (CHKS) function which is also called *the smoothed natural residual function* and has the form

$$\psi_\tau(a,b)=\frac{1}{2}\bigl(a+b-\sqrt{(a-b)^2+4\tau^2}\bigr). \tag{9}$$

For $\tau\neq 0$ one can show (cf., e.g., [3]) that $\psi_\tau(a,b)=0$ holds if and only if $ab=\tau^2,\ a>0,\ b>0$ is satisfied. Let

$$\Psi_\tau:\mathbb{R}^s\times\mathbb{R}^s\to\mathbb{R}^s,\quad (v,w)\mapsto(\psi_\tau(v_1,w_1),\ldots,\psi_\tau(v_s,w_s))^T \tag{10}$$

denote the vectorization of ψ_τ. Then, the constraints (5)–(7) hold if and only if

$$\Psi_\tau(\gamma^i,-v(x,y^i))=0, \tag{11}$$

so that MPEC is embedded in the parametric family of problems P_τ for $\tau = 0$, where

$$
\begin{aligned}
(P_\tau) \qquad & \min_{x,y^1,\gamma^1,\ldots,y^r,\gamma^r} && f(x), \\
& \text{s.t.} && g_i(x, y^i) \le 0, \quad i \in I, && (12)\\
& && D_y \mathcal{L}_i(x, y^i, \gamma^i) = 0, \quad i \in I, && (13)\\
& && \Psi_\tau(\gamma^i, -v(x, y^i)) = 0, \quad i \in I. && (14)
\end{aligned}
$$

In [3] it is shown that the degeneracy of the complementarity constraints has now been removed, and P_τ can be solved by a 'black-box' non-linear optimization software package. The optimal solution of $P_0 = \text{MPEC}$ may be iteratively found by solving a series of problems P_{τ_k}, where τ_k is a sequence converging to zero.

Algorithm 1 (Numerical GSIP Solver)

Step 1. Choose a sequence $\{\tau_k\} \subset \mathbb{R}$ with $\lim_{k\to\infty} \tau_k = 0$ and a starting point $x^0 \in \mathbb{R}^n$. Compute a starting point $(x^{0,0}, y^{1,0,0}, \gamma^{1,0,0}, \ldots, y^{r,0,0}, \gamma^{r,0,0})$ of P_{τ_0} and set $k = 0$.

Step 2. Find an optimal solution $(x^{k,*}, y^{1,k,*}, \gamma^{1,k,*}, \ldots, y^{r,k,*}, \gamma^{r,k,*})$ of P_{τ_k} using $(x^{k,0}, y^{1,k,0}, \gamma^{1,k,0}, \ldots, y^{r,k,0}, \gamma^{r,k,0})$ as the starting point.

Step 3. If some termination criterion is satisfied, terminate with the optimal solution of GSIP $x^{k,*}$.

Step 4. Set $(x^{k+1,0}, y^{1,k+1,0}, \ldots, \gamma^{r,k+1,0}) = (x^{k,*}, y^{1,k,*}, \ldots, \gamma^{r,k,*})$, replace k by $k+1$ and go to Step 2.

Since Algorithm 1 is based on the reduction of GSIP to finite non-linear problems P_τ, obviously the solutions computed by the method can be only as 'optimal' as the results delivered by the black-box non-linear solver.

Fortunately, it can be shown that the optimality conditions achieved by the finite non-linear solver can be translated into corresponding optimality conditions of GSIP. The convergence of global solutions of P_{τ_k} to a global solution of GSIP [3, Theorem 5.2.2] is rarely observed in practice due to the non-convexity of P_{τ_k}. It is more realistic to expect the non-linear solver to return merely Karush-Kuhn-Tucker or Fritz John points of P_{τ_k}. The convergence of Fritz John points of P_{τ_k} to a Fritz John point of GSIP is shown in [3, Theorem 5.1] under mild additional assumptions. The convergence of Karush-Kuhn-Tucker points is discussed in [28].

For the initialization of Algorithm 1 a certain regularization of the lower level problems is suggested in [3], which we will use for a different purpose in the sequel. In fact, for $i \in I$ consider the log-barrier problems

$$
\left(Q^i_\tau(x)\right) \quad \max_y b^i_\tau(x, y),
$$

where

$$b_\tau^i(x,y) = g_i(x,y) + \tau^2 \sum_{l=1}^{s} \log(-v_l(x,y))$$

denotes the log-barrier function of $Q^i(x)$. A necessary and sufficient optimality condition for $Q_\tau^i(x)$ is

$$0 = D_y b_\tau^i(x,y) = D_y g_i(x,y) + \sum_{l=1}^{s} \frac{\tau^2}{v_l(x,y)} D_y v_l(x,y). \tag{15}$$

By comparing this equation with the equality constraints of P_τ, it is not hard to obtain the following result.

Lemma 2.1 [3, Lemma 5.1.6(i)] *Let $i \in I$ and $\tau > 0$. The point y^i solves $Q_\tau^i(x)$ if and only if (y^i, γ^i), with*

$$\gamma_l^i = -\frac{\tau^2}{v_l(x,y^i)}, \quad l \in L, \tag{16}$$

satisfies (13), (14).

From [3, Theorem 5.2.8(iii)] it is known that for $\tau > 0$ the values of $\varphi_i(x)$, $i \in I$, are *under-estimated* by the optimal values $\varphi_\tau^i(x)$ of the lower level log-barrier problems $Q_\tau^i(x)$. This leads to a remarkable property of Algorithm 1, *the outer approximation property*.

Proposition 2.1 [3, Proposition 5.2.9] *Let $M_\tau = \mathrm{proj}_x(P_\tau)$. Then*:

(i) *For all $0 < \tau_1 < \tau_2$, $M_{\tau_1} \subset M_{\tau_2}$.*
(ii) *For all $\tau > 0$, $M \subset M_\tau$.*

A negative effect of the outer approximation property is that solutions of P_τ can be infeasible for GSIP for all $\tau > 0$, although the infeasibility vanishes in the limit. Clearly, this is a serious drawback when feasibility is a crucial issue.

3 Upper Bound on the Regularization Error

In this section, we quantify the regularization error by giving upper bounds of the lower level optimal value functions $\varphi_i(x)$. The regularization error results from the fact that P_τ is not equivalent to GSIP for $\tau > 0$. In fact, even given a feasible solution $(x, y^1, \gamma^1, \ldots, y^r, \gamma^r)$ of P_τ, $\tau > 0$, x is not necessarily feasible for GSIP.

Recall that the quantities $\varphi_i(x) = \max_{y \in Y(x)} g_i(x,y)$ play an important role for the feasibility of x for GSIP via the relationship

$$M = \{x \in \mathbb{R}^n \mid \varphi_i(x) \le 0,\ i \in I\}. \tag{17}$$

Based on this observation, we define approximate feasibility for GSIP as following.

Definition 3.1 We call x ε-feasible for GSIP if

$$\max_{i\in I}\varphi_i(x)=\max_{i\in I}\max_{y\in Y(x)}g_i(x,y)\le\varepsilon. \tag{18}$$

Consider a feasible point $(x,y^1,\gamma^1,\ldots,y^r,\gamma^r)$ of P_τ. In view of (18), we are interested in finding upper bounds of $\varphi_i(x)$, $i\in I$, that is, *over-estimates of* $\max_{y\in Y(x)}g_i(x,y)$. To this end, we rewrite the well-known weak duality theorem in terms of the lower level problems $Q^i(x)$. For a different solution approach to generalized semi-infinite programs via lower level duality see [29].

Theorem 3.1 [30, Theorem 6.2.1] *For all x and for all $i\in I$, the inequality*

$$\max_{y\in Y(x)}g_i(x,y)\le\min_{\gamma\ge0}\theta_i(x,\gamma) \tag{19}$$

holds, where

$$\theta_i(x,\gamma)=\max_y\{g_i(x,y)-\gamma^T v(x,y)\}=\max_y\mathcal{L}_i(x,y,\gamma). \tag{20}$$

The following sub-optimality estimate has been used in [30, pp. 510–513] in the formulation of the primal-dual path-following algorithm for linear programs. The validity of the result in the non-linear convex case has been shown, for example, in [31, 32]. We give a short proof in our notation for the sake of completeness.

Lemma 3.1 *For $\tau>0$ and $i\in I$, let $y_\tau^i(x)$ be a solution of $Q_\tau^i(x)$. Then,*

$$\varphi_i(x)=\max_{y\in Y(x)}g_i(x,y)\le g_i(x,y_\tau^i(x))+s\tau^2, \tag{21}$$

where $s=|L|$ denotes the number of lower level constraints v_l, $l\in L$.

Proof Let y^i be a solution of $Q_\tau^i(x)$. By Lemma 2.1, (y^i,γ^i) with

$$\gamma_l^i=-\frac{\tau^2}{v_l(x,y^i)},\quad l\in L, \tag{22}$$

solves (13), (14). Since $\gamma^i\ge0$, $\mathcal{L}_i(x,\cdot,\gamma^i)$ is concave and we can conclude from (13) that y^i is a global maximizer of $\mathcal{L}_i(x,\cdot,\gamma^i)$. By (20), this means

$$\theta_i(x,\gamma^i)=\mathcal{L}_i(x,y^i,\gamma^i). \tag{23}$$

Inserting (22) into the definition of $\mathcal{L}_i$ yields

$$\mathcal{L}_i(x,y^i,\gamma^i)=g_i(x,y^i)+s\tau^2. \tag{24}$$

From Theorem 3.1, (23) and (24), it follows that

$$\varphi_i(x) \le \theta_i(x, \gamma^i) = \mathcal{L}_i(x, y^i, \gamma^i) = g(x, y^i) + s\tau^2$$

for all $i \in I$. □

Applying Lemma 3.1 to each lower level problem and using (12) yields the following result.

Theorem 3.2 Regularization Error. *Let* $(x, y^1, \gamma^1, \ldots, y^r, \gamma^r)$ *be feasible for* P_τ. *Then* x *is* $s\tau^2$*-feasible for GSIP.*

Note that the approximation error in Theorem 3.2 *only* depends on τ and the number s of lower level constraints, but not on any entry of the variable vector $(x, y^1, \gamma^1, \ldots, y^r, \gamma^r)$. Thus the approximation error is *a-priori known*.

4 Inner Approximation of the Feasible Set

In this section, we use the upper bound derived in Lemma 3.1 to formulate a parametric *inner* approximation of M. This will significantly improve the original method described in Sect. 2 which, as discussed, only exhibits an *outer* approximation property.

Lemma 3.1 gives *over-estimates* of the optimal objective values $\varphi_i(x)$, $i \in I$, which are easy to compute for each lower level problem $Q^i(x)$. Also, the duality gap is *a-priori known* and neither depends on x nor on the auxiliary variables y^i, γ^i, $i \in I$. It is therefore natural to replace the original upper level constraints by the over-estimates *in the definition* of M. Recall that, for each $\tau > 0$, Lemma 3.1 yields

$$\varphi_i(x) \le g_i(x, y_\tau^i(x)) + s\tau^2, \quad i \in I, \tag{25}$$

whenever $y_\tau^i(x)$ solves $Q_\tau^i(x)$. This leads to the following definition of a parametric approximation of M.

Definition 4.1 We call

$$M_\tau^{\mathrm{feas}} = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{l} g_i(x, y_\tau^i(x)) + s\tau^2 \le 0, \\ y_\tau^i(x) \text{ solves } Q_\tau^i(x),\ i \in I \end{array} \right\} \tag{26}$$

the regularized inner approximation of M.

Following the lines of the proof of [3, Lemma 5.2.7], it is easily seen that M_τ^{feas} is a closed set under our regularity assumptions. The following result shows that M_τ^{feas} is in fact *an inner approximation* of M for all $\tau > 0$.

Theorem 4.1 Inner Approximation Property. *For all* $\tau > 0$,

$$M_\tau^{\mathrm{feas}} \subset M. \tag{27}$$

Proof If $M_\tau^{\text{feas}} = \emptyset$, the statement is trivial. Let $x \in M_\tau^{\text{feas}}$. Then, there exist solutions $y_\tau^i(x)$ of $Q_\tau^i(x)$, $i \in I$. Now, by Lemma 3.1,

$$\varphi_i(x) = \max_{y \in Y(x)} g_i(x, y) \leq g_i(x, y_\tau^i(x)) + s\tau^2 \leq 0$$

for all $i \in I$. Hence $x \in M$. □

The statement of Theorem 4.1 also includes the case $M_\tau^{\text{feas}} = \emptyset$, leading to the natural question under which conditions M_τ^{feas} is nonempty. The subsequent Lemma 4.1 provides a sufficient condition as well as a threshold value of τ. For its proof recall the definition of $C(x)$ from Sect. 1 and let

$$C(x)_\tau = \{y \in \mathbb{R}^m \mid g_i(x, y) + s\tau^2 \leq 0,\ i \in I\}$$

be the subset defined by the strengthened constraints. It is clear that for $0 \leq \tau_1 \leq \tau_2$ the inclusion

$$C(x)_{\tau_2} \subseteq C(x)_{\tau_1} \tag{28}$$

holds. Now $M_\tau^{\text{feas}} = \emptyset$ may hold, for instance, if $C(x)_\tau$ is void due to a large value of τ.

Lemma 4.1 *Let x be a Slater point of GSIP, that is, let*

$$\varphi_i(x) < 0, \quad i \in I. \tag{29}$$

Then, there exists a $\tau_0 > 0$, depending on x, such that $x \in M_\tau^{\text{feas}}$ for all $\tau \leq \tau_0$.

Proof Let x be a Slater point of M. Let

$$\sigma = \max_{i \in I} \varphi_i(x) < 0$$

denote the minimal feasibility of x. Define

$$\tau_0 = \sqrt{-\frac{\sigma}{s}} > 0. \tag{30}$$

Then, from

$$\varphi_i(x) + s\tau_0^2 \leq \sigma + s\tau_0^2 = 0, \quad i \in I,$$

it follows that $Y(x) \subseteq C(x)_{\tau_0}$. Now let $\tau \leq \tau_0$. By (28) we have

$$Y(x) \subseteq C(x)_{\tau_0} \subseteq C(x)_\tau. \tag{31}$$

For any solution $y_\tau^i(x)$ of $Q_\tau^i(x)$, we know that

$$y_\tau^i(x) \in Y(x) \subseteq C(x)_\tau, \tag{32}$$

and thus $x \in M_\tau^{\text{feas}}$. □

A combination of the inner approximation property from Theorem 4.1 with the outer approximation property from [3, Proposition 5.2.9] has the following consequence.

Corollary 4.1 *Let $\tau > 0$. Then*

$$M_\tau^{\text{feas}} \subset M \subset M_\tau. \tag{33}$$

Taking the limits in (33) yields a 'sandwiching' result.

Corollary 4.2 *Let $\{\tau_k\}_k$ be a zero sequence, $\tau_k > 0, k \in \mathbb{N}$. Let Sl_{GSIP} denote the set of Slater points of GSIP. Then,*

$$Sl_{GSIP} \subseteq \bigcup_{k=1}^{\infty} M_{\tau_k}^{\text{feas}} \subseteq M \subseteq \bigcap_{k=1}^{\infty} M_{\tau_k}. \tag{34}$$

Proof The first inclusion follows from Lemma 4.1, the second from Theorem 4.1, and the third is given in [3, Corollary 5.2.10]. □

Example 4.1 We consider the following disjunctive programming problem taken from [3]:

$$\begin{aligned} \text{(DP)} \qquad & \min_{x \in \mathbb{R}^2} \quad f(x), \\ & \text{s.t.} \quad y \le 0, \qquad y \in Y(x), \\ & Y(x) = \left\{ y \in \mathbb{R} \;\middle|\; \begin{array}{l} y - (x_1 - 1)^2 - x_2^2 + 1 \le 0, \\ y - x_1^2 - (x_2 - 1)^2 + 1 \le 0 \end{array} \right\}. \end{aligned}$$

The feasible set M_{DP} is easily seen to be the *union* of two unit circles centered at $(1, 0)$ and $(0, 1)$, respectively, as depicted in Fig. 1. Figures 2 and 3 illustrate the sandwiching property stated in Corollary 4.2.

Note that the 'most interior' points in M_{DP} are $x^1 = (0, 1)^T, x^2 = (1, 0)^T$ with $\max_{y \in Y(x^1)} g(x^1, y) = \max_{y \in Y(x^2)} g(x^2, y) = -1$. The threshold value of τ_0 computed according to (30) is $1/\sqrt{2} < 1$. However, as can be seen from Fig. 2, also M_1^{feas} is non-void, and M_1^{feas} contains both x^1 and x^2. This shows that the threshold value for τ from Lemma 4.1 is not necessarily best possible.

We remark that the sandwiching result in (33) may significantly improve termination criteria, depending on the problem structure. In fact, for given $\tau > 0$ the objective value of any point in M_τ^{feas} is an upper bound for the optimal value of GSIP, while the globally minimal value of f on M_τ is a corresponding lower bound. Hence, in cases where the latter globally minimal value is numerically available, the gap between the bounds may enter a termination criterion. Alternatively, in such cases the optimal value computed by the inner approximation may a-posteriori be endowed with a certificate for global optimality (up to some tolerance).

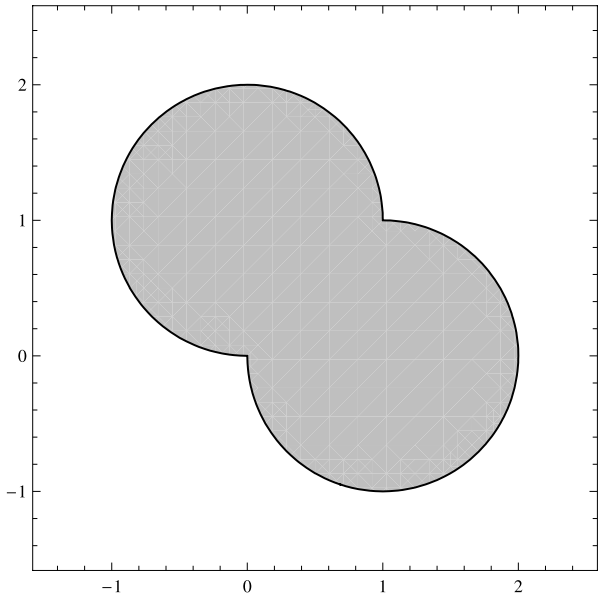


Fig. 1 The original feasible set M_{DP}

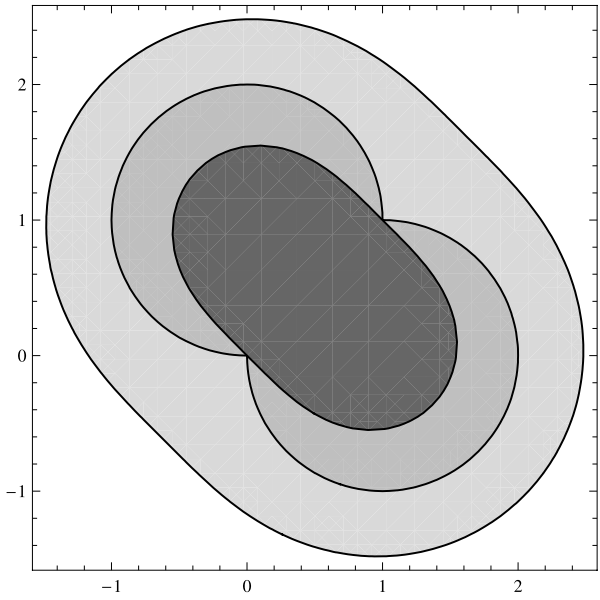
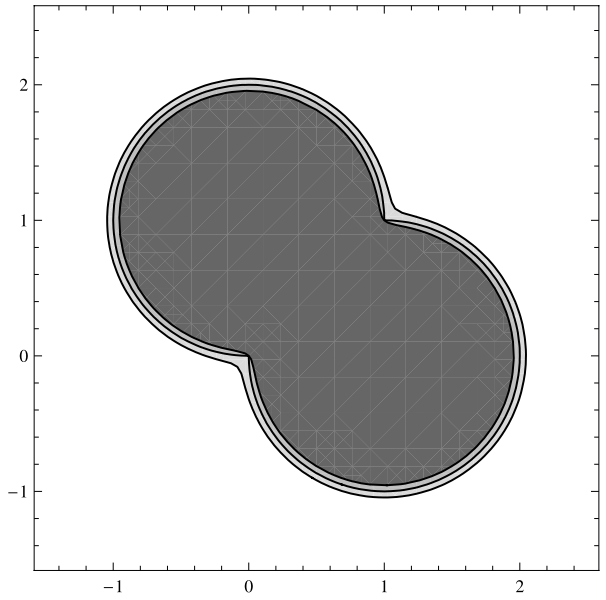


Fig. 2 Sandwiching of M_{DP} between M_1^{feas} and M_1

5 A Two-Phase Algorithm for Generalized Semi-Infinite Programming

In this section, we algorithmically exploit the inner approximation of M introduced in Sect. 4. The result is a two-phase *feasible numerical GSIP solver*. First, we replace the feasible set of P_τ according to (26) and obtain a parameterized family of *regularized*

Fig. 3 $M_{0.3}^{\text{feas}} \subset M_{DP} \subset M_{0.3}$



feasible problems $\mathrm{P}_\tau^{\text{feas}}$.

$$
\begin{aligned}
(\mathrm{P}_\tau^{\text{feas}}) \qquad & \min_{x,y^1,\gamma^1,\ldots,y^r,\gamma^r} && f(x), \\
& \text{s.t.} && g_i(x,y^i)+s\tau^2 \le 0, \\
& && D_y\mathcal{L}_i(x,y^i,\gamma^i)=0, \\
& && \Psi_\tau(\gamma^i,-v(x,y^i))=0, \quad i\in I.
\end{aligned}
$$

Analogously to Algorithm 1, we want to obtain an optimal solution of GSIP by solving $\mathrm{P}_{\tau_k}^{\text{feas}}$ for a zero sequence $\{\tau_k\}_{k\in\mathbb{N}}$. Unfortunately, for large values of τ, $\mathrm{P}_\tau^{\text{feas}}$ might not possess a feasible point. We thus formulate a parameterized family of Phase-1 problems $\mathrm{P1}_\tau^{\text{feas}}$ to determine a suitable value of τ together with a Slater point of GSIP, whenever such a point exists. For this purpose, we introduce an auxiliary variable z and consider the problem

$$
\begin{aligned}
(\mathrm{P1}_\tau^{\text{feas}}) \qquad & \min_{x,y^1,\gamma^1,\ldots,y^r,\gamma^r,z} && z, \\
& \text{s.t.} && z \ge -\varepsilon, && (35)\\
& && g_i(x,y^i)-z+2s\tau^2 \le 0, && (36)\\
& && D_y\mathcal{L}_i(x,y^i,\gamma^i)=0, && (37)\\
& && \Psi_\tau(\gamma^i,-v(x,y^i))=0, \quad i\in I, && (38)
\end{aligned}
$$

where the parameter $\varepsilon > 0$ is used to avoid unboundedness of $\mathrm{P1}_\tau^{\text{feas}}$. It is clear that $\mathrm{P1}_\tau^{\text{feas}}$ can be initialized in a feasible manner for any $\tau_0 > 0$ and for any $x^0 \in \mathbb{R}^n$ by

solving the log-barrier problems $Q^i_{\tau_0}(x^0)$ and by setting

$$z = \max\{-\varepsilon,\ 2s\tau_0^2 + \max_{i\in I} g_i(x, y^i_{\tau_0}(x^0))\}.$$

Let $z(\tau)$ denote the optimal value of $\mathrm{P1}^{\mathrm{feas}}_\tau$. Moreover, let $M_{P1^{\mathrm{feas}}_\tau}$ denote the feasible set of $\mathrm{P1}^{\mathrm{feas}}_\tau$. We now establish a connection between the set of Slater points of GSIP and $M_{P1^{\mathrm{feas}}_\tau}$.

Lemma 5.1 *The following statements are equivalent*:

(i) *x is a Slater point of GSIP.*
(ii) *There exist $\tau_0 > 0$, $z \le 0$ and $y^1, \ldots, y^r, \gamma^1, \ldots, \gamma^r$ such that*

$$(x, y^1, \ldots, y^r, \gamma^1, \ldots, \gamma^r, z) \in M_{P1^{\mathrm{feas}}_\tau}$$

for all $\tau \le \tau_0$.

Proof Let x be a Slater point of GSIP. Define

$$\tau_0 = \sqrt{-\frac{\max_{i\in I}\varphi_i(x)}{2s}} > 0. \tag{39}$$

Let $\tau \le \tau_0$. Then the solutions y^i of $Q^i_\tau(x)$ and the corresponding multipliers γ^i satisfy (37) and (38). Furthermore, set $z = 0$. Then,

$$\begin{aligned} g_i(x, y^i) - z + 2s\tau^2 &\le g_i(x, y^i) - z + 2s\tau_0^2 && (40)\\ &= g_i(x, y^i) + 2s\tau_0^2 && (41)\\ &= g_i(x, y^i) - \max_{i\in I}\varphi_i(x) && (42)\\ &\le 0, \quad i \in I. && (43) \end{aligned}$$

Thus also (35) and (36) are satisfied by $(x, y^1, \gamma^1, \ldots, y^r, \gamma^r, z)$.

For the opposite direction, let $(x, y^1, \ldots, y^r, \gamma^1, \ldots, \gamma^r, z)$ with $z \le 0$ be in $M_{P1^{\mathrm{feas}}_\tau}$ for all $\tau \le \tau_0$. By Lemma 3.1, $\varphi_i(x) \le g_i(x, y^i) + s\tau_0^2$, $i \in I$. By (36), this implies

$$\varphi_i(x) \le z - s\tau_0^2 \le -s\tau_0^2 < 0,$$

for all $i \in I$, which shows that x is a Slater point of GSIP. □

Corollary 5.1 *Let $\{\tau_k\}_{k\in\mathbb{N}}$ be a sequence with $\tau_k > 0, k \in \mathbb{N}$ and $\lim_{k\to\infty}\tau_k = 0$. If $Sl_{GSIP} \neq \emptyset$, then there exists a $k_0 \in \mathbb{N}$ such that the following assertions hold for all $k \ge k_0$:*

(i) $z(\tau_k) \le 0$.
(ii) *The x-component of any optimal solution of $P^{\mathrm{feas}}_{\tau_k}$ is a Slater point of GSIP.*

Proof Let $x \in Sl_{GSIP}$. By Lemma 5.1 there is a $\tau_0 > 0$, $z \leq 0$ and $y^i, \gamma^i, i \in I$ such that for all $\tau \leq \tau_0$ $(x, y^1, \gamma^1, \ldots, y^r, \gamma^r, z) \in M_{P1_\tau^{\text{feas}}}$. Since $\lim_{k\to\infty} \tau_k = 0$, we can choose an index k_0 such that $0 < \tau_k \leq \tau_0$ for all $k \geq k_0$. Now the first assertion follows from $z(\tau_k) \leq z \leq 0$. The second assertion also immediately follows from Lemma 5.1. □

The transition from Corollary 5.1 to a Phase-1 method is straightforward and summarized in Algorithm 2.

Algorithm 2 (Phase-1 Method)

Step 1. Choose a sequence $\{\tau_k\}_{k\in\mathbb{N}}$, $\tau_k > 0$, $\lim_{k\to\infty} \tau_k = 0$, a starting point $x^0 \in \mathbb{R}^n$ and a tolerance $\varepsilon > 0$. Compute a starting point $(x^{0,0}, y^{1,0,0}, \gamma^{1,0,0}, \ldots, y^{r,0,0}, \gamma^{r,0,0}, z^{0,0})$ of $\text{P1}_{\tau_0}^{\text{feas}}$ and set $k = 0$.

Step 2. Find an optimal solution $(x^{k,*}, y^{1,k,*}, \gamma^{1,k,*}, \ldots, y^{r,k,*}, \gamma^{r,k,*}, z^{k,*})$ of $\text{P1}_{\tau_k}^{\text{feas}}$ using $(x^{k,0}, y^{1,k,0}, \gamma^{1,k,0}, \ldots, y^{r,k,0}, \gamma^{r,k,0}, z^{k,0})$ as the starting point.

Step 3. If $z^{k,0} > 0$ and $\tau_k > \varepsilon$, then set $(x^{k+1,0}, y^{1,k+1,0}, \ldots, \gamma^{r,k+1,0}, z^{k+1,0}) = (x^{k,*}, y^{1,k,*}, \ldots, \gamma^{l,k,*}, z^{k,*})$, replace k by $k + 1$ and go to step 2.

Step 4. If $z^{k,0} \leq 0$, then terminate with $k_0 = k$ and $(x^{k,0}, y^{1,k,0}, \gamma^{1,k,0}, \ldots, y^{r,k,0}, \gamma^{r,k,0}, z^{k_0,0})$. Else terminate with failure, GSIP does not possess Slater points, or ε is too large.

Let $(x^{k_0,0}, y^{1,k_0,0}, \gamma^{1,k_0,0}, \ldots, y^{r,k_0,0}, \gamma^{r,k_0,0}, z^{k_0,0})$ be the solution of $\text{P1}_{\tau_{k_0}}^{\text{feas}}$ returned by Algorithm 2. Then

$$g_i(x, y^{i,k_0,0}) + s\tau_{k_0}^2 < g_i(x, y^{i,k_0,0}) + 2s\tau_{k_0}^2 \tag{44}$$

$$\leq z^{k_0,0} \tag{45}$$

$$\leq 0, \quad i \in I. \tag{46}$$

Thus, $(x^{k_0,0}, y^{1,k_0,0}, \gamma^{1,k_0,0}, \ldots, y^{r,k_0,0}, \gamma^{r,k_0,0}) \in M_{\tau_{k_0}}^{\text{feas}}$ is a feasible starting point for $\text{P}_{\tau_{k_0}}^{\text{feas}}$. The Phase-2 method (Algorithm 3) continues by solving $\text{P}_{\tau_k}^{\text{feas}}$ for $k \geq k_0$. We emphasize that, by Theorem 4.1, all iterates in Phase 2 are *feasible* with respect to the original problem GSIP. This essentially simplifies the termination criteria, since, in contrast to Algorithm 1, driving τ to zero is not necessary for feasibility anymore.

On the other hand, a straightforward analysis shows that the convergence results for Fritz John and Karush-Kuhn-Tucker points known for the outer approximation method described in Sect. 2 also hold for the present inner approximation approach.

Algorithm 3 (Feasible Method for GSIP)

Step 1. Choose a sequence $\{\tau_k\}_{k\in\mathbb{N}}$, $\tau_k > 0$, $\lim_{k\to\infty}\tau_k = 0$. Find an initial index k_0 and $(x^{k_0,0}, y^{1,k_0,0}, \gamma^{1,k_0,0}, \ldots, y^{r,k_0,0}, \gamma^{r,k_0,0}) \in M^{\text{feas}}_{\tau_{k_0}}$ by running Algorithm 2. Set $k = k_0$.

Step 2. Find an optimal solution $(x^{k,*}, y^{1,k,*}, \gamma^{1,k,*}, \ldots, y^{r,k,*}, \gamma^{r,k,*})$ of $\mathrm{P}^{\text{feas}}_{\tau_k}$ using $(x^{k,0}, y^{1,k,0}, \gamma^{1,k,0}, \ldots, y^{r,k,0}, \gamma^{r,k,0})$ as the starting point.

Step 3. If some termination criterion is satisfied, terminate with the optimal solution of GSIP $x^{k,0}$.

Step 4. Set $(x^{k+1,0}, y^{1,k+1,0}, \ldots, \gamma^{r,k+1,0}) = (x^{k,*}, y^{1,k,*}, \ldots, \gamma^{r,k,*})$, replace k by $k+1$ and go to step 2.

Example 5.1 We illustrate the two-phase approach with a design centering problem discussed in [3, Sect. 6.1]. The container C is defined by

$$C = \left\{ y \in \mathbb{R}^2 \;\middle|\; \begin{pmatrix} -y_1 - y_2^2 \\ 0.25y_1 + y_2 - 0.75 \\ -y_2 \end{pmatrix} \le \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}, \tag{47}$$

and shown in Fig. 4. As a design, an ellipse

$$Y(A, v) = \left\{ y \in \mathbb{R}^2 \,\middle|\, \frac{1}{2}(y - v)^T (AA^T)^{-1}(y - v) - 1 \le 0 \right\}$$

is taken, where the parameters are the center v and the matrix $A \in \mathbb{R}^{2\times 2}$. The area of $Y(A, v)$ is easily seen to be $\pi(\det(A))$, and the area is chosen as the objective function to be maximized. Summarizing, the corresponding GSIP has the form

$$\begin{aligned} \text{(EL)} \qquad & \min_{A,v} \quad -\pi(\det(A)), \\ & \text{s.t.} \quad\; Y(A, v) \subseteq C. \end{aligned}$$

The optimal value of EL is known to be approximately -3.7234. We solved EL by the two-phase approach (Algorithms 2 and 3) as well as by the classical approach (Algorithm 1) in order to compare the intermediate solutions. The sequence $\{\tau_k\}_k$ used to generate the results is $\tau_k = 1.001 - \frac{k}{50}$, and we aborted the method after 50 iterations, that is, at $\tau = 0.001$.

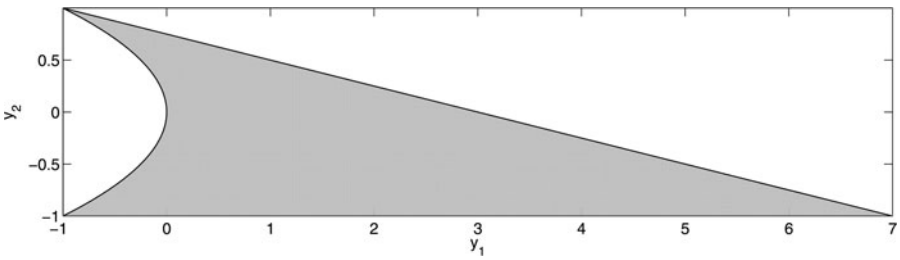


Fig. 4 The two-dimensional test container

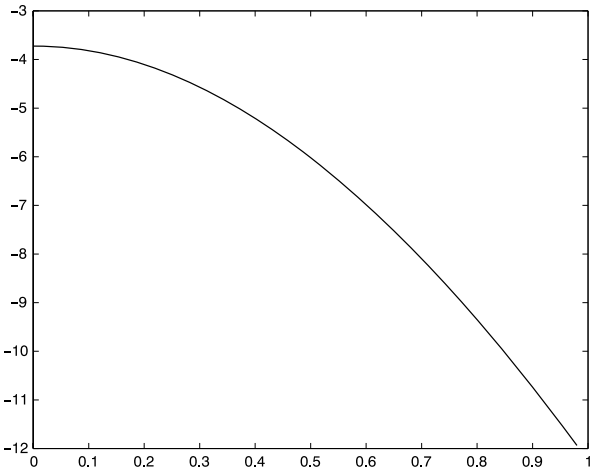


Fig. 5 Optimal value of P_τ for different values of τ

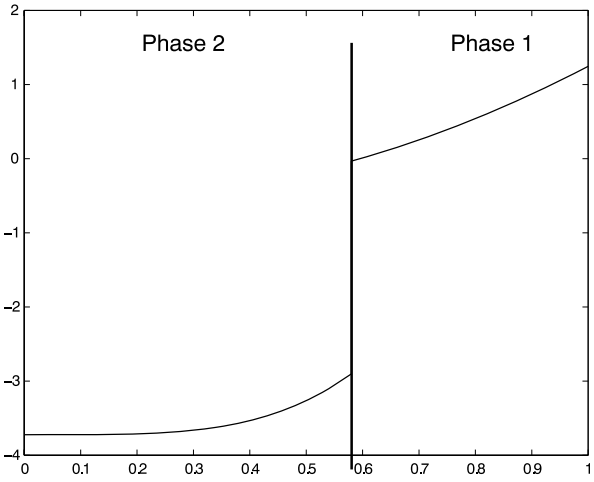


Fig. 6 Optimal value of P_τ^{feas} for different values of τ

Figure 5 shows the optimal value of P_τ as a function of τ. As expected [3, Proposition 5.2.6], it converges quadratically to the optimal objective value of EL. For the two-phase approach, the optimal value $z(\tau)$ of $P1_\tau^{\text{feas}}$ and, in Phase 2, the optimal value of P_τ^{feas} are shown in Fig. 6.

Algorithm 2 terminates after 21 iterations (with $\tau_{21} \approx 0.58$) with a Slater point of EL. From this point, Algorithm 3 continues the iterations, and the optimal value of P_τ^{feas} is strictly decreasing as τ is decreasing. This is a strong indicator that, despite the problematic discussed below in Sect. 7, for $\tau_1 \leq \tau_2$, $M_{\tau_2}^{\text{feas}} \subseteq M_{\tau_1}^{\text{feas}}$ holds for this problem.

Let us now compare some solutions of the regularized problems P_τ with those of P_τ^{feas}. Figures 7 and 8 show the solutions of $P_{0.5}$ and $P_{0.5}^{\text{feas}}$ respectively. The optimal solutions of $P_{0.2}$ and $P_{0.2}^{\text{feas}}$ are shown in Figs. 9 and 10. We note that, besides the advantage of being feasible for EL, the solution generated by Algorithm 3 in both

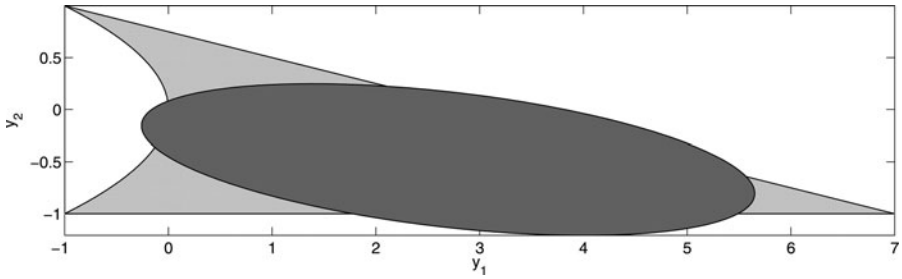


Fig. 7 Solution of $P_{0.5}$, objective value: −6.02

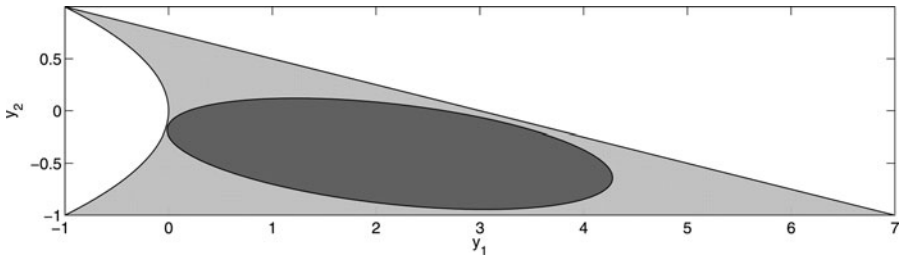


Fig. 8 Solution of $P_{0.5}^{\text{feas}}$, objective value: −3.26

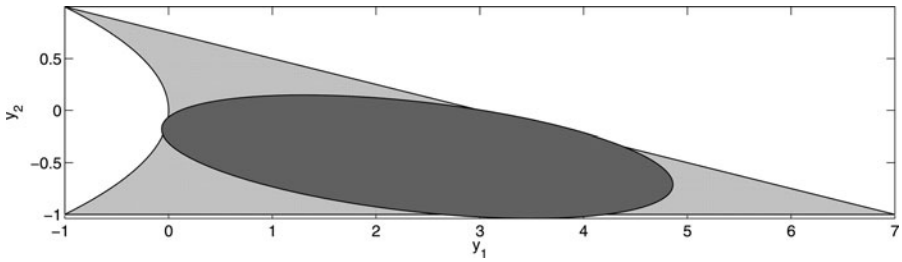


Fig. 9 Solution of $P_{0.2}$, objective value: −4.1

cases yields a better estimate of the optimal value of EL than the iterate generated by Algorithm 1.

6 Consequence for Standard Semi-Infinite Programming

Whereas in generalized semi-infinite optimization a number of real-life applications with convex lower level problems exist, the convexity assumption on the lower level problem fails for most applications of *standard* semi-infinite optimization. The recent article [2], however, presents the so-called *Adaptive Convexification Algorithm* which solves a sequence of auxiliary semi-infinite problems with convex lower levels, using the previously discussed techniques from [1, 3], to approximate a stationary point of the original problem. The convexifications are produced using the ideas of the

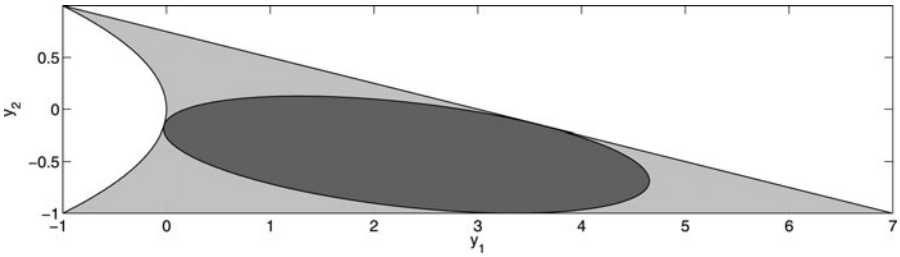


Fig. 10 Solution of $P^{\text{feas}}_{0.2}$, objective value: −3.71

αBB method for global optimization [33–36]. As αBB allows to construct concave overestimators of g, all iterates of this method are *feasible* for the original standard semi-infinite problem.

To briefly review the main idea, for simplicity consider the standard semi-infinite problem

$$\begin{aligned} \text{(SIP)} \qquad & \min_{x\in X} \quad f(x), \\ & \text{s.t.} \quad\ g(x,y)\le 0, \quad y\in Y=[0,1] \end{aligned}$$

with objective function $f\in C^2(\mathbb{R}^n,\mathbb{R})$ and constraint function $g\in C^2(\mathbb{R}^n\times\mathbb{R},\mathbb{R})$. For a convex lower level problem

$$\begin{aligned} (Q(x)) \qquad & \max_y \quad g(x,y), \\ & \text{s.t.} \quad\ y\in[0,1], \end{aligned}$$

that is, if $g(x,\cdot)$ is concave on Y for all x, the ideas of Sect. 2 lead to the equivalent reformulation of SIP as the MPEC

$$\begin{aligned} \text{(P)} \qquad & \min_{x,y,\gamma_\ell,\gamma_u} \quad f(x), \\ & \text{s.t.} \quad\ g(x,y)\le 0, \\ & \qquad\quad \nabla_y g(x,y)+\gamma_\ell-\gamma_u=0, \\ & \qquad\quad \psi(\gamma_\ell,y)=0, \\ & \qquad\quad \psi(\gamma_u,1-y)=0, \end{aligned}$$

where the straightforward description $Y=\{y\in\mathbb{R}\,|\,y\ge 0,\ \ y\le 1\}$ is used, and ψ denotes some NCP function.

If the lower level problem is *not* convex, one may convexify it with ideas from the αBB method. In fact, [2] suggests to replace $g(x,y)$ by a function $\tilde{g}(x,y)=g(x,y)+\frac{\alpha}{2}y(1-y)$ where $\alpha\ge 0$ is chosen so large that $\tilde{g}(x,\cdot)$ is a concave overestimator of $g(x,\cdot)$ on the interval $Y=[0,1]$. The latter choice of α can be performed

via techniques of interval analysis. The problem P is then replaced by

$$
\begin{array}{lll}
(\tilde{\mathrm{P}}) & \min\limits_{x,y,\gamma_\ell,\gamma_u} & f(x), \\
& \text{s.t.} & \tilde{g}(x,y) \le 0, \\
& & \nabla_y \tilde{g}(x,y) + \gamma_\ell - \gamma_u = 0, \\
& & \psi(\gamma_\ell, y) = 0, \\
& & \psi(\gamma_u, 1-y) = 0.
\end{array}
$$

The crucial idea of this approach is the fact that for any feasible point $(x, y, \gamma_\ell, \gamma_u)$ of $(\tilde{\mathrm{P}})$ the variable x is feasible for SIP.

The lower level overestimators are subsequently tightened by an adaptive refinement of Y which guarantees convergence of stationary points of the approximating problems to a stationary point of SIP under mild assumptions (cf. [2] for details). As the above arguments also hold for all iterates of the method, these iterates are feasible for SIP.

Although the numerical results for the adaptive convexification algorithm in [2] are very promising, one may well argue that a feasible point of $(\tilde{\mathrm{P}})$ has to be determined to a certain accuracy to guarantee feasibility of its x-part for SIP. Taking the presence of NCP functions into account, this may be a non-trivial task. In fact, the solution approach presented in [2] approximates $(\tilde{\mathrm{P}})$ with the techniques of Sect. 2, that is, the NCP functions are regularized, and a sequence of smooth problems

$$
\begin{array}{lll}
(\tilde{\mathrm{P}}_\tau) & \min\limits_{x,y,\gamma_\ell,\gamma_u} & f(x), \\
& \text{s.t.} & \tilde{g}(x,y) \le 0, \\
& & \nabla_y \tilde{g}(x,y) + \gamma_\ell - \gamma_u = 0, \\
& & \psi_\tau(\gamma_\ell, y) = 0, \\
& & \psi_\tau(\gamma_u, 1-y) = 0,
\end{array}
$$

with $\tau \searrow 0$ is solved. However, for the lower level optimal values this approach has a serious drawback. Let $\tilde{\varphi}(x)$ denote the optimal value of the lower level problem

$$
\begin{array}{lll}
(\tilde{Q}(x)) & \max_y & \tilde{g}(x,y), \\
& \text{s.t.} & y \in [0,1],
\end{array}
$$

and $\tilde{\varphi}_\tau(x)$ the optimal value of the corresponding log-barrier problem for $\tau > 0$ as discussed in Sect. 2. Then the optimal value $\varphi(x)$ of the original lower level problem $Q(x)$ is first *over*estimated by $\tilde{\varphi}(x)$, but the approximation of $\tilde{\varphi}(x)$ by $\tilde{\varphi}_\tau(x)$ is an *under*estimation, that is, we have $\varphi(x) \le \tilde{\varphi}(x) \ge \tilde{\varphi}_\tau(x)$. This is due to the outer approximation property of the regularization approach from Sect. 2. Consequently, to guarantee feasible iterates for SIP, the approximating problems $\tilde{P}_\tau$ have to be solved for τ very close to zero.

If, on the other hand, the inner approximation ideas of the present paper are employed, instead of $\tilde{\mathrm{P}}_\tau$ one may solve the slightly modified problems

$$(\tilde{\mathrm{P}}_\tau^{\mathrm{feas}}) \qquad \begin{array}{ll} \min\limits_{x,y,\gamma_\ell,\gamma_u} & f(x), \\ \text{s.t.} & \tilde{g}(x,y) + 2\tau^2 \le 0, \\ & \nabla_y \tilde{g}(x,y) + \gamma_\ell - \gamma_u = 0, \\ & \psi_\tau(\gamma_\ell, y) = 0, \\ & \psi_\tau(\gamma_u, 1-y) = 0, \end{array}$$

with $\tau \searrow 0$. The corresponding approximation $\tilde{\varphi}_\tau^{\mathrm{feas}}(x)$ of the lower level optimal value then satisfies $\varphi(x) \le \varphi_\tau(x) \le \tilde{\varphi}_\tau^{\mathrm{feas}}(x)$, so that feasibility of x for a feasible point $(x, y, \gamma_\ell, \gamma_u)$ of $\tilde{\mathrm{P}}_\tau^{\mathrm{feas}}$ is guaranteed for *positive* values of τ. The following example from Chebyshev approximation illustrates this effect.

Example 6.1 [2] Let the function $\sin(\pi y)$ be approximated by a quadratic function on the interval $Y = [0, 1]$ in the Chebyshev norm, that is, we wish to solve

$$\begin{array}{ll} \text{(CA)} & \min_{x\in\mathbb{R}^3} \| \sin(\pi y) - (x_3 y^2 + x_2 y + x_1)\|_{\infty,[0,1]} \\ & \text{with} \\ & \|e(x,y)\|_{\infty,[0,1]} = \max_{y\in[0,1]} |e(x,y)|, \end{array}$$

where $e(x, y) = \sin(\pi y) - (x_3 y^2 + x_2 y + x_1)$ denotes the so-called error function. The semi-infinite reformulation of CA is

$$\begin{array}{lll} (\mathrm{SIP_{CA}}) & \min_{(x,z)\in\mathbb{R}^3\times\mathbb{R}} & z, \\ & \text{s.t.} & -z \le e(x,y) \le z, \quad y \in [0,1]. \end{array}$$

Figure 11 illustrates the error function $e(x, y)$ in some (suboptimal) iterate x, together with the convex relaxations on certain refinements of Y. Note that the approximations of $e(x, y)$ are depicted for both semi-infinite constraints, where in the figure one approximation appears as piecewise concave and the other as piecewise convex. Here feasibility means that the graph of the error function is contained in the rectangle $[0, 1] \times [-z, z]$, meaning that z is a valid upper bound for the Chebyshev approximation error in the solution point.

The value of the regularization parameter τ is below machine precision so that the Fig. 11 may be interpreted as a solution of $\tilde{\mathrm{P}} = \tilde{\mathrm{P}}_0$ (with a slight abuse of notation as $\tilde{\mathrm{P}}_0$ does not take the adaptive refinements of $Y = [0, 1]$ into account). Figure 12 shows the corresponding result for the positive value $\tau = 0.1$ under the outer approximation approach from [2]. Clearly, the graph of the error function leaves the rectangle $[0, 1] \times [-z, z]$ at the right endpoint of Y, so that the corresponding point (x, z) is infeasible for $\mathrm{SIP_{CA}}$. On the other hand, Fig. 13 illustrates that for $\tau = 0.1$ the inner approximation approach yields feasibility of (x, z).

Fig. 11 Solution of $\tilde{P}$

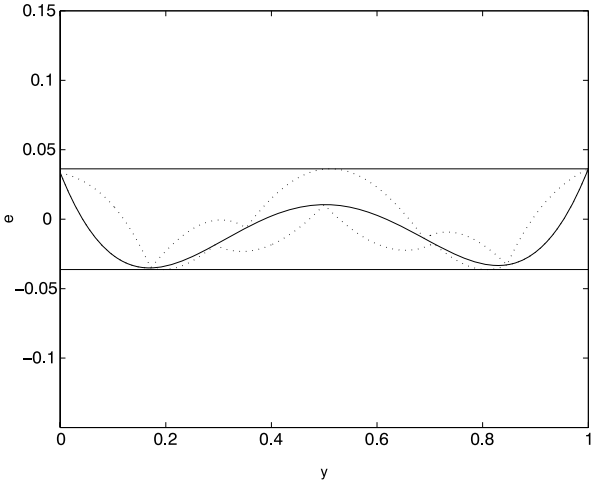
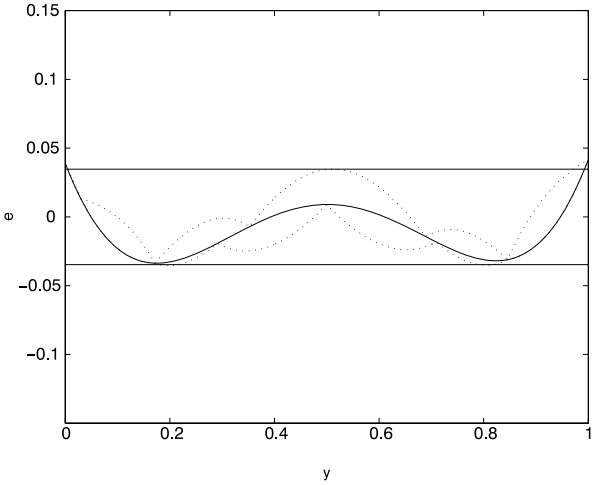


Fig. 12 Solution of $\tilde{P}_{0.1}$



7 Concluding Remarks and Open Questions

Regarding the phase 1 approach in Sect. 5 one might wonder why the inequality constraints g_i, $i \in I$, are strengthened by $2s\tau^2$ in the definition of $M_{P1_\tau^{\text{feas}}}$ rather than by $s\tau^2$ as in the case of M_τ^{feas}. In fact, if $(x, y^1, \gamma^1, \ldots, y^r, \gamma^r, z)$ with $z \leq 0$ satisfies

$$
\begin{aligned}
&g_i(x, y^i) - z + s\tau^2 \leq 0,\\
&D_y \mathcal{L}_i(x, y^i, \gamma^i) = 0,\\
&\Psi_\tau(\gamma^i, -v(x, y^i)) = 0, \quad i \in I,
\end{aligned}
$$

then x trivially belongs to M_τ^{feas}. On the other hand, if $z = 0$, then we cannot conclude that x is a Slater point of GSIP, and it is unclear whether $x \in M_\tau^{\text{feas}}$ for *all* $\tau \leq$

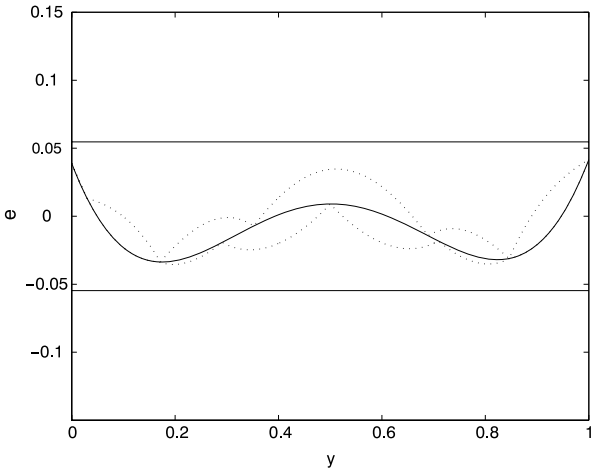


Fig. 13 Solution of $\tilde{P}_{0.1}^{\text{feas}}$

τ_0. In other words, it is unclear whether for $\tau_1 \leq \tau_2$ the inclusion $M_{\tau_2}^{\text{feas}} \subseteq M_{\tau_1}^{\text{feas}}$ holds. Recall that the converse relationship '$\tau_1 \leq \tau_2 \Rightarrow M_{\tau_1} \subseteq M_{\tau_2}$' is known for the outer-approximating sets M_τ (Proposition 2.1). The proof of this statement is largely based on the fact that, for $i \in I$, $g^i(x, y^i(x, \tau))$ is monotonously decreasing in τ. Unfortunately, this argument does *not* hold for the inner-approximating sets M_τ^{feas}. In fact, for $i \in I$, the inequality constraint $g^i(x, y^i(x, \tau)) + s\tau^2 \leq 0$ consists of the decreasing term $g^i(x, y^i(x, \tau))$, but it also contains the term $s\tau^2$, which is *increasing in* τ. It is unclear, which of the two effects dominates, and 'overtaking' situations cannot be a-priori excluded.

As we have seen in the proof of Lemma 4.1, the situation described above cannot occur if x is a Slater point of GSIP and τ is small enough. In fact, the Slater condition implies that the design $Y(x)$ is completely contained in the shrunk container $C(x)_{\tau_0}$ for some $\tau_0 > 0$. Therefore all lower level central paths and, in particular, all log-barrier maximizers $y^i(x, \tau)$, $i \in I$, are contained in $C(x)_\tau$ for $0 < \tau \leq \tau_0$. Of course, this is a rather restrictive assumption. A weaker requirement could be, e.g,

$$g_i(x, y^i(x, \tau_0)) + s\tau_0^2 \leq 0, \quad i \in I, \tag{48}$$

and

$$g_i(x, y^i(x, \tau)) \leq g_i(x, y^i(x, \tau_0)) + s(\tau_0^2 - \tau^2), \quad i \in I,\ \tau \in [0; \tau_0]. \tag{49}$$

Visually, this means that $x \in M_{\tau_0}^{\text{feas}}$, and x remains in M_τ^{feas} for all $\tau \leq \tau_0$, since 'the container boundary moves outwards faster than the log-barrier maximizers'. The reformulation of (49) in terms of the original problem data requires a deep understanding of the structure of the central path, and we leave it as an interesting open question for future research.

Another important point worth investigation is the question, under which assumptions M has a regular structure. For example, it is known [37, Theorem 7.1] that for a linear semi-infinite program the set of the Slater points and the topological interior of the feasible set coincide under the so called locally Farkas-Minkowski constraint

qualification. Generalizations of this result for GSIP, like via the Extended Mangasarian Fromovitz Constraint Qualification [4], are not only of theoretical interest, but would also have an immediate practical impact in view of the importance of Slater points for Algorithm 3.

References

1. Stein, O., Still, G.: Solving semi-infinite optimization problems with interior point techniques. SIAM J. Control Optim. **42**, 769–788 (2003)
2. Floudas, C.A., Stein, O.: The adaptive convexification algorithm: a feasible point method for semi-infinite programming. SIAM J. Optim. **18**, 1187–1208 (2007)
3. Stein, O.: Bi-level Strategies in Semi-infinite Programming. Kluwer Academic, Boston (2003)
4. Guerra Vazquez, F., Rückmann, J.-J., Stein, O., Still, G.: Generalized semi-infinite programming: a tutorial. J. Comput. Appl. Math. (2007). doi:10.1016/j.cam.2007.02.012
5. Stein, O.: A semi-infinite approach to design centering. In: Dempe, S., Kalashnikov, V. (eds.) Optimization with Multivalued Mappings, pp. 209–228. Springer, Berlin (2006)
6. Winterfeld, A.: Large-scale semi-infinite optimization applied to industrial gemstone cutting. PhD thesis, TU Kaiserslautern (2007)
7. Graettinger, T.J., Krogh, B.H.: The acceleration radius: a global performance measure for robotic manipulators. IEEE J. Robot. Autom. **4**, 60–69 (1988)
8. Horst, R., Tuy, H.: Global Optimization: Deterministic Approaches, 3rd edn. Springer, Heidelberg (1996)
9. Hogan, W.W.: Directional derivatives for extremal value functions with applications to the completely convex case. Oper. Res. **21**, 188–209 (1973)
10. Bhattacharjee, B., Lemonidis, P., Green, W.H., Barton, P.I.: Global solution of semi-infinite programs. Math. Program. **103**(2), 283–307 (2005)
11. Stein, O., Still, G.: On generalized semi-infinite optimization and bilevel optimization. Eur. J. Oper. Res. **142**, 444–462 (2002)
12. Shimizu, K., Ishizuka, Y., Bard, J.F.: Nondifferentiable and Two-Level Mathematical Programming. Kluwer Academic, Boston (1997)
13. Ferris, M.C., Pang, J.S.: Engineering and economic applications of complementarity problems. SIAM Rev. **39**(4), 669–713 (1997)
14. Dempe, S.: Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. Optimization **52**(3), 333–359 (2003)
15. Pieper, H.: Algorithms for mathematical programs with equilibrium constraints with applications to deregulated electricity markets. PhD thesis, Stanford University (2001)
16. Scheel, H., Scholtes, S.: Mathematical programs with complementarity constraints: stationarity, optimality, and sensitivity. Math. Oper. Res. **25**, 1–22 (2000)
17. Jongen, H.Th., Weber, G.-W.: Nonlinear optimization: Characterization of structural stability. J. Global Optim. **1**, 47–64 (1991)
18. Fletcher, R., Leyffer, S.: Numerical experience with solving mpecs as nlps. Technical Report NA/210, University of Dundee (2002)
19. Chen, C., Mangasarian, O.L.: Smoothing methods for convex inequalities and linear complementarity problems. Math. Program. **71**(1), 51–69 (1995)
20. Facchinei, F., Jiang, H., Qi, L.: A smoothing method for mathematical programs with equilibrium constraints. Math. Program. **85**(1), 107–134 (1999)
21. Jiang, H., Ralph, D.: Smooth SQP methods for mathematical programs with nonlinear complementarity constraints. SIAM J. Optim. **10**(3), 779–808 (2000)
22. Lin, G.-H., Fukushima, M.: A modified relaxation scheme for mathematical programs with complementarity constraints. Ann. Oper. Res. **133**(22), 63–84 (2005)
23. Yang, X.Q., Huang, X.X.: Lower-order penalty methods for mathematical programs with complementarity constraints. Optim. Methods Softw. **19**(6), 693–720 (2004)
24. Jiang, H., Ralph, D.: Extension of quasi-newton methods to mathematical programs with complementarity constraints. Comput. Optim. Appl. **25**(1–3), 123–150 (2003)
25. Stein, O.: Lifting mathematical programs with complementarity constraints. Math. Program. (2010, to appear)

26. Ferris, M.C., Kanzow, C.: Complementarity and related problems: A survey. In: Handbook of Applied Optimization, pp. 514–530 (2002)
27. Sun, D., Qi, L.: On NCP-functions. Comput. Optim. Appl. **13**, 201–220 (1999)
28. Stein, O.: On Karush-Kuhn-Tucker points for a smoothing method in semi-infinite optimization. J. Comput. Math. **24**, 719–732 (2006)
29. Levitin, E., Tichatschke, R.: A branch-and-bound approach for solving a class of generalized semi-infinite programming problems. J. Global Optim. **13**, 299–315 (1998)
30. Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear Programming: Theory and Algorithms, 2nd edn. Wiley, New York (1993)
31. Den Hertog, D., Roos, C., Terlaky, T.: On the classical logarithmic barrier function method for a class of smooth convex programming problems. J. Optim. Theory Appl. **73**(1), 1–25 (1992)
32. Izmailov, A.F., Solodov, M.V.: Smoothing methods for convex inequalities and linear complementarity problems. Lect. Not. Econ. Math. Syst. **563**, 133–145 (2006)
33. Adjiman, C.S., Androulakis, I.P., Floudas, C.A.: A global optimization method, αBB, for general twice-differentiable constrained NLPs—I: Theoretical advances. Comput. Chem. Eng. **22**, 1137–1158 (1998)
34. Adjiman, C.S., Androulakis, I.P., Floudas, C.A.: A global optimization method, αBB, for general twice-differentiable constrained NLPs—II: Implementation and computational results. Comput. Chem. Eng. **22**, 1159–1179 (1998)
35. Floudas, C.A.: Deterministic Global Optimization, Theory, Methods and Applications. Kluwer Academic, Dordrecht (2000)
36. Stein, O.: Adaptive convexification in semi-infinite optimization. In: Pardalos, P.M., Floudas, C.A. (eds.) Encyclopedia of Optimization, Part 1, 2nd edn., pp. 13–19. Springer, Berlin (2009)
37. Goberna, M.A., Lopez, M.A.: Linear Semi-Infinite Optimization. Wiley, Chichester (1998)