On the use of Augmented Lagrangians in the Solution of Generalized Semi-Infinite Min-Max Problems

E. Polak
University of California, Berkeley

J.O. Royset
Naval Postgraduate School

Abstract

We present an approach for the solution of a class of generalized semi-infinite optimization problems. Our approach uses augmented Lagrangians to transform generalized semi-infinite min-max problems into ordinary semi-infinite min-max problems, with the same set of local and global solutions as well as the same stationary points. Once the transformation is effected, the generalized semi-infinite min-max problems can be solved using any available semi-infinite optimization algorithm. We illustrate our approach with two numerical examples, one of which deals with structural design subject to reliability constraints.

1 Introduction

We consider the class of generalized semi-infinite min-max problems of the form

P $$\min_{x \in \mathbb{R}^n} \psi(x), \tag{1.1}$$

where $\psi : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$\psi(x) \stackrel{\triangle}{=} \max_{y \in Z(x)} \phi(x, y) \tag{1.2a}$$

where

$$Z(x) \stackrel{\triangle}{=} \{y \in \mathbb{R}^m \mid f(x, y) \leq 0,\ g(y) \leq 0\}, \tag{1.2b}$$

with $\phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{r_1}$, $g : \mathbb{R}^m \to \mathbb{R}^{r_2}$, and $v \leq 0$ meaning $v^1 \leq 0$, ..., $v^q \leq 0$, for any $v = (v^1, ..., v^q) \in \mathbb{R}^q$. We use superscripts to denote components of vectors.

It is the dependence of the set-valued map $Z(\cdot)$ on the design variable x that makes **P** a *generalized semi-infinite min-max problem.*

In the sequel we will need the set

$$Y \stackrel{\triangle}{=} \{y \in \mathbb{R}^m \mid g(y) \leq 0\}. \tag{1.2c}$$

Generalized semi-infinite min-max problems of the form $\mathbf{P}$ arise in various engineering applications. For example, optimal design of civil, mechanical, and aerospace structures is frequently considered in a probabilistic framework, where uncertainties in material properties, loads, and boundary conditions are taken into account. Let $x \in \mathbb{R}^n$ be a vector of deterministic design variables, e.g., physical dimensions of the structure, or parameters in the probability distribution of the random quantities. The probability of failure of a structure $p : \mathbb{R}^n \to [0,1]$ is defined by, see [6],

$$p(x) \triangleq \int_{\{y \in \mathbb{R}^m | h(x,y) \leq 0\}} \varphi(y) dy, \tag{1.3}$$

where $\varphi(\cdot)$ is the m-dimensional multi-variate standard normal probability density function, and $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is a smooth real-valued limit-state function.

The optimal design problem is typically in the form

$$\min_{x \in X} \{ \ c^0(x) + c^1(x) p(x) \ \}, \tag{1.4}$$

where $c^0 : \mathbb{R}^n \to \mathbb{R}$ is the initial cost of the structure and $c^1 : \mathbb{R}^n \to \mathbb{R}$ is the cost of structural failure. The evaluation of $p(\cdot)$ is computationally expensive. Hence a first-order approximation to the probability of failure is usually considered acceptable. Based on such approximations, it can be shown, see [24], that (1.4) can be approximated by

$$\min_{x \in X} \max_{y \in S(x)} \{ \ c^0(x) + c^1(x) \tilde{p}(x,y) \ \}, \tag{1.5}$$

where

$$\tilde{p}(x,y) \triangleq \Phi \left[\frac{\beta h(x,0)}{h(x,y) - h(x,0)} \right], \tag{1.6}$$

whenever $h(x,0) > 0$ and $h(x,y) - h(x,0) < 0$. If $h(x,0) > 0$ and $h(x,y) - h(x,0) \geq 0$, then $\tilde{p}(x,y) = 0$, and, finally, if $h(x,0) \leq 0$, then $\tilde{p}(x,y) = 1$ for all y. Above, $\Phi(\cdot)$ is the standard normal cumulative distribution function and

$$S(x) = \{ y \in \mathbb{R}^m \mid h(x,y) - h(x,0) \leq -\alpha, \|y\|^2 \leq \beta^2 \}, \tag{1.7}$$

with $\alpha, \beta > 0$. Hence, the optimal design problem (1.4) can be solved approximately by solving a generalized semi-infinite min-max problem in the form (1.1), with $f(x,y) = h(x,y) - h(x,0) + \alpha$ and $g(y) = \|y\|^2 - \beta^2$.

Theoretical results regarding the existence of and formulas for directional derivatives of generalized max-functions, such as the one in (1.2a), can be found in [3, 22]. First-order optimality conditions for generalized semi-infinite optimization problems are presented in [12, 26, 28, 29, 30, 33, 27].

In the unpublished paper [15], Levitin employs a differentiable penalty function to remove the constraints $f(x,y) \leq 0$, and shows that the sequence of global solutions of the penalized problem converges to a global solution of $\mathbf{P}$, as the penalty goes to infinity. In [33], it is shown that under the linear independence constraint qualification for the "inner problem" in (1.2a), a class of generalized semi-infinite optimization problems is equivalent to an ordinary semi-infinite optimization problem, i.e., a problem in the form $\min_{x \in \mathbb{R}^n} \{ f^0(x) \mid \phi(x, \omega) \leq 0,\ \omega \in \Omega \}$, with $\phi(\cdot, \cdot)$ smooth and Ω of infinite cardinality. However, it is not clear how to implement a procedure for constructing the equivalent problem.

There are only a few studies dealing with numerical methods for generalized semi-infinite optimization problems. Numerical methods for sub-classes of such problems, e.g., as arising in robotics, can be found in [9, 11, 13, 16]. In [30, 31] we find a conceptual algorithm for solving optimization problems with constraints in terms of generalized max-functions of the type (1.2a). In these papers it is assumed that the linear independence constraint qualification, second-order sufficient conditions, and strict complementary slackness, for the "inner-problem," in (1.2a) hold. The algorithm in [30, 31] applies a globally convergent Newton-type method to the Karush-Kuhn-Tucker system of equations for a locally reduced problem. A conceptual algorithm, based on discretization, is presented in [31]. In the recent monograph [27], an implementable algorithm for the class of generalized semi-infinite optimization problems with $\phi(x, \cdot)$ concave and $f^k(x, \cdot), g^k(\cdot)$ convex is presented. The original problem is shown to be equivalent to a Stackelberg game with inner problems replaceable by corresponding first-order optimality conditions. This leads to a sequence of finite nonlinear programming problems, which are solved by standard optimization algorithms. A review of generalized semi-infinite optimization can be found in [27].

Recently, we put forth the idea of using an exact penalty function to eliminate the inequalities in (1.2a) that depend on x, i.e., $f(x, y) \leq 0$ (see [25]). In [25], we used a standard nondifferentiable exact penalty function for this purpose. This resulted in an implementable algorithm for solving general forms of $\mathbf{P}$ under a calmness assumption. The selected approach led to an algorithm that generates sequences converging to weaker stationary points than the ones given in [28]. Moreover, the use of a nondifferentiable exact penalty function results in a semi-infinite min-max-min problem with an unknown penalty parameter and two other algorithm parameters, which are controlled by several precision-adjustment tests.

In this paper we explore an alternative approach where the inequalities in (1.2a) that depend on x, i.e., $f(x, y) \leq 0$, are eliminated by the use of an augmented Lagrangian exact penalty function. This approach requires stronger assumptions than the ones in [25], but gives rise to an equivalent ordinary semi-infinite min-max problem (without unknown parameters) which is much easier to solve. The optimality condition for the equivalent semi-infinite min-max problem appears to be stronger than that in [25], and imply that in [26]. Note that Augmented Lagrangian functions have earlier been used in connection with solving *ordinary* semi-infinite min-max problems (see [10]). However, in [10] the Augmented Lagrangian functions were associated with the "outer" minimization problem and not with the "inner" maximization as we propose.

In Section 2, we use the Rockafellar augmented Lagrangian function ([23]), and use it to remove the constraints $f(x, y) \leq 0$ from the inner problem in (1.2a). In Section 3, we show that the resulting problem is equivalent to **P**. In the process we obtain a new first-order optimality condition for **P**. The paper ends with two numerical examples and concluding remarks.

2 Augmented Lagrangian Penalty Function

We remove the constraints $f(x, y) \leq 0$ in (1.2a) by using the augmented Lagrangian exact penalty function in [23]. Let $\pi \in \mathbb{R}$ and $\eta \in \mathbb{R}^{r_1}$ be a penalty parameter and a multiplier vector, respectively. Hence, we define $\bar{\psi} : \mathbb{R}^{n+r_1+1} \to \mathbb{R}$ to be given by

$$\bar{\psi}(\bar{x}) \triangleq \max_{y \in Y} \bar{\phi}(\bar{x}, y), \tag{2.1a}$$

where $\bar{x} = (x, \eta, \pi) \in \mathbb{R}^{n+r_1+1}$, with $x \in \mathbb{R}^n, \eta \in \mathbb{R}^{r_1}, \pi \in \mathbb{R}$, and

$$\bar{\phi}(\bar{x}, y) \triangleq \phi(x, y) - \frac{1}{2e^{\pi}} \sum_{k=1}^{r_1} [(e^{\pi} f^k(x, y) + \eta^k)_+^2 - (\eta^k)^2]. \tag{2.1b}$$

In (2.1b) and the following, we use the notation $a_+ = (\max\{a^1, 0\}, \max\{a^2, 0\}, ..., \max\{a^p, 0\}) \in \mathbb{R}^p$ for any $a \in \mathbb{R}^p$. Note that the penalty associated with the augmented Lagrangian exact penalty function in (2.1b) is given by e^{π}. Hence, the penalty is positive for all values of the penalty parameter $\pi \in \mathbb{R}$. This transformation may also be advantageous from a computational point of view. Typically, the order of magnitude of π is comparable with the order of magnitude of the other components of $\bar{x}$. The use of e^{π} as a component of $\bar{x}$ would have resulted in a more ill-conditioned problem due to the potential large numerical difference between the penalty and the other components of $\bar{x}$.

Assumption 2.1. We assume that

(i) $\phi(\cdot, \cdot), f^k(\cdot, \cdot), k \in \mathbf{r_1} \triangleq \{1, ..., r_1\}$, and $g^k(\cdot), k \in \mathbf{r_2} \triangleq \{1, ..., r_2\}$, are continuous, and

(ii) $Y \subset \mathbb{R}^m$ is compact, and $Z(x) \neq \emptyset$ for all $x \in \mathbb{R}^n$.

$\square$

Theorem 2.2. *Suppose that Assumption 2.1 holds. Then, for all $\bar{x} = (x, \eta, \pi) \in \mathbb{R}^{n+r_1+1}$ with $x \in \mathbb{R}^n$, $\eta \in \mathbb{R}^{r_1}$, and $\pi \in \mathbb{R}$, we have that*

$$\bar{\psi}(\bar{x}) \geq \psi(x). \tag{2.2a}$$

Proof. By definition,

$$\bar{\psi}(\bar{x}) \geq \max_{y \in Y} \left\{ \phi(x,y) - \frac{1}{2e^{\pi}} \sum_{k=1}^{r_1} [(e^{\pi} f^k(x,y) + \eta^k)_+^2 - (\eta^k)^2] \mid f(x,y) \leq 0 \right\}. \quad (2.2b)$$

Next, suppose that $f(x,y) \leq 0$. If $\eta^k \leq 0$, then

$$(e^{\pi} f^k(x,y) + \eta^k)_+^2 - (\eta^k)^2 = -(\eta^k)^2. \quad (2.2c)$$

Furthermore,

$$e^{\pi} f^k(x,y) + \eta^k - \eta^k \leq 0. \quad (2.2d)$$

Hence, if $\eta^k > 0$, then

$$(e^{\pi} f^k(x,y) + \eta^k)_+^2 - (\eta^k)^2 \leq 0. \quad (2.2e)$$

Therefore, $f(x,y) \leq 0$ implies (2.2e) for all $k \in \mathbf{r_1}$. By (2.2b) it now follows that

$$\bar{\psi}(\bar{x}) \geq \max_{y \in Y} \{\phi(x,y) \mid f(x,y) \leq 0\} = \psi(x) \quad (2.2f)$$

for all $x \in \mathrm{I\!R}^n$, $\eta \in \mathrm{I\!R}^{r_1}$, and $\pi \in \mathrm{I\!R}$. $\square$

In [23], we find a necessary and sufficient condition for the existence of particular η and π that ensure equality in (2.2a). This condition is given in the next definition.

Definition 2.3. Consider the problem

$$\max_{y \in Z(x)} \phi(x,y) \quad (2.3a)$$

and let $v(x,u) \triangleq \max_{y \in Y} \Phi(x,y,u)$, with $\Phi(x,y,u) \triangleq \phi(x,y)$ whenever $f(x,y) \leq u$, and otherwise $\Phi(x,y,u) \triangleq -\infty$.

The problems (2.3a) is said to be stable of degree 2 at $x \in \mathrm{I\!R}^n$ if there exist an open neighborhood U around the origin in $\mathrm{I\!R}^{r_1}$ and a twice continuously differentiable function $\Gamma : U \to \mathrm{I\!R}$ such that

$$v(x,u) \leq \Gamma(u), \quad \forall u \in U, \quad (2.3b)$$

and

$$v(x,0) = \Gamma(0). \quad (2.3c)$$

$\square$

Theorem 2.4. *([23], Theorem 5) Suppose that Assumption 2.1 holds. Then,*

$$\psi(x) = \min_{\eta \in \mathrm{I\!R}^{r_1}} \min_{\pi \in \mathrm{I\!R}} \bar{\psi}((x, \eta, \pi)) \quad (2.4)$$

if and only if (2.3a) *is stable of degree 2 at* $x \in \mathbb{R}^n$. □

It is shown in [23] that a sufficient condition for stability of degree 2 is related to the standard second-order sufficiency condition for (2.3a), stated below.

Definition 2.5. A solution $\hat{y} \in Y$ of the problem in (2.3a) is said to satisfy the second-order sufficient condition at $x \in \mathbb{R}^n$ if

(i) there exists an open neighborhood $N_0 \subset \mathbb{R}^m$ of $\hat{y}$ on which $\phi(\cdot,\cdot)$, $f^k(\cdot,\cdot), k \in \mathbf{r_1}$, and $g^k(\cdot), k \in \mathbf{r_2}$, are twice continuously differentiable,

(ii) there exist multiplier vectors $\eta \in \mathbb{R}^{r_1}$ and $\lambda \in \mathbb{R}^{r_2}$ such that

$$\nabla_y \phi(x,\hat{y}) - f_y(x,\hat{y})^T \eta - g_y(\hat{y})^T \lambda = 0, \tag{2.5a}$$

$$\eta^T f(x,\hat{y}) + \lambda^T g(\hat{y}) = 0, \tag{2.5b}$$

(iii) $h^T L_{yy}(x,\hat{y},\eta,\lambda) h > 0$ for all $h \in \mathbf{H}(x,\hat{y})$, where the Hessian

$$L_{yy}(x,\hat{y},\eta,\lambda) \triangleq \phi_{yy}(x,\hat{y}) - \sum_{k=1}^{r_1} \eta^k f^k_{yy}(x,\hat{y}) - \sum_{k=1}^{r_2} \lambda^k g^k_{yy}(\hat{y}), \tag{2.5c}$$

$$\mathbf{H}(x,\hat{y}) \triangleq \left\{ \begin{array}{c|c} \begin{array}{c} h \in \mathbb{R}^m \\ h \neq 0 \end{array} & \begin{array}{l} \nabla_y f^k(x,\hat{y})^T h \le 0, k \in \hat{\mathbf{r}}_\mathbf{1}(x,\hat{y}), \\ \nabla_y f^k(x,\hat{y})^T h = 0, k \in \hat{\mathbf{r}}_{\mathbf{1}+}(x,\hat{y}), \\ \nabla_y g^k(\hat{y})^T h \le 0, k \in \hat{\mathbf{r}}_\mathbf{2}(\hat{y}), \\ \nabla_y g^k(\hat{y})^T h = 0, k \in \hat{\mathbf{r}}_{\mathbf{2}+}(\hat{y}), \end{array} \end{array} \right\}, \tag{2.5d}$$

with index sets

$$\hat{\mathbf{r}}_\mathbf{1}(x,\hat{y}) \triangleq \{k \in \mathbf{r_1} | f^k(x,\hat{y}) = 0, \eta^k = 0\}, \tag{2.5e}$$

$$\hat{\mathbf{r}}_{\mathbf{1}+}(x,\hat{y}) \triangleq \{k \in \mathbf{r_1} | f^k(x,\hat{y}) = 0, \eta^k > 0\}, \tag{2.5f}$$

$$\hat{\mathbf{r}}_\mathbf{2}(\hat{y}) \triangleq \{k \in \mathbf{r_2} | g^k(\hat{y}) = 0, \lambda^k = 0\}, \tag{2.5g}$$

$$\hat{\mathbf{r}}_{\mathbf{2}+}(\hat{y}) \triangleq \{k \in \mathbf{r_2} | g^k(\hat{y}) = 0, \lambda^k > 0\}. \tag{2.5h}$$

□

Definition 2.6. A point $\hat{y} \in Y$ is said to be the unique optimal solution of (2.3a) in the strong sense if $\hat{y}$ is the only local minimizer for (2.3a). □

Theorem 2.7. *([23], Theorem 6) Suppose that Assumption 2.1 holds. Let* $\hat{y} \in Y$ *be the unique optimal solution to* (2.3a) *in the strong sense, and assume that* $\hat{y}$ *satisfies the second-order sufficiency condition at* $x \in \mathbb{R}^n$ *with* $\eta \in \mathbb{R}^{r_1}$ *and* $\lambda \in \mathbb{R}^{r_2}$ *as the vectors of multipliers. Then* (2.3a) *is stable of degree 2, and for* π *sufficiently large we have* $\psi(x) = \bar{\psi}((x,\eta,\pi))$. □

3 Equivalent Problem and Optimality Conditions

In view of Theorems 2.4 and 2.7, we can define a problem that is equivalent to **P**. Let

$$\bar{\mathbf{P}} \qquad \min_{\bar{x}\in\mathbb{R}^{n+r_1+1}} \bar{\psi}(\bar{x}). \tag{3.1}$$

The next result gives the relations between global and local minimizers of **P** and $\bar{\mathbf{P}}$. Let for any $\hat{x} \in \mathbb{R}^n$ and $\rho > 0$, $\mathbb{B}(\hat{x}, \rho) \triangleq \{x \in \mathbb{R}^n \mid \|x - \hat{x}\| < \rho\}$.

Theorem 3.1. *Suppose that Assumption 2.1 holds and that* (2.3a) *is stable of degree 2 for all $x \in \mathbb{R}^n$. Then, the following hold:*

(a) *Global minimizers of* **P** *and* $\bar{\mathbf{P}}$ *are equivalent in the sense that*

$$\min_{x\in\mathbb{R}^n} \psi(x) = \min_{\bar{x}\in\mathbb{R}^{n+r_1+1}} \bar{\psi}(\bar{x}). \tag{3.2a}$$

(b) *If $\hat{x} \in \mathbb{R}^n$ is a local minimizer for* **P** *with domain of attraction $\mathbb{B}(\hat{x}, \rho)$, then there exist $\hat{\eta} \in \mathbb{R}^{r_1}$ and $\hat{\pi} \in \mathbb{R}$ such that $\hat{\bar{x}} = (\hat{x}, \hat{\eta}, \hat{\pi})$ is a local minimizer for* $\bar{\mathbf{P}}$ *with domain of attraction $\mathbb{B}(\hat{x}, \rho) \times \mathbb{R}^{r_1+1}$.*

(c) *If $\hat{\bar{x}} = (\hat{x}, \hat{\eta}, \hat{\pi}) \in \mathbb{R}^{n+r_1+1}$ is a local minimizer for* $\bar{\mathbf{P}}$ *with domain of attraction $\mathbb{B}(\hat{x}, \rho) \times \mathbb{R}^{r_1+1}$, then $\hat{x}$ is a local minimizer for* **P** *with domain of attraction $\mathbb{B}(\hat{x}, \rho)$.*

Proof. Part (a) follows directly from Theorem 2.4. Consider Part (b). Let $\hat{x}$, ρ be as stipulated, and let $\hat{\eta} \in \mathbb{R}^{r_1}$ and $\hat{\pi} \in \mathbb{R}$ be such that $\min_{\eta\in\mathbb{R}^{r_1}} \min_{\pi\in\mathbb{R}} \bar{\psi}((\hat{x}, \eta, \pi)) = \bar{\psi}((\hat{x}, \hat{\eta}, \hat{\pi}))$. Let $(x^*, \eta^*, \pi^*) \in \mathbb{B}(\hat{x}, \rho) \times \mathbb{R}^{r_1+1}$ be arbitrary. Then, using Theorem 2.4 and the local optimality of $\hat{x}$ we obtain that

$$\begin{aligned} \bar{\psi}((x^*, \eta^*, \pi^*)) &\geq \min_{\eta\in\mathbb{R}^{r_1}} \min_{\pi\in\mathbb{R}} \bar{\psi}((x^*, \eta, \pi)) \\ &= \psi(x^*) \\ &\geq \psi(\hat{x}) \\ &= \bar{\psi}((\hat{x}, \hat{\eta}, \hat{\pi})). \end{aligned} \tag{3.2b}$$

Next, we consider Part (c). Let $\hat{x}$, $\hat{\eta}$, $\hat{\pi}$, ρ be as stipulated, and let $x^* \in \mathbb{B}(\hat{x}, \rho)$ be arbitrary and let $\eta^* \in \mathbb{R}^{r_1}$ and $\pi^* \in \mathbb{R}$ be such that $\min_{\eta\in\mathbb{R}^{r_1}} \min_{\pi\in\mathbb{R}} \bar{\psi}((x^*, \eta, \pi)) = \bar{\psi}((x^*, \eta^*, \pi^*))$. Then, using Theorem 2.4 and the local optimality of $(\hat{x}, \hat{\eta}, \hat{\pi})$ we obtain that

$$\begin{aligned} \psi(x^*) &= \min_{\eta\in\mathbb{R}^{r_1}} \min_{\pi\in\mathbb{R}} \bar{\psi}((x^*, \eta, \pi)) \\ &= \bar{\psi}((x^*, \eta^*, \pi^*)) \\ &\geq \bar{\psi}((\hat{x}, \hat{\eta}, \hat{\pi})) \\ &= \min_{\eta\in\mathbb{R}^{r_1}} \min_{\pi\in\mathbb{R}} \bar{\psi}((\hat{x}, \eta, \pi)) \\ &= \psi(\hat{x}). \end{aligned} \tag{3.2c}$$

This completes our proof. $\square$

Before we initiate our discussion of stationary points, we need an additional assumption.

Assumption 3.2. We assume that $\phi(\cdot,\cdot)$, $f^k(\cdot,\cdot)$, $k \in \mathbf{r_1}$, and $g^k(\cdot)$, $k \in \mathbf{r_2}$, are continuously differentiable. $\square$

Since $\bar{\mathbf{P}}$ is an ordinary semi-infinite min-max problem, a corresponding optimality condition is available in the literature:

Theorem 3.3. *(e.g., [20], Theorem 3.1.5) Suppose that Assumptions 2.1 and 3.2 hold. If $\hat{\bar{x}} = (\hat{x}, \hat{\eta}, \hat{\pi}) \in \mathbb{R}^{n+r_1+1}$ is a local minimizer of $\bar{\mathbf{P}}$, then*

$$0 \in \bar{G}\bar{\psi}(\hat{\bar{x}}) \triangleq \operatorname*{conv}_{y \in Y} \left\{ \begin{pmatrix} \bar{\psi}(\hat{\bar{x}}) - \bar{\phi}(\hat{\bar{x}}, y) \\ \nabla_x \bar{\phi}(\hat{\bar{x}}, y) \\ \nabla_\eta \bar{\phi}(\hat{\bar{x}}, y) \\ \nabla_\pi \bar{\phi}(\hat{\bar{x}}, y) \end{pmatrix} \right\}. \tag{3.3a}$$

$\square$

In view of Theorems 2.4, 3.1, and 3.3, we deduce the following new optimality condition for $\mathbf{P}$.

Theorem 3.4. *Suppose that Assumptions 2.1 and 3.2 hold. If $\hat{x} \in \mathbb{R}^n$ is a local minimizer of $\mathbf{P}$ and (2.3a) is stable of degree 2 in some neighborhood of $\hat{x} \in \mathbb{R}^n$, then there exist $\hat{\eta} \in \mathbb{R}^{r_1}$ and $\hat{\pi} \in \mathbb{R}$ such that*

$$0 \in \bar{G}\bar{\psi}(\hat{\bar{x}}), \tag{3.3b}$$

where $\hat{\bar{x}} = (\hat{x}, \hat{\eta}, \hat{\pi})$. $\square$

In view of Theorems 3.3 and 3.4, it is clear that $\mathbf{P}$ and $\bar{\mathbf{P}}$ have equivalent stationary points. The optimality condition in Theorem 3.4 can be related to the following optimality condition deduced from [26]. For brevity, let

$$\hat{Y}(x) \triangleq \arg \max_{y \in Z(x)} \phi(x, y). \tag{3.4}$$

Theorem 3.5. *Suppose that Assumptions 2.1 and 3.2 hold, x is a local minimizer for $\mathbf{P}$, and the Mangasarian-Fromowitz constraint qualification holds at every $y \in \hat{Y}(x)$, i.e., for every $y \in \hat{Y}(x)$, there exists an $h \in \mathbb{R}^m$ such that $\nabla_y f^k(x, y)^T h < 0$, for all*

$k \in \mathbf{r_1}$ *such that* $f^k(x,y) = 0$, *and* $\nabla g^k(y)^T h < 0$, *for all* $k \in \mathbf{r_2}$ *such that* $g^k(y) = 0$. *Then,*

$$0 \in \underset{y \in \hat{Y}(x)}{\text{conv}} \; \underset{\alpha \in A(x,y)}{\text{conv}} \{\nabla_x \phi(x,y) - f_x(x,y)^T \alpha\}, \tag{3.5a}$$

where

$$A(x,y) \triangleq \left\{ \quad \alpha \in \mathbb{R}^{r_1} \;\middle|\; \begin{array}{l} \nabla_y \phi(x,y) - f_y(x,y)^T \alpha - g_y(y)^T \beta = 0 \\ \alpha^T f(x,y) + \beta^T g(y) = 0 \\ \alpha \geq 0; \beta \in \mathbb{R}^{r_2}, \beta \geq 0 \end{array} \right\}. \tag{3.5b}$$

□

Theorem 3.6. *Suppose that Assumptions 2.1 and 3.2 hold,* $\bar{x} = (x, \eta, \pi) \in \mathbb{R}^{n+r_1+1}$ *satisfies*

$$0 \in \bar{G}\bar{\psi}(\bar{x}), \tag{3.6}$$

and the Mangasarian-Fromowitz constraint qualification holds at every $y \in \hat{Y}(x)$. *Then,* (3.5a) *holds.*

Proof. By Caratheodory's Theorem, see, e.g., Theorem 5.2.5 in [20], (3.6) holds if and only if there exist $y_i \in Y$, $i \in \mathbf{s} \triangleq \{1, ..., s\}$, with $s \triangleq n + r_1 + 3$, and a multiplier vector $\mu \in \Sigma_s \triangleq \{\mu \in \mathbb{R}^s |\ \mu^i \geq 0, i \in \mathbf{s}, \sum_{i \in \mathbf{s}} \mu^i = 1\}$ such that

$$0 = \sum_{i \in \mathbf{s}} \mu^i [\bar{\psi}(\bar{x}) - \bar{\phi}(\bar{x}, y_i)] \tag{3.7a}$$

$$0 = \sum_{i \in \mathbf{s}} \mu^i \nabla_x \bar{\phi}(\bar{x}, y_i) \tag{3.7b}$$

$$0 = \sum_{i \in \mathbf{s}} \mu^i \nabla_\eta \bar{\phi}(\bar{x}, y_i) \tag{3.7c}$$

$$0 = \sum_{i \in \mathbf{s}} \mu^i \nabla_\pi \bar{\phi}(\bar{x}, y_i) \tag{3.7d}$$

From (3.7c), we obtain that

$$\begin{aligned} 0 &= \sum_{i \in \mathbf{s}} \mu^i \nabla_\eta \bar{\phi}(\bar{x}, y_i) \\ &= \frac{1}{e^\pi} \sum_{i \in \mathbf{s}} \mu^i [\eta - (e^\pi f(x, y_i) + \eta)_+] \\ &= \frac{1}{e^\pi} [\eta - \sum_{i \in \mathbf{s}} \mu^i (e^\pi f(x, y_i) + \eta)_+]. \end{aligned} \tag{3.7e}$$

Hence, for all $k \in \mathbf{r_1}$

$$\eta^k = \sum_{i \in \mathbf{s}} \mu^i (e^\pi f^k(x, y_i) + \eta^k)_+ \geq 0. \tag{3.7f}$$

From (3.7d), we obtain

$$\begin{aligned} 0 &= \sum_{i \in \mathbf{s}} \mu^i \nabla_\pi \bar{\phi}(\bar{x}, y_i) \\ &= \sum_{i \in \mathbf{s}} \sum_{k \in \mathbf{r_1}} \mu^i \left\{ \frac{1}{2e^\pi} [(e^\pi f^k(x, y_i) + \eta^k)_+^2 - (\eta^k)^2] - (e^\pi f^k(x, y_i) + \eta^k)_+ f^k(x, y_i) \right\}. \end{aligned} \tag{3.7g}$$

We now look at an individual term in the summation in (3.7g). We have five cases. (**i**) Suppose that $f^k(x, y_i) > 0$. Then, because $\eta^k \geq 0$ by (3.7f),

$$\begin{aligned} & \frac{1}{2e^\pi} [(e^\pi f^k(x, y_i) + \eta^k)_+^2 - (\eta^k)^2] - (e^\pi f^k(x, y_i) + \eta^k)_+ f^k(x, y_i) \\ = & \frac{1}{2e^\pi} [(e^\pi f^k(x, y_i))^2 + 2e^\pi f^k(x, y_i)\eta^k] - e^\pi (f^k(x, y_i))^2 - f^k(x, y_i)\eta^k \\ = & -\frac{e^\pi}{2} (f^k(x, y_i))^2. \end{aligned} \tag{3.7h}$$

(**ii**) Suppose that $f^k(x, y_i) = 0$. Then, because $\eta^k \geq 0$, we obtain

$$\frac{1}{2e^\pi} [(e^\pi f^k(x, y_i) + \eta^k)_+^2 - (\eta^k)^2] - (e^\pi f^k(x, y_i) + \eta^k)_+ f^k(x, y_i) = 0. \tag{3.7i}$$

(**iii**) Suppose that $f^k(x, y_i) < 0$ and $\eta^k = 0$. Then,

$$\frac{1}{2e^\pi} [(e^\pi f^k(x, y_i) + \eta^k)_+^2 - (\eta^k)^2] - (e^\pi f^k(x, y_i) + \eta^k)_+ f^k(x, y_i) = 0. \tag{3.7j}$$

(**iv**) Suppose that $f^k(x, y_i) < 0$, $\eta^k > 0$, and $e^\pi f^k(x, y_i) + \eta^k \leq 0$. Then,

$$\begin{aligned} & \frac{1}{2e^\pi} [(e^\pi f^k(x, y_i) + \eta^k)_+^2 - (\eta^k)^2] - (e^\pi f^k(x, y_i) + \eta^k)_+ f^k(x, y_i) \\ = & -\frac{(\eta^k)^2}{2e^\pi}. \end{aligned} \tag{3.7k}$$

(v) Suppose that $f^k(x, y_i) < 0$, $\eta^k > 0$, and $e^\pi f^k(x, y_i) + \eta^k > 0$. Then, (3.7$h$) holds. Next, we split (3.7g) into sub-summations that corresponds to the results from the five cases. This gives,

$$\begin{aligned} 0 &= \sum_{i \in \mathbf{s}} \sum_{k \in \mathbf{r_1}} \mu^i \left\{ \frac{1}{2e^\pi} [(e^\pi f^k(x, y_i) + \eta^k)_+^2 - (\eta^k)^2] - (e^\pi f^k(x, y_i) + \eta^k)_+ f^k(x, y_i) \right\} \\ &= -\sum_{(i,k) \in \mathbf{I_1}(x)} \mu^i \frac{e^\pi}{2} (f^k(x, y_i))^2 - \sum_{(i,k) \in \mathbf{I_4}(x)} \mu^i \frac{(n^k)^2}{2e^\pi} - \sum_{(i,k) \in \mathbf{I_5}(x)} \mu^i \frac{e^\pi}{2} (f^k(x, y_i))^2, \end{aligned} \tag{3.7l}$$

where

$$
\begin{aligned}
\mathbf{I_1}(x) &\stackrel{\triangle}{=} \{(i,k) \in \mathbf{s} \times \mathbf{r_1} \mid f^k(x,y_i) > 0\} \\
\mathbf{I_4}(x) &\stackrel{\triangle}{=} \{(i,k) \in \mathbf{s} \times \mathbf{r_1} \mid f^k(x,y_i) < 0,\ \eta^k > 0,\ e^\pi f^k(x,y_i) + \eta^k \le 0\} \qquad (3.7m) \\
\mathbf{I_5}(x) &\stackrel{\triangle}{=} \{(i,k) \in \mathbf{s} \times \mathbf{r_1} \mid f^k(x,y_i) < 0,\ \eta^k > 0,\ e^\pi f^k(x,y_i) + \eta^k > 0\}.
\end{aligned}
$$

We see from (3.7l) that $\mu^i = 0$ for all i appearing in the summation over $\mathbf{I_1}(x)$, $\mathbf{I_4}(x)$, and $\mathbf{I_5}(x)$. Hence, the subset $\mathbf{t}' \subset \mathbf{s}$, defined by

$$
\mathbf{t}' \stackrel{\triangle}{=} \{i \in \mathbf{s} \mid (i,k) \notin \mathbf{I_1}(x) \bigcup \mathbf{I_4}(x) \bigcup \mathbf{I_5}(x), \forall k \in \mathbf{r_1}\}, \qquad (3.7n)
$$

has the property that $i \notin \mathbf{t}'$ implies $\mu^i = 0$. Since $\sum_{i \in \mathbf{s}} \mu^i = 1$, we must have $\mathbf{t}' \neq \emptyset$. Moreover, if $i \in \mathbf{t}'$, then $f^k(x,y_i) \le 0$ for all $k \in \mathbf{r_1}$ and, additionally, $f^k(x,y_i) < 0$ implies $\eta^k = 0$.

From (3.7a), we observe that if $\bar{\psi}(\bar{x}) - \bar{\phi}(\bar{x}, y_i) > 0$, then $\mu^i = 0$. Let

$$
\mathbf{t} \stackrel{\triangle}{=} \{i \in \mathbf{t}' \mid \bar{\psi}(\bar{x}) - \bar{\phi}(\bar{x}, y_i) = 0\}. \qquad (3.7o)
$$

Since $\sum_{i \in \mathbf{s}} \mu^i = 1$, $i \notin \mathbf{t}'$ implies $\mu^i = 0$, and (3.7a) holds, we must have $\mathbf{t} \neq \emptyset$.

Since $f^k(x,y_i) \le 0$ for all $k \in \mathbf{r_1}, i \in \mathbf{t}$ and $f^k(x,y_i) < 0$ implies $\eta^k = 0$, we have that

$$
(e^\pi f^k(x,y_i) + \eta^k)_+ = \eta^k \qquad (3.7p)
$$

for all $k \in \mathbf{r_1}, i \in \mathbf{t}$.

In view of (3.7o) and (3.7p), we have

$$
\bar{\psi}(\bar{x}) = \bar{\phi}(\bar{x}, y_i) = \phi(x, y_i) \quad \forall i \in \mathbf{t}. \qquad (3.7q)
$$

Hence, it follows from Theorem 2.2 that for every $i \in \mathbf{t}$,

$$
\psi(x) = \max_{y \in Z(x)} \phi(x,y) \ge \phi(x,y_i) = \bar{\psi}(\bar{x}) \ge \psi(x). \qquad (3.7r)
$$

Consequently, $\psi(x) = \bar{\psi}(\bar{x})$ and $y_i \in \hat{Y}(x)$ for all $i \in \mathbf{t}$.

From (3.7b) and (3.7p), we obtain

$$
\begin{aligned}
0 &= \sum_{i \in \mathbf{t}} \mu^i [\nabla_x \phi(x,y_i) - \sum_{k \in \mathbf{r_1}} (e^\pi f^k(x,y_i) + \eta^k)_+ \nabla_x f^k(x,y_i)] \\
&= \sum_{i \in \mathbf{t}} \mu^i [\nabla_x \phi(x,y_i) - \sum_{k \in \mathbf{r_1}} \eta^k \nabla_x f^k(x,y_i)].
\end{aligned} \qquad (3.7s)
$$

Since $y_i \in \hat{Y}(x)$ is also a maximizer of $\bar{\phi}(\bar{x}, y)$ over Y, it follows from the Mangasarian-Fromowitz constraint qualification (see, e.g., Chapter 5 in [1]) that there exists a $\lambda \in \mathbb{R}^{r_2}, \lambda \ge 0$ such that

$$
\nabla_y \bar{\phi}(\bar{x}, y_i) - g_y(y_i)^T \lambda = 0 \qquad (3.7t)
$$

$$\lambda^T g(y_i) = 0. \tag{3.7u}$$

Hence, it follows from (3.7t) and (3.7p) that

$$\begin{aligned}
0 &= \nabla_y \bar{\phi}(\bar{x}, y_i) - g_y(y_i)^T \lambda \\
&= \nabla_y \phi(x, y_i) - \sum_{k \in \mathbf{r_1}} (e^\pi f^k(x, y_i) + \eta^k)_+ \nabla_y f^k(x, y_i) - g_y(y_i)^T \lambda \\
&= \nabla_y \phi(x, y_i) - \sum_{k \in \mathbf{r_1}} \eta^k \nabla_y f^k(x, y_i) - g_y(y_i)^T \lambda \\
&= \nabla_y \phi(x, y_i) - f_y(x, y_i)^T \eta - g_y(y_i)^T \lambda.
\end{aligned} \tag{3.7v}$$

In view of (3.7v) and (3.7s), we conclude that (3.5a) holds. This completes the proof.□

We are able to show the reverse relation only under strong assumptions:

Theorem 3.7. *Suppose that Assumptions 2.1 and 3.2 hold and $x \in \mathbb{R}^n$ satisfies (3.5a). Furthermore, suppose that $y \in Y$ is the unique optimal solution of (2.3a) in the strong sense, that y satisfies the second-order sufficiency conditions at x with $\eta \in \mathbb{R}^{r_1}$ and $\lambda \in \mathbb{R}^{r_2}$ as the vectors of multipliers, and that the linear independence constraint qualification*[1] *is satisfied at (x, y).*

Then, (3.6) holds, with $\bar{x} = (x, \eta, \pi)$, for some π sufficiently large.

Proof. Under the given assumptions, (3.5a) simplifies to

$$0 = \nabla_x \phi(x, y) - f_x(x, y)^T \eta. \tag{3.8a}$$

It follows from Theorem 2.7 that there exists $\pi \in \mathbb{R}$ such that

$$\psi(x) = \bar{\psi}(\bar{x}), \tag{3.8b}$$

where $\bar{x} = (x, \eta, \pi)$. Using the fact that $f^k(x, y) \leq 0$ for all $k \in \mathbf{r_1}$ and $f^k(x, y) < 0$ implies $\eta^k = 0$, we obtain

$$(e^\pi f^k(x, y) + \eta^k)_+ = \eta^k \tag{3.8c}$$

for all $k \in \mathbf{r_1}$. Hence, $\bar{\phi}(\bar{x}, y) = \phi(x, y)$ and, by (3.8b) and the optimality of y, $\bar{\psi}(\bar{x}) - \bar{\phi}(\bar{x}, y) = 0$. From (3.8$a$) and (3.8$c$), we find $\nabla_x \bar{\phi}(\bar{x}, y) = 0$. Using (3.8$c$), we obtain $\nabla_\eta \bar{\phi}(\bar{x}, y) = 0$ and $\nabla_\pi \bar{\phi}(\bar{x}, y) = 0$. This completes the proof. □

4 Algorithms and Numerical Examples

In view of Theorem 3.1 and its assumptions, many problems of the form $\mathbf{P}$ can be addressed by solving the ordinary semi-infinite min-max problem $\bar{\mathbf{P}}$. Any semi-infinite

[1] Can be replaced by the strict Mangasarian-Fromowitz constraint qualification [7, 14], or other conditions which imply uniqueness of multipliers

min-max algorithm can be used for this purpose. In particular, Algorithm 3.4.6, 3.4.9, and 3.4.16 in [20], which are based on discretization of the set Y, Algorithm 3.4.26 in [20], which uses the outer approximation method, and the outer approximation algorithms in [8], which includes constraint dropping schemes, are suitable. When applied to $\bar{\mathbf{P}}$, these algorithms are known to converge to points satisfying (3.3a) [8, 20]. Consequently, these algorithms converge to stationary points (in the sense of Theorem 3.4) for $\mathbf{P}$ under the stability of degree 2 assumption. In view of Theorem 3.6, the algorithms also converge to points satisfying the optimality condition for $\mathbf{P}$ in Theorem 3.5. Hence, implementable algorithms for computing stationary points of $\mathbf{P}$ are available under the stability of degree 2 assumption.

We illustrate the solution strategy for solving $\mathbf{P}$ by considering the following numerical examples using Matlab [18] and a 500 MHz PC.

4.1 Example 1

Let $x = (x^1, x^2, x^3) \in \mathbb{R}^3$, $y \in \mathbb{R}$, and

$$\phi(x,y) = 3(x^1 - y)^2 + (2-y)(x^2)^2 + 5(x^3+y)^2 + 2x^1 + 3x^2 - x^3 + e^{4y^2}, \tag{4.1a}$$

$$f(x,y) = \frac{1}{4}\sin(x^1x^2) + y - \frac{1}{2}, \tag{4.1b}$$

$$g^1(y) = -y, \tag{4.1c}$$

$$g^2(y) = y - 1, \tag{4.1d}$$

i.e., $r_1 = 1$, $r_2 = 2$, and $Y = [0,1] \subset \mathbb{R}$.

We use the Pironneau-Polak-Pshenichnyi (PPP) min-max algorithm (i.e., Algorithm 2.4.1 in [20]) to solve the finite min-max problem obtained from $\bar{\mathbf{P}}$ by discretizing Y with 1113 equally spaced points. This number of discretization points is selected to facilitate comparison with the results in [25]. The algorithm in [25] uses an adaptive scheme to determine the number of discretization points of Y. At the termination point of the run of this example in [25], the discretization of Y consisted of 1113 points. Additionally, we set the parameters in Algorithm 2.4.1 in [20] to be $\alpha_a = 0.5$, $\beta_a = 0.8$, and $\delta_a = 1$.

Initial numerical testing revealed that the graph of the function $\bar{\psi}(\cdot)$ appears to be fairly flat in the direction of η. As illustrated in Figure 1, this results in significantly slower convergence in η than in x as the algorithm progresses on $\bar{\mathbf{P}}$. In Figure 1, the abscissa axis gives the discrepancy between the current iterate (x_i, η_i) and the solution of the problem (x^*, η^*). To compensate for this effect, we recommend to scale η by a factor $\sigma = \|\nabla_x\bar{\phi}(\bar{x}_0, y_0)\|^2/\|\nabla_\eta\bar{\phi}(\bar{x}_0, y_0)\|^2$, where $\bar{x}_0$ is the value used for initialization of the algorithm and y_0 is an approximate maximizer of $\bar{\phi}(\bar{x}_0, y)$. Hence, we replace η by $\sigma\eta$ in the algorithm. However, the numbers reported in the following are scaled back to the original η for consistency. For the various initial points in Examples 1 and 2, the recommended formula gives $\sigma \approx 10^2$. Hence, we set in this example and

Example 2, $\sigma = 100$. As seen from Figure 2, scaling creates initial oscillations (in fact with a discrepancy larger than 9 for iterations 1-70), but convergence is reached faster than without scaling.

The numerical results for this example are summarized in Tables 1 and 2, where the performance of the proposed approach is compared with that of the algorithm in [25]. In Table 1, x_0, η_0 and π_0 denote the initialization values of x, η and π, respectively, for the Pironneau-Polak-Pshenichnyi algorithm in solving $\bar{\mathbf{P}}$. The different runs are denoted PPP-1 to PPP-4. In Table 2, $y_m \in \arg\max_{y\in Y} \bar{\phi}(\bar{x}, y)$. Hence, a positive value of $f(x, y_m)$ indicates a constraint violation in the inner problem in (1.2a).

We observe from Table 1 that the proposed approach is significantly faster than the one in [25]. It should be noted that the current implementation of the approach is rather unsophisticated. Other semi-infinite min-max algorithms, such as the ones in [20] involving adaptive discretization schemes or the efficient SQP-based algorithm in [32], are expected to yield even better results. SQP-based algorithms may also reduce the need for scaling.

From Tables 1 and 2, we see that there is discrepancy between $\bar{\psi}((x, \eta, \pi))$ and $\psi(x)$ when there is a constraint violation, i.e., $f(x, y_m) > 0$. Otherwise, we have $\bar{\psi}((x, \eta, \pi)) = \psi(x)$ as stated in Theorem 2.4.

In a neighborhood of the solution x, it can be shown that

$$v(x, u) = \phi\left(x, u - \frac{1}{4}\sin(x_1 x_2) + \frac{1}{2}\right) \tag{4.2}$$

for sufficiently small $|u|$. For a given $x^* \in \mathbb{R}^n$, the right-hand side of (4.2) is a twice continuously differentiable function of u. Hence, we can set $\Gamma(u) = \phi(x^*, u - (1/4)\sin(x_1^* x_2^*) + 1/2)$. Consequently, it follows directly from Definition 2.3 that (2.3a) is stable of degree 2 for values close to the solution x in this example.

4.2 Example 2

A second example arises in the optimal design of a short structural column with a rectangular cross section of dimensions $b \times h$. As discussed in the introduction, an approximation to the optimal design of the column can be obtained by solving (1.5). To avoid the unrealistic case of negative dimensions of the column, we set $b = (x^1)^2$ and $h = (x^2)^2$, where $x = (x^1, x^2) \in \mathbb{R}^2$ is the design vector in the optimization problem. Suppose that the initial cost of the design is $c^0(x) = bh$, and the cost of failure $c^1(x) = 100c^0(x)$.

The column is subjected to bi-axial bending moments and an axial force, which, together with the yield strength of the material, are considered to be random variables. This gives rise to a limit-state function $h : \mathbb{R}^2 \times \mathbb{R}^4 \to \mathbb{R}$ defining the probability of failure, see (1.3), where

$$h(x, y) \triangleq 1 - \frac{4v^1(y)}{(x^1)^2(x^2)^4 v^4(y)} - \frac{4v^2(y)}{(x^1)^4(x^1)^2 v^4(y)} - \left(\frac{v^3(y)}{(x^1)^2(x^2)^2 v^4(y)}\right)^2, \tag{4.3}$$

with $v(y) \stackrel{\triangle}{=} (v^1(y), v^2(y), v^3(y), v^4(y)) \in \mathbb{R}^4$ being given by

$$v^1(y) = \exp(12.386 + 0.29356 z^1(y)), \tag{4.4a}$$

$$v^2(y) = \exp(11.693 + 0.29356 z^2(y)), \tag{4.4b}$$

$$v^3(y) = \exp(14.712 + 0.19804 z^3(y)), \tag{4.4c}$$

$$v^4(y) = \exp(17.499 + 0.09975 z^4(y)), \tag{4.4d}$$

$z(y) \stackrel{\triangle}{=} Ly$, and the 4×4 matrix L may have different structures.

The optimal design of the column is computed by solving (1.5). By inspection, (1.5) is of the form **P**, with $Y = \{y \in \mathbb{R}^4 \mid \|y\| \leq \beta\}$. We set $\alpha = 0.0001$ and $\beta = 3$ in (1.5).

Suppose L is the unit diagonal matrix, which is the case corresponding to statistically independent random variables. By using an outer approximation algorithm (OAA) (see Chapter 3 of [20]), we obtain from the initial point $(x_0, \eta_0, \pi_0) = (\sqrt{0.75}, \sqrt{0.75}, 1, 1)$, with $\psi(x_0) = 0.5625$, the result in the first rows of Tables 3 and 4. The second rows of Tables 3 and 4 contain results using the algorithm in [25] with the same initialization. The discretization of the set Y required by the algorithm in [25] could be constructed by using a uniform grid. However, to reduce the computing time and to be more comparable with the outer approximation algorithm used with the approach in this paper, we adopted an heuristic approach. The discretization of Y was constructed as the algorithm progressed by solving the maximization problem in (1.5) approximately at each iteration. The discretization of Y was defined as the set of all such maximizers accumulated up to the present iteration.

The results obtained by the outer approximation algorithm (OAA) and by the algorithm in [25] correspond to columns with cross-section dimensions $b \times h = 0.266 \times 0.715$ and $b \times h = 0.315 \times 0.631$, respectively. It appears that our new approach is significantly faster than the one in [25].

Since $f(x, y_m) < 0$ at the final iterate (see Table 4), the constraint $f(x, y) \leq 0$ in the inner problem (1.2a) is not active. This is also indicated by the fact that $\eta \approx 0$ and π is essentially equal to its initial value. Since $f(x, y) \leq 0$ is not active at the solution x, $v(x, u)$ is constant with respect to u for sufficiently small perturbations in u. Hence, $\Gamma(u)$ in Definition 2.3 can be set equal to $v(x, u)$. The same argument holds in a sufficiently small neighborhood of the solution x. Consequently, (2.3a) is stable of degree 2 in a neighborhood of the solution in this example.

Obviously, the situation with an inactive constraint can rarely be identified a priori. A case where it is difficult to identify whether $f(x, y) \leq 0$ is active or not arises in optimal design with correlated random variables. Let the matrix L involved in the definition of the limit state function in (4.3) be defined by

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.7580 & 0.6523 & 0 & 0 \\ -0.5239 & -0.1944 & 0.8293 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.5}$$

This corresponds to a case when the random variables are correlated. By using the approach suggested in this paper with the same outer approximation algorithm as above (see Chapter 3 of [20]), we obtain after 12 iterations and 562 seconds from the initial point $(\sqrt{0.75}, \sqrt{0.75}, 1, 1)$ the local minimizer $\bar{x} = (x, \eta, \pi) = (0.5810, 0.7828, 0.0000, 1.0001)$, with $\bar{\psi}(\bar{x}) = \psi(x) = 0.2129$. This corresponds to a column with cross-section dimensions $b \times h = 0.338 \times 0.613$. Note that the constraint $f(x, y) \leq 0$ is inactive in this case as well, with $\eta \approx 0$ and π close to its initial value. Since $f(x, y) \leq 0$ is not active in this example, $(2.3a)$ is stable of degree 2 in a neighborhood of the solution by the same argument as above.

5 Conclusions

We have developed an implementable approach for solving generalized semi-infinite min-max problems based on the use of an augmented Lagrangian exact penalty function. The augmented Lagrangian is used to convert the original problem into a ordinary semi-infinite min-max problem. The associated multipliers and penalty parameter in the augmented Lagrangian function are added as auxiliary optimization variables in the ordinary semi-infinite min-max problem. The resulting ordinary semi-infinite min-max problem has the same local/global solutions and stationary points as the original problem. Using this fact, we have derived a new first-order optimality condition for the generalized semi-infinite min-max problem, which is shown to imply an existing first-order optimality condition.

The new approach consists of solving the equivalent ordinary semi-infinite min-max problem. Any semi-infinite optimization algorithm can be used for this purpose. The approach is limited to cases where the inner maximization problem is stable of degree 2. We expect this assumption to be satisfied in most practical cases. The approach was tested numerically on one artificial and one engineering design example and found to compute better than an alternative algorithm.

References

[1] M. S. Bazaraa, H. D. Sherall, and C. M. Shetty, *Nonlinear Programming. Theory and Algorithms*, Second Edition, Wiley, New York, NY, 1993.

[2] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, NY, 1982.

[3] J.F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer Verlag, New York, NY, 2000.

[4] J.V. Burke, "Calmness and exact penalization," *SIAM J. Control and Optimization*, Vol. 29, No. 2, pp. 493-497, 1991.

[5] F. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, NY, 1983.

[6] O. Ditlevsen and H.O. Madsen, *Structural Reliability Methods*, Wiley, New York, NY, 1996.

[7] O. Fujiwara, S.-P. Han, and O.L. Mangasarian, "Local Duality of Nonlinear Programs," *SIAM J. Control and Optimization*, Vol. 22, pp. 162-169, 1984.

[8] C. Gonzaga and E. Polak, "On Constraint Dropping Schemes and Optimality Functions for a Class of Outer Approximations Algorithms," *SIAM J. Control and Optimization*, Vol. 17, No. 4, pp. 477-493, 1979.

[9] T.J. Graettinger and B.H. Krogh, "The Acceleration Radius: A Global Performance Measure for Robotic Manipulators," *IEEE J. of Robotics and Automation*, Vol. 4, pp. 60-69, 1988.

[10] W.W. Hager and D.L. Presler, "Dual Techniques for Minimax," *SIAM J. Control and Optimization*, Vol. 25, pp. 660-685, 1987.

[11] R. Hettich and G. Still, "Semi-infinite Programming Models in Robotics," *Parametric Optimization and Related Topics II*, J. Goddat et al. (eds.), Akademie Verlag, Berlin, pp. 112-118, 1991.

[12] H.Th. Jongen, J.-J. Ruckmann and O. Stein, "Generalized Semi-infinite Optimization: a first order optimality condition and examples," *Mathematical Programming*, Vol. 83, pp. 145-158, 1998.

[13] A. Kaplan and R. Tichatschke, "On the Numerical Treatment of a Class of Semi-infinite Terminal Problems," *Optimization*, Vol. 41, pp. 1-36, 1997.

[14] J. Kyparisis, "On the Uniquenss of Kuhn-Tucker Multipliers in Nonlinear Programming," *Mathematical Programming*, Vol. 32, pp. 242-246, 1985.

[15] 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, Germany, 2001.

[16] E. Levitin and R. Tichatschke, "A Branch-and-Bound Approach for Solving a Class of Generalized Semi-infinite Programming Problems," *J. Global Optimization*, Vol. 13, pp. 299-315, 1998.

[17] X. Li, "An entropy-based aggregate method for minimax optimization," *Engineering Optimization*, Vol. 18, pp. 277-285, 1997.

[18] Matlab reference manual, version 5.3, (R11), MathWorks, Inc., Natick, Mass., 1999.

[19] G. Di Pillo, "Exact penalty methods," *Algorithms for Continuous Optimization: the State of the Art*, E. Spedicato (Ed.), Kluwer Academic Pub., Dordrecht, 1994.

[20] E. Polak, *Optimization. Algorithms and Consistent Approximations*, Springer Verlag, New York, NY, 1997.

[21] B. N. Pshenichnyi, and Yu. M. Danilin, *Numerical Methods in Extremal Problems (Chislennye Metody v Ekstremal'nykh Zadachakh)*, Nauka, Moscow, 1975.

[22] R. T. Rockafellar and R. J-B. Wets, *Variational Analysis*, Springer-Verlag, New York, NY, 1997.

[23] R. T. Rockafellar, "Augmented Lagrange multiplier functions and duality in nonconvex programming," *SIAM J. Control*, Vol. 12. No. 2, pp. 268-285, 1974.

[24] J. O. Royset, A. Der Kiureghian and E. Polak, "Reliability-Based Optimal Structural Design by the Decoupling Approach," *Reliability Engineering and System Safety*, Elsevier Science, Vol. 73, No. 3, 2001, pp. 213-221.

[25] J.O. Royset, E. Polak and A. Der Kiureghian, "Adaptive approximations and exact penalization for the solution of generalized semi-infinite min-max problems," *SIAM J. Optimization*, Vol. 14, No. 1, pp. 1-34, 2003.

[26] J.-J. Ruckmann and A. Shapiro, "On First-Order Optimality Conditions in Generalized Semi-infinite Programming," *J. Optimization Theory and Applications*, Vol. 101, No. 3, pp. 677-91, 1999.

[27] O. Stein, *Bi-Level Strategies in Semi-Infinite Programming*, Kluwer Academic, Boston, 2003.

[28] O. Stein, "First Order Optimality Conditions for Degenerate Index Sets in Generalized Semi-infinite Programming," *Mathematics of Operations Research*, Vol. 26, pp. 565-582, 2001.

[29] O. Stein and G. Still, "On Optimality Conditions for Generalized Semi-infinite Programming Problems," *J. Optimization Theory and Applications*, Vol. 104, pp. 443-458, 2000.

[30] G. Still, "Generalized Semi-infinite Programming: Theory and Methods," *European J. of Operations Research*, Vol. 119, pp. 301-313, 1999.

[31] G. Still, "Generalized Semi-infinite Programming: Numerical Aspects," *Optimization*, Vol. 49, No. 3, pp. 223-242, 2001.

[32] J.L. Zhou and A.L. Tits, "An SQP Algorithm for Finely Discretized Continuous Minimax Problems and Other Minimax Problems with Many Objective Functions," *SIAM Journal on Optimization*, Vol. 6, pp. 461-487, 1996.

[33] G.-W. Weber, "Generalized Semi-infinite Optimization: On some Foundations," *Vychislitel'nye Tekhnologii*, Vol. 4, pp. 41-61, 1999.

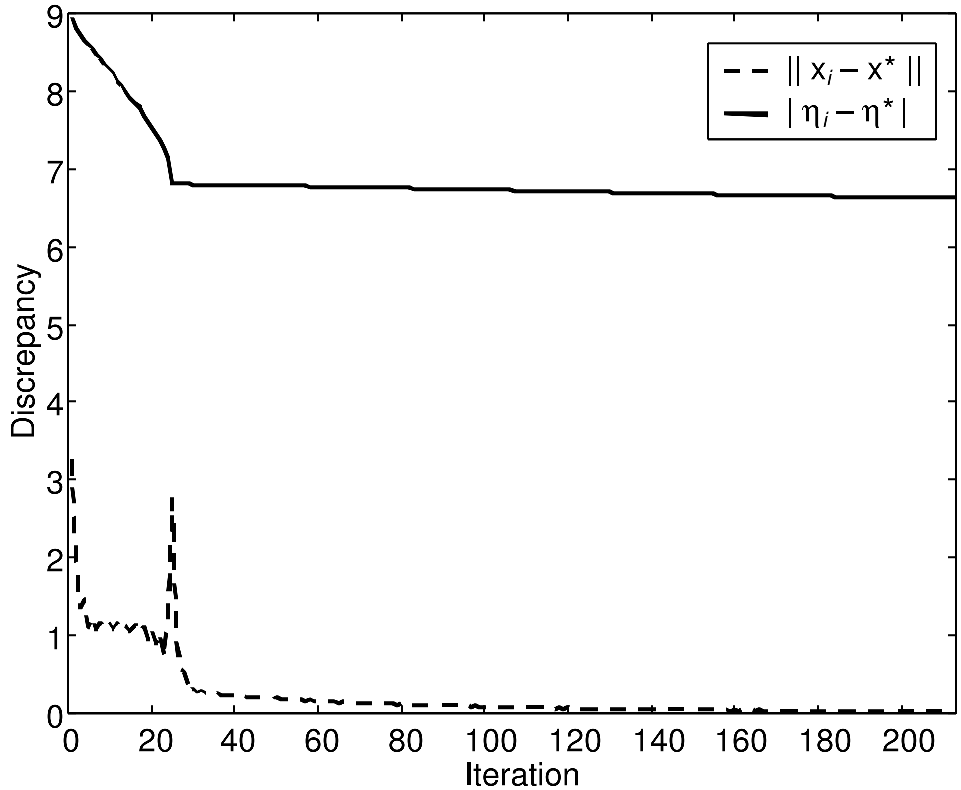

Figure 1: Computing without scaling of η.

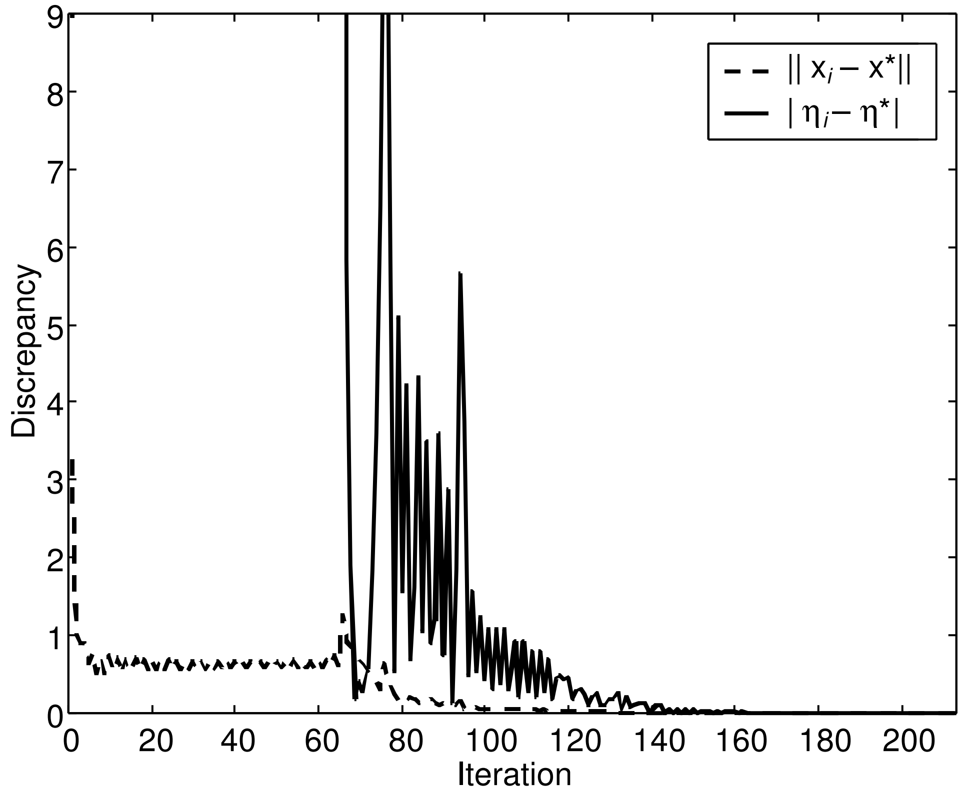

Figure 2: Computing with scaling of η.

Table 1: Numerical Results for Example 1.

Algo.	(x_0, η_0, π_0)	$\psi(x_0)$	Time to reach $\psi(x) \leq 2.41$	$\psi(x)$ after 625 sec.
PPP-1	(2, 1, 0, 1, 1)	22.00	291 sec.	1.9134
PPP-2	(2, 1, 0, 5, 5)	22.00	37 sec.	1.9135
PPP-3	(1, 1, 1, 1, 1)	16.94	203 sec.	1.9202
PPP-4	(1, 1, 1, 5, 5)	16.94	13 sec.	1.9135
[25]	(2, 1, 0, -, -)	22.00	625 sec.	2.4100

Table 2: Numerical Results for Example 1.

Algo.	(x, η, π) after 625 sec.	$\psi((x, \eta, \pi))$	$f(x, y_m)$
PPP-1	$(-0.3896, -1.1985, -0.2875,\ 9.9495,\ 5.6299)$	1.9135	-0.0003
PPP-2	$(-0.3896, -1.1985, -0.2875,\ 9.9505,\ 5.6302)$	1.9135	-0.0003
PPP-3	$(-0.3465, -1.1686, -0.3092,\ 9.2709,\ 7.1959)$	1.9205	0.0007
PPP-4	$(-0.3896, -1.1985, -0.2875,\ 9.9505,\ 5.6300)$	1.9135	-0.0003
[25]	$(-0.0033, -1.0002, -0.3928,$ - , - $)$	-	-

Table 3: Numerical Results for Example 2.

Algo.	Time to reach $\psi(x) \leq 0.2039$	$\psi(x)$ after 645 sec.
OAA	104 sec.	0.1950
[25]	645 sec.	0.2039

Table 4: Numerical Results For Example 2.

Algo.	(x,η,π) after 625 sec.	$\psi((x,\eta,\pi))$	$f(x,y_m)$
OAA	$(0.5161, 0.8454, -0.0000, 1.0001)$	0.1950	-0.3741
[25]	(0.5613, 0.7944, - , -)	-	-