AN IMPLEMENTABLE ALGORITHM FOR THE OPTIMAL DESIGN CENTERING, TOLERANCING AND TUNING PROBLEM

by

E. Polak

Memorandum No. UCB/ERL M79/33

28 February 1980
REVISED

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

AN IMPLEMENTABLE ALGORITHM FOR THE OPTIMAL DESIGN CENTERING, TOLERANCING AND TUNING PROBLEM

E. Polak

Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory
University of California, Berkeley, California 94720

ABSTRACT

An implementable master algorithm for solving optimal DCTT problems is presented. This master algorithm decomposes the original non-differentiable optimization problem into a sequence of ordinary nonlinear programming problems. The master algorithm generates sequences with accumulation points that are feasible and satisfy a new optimality condition which is shown to be stronger than the one previously used for these problems.

Research sponsored by National Science Foundation (RANN) Grant ENV 76-04264 and the Joint Services Electronic Program Contract F44620-76-C-0100.

1. Introduction

Quite commonly, the engineering designer has to take into account the fact that the parameter values of the actual system, structure, or device, will be different from the nominal values in the design. In control system design, this discrepancy is largely due to identification errors; in steel structures and in electronic circuit design it is due to production tolerances. Recently, optimization algorithms have been proposed [1,2,14] which enable the designer to ensure satisfaction of specifications not only by the nominal design but also by all possible system or device realizations within a prescribed tolerance range. When using such algorithms, one can make the tolerance range a design parameter and maximize it while minimizing some other cost function of the other parameters by constructing an aggregate cost by means of weighted combinations. It has been known for some time that the requirement of 100% yield, i.e., that all realizations within tolerance range satisfy specifications, may result in very tight tolerances or very high precision of identification requirements [3,4]. To overcome this difficulty, it has become common to tune a control system or to trim, by laser beam, electronic devices after manufacture [5,6,7,8,9]. Empirically, it has been found that tuning and trimming permits the relaxation of the error or tolerance range to acceptable levels and hence results in considerably increased yield. Quite recently, two conceptual† algorithms have been proposed for solving design problems with tolerances and tuning or trimming [3,10]. In the electronics literature such problems are referred to as design centering, tolerancing

†We say that an algorithm is conceptual if it contains infinite operations which cannot be easily approximated. We say that an algorithm is implementable if specific truncation rules are given for all such infinite operations.

and trimming problems (DCTT).

Typically, a specification on the nominal value of the design parameter $x \in \mathbb{R}^n$ takes the form

$$f^i(x) \leq 0, \qquad (1)$$

with $f^i : \mathbb{R}^n \to \mathbb{R}^1$ continuously differentiable. When identification errors, or tolerances, $\varepsilon \in \mathbb{R}^{n_\varepsilon}$, ranging over a compact set $E \subset \mathbb{R}^m$, need to be taken into account, the constraint (1) becomes modified to

$$\max_{\varepsilon \in E} \phi^i (x,\varepsilon) \leq 0, \qquad (2)$$

with $\phi^i : \mathbb{R}^n \times \mathbb{R}^{n_\varepsilon} \to \mathbb{R}^1$. Finally, assuming that tuning or trimming will be performed by means of a parameter $\tau \in \mathbb{R}^{n_\tau}$, ranging over a compact set $T \subset \mathbb{R}^{n_\tau}$, after the process picks an ε, we modify (2) to

$$\psi_{E,T}(x) = \max_{\varepsilon \in E} \min_{\tau \in T} \max_{k \in K} \zeta^k(x,\varepsilon,\tau) \leq 0, \qquad (3)$$

since τ must work for all the specifications $k \in K$, with $K = \{1,2,\ldots,m\}$. For a more detailed exposition of the design centering, tolerancing and trimming problem formulation as an optimization problem, the reader is referred to [3,4].

Although Lipschitz continuous [3], the function $\psi_{E,T}(\cdot)$ is, in general, not differentiable; in fact, it may even fail to have directional derivatives. In [3] we find a conceptual algorithm for optimization problems with constraints of the form (3), based on the concept of generalized gradients [11,12,13]. The algorithm in [3] consists of two parts: a master outer approximations algorithm, which replaces the set E in (3) with a discrete subsets $E_i \subset E$ (as in [14]), and an inner, non-differentiable optimization subalgorithm which solves the resulting simpler problems. It is shown that the solutions of the simpler problems converge to a solution of the original problem. The inner, nondifferentiable optimization subalgorithm in [3] has two serious drawbacks. The first, common to many nondifferentiable optimization algorithms, is that it utilizes a very expensive bisection procedure to get adequate approximations to required bundles of subgradients (see, for example, [18,15,16]). The second, because of a requirement of semi-smoothness [13], is that it is applicable only to the case where there is only one constraint function ζ^k (i.e., $K = \{k\}$).

In this paper we present a new, implementable algorithm for solving optimization problems with constraints of the form (3). Just as the algorithm in [3,14], it makes use of outer approximations to the feasible set to decompose the original problem into an infinite sequence of simpler problems, by replacing the set E with discrete subsets E_i. However, because of the use of certain transformations, the resulting simpler problems are ordinary, differentiable, constrained optimization problems, solvable by a large number of existing, efficient algorithms. Thus, as a consequence of these simple, but not immediately obvious transformations, all the computational difficulties caused by the need for a nondifferentiable optimization subalgorithm in [3] have been removed. Truncation rules are given for all the major infinite operations in the new algorithm, making it implementable and hence easily programmable. Finally, it is shown in the Appendix that the optimality conditions, on which the present algorithm is based, are sharper than the ones used in [3].

It is our hope that this new algorithm will become a valuable tool in the arsenal of the engineering designer.

2. Problem Decomposition via Outer Approximations to the Feasible Set

The most general form of an optimization problem arising in engineering design, that we shall consider in this paper, is as follows:

$$
\begin{aligned}
P_g:\ \min_{x}\ \{f^0(x)\,|\,f^i(x) &\le 0,\ i \in I;\\
\max_{\omega^j \in \Omega^j}\ \phi^j\,(x,\omega^j) &\le 0,\ j \in J;\\
\max_{\varepsilon \in E}\ \min_{\tau \in T}\ \max_{k \in K}\ \zeta^k\,(x,\varepsilon,\tau) &\le 0\}
\end{aligned}
\tag{4}
$$

where I,J,K are sets of integers (e.g., $K=\{1,2,\dots,m\}$) $f^0:\mathbb{R}^n \to \mathbb{R}^1$, $f^i:\mathbb{R}^n \to \mathbb{R}^1$, $i \in I$, $\phi^j:\mathbb{R}^n \times \mathbb{R}^{(n_\omega)_j} \to \mathbb{R}^j$, $j \in J$, and $\zeta^k:\mathbb{R}^n \times \mathbb{R}^{n_\varepsilon} \times \mathbb{R}^{n_\tau} \to \mathbb{R}^1$ $k \in K$, are all continuously differentiable. In this context, x is the nominal design vector; the ω^j are tolerances, errors, or variables, such as frequency, or time, or temperature, which must be considered over a continuum of values; ε is an error or tolerance to be overcome by tuning, and τ is the tuning parameter. We shall assume that the sets Ω^j, E and T are <u>compact</u> and specified by differentiable inequalities, which we shall introduce, as needed, later.

There is no essential loss of generality, as far as the exposition of our method is concerned, in assuming that I and J contain only one index each, i.e., in considering only one constraint of each kind. On the other hand, there is a considerable simplification in notation when the indices i and j are eliminated. We shall therefore restrict ourselves to the simpler problem

$$
\begin{aligned}
P:\ \min_{x}\ \{f^0(x)\,|\,f(x) \le 0,\ &\max_{\omega \in \Omega}\ \phi(x,\omega) \le 0,\\
\max_{\varepsilon \in E}\ \min_{\tau \in T}\ \max_{k \in K}\ &\zeta^k(x,\varepsilon,\tau) \le 0\}
\end{aligned}
\tag{5}
$$

where $f:\mathbb{R}^n \to \mathbb{R}^1$, $\phi:\mathbb{R}^n \times \mathbb{R}^{n_\omega} \to \mathbb{R}^1$, are continuously differentiable, $\Omega \subset \mathbb{R}^{n_\omega}$, and $E \subset \mathbb{R}^{n_\varepsilon}$ and $T \subset \mathbb{R}^{n_\tau}$ are compact and the remaining quantities are as in P_g.

Now, let $\Omega_i \subset \Omega$ and $E_i \subset E$ be discrete sets and consider the problem

$$P_i: \min_x \{f^0(x) \mid f(x) \leq 0, \max_{\omega \in \Omega_i} \phi(x,\omega) \leq 0,$$

$$\max_{\varepsilon \in E_i} \min_{\tau \in T} \max_{k \in K} \zeta^k(x,\varepsilon,\tau) \leq 0\}. \tag{6}$$

We shall show that unlike the problem P, P_i is an ordinary nonlinear programming problem. Thus, suppose that

$$\Omega_i = \{\omega_j \in \Omega \mid j \in J_i\},\ E_i = \{\varepsilon_\ell \in E \mid \ell \in L_i\}, \tag{6a}$$

with J_i, L_i <u>finite</u> sets of integers and let

$$\phi^j(x) \triangleq \phi(x,\omega_j),\ j \in J_i \tag{6b}$$

$$\zeta^{k\ell}(x,\tau) \triangleq \zeta^k(x,\varepsilon_\ell,\tau),\ \ell \in L_i,\ k \in K\ . \tag{6c}$$

Now suppose that

$$T \triangleq \{\tau \mid g^s(\tau) \leq 0,\ s \in S\} \tag{6d}$$

with $S \triangleq \{1,2,\ldots,\sigma\}$, $g^s:\mathbb{R}^{n_\tau} \to \mathbb{R}^1$ continuously differentiable, and consider the problem

$$\hat{P}_i: \min_{(x,\tau_\ell)} \{f^0(x) \mid f(x) \leq 0;\ \phi^j(x) \leq 0,\ j \in J_i;$$

$$\zeta^{k\ell}(x,\tau_\ell) \leq 0,\ k \in K,\ \ell \in L_i;\ g^s(\tau_\ell) \leq 0,$$

$$s \in S,\ \ell \in L_i\} \tag{6e}$$

<u>Proposition 1</u>: $\hat{x}$ solves P_i if and only if $\{\hat{x};\hat{\tau}_\ell,\ell\in L_i\}$ solves $\hat{P}_i$, for some $\hat{\tau}_\ell\in T$, $\ell\in L_i$.

<u>Proof</u>: $\Rightarrow$ Suppose $\hat{x}$ solves P_i. Then $\phi^j(\hat{x})\le 0$ for $j\in J_i$ and there exist $\hat{\tau}_\ell\in T$, $\ell\in L_i$ such that $\max_{\ell\in L_i}\max_{k\in K}\zeta^{k\ell}(\hat{x},\hat{\tau}_\ell)\le 0$, i.e., $\{\hat{x},\hat{\tau}_\ell,\ell\in L_i\}$ are feasible for $\hat{P}_i$. Now suppose that that this triplet is not optimal for $\hat{P}_i$. Then there exist $\{\tilde{x};\tilde{\tau}_\ell,\ell\in L_i\}$ feasible for $\hat{P}_i$, and such that $f^0(\tilde{x})<f^0(\hat{x})$. Now, $\tilde{x}$ satisfies $f(\tilde{x})\le 0$, $\phi^j(\tilde{x})\le 0$, $j\in J_i$ and

$$\max_{\varepsilon\in E_i}\ \min_{\tau\in T}\ \max_{k\in K}\ \zeta^k(\tilde{x},\varepsilon,\tau)$$

$$=\max_{\ell\in L_i}\ \min_{\tau\in T}\ \max_{k\in K}\ \zeta^{k\ell}(\tilde{x},\tau)$$

$$\le\max_{\ell\in L_i}\ \max_{k\in K}\ \zeta^{k\ell}(\tilde{x},\tilde{\tau}_\ell)\le 0 \tag{6d}$$

i.e., $\tilde{x}$ is feasible for P_i. But then $f^0(\tilde{x})<f^0(\hat{x})$ contradicts the optimality of $\hat{x}$.

$\Leftarrow$ Now suppose that $\{\hat{x};\hat{\tau}_\ell,\ell\in L_i\}$ solves $\hat{P}_i$. Then, because of (6d), $\hat{x}$ is feasible for P_i. Suppose that $\hat{x}$ is not optimal for P_i. Then there exists an $\tilde{x}$ and corresponding $\tilde{\tau}_\ell,\ell\in L_i$ such that $\tilde{x}$ is feasible for P_i, $f^0(\tilde{x})<f^0(\hat{x})$ and $\max_{\ell\in L_i}\ \min_{\tau\in T}\ \max_{k\in K}\ \zeta^{k\ell}(x,\tau)=$ $\max_{\ell\in L_i}\ \max_{k\in K}\ \zeta^{k\ell}(\hat{x},\tilde{\tau}_\ell)\le 0$, so that $\{\tilde{x};\tilde{\tau}_\ell,\ell\in L_i\}$ is feasible also for $\hat{P}_i$. But this contradicts the optimality of $\hat{x}$ and we are done.

¤

Next, let $\chi:\mathbb{R}^n\times\mathbb{R}^{n_\varepsilon}\to\mathbb{R}$ be defined by

$$\chi(x,\varepsilon)\triangleq\min_{\tau\in T}\ \max_{k\in K}\ \zeta^k(x,\varepsilon,\tau) \tag{7}$$

It was shown in [3] that $\chi(\cdot,\cdot)$ is Lipschitz continuous. Since

$\Omega_i \subset \Omega$ and $E_i \subset E$,

$$\{x \mid \max_{\omega \in \Omega} \phi(x,\omega) \le 0\} \subset \{x \mid \max_{\omega \in \Omega_i} \phi(x,\omega) \le 0\} \tag{8}$$

and

$$\{x \mid \max_{\varepsilon \in E} \chi(x,\varepsilon) \le 0\} \subset \{x \mid \max_{\varepsilon \in E_i} \chi(x,\varepsilon) \le 0\}. \tag{9}$$

Consequently, the feasible set in (6) contains the feasible set in (5), and therefore, if $\hat{x}$ solves P and $\hat{x}_i$ solves P_i, we must have

$$f^0(\hat{x}_i) \le f^0(\hat{x}) \tag{10}$$

Now suppose that we construct an infinite sequence of problems P_i, with solutions $\hat{x}_i$ such that $\hat{x}_i \to \tilde{x}$ as $i \to \infty$ and $\tilde{x}$ is feasible for P, i.e.,

$$\tilde{x} \in F \triangleq \{x \mid f(x) \le 0,\ \max_{\omega \in \Omega} \phi(x,\omega) \le 0,\ \max_{\varepsilon \in E} \chi(x,\varepsilon) \le 0\}. \tag{11}$$

Then, by continuity of f^0 and (10)

$$f^0(\tilde{x}) \le f^0(\hat{x}) \tag{12}$$

and hence $\tilde{x}$ is optimal for P. Thus, since we can solve problems of the form $\hat{P}_i$, the solution of problem P can be assured by constructing discrete sets Ω_i, E_i such that any accumulation point, $\hat{x}$, of a solution sequence $\{\hat{x}_i\}$ satisfies $\hat{x} \in F$. In [14] we find two specific schemes for achieving this result. We shall now summarize the most relevant results from [14].

<u>Proposition 1</u>: Let $\{\delta_i\}_{i=0}^{\infty}$ and $\{\hat{\delta}_i\}_{i=0}^{\infty}$ be a positive, monotonically decreasing sequences, with $\delta_i \to 0$, $\hat{\delta}_i \to 0$ as $i \to \infty$. Let $E_0 \subset E$ and $\Omega_0 \subset \Omega$ be discrete sets. For $i = 0,1,2,\ldots$ let $x_i \in \mathbb{R}^n$, $\Omega_i \subset \Omega$ and $E_i \subset E$ be defined recursively as follows: let x_0 be given and let x_i be such that

$$f(x_i) \leq \delta_i, \tag{13a}$$

$$\max_{\omega \in \Omega_i} \phi(x_i,\omega) \leq \delta_i, \tag{13b}$$

$$\max_{\varepsilon \in E_i} \chi(x_i,\varepsilon) \leq \delta_i. \tag{13c}$$

Let ω_i, ε_i be such that

$$\phi(x_i,\omega_i) = \max_{\omega \in \Omega_i} \phi(x_i,\omega); \tag{14a}$$

$$\chi(x_i,\varepsilon_i) = \max_{\varepsilon \in E_i} \chi(x_i,\varepsilon) \tag{14b}$$

and let

$$\begin{aligned} \Omega_{i+1} &= \Omega_i \cup \{\omega_i\} \quad \text{if} \quad \phi(x_i,\omega_i) > \hat{\delta}_i \\ &= \Omega_i \text{ otherwise,} \end{aligned} \tag{15a}$$

$$\begin{aligned} E_{i+1} &= E_i \cup \{\varepsilon_i\} \quad \text{if} \quad \chi(x_i,\varepsilon_i) > \hat{\delta}_i \\ &= E_i \text{ otherwise.} \end{aligned} \tag{15b}$$

Then any accumulation point $\hat{x}$ of $\{x_i\}_{i=0}^{\infty}$ is in F.

<u>Proof</u>: Suppose that $x_i \xrightarrow{I} \hat{x}$, with $I \subset \{1,2,3,\ldots\}$. Then, since $f(\cdot)$ is continuous and $\delta_i \to 0$,

$$f(\hat{x}) \leq 0 \tag{16}$$

Next, since the functions

$$\psi_\Omega(x) \triangleq \max_{\omega \in \Omega} \phi(x,\omega) \tag{17a}$$

and

$$\psi_{E,T}(x) \triangleq \max_{\varepsilon \in E} \chi(x,\varepsilon) \tag{17b}$$

are both continuous, $\psi_{\Omega}(x_i) \overset{I}{\to} \psi_{\Omega}(\hat{x})$ and $\psi_{E,T}(x_i) \to \psi_{E,T}(\hat{x})$. For the sake of contradiction, suppose that $\psi_{\Omega}(\hat{x}) > 0$. Then, since $x_i \overset{I}{\to} \hat{x}$ and $\delta_i \to 0$, there exists an $i_0 \in I$ such that $\psi_{\Omega}(x_i) > \hat{\delta}_i$ for all $i \in I$, $i \geq i_0$ and hence for any $i_2 > i_1 > i_0$ in I,

$$\Omega_{i_2} \supset \{\omega_i \mid i \in I,\ i_0 \leq i \leq i_1\} \tag{18}$$

Because of (13b), we then have

$$\phi(x_{i_2},\omega_i) \leq \delta_{i_2} \text{ for all } i \in I,\ i_0 \leq i \leq i_1 \tag{19}$$

and therefore, in particular, $\lim \phi(x_{i_2},\omega_{i_1}) \leq 0$ as $i_2 > i_1 \to \infty$, $i_2, i_1 \in I$. But, because Ω is compact and $x_i \overset{I}{\to} \hat{x}$,

$$|\phi(x_{i_2},\omega_{i_1}) - \phi(x_{i_1},\omega_{i_1})| \to 0 \tag{20}$$

as $i_2 > i_1 \to \infty$, $i_2, i_1 \in I$. Because of (19) and (20) we now obtain that

$$\psi_{\Omega}(\hat{x}) = \lim_{i_1 \in I} \phi(x_{i_1},\omega_{i_1}) \leq 0 \tag{21}$$

and we have a contradiction. In exactly the same way it can be shown that $\psi_{E,T}(\hat{x}) \leq 0$ and hence the proof is complete. ¤

The simplest way of incorporating the construction of the Ω_i, E_i, described in Proposition 2, into a master algorithm for solving P is as follows (cf. [14]).

<u>Master Algorithm 1</u>

<u>Parameters</u>: $\{\hat{\delta}_i\}_{i=0}^{\infty}$, $\hat{\delta}_i > 0$, $\hat{\delta}_i \to 0$ as $i \to \infty$.

<u>Data</u>: $\Omega_0 \subset \Omega$, $E_0 \subset E$, discrete sets.

Step 0: Set $i = 0$.

Step 1: Solve $\hat{P}_i$ for $\{\hat{x}_i;\hat{\tau}_\ell,\ell\in L_i\}$.

Step 2: Compute $\psi_\Omega(\hat{x}_i)$, $\psi_{E,T}(\hat{x}_i)$ and corresponding $\omega_i\in\Omega$ and $\varepsilon_i\in E$, satisfying (14a,b) for $x_i = \hat{x}_i$.

Step 3: a) If $\psi_\Omega(\hat{x}_i) > \hat{\delta}_i$, set $\Omega_{i+1} = \{\omega_i\}\cup\Omega_i$. Else set $\Omega_{i+1} = \Omega_i$.

b) If $\psi_{E,T}(\hat{x}_i) > \hat{\delta}_i$, set $E_{i+1} = \{\varepsilon_i\}\cup E_i$. Else set $E_{i+1} = E_i$.

Step 4: If $\psi_\Omega(\hat{x}_i) \leq 0$ and $\psi_{E,T}(x_i) \leq 0$, stop; else proceed.

Step 5: Set $i = i+1$ and go to Step 1. ¤

The following result is obvious in view of Proposition 2:

Theorem 1: If Master Algorithm 1 stops in Step 4, then $\hat{x}_i$ is optimal for P. If Master Algorithm 1 constructs an infinite sequence $\{\hat{x}_i\}$, then any accumulation point $\hat{x}$ of $\{\hat{x}_i\}$ is optimal for P. ¤

Master Algorithm 1 has the disadvantage that the sets Ω_i and E_i grow without bound. To see this, suppose that $\psi_\Omega(\hat{x}_i) \leq \hat{\delta}_i$ and $\psi_{E,T}(\hat{x}_i) \leq \hat{\delta}_i$. Then the algorithm simply increases the index i to i+1, sets $\hat{x}_{i+1} = \hat{x}_i$ and continues doing so until $\hat{\delta}_i$ declines enough for either $\psi_\Omega(\hat{x}_i) > \hat{\delta}_i$ or $\psi_{E,T}(\hat{x}_i) > \hat{\delta}_i$ to take place. In [14] we find a way of circumventing this undesirable phenomenon. Let $\{\delta_{ij}\}$ be a sequence such that $\delta_{ii} = 0$, $\delta_{ij} > 0$ for $i > j$ and let $\delta_{ij} \to \hat{\delta}_j$, as $i\to\infty$, (e.g., $\delta_{ij} = 10 \max \{0, 1/(i+j)-1/(1+i)\}$) and suppose that we include ω_j in E_i and ε_j in E_i for all $i \geq j+1$ such that $\psi_\Omega(x_j) > \delta_{i-1,j}$ or $\psi_{E,T}(x_j) > \delta_{i-1,j}$, respectively. Now, suppose that $\psi_\Omega(x_j) \leq \hat{\delta}_j$. Then ω_j would not be included in Ω_{j+1} in Master Algorithm 1, but under the new scheme, ω_j would be included in Ω_{j+1}, if $\psi_\Omega(x_j) > 0$ and, furthermore, it would be retained for a certain number of iterations in Ω_i, $i = j+1,j+2,\ldots$, until $\psi_\Omega(x_j) \leq \delta_{i-1,j}$ took place. It would then be dropped, never to be used again. As a result, we get a scheme which tends to keep the cardinality of the Ω_i and E_i small. To make the cardinality of the Ω_i and E_i as small as possible, the elements of the sequence $\{\hat{\delta}_i\}$ should be large and decay to zero as slowly as possible. To formalize this discussion we summarize it in the form of

<u>Master Algorithm 2 [14]</u>

<u>Parameters</u>: $\{\delta_{ij}\}_{i,j=0}^{\infty}$ such that

a) $\delta_{ij} = 0$ for all $i \leq j$, and $\delta_{ij} > 0$ otherwise.

b) $\delta_{ij} \nearrow \hat{\delta}_j$ as $i \to \infty$.

c) $\delta_j \to 0$ as $j \to \infty$.

<u>Data</u>: $\Omega_0 \subset \Omega$, $E_0 \subset E$, discrete sets.

<u>Step 0</u>: Set $i = 0$.

<u>Step 1</u>: Solve $\hat{P}_i$ for $(\hat{x}_i, \hat{\tau}_{i_1}, \ldots \hat{\tau}_{i_{t_i}})$.

<u>Step 2</u>: Compute $\psi_\Omega(\hat{x}_i)$, $\psi_{E,T}(\hat{x}_i)$ and corresponding $\omega_i \in \Omega$, $\varepsilon_i \in E$ satisfying (14a,b) for $x_i = \hat{x}_i$.

<u>Step 3</u>: Include ω_j in Ω_{i+1} for all $0 \leq j \leq i$ such that $\psi_\Omega(\hat{x}_j) > \delta_{ij}$, and include ε_j in E_{i+1}, for all $0 \leq j \leq i$ such that $\psi_{E,T}(\hat{x}_j) > \delta_{ij}$.

<u>Step 4</u>: If $\psi_\Omega(\hat{x}_i) \leq 0$ and $\psi_{E,T}(\hat{x}_i) \leq 0$ stop. Else proceed.

<u>Step 5</u>: Set $i = i+1$ and go to Step 1. ¤

Since $\delta_{ij} \leq \delta_j$ for all $i \geq j$, the following theorem is a trivial consequence of Theorem 1.

<u>Theorem 2</u>: If Master Algorithm 2 stops in Step 4, then $\hat{x}_i$ is optimal for P. If Master Algorithm 2 constructs an infinite sequence $\{\hat{x}_i\}$, then any accumulation point $\hat{x}$ of $\{\hat{x}_i\}$ is optimal for P. ¤

So far we have assumed that we can solve the problems $\hat{P}_i$ exactly and also evaluate the functions $\psi_\Omega(\cdot)$ and $\psi_{E,T}(\cdot)$ exactly. In the next section we shall consider what happens when one solves the problems $\hat{P}_i$ only to the extent of finding approximate stationary points and when one evaluates the functions $\psi_\Omega(\cdot)$ and $\psi_{E,T}(\cdot)$ approximately, as will be the case in practice. We shall then summarize our findings in the form of an implementable algorithm.

3. The Implementable Algorithm

First we shall discuss the evaluation of the functions $\psi_\Omega(\cdot)$ and $\psi_{E,T}(\cdot)$, defined in (17a,b). We shall assume that

$$\Omega \triangleq \{\omega \mid h^j(\omega) \leq 0,\ j \in H_\omega\} \tag{22}$$

with $h^j : \mathbb{R}^{n_\omega} \to \mathbb{R}^1$ continuously differentiable and H_ω a finite set of integers. We then see that

$$\psi_\Omega(x) = \max\{\phi(x,\omega) \mid h^j(\omega) \leq 0,\ j \in H_\omega\} \tag{23}$$

is defined by an ordinary nonlinear programming problem, the difficulty of which depends entirely on the nature of the functions $\phi(x,\cdot)$ and $h^j(\cdot)$. In engineering design situations, the constraints $h^j(\omega) \leq 0$, $j \in H_\omega$, tend to be very simple and (23) is not too difficult to solve, at least approximately.

Next, we consider $\psi_{E,T}(x)$. First, we note that

$$\begin{aligned}\psi_{E,T}(x) &= \max_{\varepsilon \in E}\{\min_{\tau \in T} \max_{k \in K} \zeta^k(x,\varepsilon,\tau)\}\\ &= \max_{(\varepsilon^0,\varepsilon)} \{\varepsilon^0 \mid \min_{\tau \in T} \max_{k \in K} \zeta^k(x,\varepsilon,\tau) - \varepsilon^0 \geq 0\ ,\ \varepsilon \in E\}\end{aligned} \tag{24}$$

If we substitute a discrete set $T_i \subset T$ for T in (24), we get a larger feasible set and hence (24) can be evaluated by a straightforward modification of Master Algorithm 2, such as the one to be presented shortly. The use of this subalgorithm requires the development of the following details. Suppose that

$$E \triangleq \{\varepsilon | p^j(\varepsilon) \leq 0,\ j \in H_\varepsilon\} \tag{25}$$

with $p^j : \mathbb{R}^{n_\varepsilon} \to \mathbb{R}^1$ all continuously differentiable and H_ε a finite set of integers, and that $T_i = \{\tau_j \in T | j \in G_i\}$, with G_i a finite set of integers. Then the discretized problem which the outer approximations subalgorithm 1, below requires to be solved at each iteration is

$$\max_{(\varepsilon^0,\varepsilon)} \{\varepsilon^0 | \min_{\tau \in T_i} \max_{k \in K} \zeta^k(x,\varepsilon,\tau) - \varepsilon^0 \geq 0;\ p^j(\varepsilon) \leq 0,\ j \in H_\varepsilon\}$$

$$= \max_{(\varepsilon^0,\varepsilon)} \{\varepsilon^0 | \max_{k \in K} \zeta^k(x,\varepsilon,\tau_\ell) - \varepsilon^0 \geq 0, \ell \in G_i;\ p^j(\varepsilon) \leq 0,\ j \in H_\varepsilon\}$$

$$= \max_{k \in K} \max_{(\varepsilon^0,\varepsilon)} \{\varepsilon^0 | \nu - \varepsilon^0 \geq 0;\ \zeta^k(x,\varepsilon,\tau_\ell) - \nu \leq 0,\ \ell \in G_i;\ p^j(\varepsilon) \leq 0,$$

$$j \in H_\varepsilon\} \tag{26}$$

which is seen to be a set of ordinary nonlinear programming problems. However, since one has to solve a number of such problems to get a reasonable approximation to $\psi_{E,T}(x)$, it is clear that the computation of approximations to $\psi_{E,T}(x)$ will be the most time consuming operation at each iteration of our outer approximations algorithm for solving P. For the sake of clarity, it seems preferable to state the subalgorithm for computing $\psi_{E,T}(x)$ in conceptual form.

<u>Subalgorithm 1</u>: Evaluates $\psi_{E,T}(x)$.

<u>Parameters</u>: $\{\delta_{ij}\}$ as in Master Algorithm 2.

<u>Data</u>: $x \in \mathbb{R}^n$, $T_0 \subset T$, a discrete set.

Step 0: Set $i = 0$.

Step 1: Solve (26) for $(\varepsilon_i^0, \varepsilon_i)$.

Step 2: Compute $\chi(x,\varepsilon_i)$ and τ_i by solving

$$\chi(x,\varepsilon_i) = \min_{\tau\in T} \max_{k\in K} \zeta^k(x,\varepsilon_i,\tau)$$

$$= \min_{(\tau^0,\tau)} \{\tau^0 | \zeta^k(x,\varepsilon_i,\tau) - \tau^0 \leq 0,\ k\in K,\ g^s(\tau) \leq 0,\ s\in S\} \quad (27)$$

Step 3: Include τ_j in T_{i+1} for all $0 \leq j \leq i$ such that $\varepsilon_j^0 - \chi(x,\hat{\varepsilon}_j) > \delta_{ij}$.

Step 4: If $(\chi(x,\varepsilon_j) - \varepsilon_j^0) \geq 0$, stop; else proceed.

Step 5: Set $i = i+1$ and go to Step 1. ¤

Since Subalgorithm 1 is merely a transcription of the Master Algorithm 2 to the solution of problem (24), the following result is a direct consequence of Theorem 2.

Theorem 3: If Subalgorithm 1 stops in Step 4, then $\psi_{E,T_\infty}(x) = \max_{k\in K} \zeta^k(x,\varepsilon_i,\tau_i)$. If Subalgorithm 1 constructs an infinite sequence $\{(\hat{\varepsilon}_i,\hat{\tau}_i)\}_{i=0}$, then any accumulation point $(\hat{\varepsilon},\hat{\tau})$ of this sequence satisfies
$\psi_{E,T}(x) = \max_{k\in K} \zeta^k(x,\hat{\varepsilon},\hat{\tau})$. ¤

Next we turn towards constructing an implementation for Master Algorithm 2 which will, by the same token, yield an implementation for Subalgorithm 1. For this purpose we must establish appropriate optimality conditions for P and the subproblems $\hat{P}_i$. Returning to problem $\hat{P}_i$ in (6e), it is clear that it can be rewritten in the following more convenient form:

$$\hat{P}_i: \min_{(x,y_{E_i})} \{f^0(x) | f(x) \leq 0;\ \phi(x,\omega) \leq 0,\ \omega\in\Omega_i;$$

$$\zeta^k(x,\varepsilon,\tau_\varepsilon) \le 0,\ k\in K,\ \varepsilon\in E_i\ \ g^s(\tau_\varepsilon) \le 0,$$
$$s\in S,\ \varepsilon\in E_i\}, \tag{28}$$

where

$$y_{E_i} \triangleq \{\tau_\varepsilon | \varepsilon\in E_i\}\ . \tag{28a}$$

Next, we define

$$\psi_{\Omega_i}(x) \triangleq \max_{\omega\in\Omega_i} \phi(x,\omega),$$

$$\psi_{E_i}(x,y_{E_i}) \triangleq \max_{\substack{\varepsilon\in E_i\\ k\in K}} \zeta^k(x,\varepsilon,\tau_\varepsilon), \tag{28b}$$

$$\psi_{T,E_i}(y_{E_i}) \triangleq \max_{\substack{\varepsilon\in E_i\\ s\in S}} g^s(\tau_\varepsilon), \tag{28c}$$

and finally we define

$$\psi_{\hat{P}_i}(x,y_{E_i}) \triangleq \max\{0, f(x), \psi_{\Omega_i}(x),\ \psi_{E_i}(x,y_{E_i}),\ \psi_{T,E_i}(y_{E_i})\}. \tag{28d}$$

<u>Proposition 3</u>: Let $E_i = \{\varepsilon_\ell \in E | \ell\in L_i\}$ and let

$$\theta_{\Omega_i,E_i}(x,y_{E_i}) \triangleq \min_{\substack{\|h\|_\infty\le 1\\ \|\eta_\ell\|_\infty \le 1}} \max\{\langle \nabla f^0(x),h\rangle;$$

$$f(x) + \langle \nabla f(x),h\rangle;\ \phi(x,\omega) + \langle \nabla_x\phi(x,\omega),h\rangle,$$

$$\omega\in\Omega_i;\ \zeta^k(x,\varepsilon_\ell,\tau_\ell) + \langle \nabla_x\zeta^k(x,\varepsilon_\ell,\tau_\ell),h\rangle$$

$$+\ \langle \nabla_\tau\zeta^k(x,\varepsilon_\ell,\tau_\ell),\eta_\ell\rangle\ \ \ell\in L_i;$$

$$g^s(\tau_\ell) + \langle \nabla_\tau g^s(\tau_\ell),\eta_\ell\rangle,\ s\in S,\ \ell\in L_i\}$$

$$-\ \psi_{\hat{P}_i}(x,y_{E_i}) \tag{28e}$$

If $(\hat{x},\hat{y}_{E_i})$ is optimal for $\hat{P}_i$ then $\theta_{\Omega_i,E_i}(\hat{x},\hat{y}_{E_i}) = 0$.

Proof: $\theta_{\Omega_i,E_i}(\hat{x},\hat{y}_{E_i}) = 0$ is the Topkis-Veinott necessary optimality condition which was shown in [17] to be equivalent to the F. John optimality condition. ¤

To extend this optimality condition to problem P, we proceed as follows. For every $\varepsilon \in E$, let τ_ε be some vector in $\mathbb{R}^{n_\tau}$. Then we see that P is equivalent to the following problem

$$\hat{P}: \min_{(x,y_E)} \{f^0(x) \,|\, f(x) \leq 0;\ \phi(x,\omega) \leq 0\ \forall\omega\in\Omega;$$

$$\zeta^k(x,\varepsilon,\tau_\varepsilon) \leq 0,\ \forall\ k\in K,\ \forall\ \varepsilon\in E;$$

$$g^s(\tau_\varepsilon) \leq 0,\ \text{for } s\in S,\ \forall\ \varepsilon\in E\} \tag{29}$$

with

$$y_E \triangleq \{\tau_\varepsilon \,|\, \varepsilon\in E\}. \tag{29a}$$

The equivalence between P and $\hat{P}$ is in the sense $\hat{x}$ solves P $\Longleftrightarrow$ $(\hat{x},\hat{y}_E)$ solves $\hat{P}$. Next, let $\psi_{\hat{P}}(x,y_E)$ be defined as in (28d), with E replacing E_i and let

$$\theta_{\Omega,E}(x,y_E) \triangleq \min_{\substack{\|h\|_\infty\leq 1\\ \|\eta_\varepsilon\|_\infty\leq 1}} \max\ \{\langle \nabla f^0(x),h\rangle;$$

$$f(x) + \langle \nabla f(x),h\rangle;\ \phi(x,\omega) + \langle \nabla_x\phi(x,\omega),h\rangle, \omega\in\Omega\ ;$$

$$\zeta^k(x,\varepsilon,\tau_\varepsilon) + \langle \nabla_x\zeta^k(x,\varepsilon,\tau_\varepsilon),h\rangle + \langle \nabla_\tau\zeta^k(x,\varepsilon,\tau_\varepsilon),\eta_\varepsilon\rangle,\ k\in K,\varepsilon\in E;$$

$$g^s(\tau_\varepsilon) + \langle \nabla_\tau g^s(\tau_\varepsilon),\eta_\varepsilon\rangle\ ,\ s\in S,$$

$$\varepsilon\in E\,\} - \psi_{\hat{P}}(x,y_E), \tag{29b}$$

<u>Theorem 3</u>: Suppose that $(\hat{x},\hat{y}_E)$ is optimal for $\hat{P}$ (29), then $\theta_{\Omega,E}(\hat{x},\hat{y}_E) = 0$.

<u>Proof</u>: By construction, $\theta_{\Omega,E}(x,y_E) \leq 0$ for any (x,y_E), and since $(\hat{x},\hat{y}_E)$ are feasible for $\hat{P}$, $\psi_{\hat{P}}(\hat{x},\hat{y}_E) = 0$. To obtain a contradiction, suppose that $\theta_{\Omega,E}(\hat{x},\hat{y}_E) = -\hat{\delta} < 0$. Then there exists a vector $\hat{h}$, with $\|\hat{h}\|_\infty \leq 1$, such that

$$\langle \nabla f^0(\hat{x}),\hat{h}\rangle \leq -\hat{\delta} \tag{30}$$

and hence there exists a $\lambda_0 > 0$ such that

$$f^0(\hat{x}+\lambda\hat{h}) - f^0(\hat{x}) \leq -\lambda\hat{\delta}/2 \tag{31}$$

for all $\lambda \in (0,\lambda_0]$. Next, we have

$$f(\hat{x}) + \langle \nabla f(\hat{x}),h\rangle \leq -\hat{\delta} \tag{32}$$

and hence, for $\lambda \in [0,1]$,

$$f(\hat{x}) + \lambda\langle \nabla f(\hat{x}),h\rangle \leq -\lambda\hat{\delta} + (1-\lambda)f(\hat{x}) \leq -\lambda\hat{\delta}, \tag{33}$$

since $f(\hat{x}) \leq 0$. Consequently, there exists a $\lambda_1 \in (0,\lambda_0]$ such that

$$f(\hat{x}+\lambda\hat{h}) \leq 0, \text{for all } \lambda \in (0,\lambda_1]. \tag{34}$$

Similarly,

$$\phi(\hat{x},\omega) + \langle \nabla_x\phi(\hat{x},\omega),\hat{h}\rangle \leq -\hat{\delta} \tag{35}$$

implies that for $\lambda \in [0,1]$, since $\phi(\hat{x},\omega) \leq 0$ for all $\omega \in \Omega$,

$$\begin{aligned}\phi(\hat{x},\omega) + \lambda\langle \nabla_x\phi(\hat{x},\omega),\hat{h}\rangle &\leq -\lambda\hat{\delta} + (1-\lambda)\phi(\hat{x},\omega)\\ &\leq -\lambda\hat{\delta},\end{aligned} \tag{36}$$

and hence, since Ω is compact, there exists a $\lambda_2 \in (0,\lambda_1]$ such that

$$\max_{\omega\in\Omega} \phi(\hat{x}+\lambda\hat{h},\omega) = \max_{\omega\in\Omega} \phi(\hat{x},\omega) + \lambda\langle \nabla_x\phi(\hat{x},\omega),\hat{h}\rangle$$

$$+ \lambda\int_0^1 \langle (\nabla_x\phi(\hat{x}+s\lambda\hat{h},\omega) - \nabla\phi(\hat{x},\omega)),\hat{h}\rangle\, ds$$

$$\leq - \lambda[\hat{\delta}+\max_{\omega\in\Omega}\int_0^1 \|\nabla_x\phi(\hat{x}+s\lambda\hat{h}) - \nabla\phi(\hat{x},\omega)\| ds]$$

$$\leq 0, \text{ for all } \lambda \in [0,\lambda_2]. \tag{37}$$

Next, there exist $\hat{\eta}_\varepsilon$, with $\|\hat{\eta}_\varepsilon\|_\infty \leq 1$, such that

$$\zeta^k(\hat{x},\varepsilon,\hat{\tau}_\varepsilon) + \langle \nabla_x\zeta^k(\hat{x},\varepsilon,\hat{\tau}_\varepsilon),\hat{h}\rangle + \langle \nabla_\tau\zeta^k(\hat{x},\varepsilon,\hat{\tau}_\varepsilon),\hat{\eta}_\varepsilon\rangle \leq -\hat{\delta}, \tag{38}$$

and hence for $\lambda\in[0,1]$,

$$\zeta^k(\hat{x},\varepsilon,\hat{\tau}_\varepsilon) + \lambda\{\langle \nabla_x\zeta^k(\hat{x},\varepsilon,\hat{\tau}_\varepsilon),\hat{h}\rangle +$$
$$\langle \nabla_\tau\zeta^k(\hat{x},\varepsilon,\hat{\tau}_\varepsilon),\hat{\eta}_\varepsilon\rangle\} \leq - \lambda\hat{\delta} + (1-\lambda)\zeta^k(\hat{x},\varepsilon,\hat{\tau}_\varepsilon)$$
$$\leq - \lambda\hat{\delta}, \text{ for all } k\in K, \text{ and } \varepsilon\in E, \tag{39}$$

because $\zeta^k(\hat{x},\varepsilon,\hat{\tau}_\varepsilon) \leq 0$ for all $k\in K$ and $\varepsilon\in E$. Hence, since E and T are compact and $\hat{h},\hat{\eta}_\varepsilon$ are bounded, we conclude, as in (37), that there exists a $\lambda_3 \in (0,\lambda_2]$ such that

$$\max_{\varepsilon\in E}\min_{\tau\in T}\max_{k\in K} \zeta^k(\hat{x}+\lambda\hat{h},\varepsilon,\tau) \leq$$
$$\max_{\varepsilon\in E}\max_{k\in K} \zeta^k(\hat{x}+\lambda\hat{h},\varepsilon,\hat{\tau}_\varepsilon+\lambda\hat{\eta}_\varepsilon) \leq 0, \text{ for all } \lambda\in[0,\lambda_3]. \tag{40}$$

Finally,

$$g^s(\hat{\tau}_\varepsilon) + \langle \nabla g^s(\hat{\tau}_\varepsilon),\hat{\eta}_\varepsilon\rangle \leq - \hat{\delta}, \text{ for all } \varepsilon\in E \text{ and } s\in S, \tag{41}$$

and hence for $\lambda \in [0,1]$

$$g^s(\hat{\tau}_\varepsilon) + \lambda\langle \nabla g^s(\hat{\tau}_\varepsilon),\hat{\eta}_\varepsilon\rangle \leq -\lambda\hat{\delta}, \text{ for all } \varepsilon \in E,\ s \in S \tag{42}$$

Since T is compact, we conclude that there exist a $\lambda_4 \in (0,\lambda_3]$ such that

$$\max_{\substack{\varepsilon \in E \\ s \in S}} g^s(\hat{\tau}_\varepsilon + \lambda\hat{\eta}_\varepsilon) \leq 0, \text{ for all } \lambda \in [0,\lambda_4]. \tag{43}$$

We thus have shown that $\{(\hat{x} + \lambda_4\hat{h}), (\hat{\tau}_\varepsilon + \lambda_4\hat{\eta}_\varepsilon, \varepsilon \in E)\}$ is feasible for $\hat{P}$ and results in a lower cost, which contradicts the optimality of $\hat{x}$. Hence the theorem must be true. ◻

We now establish an important relationship between the stationary points for the problems $\hat{P}_i$ and those of problem $\hat{P}$.

<u>Theorem 4</u>: Let $\Omega_i \subset \Omega$, $E_i \subset E$, $i = 0,1,2,\ldots$, be infinite sequences of discrete subsets and let $x_i \in \mathbb{R}^n$, $y_{E_i} \triangleq \{\tau_{\varepsilon,i} | \tau_{\varepsilon,i} \in E_i\}$, $i = 0,1,2,\ldots$, be such that

$$\psi_{\hat{P}_i}(x_i, y_{E_i}) \to 0 \quad \text{as } i \to \infty, \tag{44a}$$

$$\theta_{\Omega_i,E_i}(x_i, y_{E_i}) \to 0 \quad \text{as } i \to \infty, \tag{44b}$$

and $x_i \to \hat{x}$ as $i \to \infty$, with

$$\max\{f(\hat{x}), \psi_\Omega(\hat{x}), \psi_{E,T}(\hat{x})\} \leq 0. \tag{44c}$$

Then there exists a $\hat{y}_E = \{\tau_\varepsilon | \varepsilon \in E\}$ such that

$$\psi_{\hat{P}}(\hat{x}, \hat{y}_E) = 0 \tag{44d}$$

and

$$\theta_{\Omega,E}(\hat{x}, \hat{y}_E) = 0. \tag{44e}$$

Proof: For every $\varepsilon \in E$ and $i = 0,1,2,\ldots,$ let

$$y_i \triangleq \{\tau_{\varepsilon,i} \mid \tau_{\varepsilon,i} \in y_{E_i} \ \forall \ \varepsilon \in E_i,$$

$$\tau_{\varepsilon,i} \in \arg \min_{\tau \in T} \max_{k \in K} \zeta^k(x_i,\varepsilon,\tau),$$

$$\forall \ \varepsilon \in E \sim E_i\}. \tag{45}$$

(the definition of y_i is not necessarily unique). Then, because $\psi_{E,T}(\cdot)$ is continuous, and because $\psi_{E,T}(\hat{x}) \leq 0$, we have

$$\lim_{i \to \infty} \min_{\tau \in T} \max_{k \in K} \zeta^k(x_i,\varepsilon,\tau) \leq 0, \forall \ \varepsilon \in E, \tag{46}$$

and hence (with E replacing E_i in (28b))

$$\lim_{i \to \infty} \psi_E(x_i,y_i) \triangleq \lim_{i \to \infty} \max_{\substack{\varepsilon \in E \\ k \in K}} \zeta^k(x_i,\varepsilon,\tau_{\varepsilon,i})$$

$$= \lim_{i \to \infty} \max \{\max_{\substack{\varepsilon \in E \sim E_i \\ k \in K}} \zeta^k(x_i,\varepsilon,\tau_{\varepsilon,i}), \max_{\substack{\varepsilon \in E_i \\ k \in K}} \zeta^k(x_i,\varepsilon,\tau_{\varepsilon,i})\} \leq 0. \tag{47}$$

For every $\varepsilon \in E$, $\{\tau_{\varepsilon,i}\}_{i=0}^{\infty}$ is contained in a compact set and hence has at least one accumulation point $\hat{\tau}_\varepsilon$. Let

$$\hat{y}_E \triangleq \{\hat{\tau}_\varepsilon \mid \varepsilon \in E\}, \tag{48}$$

with $\hat{\tau}_\varepsilon$ an accumulation point of $\{\tau_{\varepsilon,i}\}_{i=0}^{\infty}$. Then, because of (47),

$$\psi_E(\hat{x},\hat{y}_E) \leq 0 \tag{49}$$

and hence, because of (44c), we see that (44d) holds.

Next, by construction of y_i,

$$\theta_{\Omega_i,E_i}(x_i,y_{E_i}) \leq \theta_{\Omega,E}(x_i,y_i)-\psi_{\hat{P}_i}(x_i,y_{E_i}) + \psi_{\hat{P}}(x_i,y_i). \quad (50)$$

Since $\theta_{\Omega,E}(x_i,y_i) \leq 0$ for all i, and since $\theta_{\Omega_i,E_i}(x_i,y_{E_i}) \to 0$, $\psi_{\hat{P}_i}(x_i,y_{E_i}) \to 0$ and $\psi_{\hat{P}}(x_i,y_i) \to 0$ as $i \to \infty$, it now follows that (44e) is true, which completes the proof. ¤

We can now state our implementable algorithm.

<u>Master Algorithm 3</u>

<u>Parameters</u>: $\{\alpha_i\}_{i=0}^{\infty}$ such that $\alpha_i \searrow 0$ as $i \to \infty$, and $\{\delta_{ij}\}_{i,j=0}^{\infty}$ such that

a) $\delta_{ij} = 0$ for all $i \leq j$, and $\delta_{ij} > 0$ otherwise.

b) $\delta_{ij} \nearrow \hat{\delta}_j$ as $i \to \infty$.

c) $\hat{\delta}_j \searrow 0$ as $j \to \infty$.

<u>Data</u>: $\Omega_0 \subset \Omega$, $E_0 \subset E$ discrete sets.

<u>Step 1</u>: Apply iterations of a nonlinear programming algorithm to $\hat{P}_i$ until an (x_i,y_{E_i}) pair is computed satisfying

$$-\alpha_i \leq \theta_{\Omega_i,E_i}(x_i,y_{E_i}), \quad (51a)$$

$$\psi_{\hat{P}_i}(x_i,y_{E_i}) \leq \alpha_i. \quad (51b)$$

<u>Step 2</u>: Apply iterations of a nonlinear programming algorithm to (23) and iterations of Subalgorithm 1 to (24) to obtain an $\omega_i \in \Omega$ and an

$\varepsilon_i \in E$, with the property that[†]

$$|\psi_\Omega(x_i) - \phi(x_i,\omega_i)| \to 0 \text{ as } i \to \infty, \tag{51c}$$

$$|\psi_{E,T}(x_i) - \chi(x_i,\varepsilon_i)| \to 0 \text{ as } i \to \infty. \tag{51d}$$

Step 3: Include ω_j in Ω_{i+1} for all $0 \le j \le i$ such that $\phi(x_j,\omega_j) > \delta_{ij}$, and include ε_j in E_{i+1} for all $0 \le j \le i$ such that $\chi(x_j,\varepsilon_j) > \delta_{ij}$.

Step 4: Set $i = i+1$ and go to Step 1. ¤

Since we want δ_{ij} to decay very slowly, a good choice for δ_{ij} seems to be $\delta_{ij} = M \max\{0, \frac{1}{(1+j)^{1/L}} - \frac{1}{(1+i)^{1/L}}\}$, with $M >> 1$ and $L \ge 10$ say.

Master Algorithm 3 has properties which are quite analogous to those of Master Algorithm 2, as we see from:

Theorem 5: Consider the sequence $\{x_i\}_{i=0}^{\infty}$ constructed by Master Algorithm 3. If $x_i \xrightarrow{I} \hat{x}$ as $i \to \infty$, $I \subset \{0,1,2,\ldots\}$, then there exists a $\hat{y}_E = \{\hat{\tau}_\varepsilon | \varepsilon \in E\}$ such that $\psi_{\hat{P}}(\hat{x},\hat{y}_E) = 0$ and $\theta_{\Omega,E}(\hat{x},\hat{y}_E) = 0$.

Proof: Because of (51b-d) and because $\psi_{E_i}(x_i) \le \psi_{E_i}(x_i,y_{E_i}) \le \psi_{\hat{P}_i}(x_i,y_{E_i})$, it follows by a trivial extension of Proposition 2 that $\max\{f(\hat{x}),\psi_\Omega(\hat{x}),\psi_E(\hat{x})\} \le 0$ (i.e., $\hat{x} \in F$ (see (11)). The theorem now follows from (51a) and Theorem 4. ¤

4. Conclusion

The main contribution of this paper is to show that design centering, tolerancing and tuning (DCTT) problems can be treated outside of

[†]This property is achieved by applying the appropriate algorithm for a progressively larger and larger number of iterations as $i \to \infty$.

the framework of nondifferentiable optimization algorithms. As a result a number of major obstacles to obtaining an implementable algorithm have been overcome and the first implementable algorithm for solving DCTT problems has been constructed. We hope that this algorithm will have practical impact.

Finally, in the Appendix the optimality conditions used in this paper are compared to the ones based on generalized gradients, used in [3]. It seems that the optimality conditions in this paper are somewhat sharper than the ones used in [3].

APPENDIX

A COMPARISON OF OPTIMALITY CONDITIONS

In [3], problems such as (6) were treated as nondifferentiable optimization problems. Mifflin [13] has developed a necessary condition of optimality for such problems based on the theory of subgradients (see [12]). Unfortunately, Mifflin's condition is not verifiable for problems such as (6), with $K = \{1\}$ and hence his conditions were somewhat relaxed in [3], as follows:

Theorem A1 [3]: Suppose that x_i is optimal for P_i in (6), then $f(x_i) \leq 0$, $\psi_{\Omega_i}(x_i) \leq 0$, $\psi_{E_i,T}(x_i) \leq 0$ and

$$0 \in M(x_i), \tag{A1}$$

where

$$\begin{aligned} M(x_i) = Co\ \{&\nabla f^0(x_i); \delta_f \nabla f(x_i); \\ &\delta_{\Omega_i} \nabla_x \phi(x_i,\omega), \omega \in \Omega_i(x_i); \\ &\delta_{E_i} \nabla_x \zeta^1(x_i,\varepsilon,\tau), \\ &\varepsilon \in E_i(x_i,\varepsilon,\tau),\ \varepsilon \in E_i, \tau \in T(x_i,\varepsilon)\} \end{aligned} \tag{A2}$$

where δ_{Ω_i} (δ_{E_i}) is zero if $\psi_{\Omega_i}(x_i)<0$ ($\psi_{E_i,T}(x_i)<0$) and is equal to one otherwise,

$$\Omega_i(x_i) \triangleq \arg\max_{\omega \in \Omega_i} \phi(x_i,\omega) \tag{A3}$$

$$E_i(x) \triangleq \arg\max_{\varepsilon \in E_i} \chi(x_i,\varepsilon) \tag{A4}$$

$$T(x_i,\varepsilon) \triangleq \arg\min_{\tau \in T} \zeta^1(x_i,\varepsilon,\tau) \tag{A5}$$

¤

Now, by Proposition 3, if (x_i, y_{E_i}) are optimal for problem $\hat{P}_i$, (6e), then $x_i \in F$ and $\theta_{\Omega_i,E_i}(x_i,y_{E_i}) = 0$. Since y_{E_i} is not necessarily unique, to obtain a comparison with Theorem A1 we shall assume that $\tau_\varepsilon \in y_{E_i} \Rightarrow \tau_\varepsilon \in T(x_i,\varepsilon)$ for all $\varepsilon \in E_i$. Now, it is shown in Sections 1.2 and 4.4 of [17], that for $x_i \in F$, $\theta_{\Omega_i,E_i}(x_i,y_{E_i}) = 0$ if and only if the F. John condition [17] is satisfied at (x_i,y_{E_i}), i.e., there exist multipliers μ^0, μ_f, μ_ω, $\mu_{\varepsilon,\tau}$, $\mu_\varepsilon^s \geq 0$, not all zero such that

$$\mu^0\nabla f^0(x_i) + \mu_f\nabla f(x_i) + \sum_{\omega\in\Omega(x_i)} \mu_\omega \nabla_x\phi(x_i,\omega) + \sum_{\substack{\varepsilon\in E(x_i)\\ \tau\in T(x_i,\varepsilon)}} \mu_{\varepsilon,\tau}\nabla_x\zeta^1(x_i,\varepsilon,\tau) = 0, \tag{A6}$$

$$\mu_{\varepsilon,\tau}\nabla_\tau\zeta^1(x_i,\varepsilon,\tau) + \sum{}' \mu_\varepsilon^s\nabla g^s(\tau) = 0, \quad \forall\ \varepsilon\in E(x_i), \tau\in T(x_i,\varepsilon) \tag{A7}$$

and

$$\mu_f f(x_i) = 0, \tag{A8}$$

$$\mu_\omega\phi(x_i,\omega) = 0, \ \forall\ \omega\in\Omega(x_i) \tag{A9}$$

$$\mu_{\varepsilon,\tau}\zeta^1(x_i,\varepsilon,\tau) = 0, \ \forall\ \varepsilon\in E(x_i), \ \forall\ \tau\in T(x_i,\varepsilon) \tag{A10}$$

$$\mu_\varepsilon^s g^s(\tau) = 0, \ \forall\ \varepsilon\in E(x_i), \forall \tau\in T(x_i,\varepsilon). \tag{A11}$$

We note that (A7) together with (A11) is merely the F. John condition for the problems

$$\min_{\tau\in T} \ \zeta(x_i,\varepsilon,\tau), \ \varepsilon\in E(x_i). \tag{A12}$$

Next, (A6) sums a smaller number of vectors to zero than (A1) and hence, together with (A8-A10), is a sharper condition of optimality then Theorem A1. By the same token, the optimality condition $\theta(\hat{x},\hat{y}_E) = 0$ used in this paper for $\hat{P}$ is sharper than the optimality condition for P used in [3].

REFERENCES

[1] D.Q. Mayne, E. Polak, and R. Trahan, "An Outer Approximations Algorithm for Computer Aided Design Problems," JOTA, in press.

[2] E. Polak, "On a Class of Computer Aided Design Problems," Proc. of 7th IFAC World Congress, Helsinki, Finland, June 1978.

[3] E. Polak and A. Sangiovanni-Vincentelli, "Theoretical and Computational Aspects of the Optimal Design Centering, Tolerancing and Tuning Problem," IEEE CAS Transactions, special issue on Computer Aided Design, in press.

[4] J.W. Bandler, P.C. Liu and H. Tromp, "A Nonlinear Programming Approach to Optimal Design Centering, Tolerancing and Tuning," IEEE Trans. on Circuits and Systems, vol. CAS-23, n. 3, March 1976.

[5] P.V. Lopresti, "Optimum Design of Linear Tuning Algorithms," IEEE Trans. on Circuits and Systems, vol. CAS-24, n. 3, March 1977.

[6] J.F. Pinel, "Computer-Aided Network Tuning," IEEE Trans. on Circuit Theory, vol. CT-17, n. 1, Jan. 1971, pp. 192-194.

[7] G. Müller, "On Computer-Aided Tuning of Microwave Filters," IEEE Proc. ISCAS 1976, Münich, May 1976, pp. 209-211.

[8] E. Lüder and B. Kaiser, "Precision Tuning of Miniaturized Circuits," IEEE Proc. ISCAS 1976, Münich, May 1976, pp. 722-725.

[9] R.L. Adams and V.K. Manaktala, "An Optimization Algorithm Suitable for Computer-Assisted Network Tuning," IEEE Proc. 1975 ISCAS, Boston 1975, pp. 210-212.

[10] E. Polak and A. Sangiovanni-Vincentelli, "On Optimization Algorithms for Engineering Design Problems with Distributed Constraints, Tolerances and Tuning," Proc. 1978 JACC, Philadelphia, Oct. 1978.

[11] F.H. Clarke, "Generalized Gradients and Applications," Trans. Amer. Math. Soc., vol. 205, 1975, pp. 247-262.

[12] F.H. Clarke, "A New Approach to Lagrange Multipliers," Math. of Oper. Research, vol. 1, 1976, pp. 165-174.

[13] R. Mifflin, "Semismooth and Semiconvex Functions in Constrained Optimization," SIAM J. on Control and Optimization, Nov. 1977, pp. 959-972.

[14] C. Gonzaga and E. Polak, "On Constraint Dropping Schemes and Optimality Functions for a Class of Outer Approximation Algorithms," Memo. #M77-68, ERL, University of California, Berkeley, November 1977, also SIAM J. on Control and Optimization, in press.

[15] C. Lemarechal, "An Extension of Davidon Methods to Nondifferential Problems," in M.L. Balinski and P. Wolfe, eds., Nondifferentiable Optimization, Mathematical Programming Study 3, North Holland, Amsterdam, 1975, pp. 95-109.

[16] P. Wolfe, "A Method of Conjugate Subgradients for Minimizing Nondifferentiable Functions," ibid, pp. 145-173.

[17] E. Polak, Computational Methods in Optimization, Academic Press, 1971.

[18] R. Mifflin, "An Algorithm for Constrained Optimization with Semismooth Functions," Math. of Oper. Research, vol. 2, n. 2, May 1977, pp. 181-207.