Minimax Optimization of Networks by Grazor Search

JOHN W. BANDLER, MEMBER, IEEE, THANDANGORAI V. SRINIVASAN, AND CHRISTAKIS CHARALAMBOUS

Abstract—A new optimization method called *grazor search* has been developed. This method is suitable for nonlinear minimax optimization of network and system responses. A linear programming problem using gradient information of one or more highest ripples in the response error function to produce a downhill direction followed by a linear search to find a minimum in that direction is central to the algorithm. Unlike the razor search method due to Bandler and Macdonald, the present method overcomes the problem of discontinuous derivatives characteristic of minimax objectives without using random moves. It can fully exploit the advantages of the adjoint network method of evaluating partial derivatives of the response function with respect to the variable parameters. Sufficient details are given to enable the grazor search method to be readily programmed and used. Although the method is intended for the computer-aided solution of an extremely wide range of design problems, it is largely compared with other methods on microwave network design problems, for which the solutions are known. Its reliability and efficiency on more arbitrary problems, examples of which are also included, is thereby established.

INTRODUCTION

THE MINIMAX algorithm due to Osborne and Watson [1]–[3] deals with minimax formulations by following two steps—a linear programming part that provides a given step in the parameter space, followed by a linear search along the direction of the step. This algorithm is very similar to the one proposed by Ishizaki and Watanabe [4] and works very well if the objective function is not highly nonlinear in the vicinity of the optimum. In cases when the linear approximation is not very good in the vicinity of the optimum, this method may fail to converge toward the optimum for successive iterations.

The razor search method of Bandler and Macdonald [5], [6], which has been used to optimize microwave networks by computer [7], was developed to minimize the maximum deviation of some network response from an ideal response specification. The direct minimax formulation that they employed gives rise, in general, to discontinuous partial derivatives of the objective function with respect to the variable parameters. For this reason and because general and efficient methods of evaluating derivatives of distributed network responses with respect to network parameters did not at that time appear to be available, only methods not requiring derivatives were considered. The razor search method as presented by Bandler and Macdonald is based on pattern search [8]. A few random moves are used in an effort to negotiate certain kinds of "razor sharp" valleys in multidimensional space [5].

A more recent algorithm due to Bandler and Lee-Chan [9] exploits the gradient information of the extrema of the error function to get a downhill direction by solving a set of simultaneous equations. The method works well except that in the case of linear dependence of the equations, some problems may arise in the convergence toward the optimum. Another method proposed by Heller [10] uses a quadratic programming approach to solve the minimax problem, but consumes a considerable amount of computer time.

A new algorithm called the *grazor search* method has been developed in which gradient information of one or more of the highest ripples in the error function is used to produce a downhill direction by solving a suitable linear programming problem. A linear search follows to find the minimum in that direction, and the procedure is repeated. This type of descent process is repeated with as many ripples as necessary until a minimax solution is reached to some desired accuracy. The algorithm is compared numerically with the razor search method and another based on the Osborne and Watson algorithm on the optimization of commensurate and noncommensurate transmission-line matching networks, for which the optima are known. Examples of the application of the grazor search method in this paper also include the design of microwave filters. Response gradients are evaluated using the results of one network analysis by the adjoint network method [11], [12].

THE GRAZOR SEARCH STRATEGY

The grazor search algorithm is a generalization of the gradient razor search method [9] and is basically of the steepest descent type. In this case, suppose we have the problem of minimizing

$$U(\boldsymbol{\phi}) = \max_{i \in I} y_i(\boldsymbol{\phi}) \tag{1}$$

where $\boldsymbol{\phi}$ denotes the k independent parameters, I is an index set relating to discrete elements corresponding to

Manuscript received November 29, 1971; revised April 17, 1972. This work was supported by the National Research Council of Canada under Grant A7239. The paper is based on papers presented at the 14th Midwest Symposium on Circuit Theory, Denver, Colo., May 1971 and the 1971 IEEE G-MTT International Microwave Symposium, Washington, D. C., May 1971. The contributions of C. Charalambous are the theoretical considerations of the downhill properties of the algorithm.

The authors are with the Communications Research Laboratory and Department of Electrical Engineering, McMaster University, Hamilton, Ont., Canada.

the i, and the y_i are real nonlinear differentiable functions generally.

This minimax approximation problem consists of finding a point $\check{\phi}$ such that

$$U(\check{\phi}) = \min_{\phi} \max_{i \in I} y_i(\phi). \tag{2}$$

Theoretical Considerations

Define a subset $J \subset I$ such that

$$J(\phi^j, \epsilon^j) \triangleq \{ i \mid \max_i y_i(\phi^j) - y_i(\phi^j) \leq \epsilon^j, i \in I \} \tag{3}$$

$$\epsilon^j \geq 0 \tag{4}$$

where ϕ^j denotes a feasible point at the beginning of the jth iteration and ϵ^j is the tolerance with respect to the current $\max_{i \in I} y_i(\phi^j)$ within which the y_i for $i \in J$ lie.

Linearizing y_i at ϕ^j, we can consider the first-order changes

$$\delta y_i(\phi^j) = \nabla^T y_i(\phi^j) \Delta \phi^j, \qquad i \in J(\phi^j, \epsilon^j) \tag{5}$$

where Δ denotes incremental changes and ∇ denotes the first partial derivative operator with respect to the parameter vector ϕ.

A sufficient condition for $\Delta\phi^j$ to define a descent direction for $U(\phi^j)$ is

$$\nabla^T y_i(\phi^j) \Delta \phi^j < 0, \qquad i \in J(\phi^j, \epsilon^j). \tag{6}$$

Consider

$$\Delta \phi^j = - \sum_{i \in J} \alpha_i^j \nabla y_i(\phi^j) \tag{7}$$

$$\sum_{i \in J} \alpha_i^j = 1 \tag{8}$$

$$\alpha_i^j \geq 0. \tag{9}$$

Equation (6) may now be written as

$$-\nabla^T y_i(\phi^j) \sum_{i \in J} \alpha_i^j \nabla y_i(\phi^j) < 0, \qquad i \in J(\phi^j, \epsilon^j) \tag{10}$$

which suggests the linear program: maximize

$$\alpha_{k_r+1}^j(\phi^j, \epsilon^j) \geq 0 \tag{11}$$

subject to

$$-\nabla^T y_i(\phi^j) \sum_{i \in J} \alpha_i^j \nabla y_i(\phi^j) \leq -\alpha_{k_r+1}^j, \quad i \in J(\phi^j, \epsilon^j) \tag{12}$$

plus (8) and (9), where k_r denotes the number of elements of $J(\phi^j, \epsilon^j)$. Note that if $\Delta\phi^j = \mathbf{0}$ for $\epsilon^j = 0$, the necessary conditions for a minimax optimum are satisfied at ϕ^j (see Bandler [13]). Observe that J is nonempty and that if J has only one element, we obtain the steepest descent direction for the corresponding maximum of the $y_i(\phi)$.

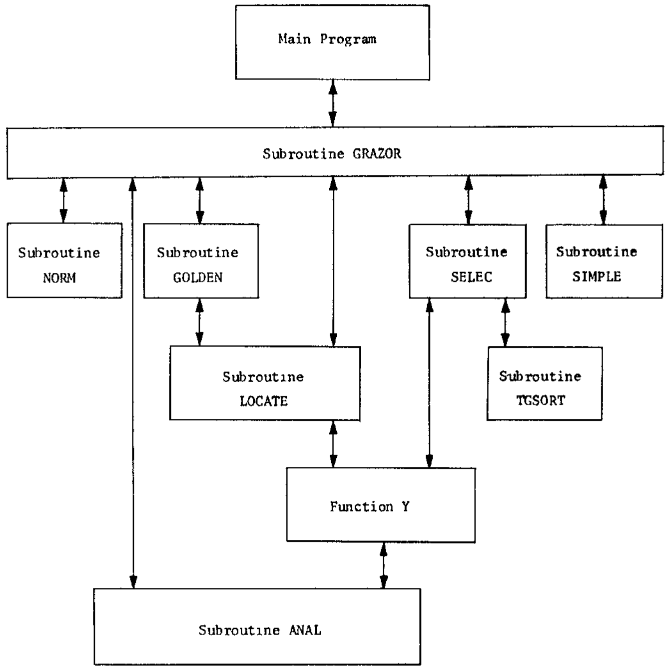

Fig. 1. Block diagram summarizing the computer program structure and illustrating the relative hierarchy of the subprograms.

Before proving the convergence of the algorithm it may be worth restating the following lemma due to Farkas (see, for example, Lasdon [14]).

Let $\{p_0, p_1, \cdots, p_n\}$ be an arbitrary set of vectors. There exist $\beta_i \geq 0$ such that

$$p_0 = \sum_{i=1}^{n} \beta_i p_i \tag{13}$$

if and only if

$$p_0^T q \geq 0 \tag{14}$$

for all q satisfying

$$p_i^T q \geq 0, \qquad i = 1, 2, \cdots, n. \tag{15}$$

It therefore is possible to find nonnegative values of α_i^j in the expression for (7) if and only if

$$(-\Delta\phi^j)^T(-\Delta\phi^j) \geq 0 \tag{16}$$

for all $\Delta\phi^j$ satisfying

$$\nabla^T y_i(\phi^j)(-\Delta\phi^j) \geq 0, \qquad i \in J(\phi^j, \epsilon^j) \tag{17}$$

where (14) and (15) correspond to (16) and (17), respectively, and $-\Delta\phi^j$, $\nabla y_i(\phi^j)$, $-\Delta\phi^j$ take the place of p_0, p_i, q.

Now (16) is always satisfied, though it may not be possible to satisfy (17) if ϵ^j is too large. By suitably decreasing ϵ^j, (17) may be forced to hold.

Practical Implementation

Fig. 1 illustrates how the different subroutines are called and their relative hierarchy. Flow charts of sub-

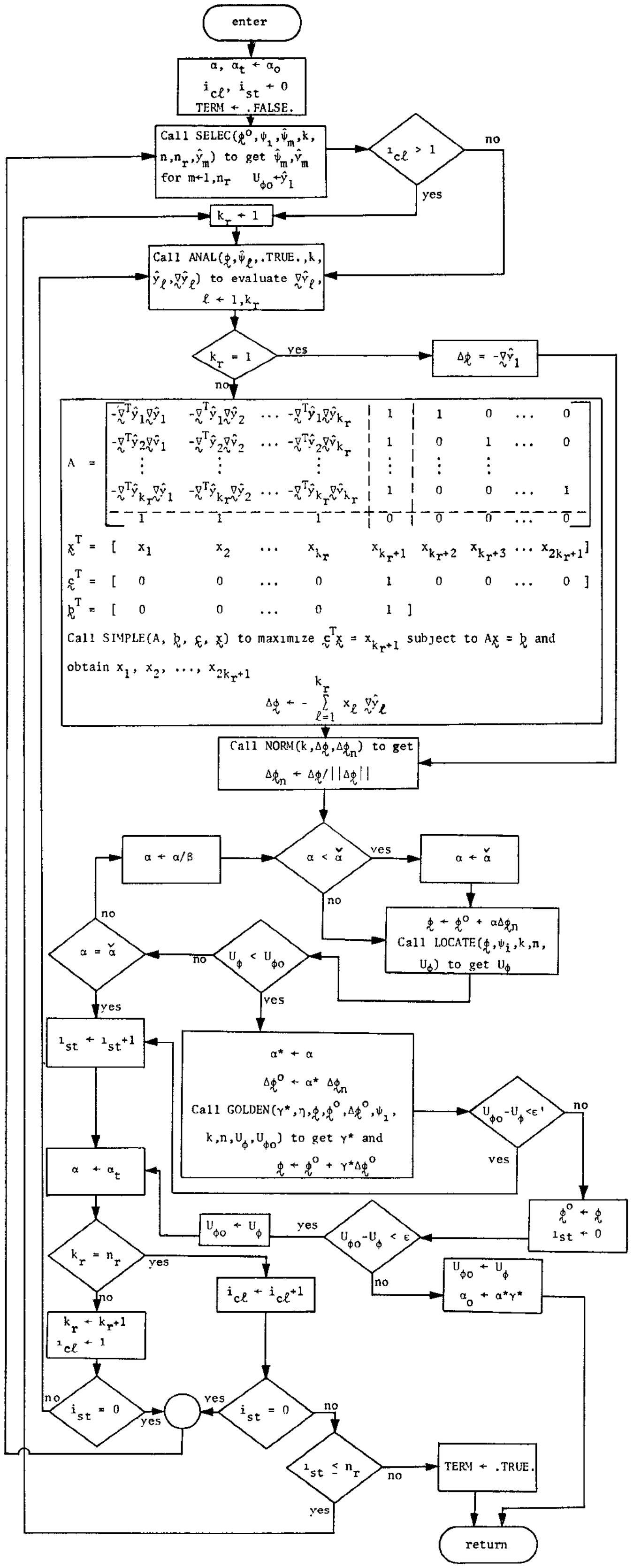

Fig. 2. Mathematical flow diagram of subroutine GRAZOR (α_o, $\check{\alpha}$, β, ϵ, ϵ', η, ϕ^o, ψ_i, k, k_r, n, n_r, $U_{\phi o}$, TERM).

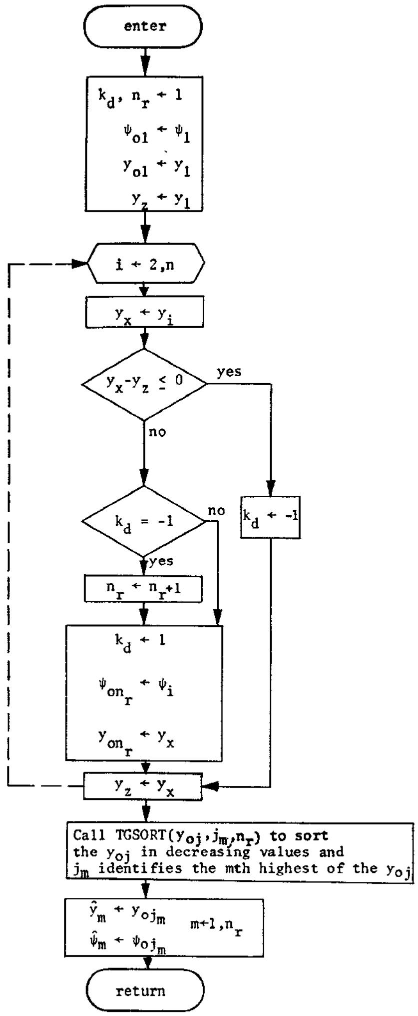

Fig. 3. Mathematical flow diagram of subroutine SELEC (ϕ^o, ψ_i, $\hat{\psi}_m$, k, n, n_r, $\hat{y}_m$).

routines GRAZOR, SELEC, and GOLDEN appear in Figs. 2–4. See Appendix for further details and definitions. $U(\phi^j)$ is calculated by subroutine LOCATE.

As given by linear programming [15], $\Delta\phi^j$ is normalized to $\Delta\phi_n{}^j = \Delta\phi^j/\|\Delta\phi^j\|$ (subroutine NORM). Starting at ϕ^j, a step $\alpha^j\Delta\phi_n{}^j$ is taken for $\alpha^j = \alpha_o{}^j$; if no improvement in U results, α^j is reduced by factors of β until a better point is obtained or $\alpha^j < \check{\alpha}$. Let α^{j*} produce the first improved point from ϕ^j. Then $\Delta\phi^o = \alpha^{j*}\Delta\phi_n{}^j$ is defined.

Next a method based on golden section search (suggested by Temes [16]) is used to find the γ^{j*} corresponding to the constrained minimum value of $\max_{i \in I} y_i(\phi^j + \gamma^j\Delta\phi^o)$. The jth iteration ends by setting $\phi^{j+1} = \phi^j + \gamma^{j*}\Delta\phi^o$ and $\alpha_o{}^{j+1} = \alpha^{j*}\gamma^{j*}$.

In Fig. 4, $\tau = \frac{1}{2}(1+\sqrt{5})$ is the factor associated with the golden section. Subscripts l and u denote lower and upper limits, respectively, and a and b denote interior points of the interval of search. An attempt to bound the minimum is made. Then golden section search is used to locate the minimum to a desired accuracy. The search is terminated when the resolution between two

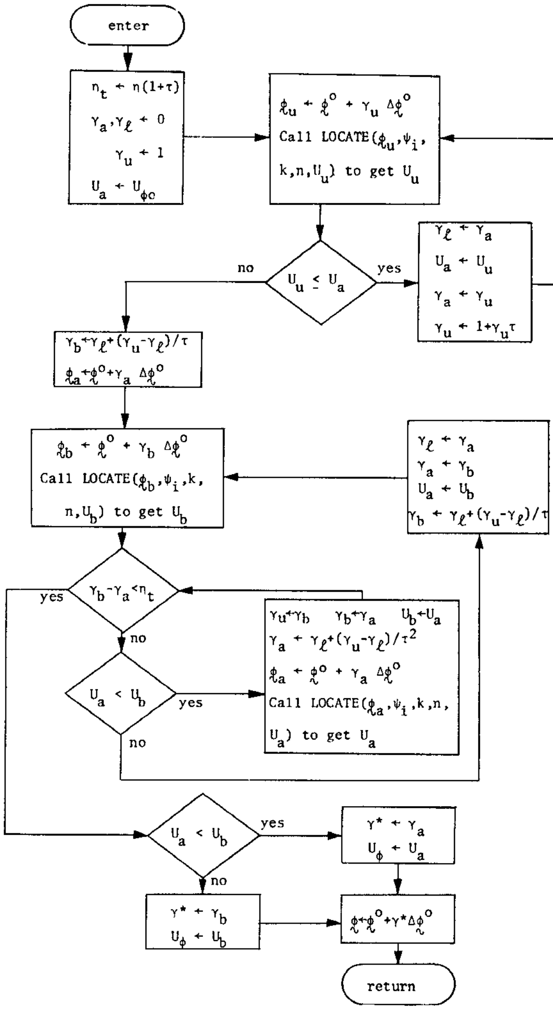

Fig. 4. Mathematical flow diagram of subroutine GOLDEN (γ^*, η, ϕ, ϕ^o, $\Delta\phi^o$, ψ_i, k, n, U_ϕ, $U_{\phi o}$).

interior points falls below a factor η of the initial interval.

In Fig. 3, the maxima implied by the functions y_i, sampled in a certain order, are located and sorted out in decreasing magnitude [17].

Fig. 2 shows the grazor search strategy. Note that in setting up $Ax=b$, slack variables (x_{k_r+2}, x_{k_r+3}, $\cdots$, x_{2k_r+1}) are introduced. We try to generate a descent direction based on the gradient of the maximum function ($k_r=1$), proceed to the minimum of U in that direction, and repeat the process. If, at any stage, this process or the linear program does not yield a direction of decreasing U, or does not provide an improvement greater than ϵ, the procedure is repeated after including the function corresponding to the next largest of the current n_r discrete local maxima (i.e., ripples) if one exists. When all local maxima have been included and U can still not be reduced or improved satisfactorily by a value greater than ϵ, we repeat the procedure with k_r functions corresponding to the first k_r largest of the candidates, beginning with $k_r=1$, in another series of attempts to reduce U. The algorithm terminates only when there are no

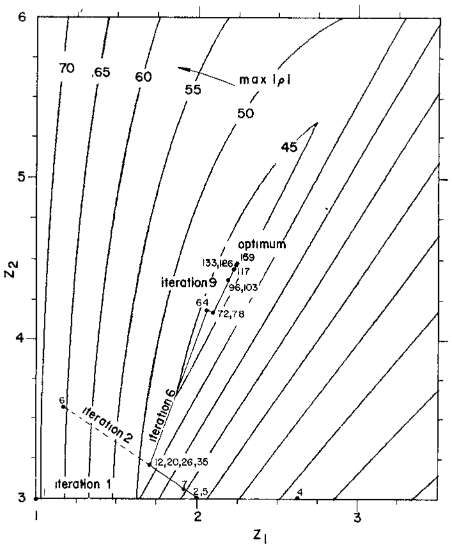

Fig. 5. Example illustrating how the grazor search strategy follows the narrow path of discontinuous derivatives.

more suitable functions left and when there are either no improvements or improvements less than ϵ' over one complete cycle of k_r, starting from 1 and ending with n_r.

Example: Table I, in association with Fig. 5, illustrates how the grazor search strategy effectively follows the path of discontinuous derivatives to locate the optimum in the course of minimax optimization of a 2-section cascaded transmission-line network [5], [18] (illustrated in Fig. 6) when the lengths are fixed at quarter-wavelength values and the impedances are varied. We let $y_i(\phi)=\frac{1}{2}|\rho(\phi,\psi_i)|^2$ and define $U'(\phi)=\max_i|\rho(\phi,\psi_i)|$, where ρ is the reflection coefficient on 11 uniformly spaced frequencies ψ_i in the band 0.5 to 1.5 GHz. The grazor search strategy starts at $\phi^1(1.0, 3.0)$, $U'(\phi^1)=0.70954$, and the values of parameters used are $\alpha_o=1$ (at start), $\check{\alpha}=10^{-6}$, $\beta=10$, $\eta=0.5$, $\epsilon=10^{-4}$, $\epsilon'=10^{-6}$.

The first iteration extends from ϕ^1 to ϕ^5; ϕ^2 is the new point obtained when taking a unit step along the direction suggested by the negative gradient. Since ϕ^2 is a satisfactory improvement, a golden section search is initiated, yielding $\phi^3(\gamma=1+\tau)$ which is not an improvement over ϕ^2. The interval of search is thus found. $\phi^4(\gamma=\tau)$ is found to be no improvement over ϕ^2. The golden section search is now terminated, since the current resolution between two interior points of search falls below the minimum allowed value. $\phi^5=\phi^2$ is thus the best point attained at the end of iteration 1. At the end of iteration 5, $U(\phi^{26})-U(\phi^{35})<\epsilon$, so k_r is increased from 1 to 2 in the next iteration. For a similar reason, k_r is increased from 2 to 3 for iteration 12, and reset to 1 from 3 for iteration 13. During iteration 18, the parameter values remain the same to 5 significant digits, and the improvement in U at the end is less than ϵ'; all successive attempts to achieve a better point with an improvement greater than ϵ' (by considering 1, 2, and 3 ripples) fail, and the procedure is terminated.

TABLE I

SUMMARY OF IMPORTANT STEPS IN THE EXAMPLE ILLUSTRATING THE GRAZOR SEARCH STRATEGY
$\check{\phi} = (2.23605, 4.47210)$, $U'(\check{\phi}) = 0.42857$

Iteration Number	Points of Iteration	Starting Point of Iteration	Values of Scale Factors		Number of Ripples Considered
			Point	Scale Factor	
1	1–5	$\phi^1 = (1.0, 3.0)$ $U'(\phi^1) = 0.70954$	ϕ^2 ϕ^4 $\phi^5 = \phi^2$	$\alpha^* = 1.00$ $\gamma = 1+\tau$ $\gamma^* = 1.000$	1
2	5–12	$\phi^5 = (1.99996, 3.00893)$ $U'(\phi^5) = 0.63086$	ϕ^6 ϕ^7 ϕ^{12}	$\alpha = 1.00$ $\alpha^* = 0.10$ $\gamma^* = 2+\tau$	1
3	12–20	$\phi^{12} = (1.69865, 3.20921)$ $U'(\phi^{12}) = 0.48073$	ϕ^{13} ϕ^{14} ϕ^{15} ϕ^{20}	$\alpha = 0.1(\tau+2)$ $\alpha = 0.01(\tau+2)$ $\alpha^* = 0.001(\tau+2)$ $\gamma^* = \tau+1$	1
4	20–26	$\phi^{20} = (1.70806, 3.20821)$ $U'(\phi^{20}) = 0.47843$	ϕ^{21} ϕ^{22} ϕ^{26}	$\alpha = 9.472\times10^{-3}$ $\alpha^* = 9.472\times10^{-4}$ $\gamma^* = 1.000$	1
5	26–35	$\phi^{26} = (1.70723, 3.20865)$ $U'(\phi^{26}) = 0.47794$	ϕ^{30} ϕ^{35}	$\alpha^* = 1.0\times10^{-6}$ $\gamma^* = \tau+1$	1
6	35–64	$\phi^{35} = (1.70723, 3.20866)$ $U'(\phi^{35}) = 0.47794$	ϕ^{36} ϕ^{64}	$\alpha^* = 9.472\times10^{-4}$ $\gamma^* = 1.096\times10^{3}$	2
7	64–72	$\phi^{64} = (2.05489, 4.18669)$ $U'(\phi^{64}) = 0.44084$	ϕ^{64}	$\gamma^* = \tau+2$	2
8	72–78	$\phi^{72} = (2.09028, 4.17411)$ $U'(\phi^{72}) = 0.43199$	ϕ^{78}	$\gamma^* = 1.000$	2
9	78–96	$\phi^{78} = (2.09380, 4.17280)$ $U'(\phi^{78}) = 0.43146$	ϕ^{96}	$\gamma^* = 60.69$	2
10	96–103	$\phi^{96} = (2.18832, 4.38018)$ $U'(\phi^{96}) = 0.42929$	ϕ^{98} ϕ^{103}	$\alpha^* = 2.279\times10^{-3}$ $\gamma^* = 1.000$	2
11	103–117	$\phi^{103} = (2.19040, 4.37924)$ $U'(\phi^{103}) = 0.42886$	ϕ^{117}	$\gamma^* = 30.03$	2
12	117–126	$\phi^{117} = (2.22029, 4.44082)$ $U'(\phi^{117}) = 0.42864$	ϕ^{126}	$\gamma^* = 10.47$	3
13	126–132	$\phi^{126} = (2.23088, 4.46221)$ $U'(\phi^{126}) = 0.42862$			1
13	133–136	$\phi^{133} = \phi^{126}$ $U'(\phi^{133}) = U'(\phi^{126})$	ϕ^{134} ϕ^{136}	$\alpha^* = 2.279\times10^{-3}$ $\gamma^* = \tau+2$	2
18	169–176	$\phi^{169} = (2.23595, 4.47237)$ $U'(\phi^{169}) = 0.42861$	ϕ^{176} $= \phi^{169}$	$\gamma^* = 1.000$	3

EXAMPLES

Example 1

The grazor search algorithm has been compared numerically with the razor search method and the Osborne and Watson algorithm on the problem of minimizing $\max|\rho|$ on 11 frequencies ψ_i in the band 0.5 to 1.5 GHz for the network shown in Fig. 6.

For the grazor search and Osborne and Watson algorithms we took $y_i(\phi) = \frac{1}{2}|\rho(\phi, \psi_i)|^2$. Gradients were evaluated using the adjoint network method [11], [12]. In the 2-section examples, the 11 frequencies were uniformly spaced. In the 3-section examples, the frequencies were 0.5, 0.6, 0.7, 0.77, 0.9, 1.0, 1.1, 1.23, 1.30, 1.40, and 1.50 GHz. The progress of the algorithms from identical starting points with respect to the number of function evaluations (one corresponding to 11 evalua-

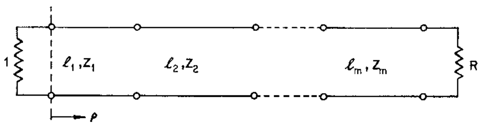

Fig. 6. The m-section resistively terminated cascade of transmission lines. Optimum matching over 100-percent band centered at 1 GHz for $R=10$ occurs for the following parameter values. 2-section: $l_1=l_2=l_q$, $Z_1=2.23605$, $Z_2=4.4721$. 3-section: $l_1=l_2=l_3=l_q$, $Z_1=1.63471$, $Z_2=3.16228$, $Z_3=6.11729$. $l_q=7.49481$ cm is the quarter wavelength at center frequency.

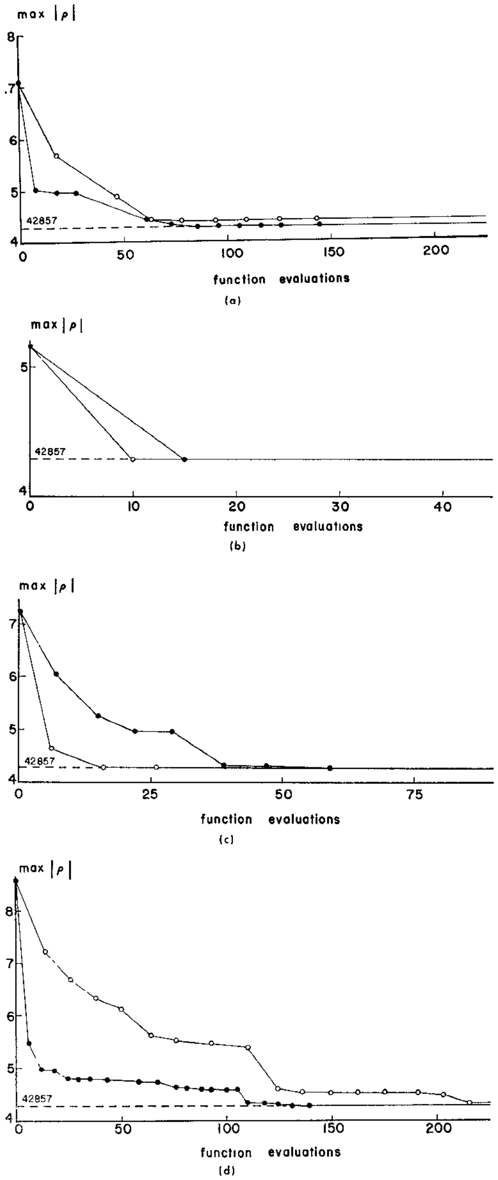

Fig. 7. The 2-section transformer problem. Solid points distinguish the grazor search algorithm from the algorithm based on the Osborne and Watson method. (a) $l_1=l_2=l_q$. Starting point $Z_1=1.0$, $Z_2=3.0$. (b) Z_1 and Z_2 fixed at optimum values. Starting point $l_1/l_q=0.8$, $l_2/l_q=1.2$. (c) l_2 and Z_2 fixed at optimum values. Starting point $l_1/l_q=1.2$, $Z_1=3.5$. (d) Starting point $l_1/l_q=1.2$, $l_2/l_q=0.8$, $Z_1=3.5$, $Z_2=3.0$.

tions of ρ) is recorded in Figs. 7 and 8. The points shown mark the successful end of a linear search or the beginning of linear programming.

Tables II and III compare the two algorithms with the razor search method. From Table II, it is clear that

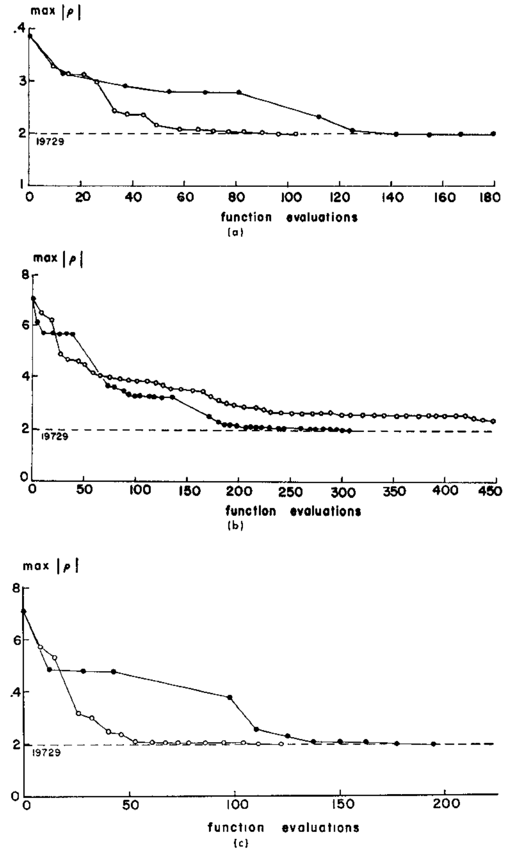

Fig. 8. The 3-section transformer problem. Solid points distinguish the grazor search algorithm from the algorithm based on the Osborne and Watson method. (a) $l_1=l_2=l_3=l_q$. Starting point $Z_1=1.0$, $Z_2=3.16228$, $Z_3=10.0$. (b) Starting point $l_1/l_q=0.8$, $l_2=l_q=1.2$, $l_3/l_q=0.8$, $Z_1=1.5$, $Z_2=3.0$, $Z_3=6.0$. (c) Starting point $l_1/l_q=l_2/l_q=l_3/l_q=1.0$, $Z_1=1.0$, $Z_2=3.16228$, $Z_3=10.0$.

TABLE II
OPTIMIZATION OF A 2-SECTION 10:1 QUARTER-WAVE TRANSFORMER OVER 100-PERCENT BANDWIDTH

Starting Point		Function Evaluations[a]	
Z_1	Z_2	Razor Search	Grazor Search
1.0	3.0	157 207	126
1.0	6.0	34 152	83
3.5	6.0	223 100	52
3.5	3.0	210 163	29

[a] Number of function evaluations required to bring the reflection coefficient within 0.01 percent of its optimum value.

the grazor search algorithm is, in general, faster than the razor search technique for the 2-section case when the lengths are kept fixed and the impedances are varied. From Table III, it is clear that the grazor search algorithm is the best. The Osborne and Watson algorithm, though fairly fast initially, may in some cases fail or slow down near the optimum.

The grazor search method and the Osborne and Watson algorithm were further compared on the 3-section transformer problem when the lengths were fixed at quarter-wavelength values and the impedances were

TABLE III
OPTIMIZATION OF A 3-SECTION 10:1 TRANSFORMER OVER A 100-PERCENT BANDWIDTH

	Fixed Lengths		Variable Lengths and Impedances			
Parameters ϕ_i	Starting Point	Maximum Reflection Coefficient at Start	Starting Point	Maximum Reflection Coefficient at Start	Starting Point	Maximum Reflection Coefficient at Start
l_1/l_q	1.0		1.0		0.8	
Z_1	1.0		1.0		1.5	
l_2/l_q	1.0		1.0		1.2	
Z_2	3.16228	0.70930	3.16228	0.70930	3.0	0.38865
l_3/l_q	1.0		1.0		0.8	
Z_3	10.0		10.0		6.0	
Razor search algorithm	final maximum reflection coefficient	0.19729		0.19733		0.19731
	number of function evaluations	406		1300		1250
Grazor search algorithm	final maximum reflection coefficient	0.19729		0.19729		0.19729
	number of function evaluations	219		696		498
Algorithm due to Osborne and Watson [1]	final maximum reflection coefficient	0.19729		0.20831		0.19788
	number of function evaluations	199		860		237

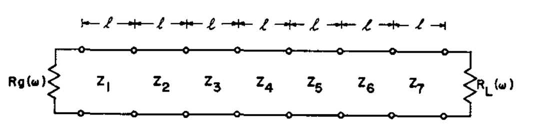

Fig. 9. Cascaded transmission-line filter operating between $R_g(\omega) = R_L(\omega) = 377/\sqrt{1-(f_c/f)^2}$, where $f_c = 2.077$ GHz and $l = 1.5$ cm.

varied. For a starting point of $Z_1 = 3.16228$, $Z_2 = 1.0$, and $Z_3 = 10.0$, the former took 184 and 218 function evaluations, while the latter consumed 151 and 219 function evaluations to reach within 0.01 and 0.001 percent of the optimum value of the maximum reflection coefficient, respectively. This case illustrates how the two algorithms compare when both methods work efficiently.

Example 2

The grazor search method was used in the optimization of a 7-section cascaded transmission-line filter as shown in Fig. 9. This interesting problem has been considered by Carlin and Gupta [19]. The frequency variation of the terminations is like that of rectangular waveguides operating in the H_{10} mode with cutoff frequency 2.077 GHz. All section lengths were kept fixed at 1.5 cm so that the maximum stopband insertion loss would occur at about 5 GHz. The passband 2.16 to 3 GHz was selected, for which a maximum passband insertion loss of 0.4 dB was specified.

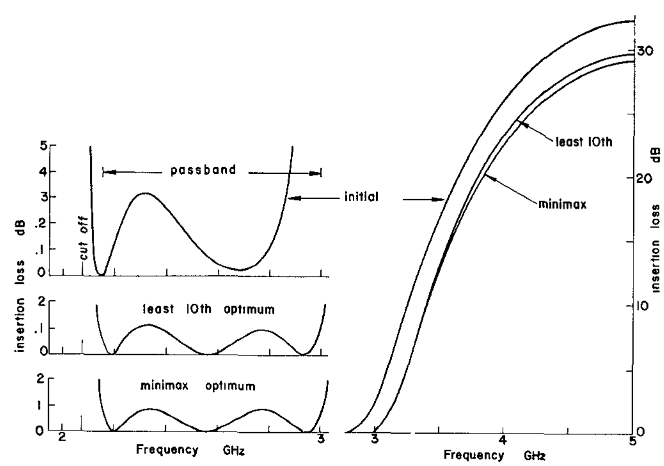

Fig. 10. Responses of the network of Fig. 9. The response of Carlin and Gupta [19] is the initial one. The least 10th response was obtained by Bandler and Seviora [12]. The minimax response was produced by the grazor search method.

Fig. 10 shows the response of Carlin and Gupta which was used as an initial design. The other responses in Fig. 10 are a least 10th optimum obtained by Bandler and Seviora [12] and a minimax optimum obtained by the grazor search strategy. In both cases only the passband was optimized. The minimax response has a maximum passband insertion loss of 0.086 dB. Table IV gives the appropriate parameter values.

Fig. 11 shows the results of applying the grazor search method to optimize the sections in a filtering

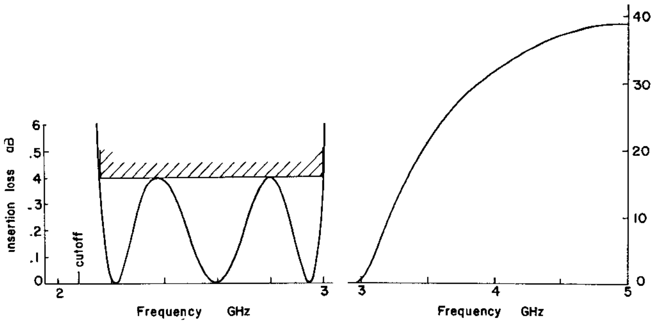

Fig. 11. Response of the minimax design of the network of Fig. 9 with 0.4-dB passband insertion loss produced by the grazor search method.

TABLE IV

COMPARISON OF PARAMETER VALUES FOR THE 7-SECTION FILTER

Characteristic Impedances (Normalized)	Carlin and Gupta [19]	Minimax Design (Fig. 10)	Minimax Design (Fig. 11)
Z_1	1476.5	1305.2	3069.4
Z_2	733.6	607.8	2856.4
Z_3	1963.6	1323.3	25871.2
Z_4	461.8	362.7	10573.3
Z_5	1963.6	1323.2	25874.0
Z_6	733.6	607.9	2856.7
Z_7	1476.5	1305.2	3069.8

sense. Thus it was desired to meet the 0.4-dB passband insertion loss while maximizing the stopband insertion loss at a single frequency (5 GHz). We let [20]

$$y_i(\phi) = \begin{cases} \frac{1}{2}[\,|\rho_i(\phi)|^2 - r^2], & \text{in the passband} \\ \frac{1}{2}[1 - |\rho_i(\phi)|^2], & \text{in the stopband} \end{cases} \tag{18}$$

where

$$\phi = [Z_1 Z_2 \cdots Z_7]^T$$

and r is the reflection coefficient magnitude corresponding to an insertion loss of 0.4 dB and $\rho_i(\phi)$ is the reflection coefficient of the filter at the ith frequency. Here 22 uniformly spaced points were selected from the passband. Table IV gives the resulting parameter values. A similar response was attained by the grazor search technique when the section impedances were assumed symmetrical, i.e., $Z_5 = Z_3$, $Z_6 = Z_2$, $Z_7 = Z_1$.

CONCLUSIONS

The results indicate that the grazor search algorithm is generally more reliable in reaching an optimal minimax solution than the Osborne and Watson algorithm, and is faster than the razor search technique. Typically, 1 min is sufficient time to optimize a six-parameter design, and 2 to 3 min are sufficient to optimize a ten-parameter problem, depending on how far from the optimum one starts and how close one wishes to get, on a CDC 6400 computer.

The grazor search algorithm has been successfully applied to problems of cascaded lumped LC filter designs, antenna modeling circuit designs, and modeling high-order control systems [21]. The algorithm should be able to handle, without much difficulty, filter design problems with upper and lower specifications over many frequency bands. It is felt, however, that in the microwave area, the algorithm will find most use in design problems for which exact methods are not available.

APPENDIX

NOMENCLATURE

The following is a list of some of the arguments and important variables of the grazor search package as indicated in the flow charts of Figs. 2–4.

α Scale factor for determining the magnitude of the parameter step to be taken at the end of linear program.

α_o Initial specified value of α, previous value of α which gave a satisfactory improvement.

$\breve{\alpha}$ Minimum allowable α.

β Reduction factor for α.

γ^* Factor of the step $\Delta\phi^o$ which gives the best new point, when starting from ϕ^o.

ϵ Number of discrete maxima under consideration k_r is increased by one (if $k_r \leq n_r - 1$) or set equal to one (if $k_r = n_r$) if the improvement of the objective function at the new best point as compared to the value at the previous point is less than this quantity.

ϵ' Main program is eventually terminated if the improvement of objective function at the new best point as compared to the value at the previous point is repeatedly less than this quantity.

η Specified factor of the initial interval of linear search which determines the final resolution between two internal points of the search.

ϕ Current point.

ϕ^o Starting point, current best point.

$\Delta\phi^o$ Increment from ϕ^o which gives the first improved point obtained in each iteration on entering the linear search.

ψ_i ith sample point.

$\hat{\psi}_i$ Sample points corresponding to the $\hat{y}_i$.

ψ_{oi} Sample points corresponding to the y_{oi}.

DERIV Logical variable; if .TRUE. the ∇y_i are calculated, otherwise they are not calculated.

j_i Identifies the ith highest of the y_{oj}.

k Dimensionality of parameter space.

k_r Number of local discrete maxima $\hat{y}_i$ under consideration.

n Number of sample points ψ_i.

n_r Available number of discrete local maxima $\hat{y}_i$.

U_ϕ Value of the objective function at ϕ.

$U_{\phi o}$ Value of the objective function at ϕ^o.

y_i Function value at ψ_i for a given ϕ.

$\hat{y}_i$ ith highest discrete local maximum.

∇y_i Gradient of y_i with respect to ϕ.

y_{oj} Discrete local maxima implied by the y_i.

TERM Logical variable, initially set to .FALSE., is reset to .TRUE. only if there are failures or improvements in objective function value; less than ϵ' after considering values of k_r from 1 to n_r in one complete cycle.

The variables α_o, $\breve{\alpha}$, β, ϵ, ϵ', η, ϕ^o, ψ_i, k, k_r, and n are initially assigned values on entry to the grazor search package. The subroutine ANAL(ϕ, ψ_i, DERIV, k, y_i, ∇y_i) is an analysis program to evaluate y_i and/or ∇y_i at a given point ϕ. The function subprogram $Y(\phi, \psi_i, k)$ calculates the y_i corresponding to the point ϕ by calling ANAL. The subroutine LOCATE (ϕ, ψ_i, k, n, U_ϕ) evaluates the objective function U_ϕ by calling $Y(\phi, \psi_i, k)$ for $i = 1, 2, \cdots, n$.

ACKNOWLEDGMENT

The authors wish to thank A. Lee-Chan of the Data Processing and Computing Centre of McMaster University, Hamilton, Ont., Canada, for his programming assistance.

REFERENCES

[1] M. R. Osborne and G. A. Watson, "An algorithm for minimax approximation in the non-linear case," *Comput. J.*, vol. 12, pp. 63–68, Feb. 1969.
[2] G. A. Watson, "On an algorithm for nonlinear minimax approximation," *Commun. Ass. Comput. Mach.*, vol. 13, pp. 160–162, Mar. 1970.
[3] J. W. Bandler and T. V. Srinivasan, "A new gradient algorithm for minimax optimization of networks and systems," in *Proc. 14th Midwest Symp. Circuit Theory* (Denver, Colo., May 1971), pp. 16.5.1–16.5.11.
[4] Y. Ishizaki and H. Watanabe, "An iterative Chebyshev approximation method for network design," *IEEE Trans. Circuit Theory*, vol. CT-15, pp. 326–336, Dec. 1968.
[5] J. W. Bandler and P. A. Macdonald, "Optimization of microwave networks by razor search," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 552–562, Aug. 1969.
[6] ——, "The razor search program," *IEEE Trans. Microwave Theory Tech.* (Comput. Program Descr.), vol. MTT-19, p. 667, July 1971.
[7] J. W. Bandler, "Computer optimization of inhomogeneous waveguide transformers," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 563–571, Aug. 1969.
[8] R. Hooke and T. A. Jeeves, " 'Direct search' solution of numerical and statistical problems," *J. Ass. Comput. Mach.*, vol. 8, pp. 212–229, Apr. 1961.
[9] J. W. Bandler and A. G. Lee-Chan, "Gradient razor search method for optimization," in *IEEE Int. Microwave Symp. Dig.* (Washington, D. C., May 1971), pp. 118–119.
[10] J. E. Heller, "A gradient algorithm for minimax design," Coordinated Science Lab., Univ. of Illinois, Urbana, Rep. R-406, Jan. 1969.
[11] S. W. Director and R. A. Rohrer, "The generalized adjoint network and network sensitivities," *IEEE Trans. Circuit Theory*, vol. CT-16, pp. 318–323, Aug. 1969.
[12] J. W. Bandler and R. E. Seviora, "Current trends in network optimization," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-18, pp. 1159–1170, Dec. 1970.
[13] J. W. Bandler, "Conditions for a minimax optimum," *IEEE Trans. Circuit Theory* (Corresp.), vol. CT-18, pp. 476–479, July 1971.
[14] L. S. Lasdon, *Optimization Theory for Large Systems.* New York: Macmillan, 1970, ch. 1.
[15] "Subroutine SIMPLE" Data Processing and Computing Centre, Library Information Sheet MILIS 5.3.130, McMaster Univ., Hamilton, Ont., Canada.
[16] G. C. Temes, "Optimization methods in circuit design," in *Computer Oriented Circuit Design*, F. F. Kuo and W. G. Magnuson, Jr., Eds. Englewood Cliffs, N. J.: Prentice-Hall, 1969.
[17] "Subroutine TGSORT" Data Processing and Computing Centre, Library Information Sheet MILIS 5.3.34, McMaster Univ., Hamilton, Ont., Canada.
[18] J. W. Bandler and P. A. Macdonald, "Cascaded noncommensurate transmission-line networks as optimization problems," *IEEE Trans. Circuit Theory* (Corresp.), vol. CT-16, pp. 391–394, Aug. 1969.
[19] H. J. Carlin and O. P. Gupta, "Computer design of filters with lumped-distributed elements or frequency variable terminations," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 598–604, Aug. 1969.
[20] J. W. Bandler, "Optimization methods for computer-aided design," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 533–552, Aug. 1969.
[21] J. W. Bandler, N. D. Markettos, and T. V. Srinivasan, "A comparison of recent minimax techniques on optimum system modelling," in *Proc. 6th Princeton Conf. Information Sciences and Systems* (Princeton, N. J., Mar. 1972).

Anomalous Loss at a Ferrite Boundary

LEONARD LEWIN, ASSOCIATE MEMBER, IEEE

Abstract—The occurrence of anomalous loss and its explanation in terms of surface waves is discussed. For this type of explanation to be possible the region of occurrence of the surface wave must at least straddle the region of anomalous loss. It is shown that this is so, particularly for the case when there is a mixed air–ferrite surface layer for which this result is not obvious: as the air content decreases, a ferrite–metal surface wave appears and takes over the function of the layer wave. The means by which these waves are generated, and the determination of their amplitudes, appear to require a new physical principle to be applied. A new type of "edge condition" is postulated.

Manuscript received December 15, 1971; revised March 10, 1972.
The author is with the Department of Electrical Engineering, University of Colorado, Boulder, Colo. 80302.

I. INTRODUCTION

THE FIRST intimation that something peculiar could be happening in a waveguide–ferrite configuration appeared in a paper by Lax and Button [1] and led to the so-called "thermodynamic paradox," in which energy could apparently be conveyed in only one direction in a lossless medium. Bresler [2] at-