

Practical Least p th Optimization of Networks

JOHN W. BANDLER AND CHRISTAKIS CHARALAMBOUS

Abstract—A new and practical approach to computer-aided design optimization is presented. Central to the process is the application of least p th approximation using extremely large values of p , typically 1000 to 1 000 000. It is shown how suitable and reasonably well-conditioned objective functions can be formulated, giving particular emphasis to more general approximation problems as, for example, in filter design. It is demonstrated how easily and efficiently extremely near minimax results can be achieved on a discrete set of sample points. Highly efficient gradient methods can be employed and, in network design problems, the use of the adjoint network approach for evaluating gradients results in greater savings in computer effort. A comparison between the Fletcher–Powell method and the more recent Fletcher method is made on the application of least p th approximation, using a range of values of p up to 1 000 000 000 000 on transmission-line transformer problems for which optimal minimax solutions are known. This is followed by filter design examples subject to certain constraints.

I. INTRODUCTION

LEAST p th approximation with a sufficiently large value of p can, in principle, be used to achieve near minimax approximations for a wide class of circuit- and system-design problems. Early reports [1]–[4] simply suggested that appropriate error functions be raised to a power p . This approach, in practice, can lead to ill-conditioning of the objective function for values of p greater than or equal to about 10. In certain design problems the unwary designer may be led to the conclusion that his problem has many local minima (see, for example, [1] and [2]) in a region of the parameter space where, in fact, a unique minimum exists.

Bandler and Charalambous [5] have shown how to apply least p th approximation to design problems having upper and lower response specifications, e.g., as in filter design. However, the same ill-conditioning could arise in that particular formulation. More recent theoretical work has been published on conditions for optimality in least p th approximation with $p \rightarrow \infty$ [6], from which conditions for a minimax approximation [7] fall out.

It is the purpose of the present paper to present a computationally practical approach to least p th approximation for use in design problems. The important

feature of the approach is the use that can be made of efficient gradient minimization techniques, such as the Fletcher–Powell method [8] and the more recent Fletcher method [9], in conjunction with least p th objective functions employing extremely large values of p , typically 1000–1 000 000. It is demonstrated how easily and efficiently extremely near minimax results can be achieved on a discrete set of sample points.

A comparison between the Fletcher–Powell and Fletcher methods is made using a range of values of p up to 1 000 000 000 000 on transmission-line transformer problems for which optimal minimax solutions are known. Filter-design examples with constraints are also provided. In all cases the adjoint network method [4] is used to obtain all the required partial derivatives at a given point in the parameter space from the results of one network analysis.

II. THEORY

Definitions

The notation to be used in this paper largely follows that used previously by the authors [3], [5].

$F(\phi, \psi)$	The approximating function (actual response).
$S_u(\psi)$	An upper specified function (desired response bound).
$S_u'(\psi, \xi)$	An artificial upper specified function.
$S_l(\psi)$	A lower specified function (desired response bound).
$S_l'(\psi, \xi)$	An artificial lower specified function.
$w_u(\psi)$	An upper positive weighting function.
$w_l(\psi)$	A lower positive weighting function.
ϕ	A vector containing the k independent parameters.
ψ	An independent variable (e.g., frequency or time).
ξ	Margin of errors with respect to the artificial and desired specifications.

The introduction of the artificial margin ξ , which is a constant during optimization, allows for certain flexibility in formulating the optimization problem. Its advantage will become evident at a later stage.

Now we can define real error functions related to the upper and lower specifications as

$$e_u(\phi, \psi) \triangleq w_u(\psi)(F(\phi, \psi) - S_u(\psi)) \quad (1)$$

$$e_u'(\phi, \psi, \xi) \triangleq w_u(\psi)(F(\phi, \psi) - S_u'(\psi, \xi)) = e_u(\phi, \psi) - \xi \quad (2)$$

Manuscript received June 9, 1972; revised August 4, 1972. This work was supported by the National Research Council of Canada under Grants A7239 and C154, and by a Frederick Gardner Cottrell Grant from the Research Corporation. This paper is based on papers presented at the 5th Asilomar Conference on Circuits and Systems, Pacific Grove, Calif., November 1971, and the 1972 IEEE International Microwave Symposium, Chicago, Ill., May 22–24, 1972.

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

$$e_l(\phi, \psi) \triangleq w_l(\psi)(F(\phi, \psi) - S_l(\psi)) \quad (3)$$

$$e_l'(\phi, \psi, \xi) \triangleq w_l(\psi)(F(\phi, \psi) - S_l'(\psi, \xi)) = e_l(\phi, \psi) + \xi \quad (4)$$

where $S_u'(\psi, \xi)$ and $S_l'(\psi, \xi)$ are taken, respectively, as

$$S_u'(\psi, \xi) = S_u(\psi) + \frac{\xi}{w_u(\psi)} \quad (5)$$

$$S_l'(\psi, \xi) = S_l(\psi) - \frac{\xi}{w_l(\psi)}. \quad (6)$$

In practice, we will evaluate all the functions at a finite discrete set of values of ψ taken from one or more closed intervals. Therefore, we will define the functions

$$e_{ui}'(\phi, \xi) \triangleq e_{ui}'(\phi, \psi_i, \xi), \quad i \in I_u \quad (7)$$

$$e_{li}'(\phi, \xi) \triangleq e_{li}'(\phi, \psi_i, \xi), \quad i \in I_l \quad (8)$$

where it is assumed that a sufficient number of sample points have been chosen so that the discrete approximation problem adequately approximates the continuous problem. I_u and I_l are appropriate index sets. We assume that we can choose ψ , such that the corresponding $e_{ui}'(\phi, \xi)$ and $e_{li}'(\phi, \xi)$ are continuous with continuous derivatives with respect to ϕ .

The Objective Function

Here we have to consider two separate cases, the first one when the specification is violated and the second one when the specification is satisfied.

In the first case some of the $e_{ui}'(\phi, \xi)$ or $-e_{li}'(\phi, \xi)$ are positive. To meet the artificial specification [same as original if $\xi = 0$ as indicated by (5) and (6)] we might propose the following objective function to be minimized:

$$U(\phi, \xi) = \left(\sum_{i \in J_u(\phi, \xi)} [e_{ui}'(\phi, \xi)]^p + \sum_{i \in J_l(\phi, \xi)} [-e_{li}'(\phi, \xi)]^p \right)^{1/p} \quad (9)$$

where¹

$$J_u(\phi, \xi) \triangleq \{i \mid e_{ui}'(\phi, \xi) \geq 0, \quad i \in I_u\} \quad (10)$$

$$J_l(\phi, \xi) \triangleq \{i \mid -e_{li}'(\phi, \xi) \geq 0, \quad i \in I_l\} \quad (11)$$

and

$$p > 1.$$

For larger values of p we would expect the maximum of the functions to be emphasized, since

¹ It is important to note that the sets J_u and J_l are dependent on ϕ and ξ . Thus temporarily excluded sample points are immediately included when the corresponding errors violate the specification, and temporarily included sample points are immediately excluded when the corresponding errors satisfy the specification.

$$\begin{aligned} \max_{i,j} [e_{ui}'(\phi, \xi), -e_{lj}'(\phi, \xi)] \\ = \lim_{p \rightarrow \infty} U(\phi, \xi), \quad i \in J_u(\phi, \xi) \\ i \in J_l(\phi, \xi). \end{aligned} \quad (12)$$

Letting

$$\nabla \triangleq \begin{pmatrix} \frac{\partial}{\partial \phi_1} \\ \frac{\partial}{\partial \phi_2} \\ \vdots \\ \frac{\partial}{\partial \phi_k} \end{pmatrix} \quad (13)$$

we have

$$\begin{aligned} \nabla U(\phi, \xi) &= \left(\sum_{i \in J_u(\phi, \xi)} [e_{ui}'(\phi, \xi)]^p + \sum_{i \in J_l(\phi, \xi)} [-e_{li}'(\phi, \xi)]^p \right)^{(1/p)-1} \\ &\cdot \left(\sum_{i \in J_u(\phi, \xi)} [e_{ui}'(\phi, \xi)]^{p-1} \nabla e_{ui}'(\phi, \xi) - \sum_{i \in J_l(\phi, \xi)} [-e_{li}'(\phi, \xi)]^{p-1} \nabla e_{li}'(\phi, \xi) \right). \end{aligned} \quad (14)$$

For the second case all the $-e_{ui}'(\phi, \xi)$ and $e_{li}'(\phi, \xi)$ will be positive. To exceed the specification by the greatest amount, we might propose the following objective function to be minimized:

$$U(\phi, \xi) = - \left(\sum_{i \in I_u} [-e_{ui}'(\phi, \xi)]^{-p} + \sum_{i \in I_l} [e_{li}'(\phi, \xi)]^{-p} \right)^{-(1/p)} \quad (15)$$

for

$$-e_{ui}'(\phi, \xi) > 0, \quad i \in I_u \quad (16)$$

$$e_{li}'(\phi, \xi) > 0, \quad i \in I_l \quad (17)$$

and

$$p \geq 1.$$

Again, for larger values of p we would expect the maximum of the functions to be emphasized, since

$$\begin{aligned} \max_{i,j} [e_{ui}'(\phi, \xi), -e_{lj}'(\phi, \xi)] \\ = \lim_{p \rightarrow \infty} U(\phi, \xi), \quad i \in I_u \\ j \in I_l. \end{aligned} \quad (18)$$

Differentiating (15) we obtain

$$\begin{aligned} \nabla U(\phi, \xi) = & \left(\sum_{i \in I_u} [-e_{ui}'(\phi, \xi)]^{-p} \right. \\ & \left. + \sum_{i \in I_l} [e_{li}'(\phi, \xi)]^{-p} \right)^{-(1/p)-1} \\ & \cdot \left(\sum_{i \in I_u} [-e_{ui}'(\phi, \xi)]^{-p-1} \nabla e_{ui}'(\phi, \xi) \right. \\ & \left. - \sum_{i \in I_l} [e_{li}'(\phi, \xi)]^{-p-1} \nabla e_{li}'(\phi, \xi) \right). \quad (19) \end{aligned}$$

In an effort to alleviate the ill-conditioning resulting from the numerical evaluation of $[\pm e_{ui}']^{\pm p}$ and $[\mp e_{li}']^{\pm p}$ for very large values of p , we use

$$M(\phi, \xi) \triangleq \max_{i,j} [e_{ui}'(\phi, \xi), -e_{li}'(\phi, \xi)], \quad \begin{array}{l} i \in I_u \\ j \in I_l \end{array} \quad (20)$$

as follows. Let

$$\begin{aligned} U(\phi, \xi) = M(\phi, \xi) & \left(\sum_{i \in K_u} \left[\frac{e_{ui}'(\phi, \xi)}{M(\phi, \xi)} \right]^q \right. \\ & \left. + \sum_{i \in K_l} \left[\frac{-e_{li}'(\phi, \xi)}{M(\phi, \xi)} \right]^q \right)^{1/q} \quad (21) \end{aligned}$$

and

$$\begin{aligned} \nabla U(\phi, \xi) = & \left(\sum_{i \in K_u} \left[\frac{e_{ui}'(\phi, \xi)}{M(\phi, \xi)} \right]^q \right. \\ & \left. + \sum_{i \in K_l} \left[\frac{-e_{li}'(\phi, \xi)}{M(\phi, \xi)} \right]^q \right)^{(1/q)-1} \\ & \cdot \left(\sum_{i \in K_u} \left[\frac{e_{ui}'(\phi, \xi)}{M(\phi, \xi)} \right]^{q-1} \nabla e_{ui}'(\phi, \xi) \right. \\ & \left. - \sum_{i \in K_l} \left[\frac{-e_{li}'(\phi, \xi)}{M(\phi, \xi)} \right]^{q-1} \nabla e_{li}'(\phi, \xi) \right) \quad (22) \end{aligned}$$

where

$$K_u \triangleq \begin{cases} J_u(\phi, \xi) & \text{if } M(\phi, \xi) > 0 \\ I_u & \text{if } M(\phi, \xi) < 0 \end{cases} \quad (23)$$

$$K_l \triangleq \begin{cases} J_l(\phi, \xi) & \text{if } M(\phi, \xi) > 0 \\ I_l & \text{if } M(\phi, \xi) < 0 \end{cases} \quad (24)$$

and

$$q \triangleq \frac{M(\phi, \xi)}{|M(\phi, \xi)|} p \begin{cases} 1 < p < \infty & \text{for } M > 0 \\ 1 \leq p < \infty & \text{for } M < 0. \end{cases} \quad (25)$$

If $e_{ui}'(\phi, \xi)$ for $i \in I_u$ and $e_{li}'(\phi, \xi)$ for $i \in I_l$ are continuous with continuous partial derivatives, then, under the stated conditions, the proposed objective function is continuous everywhere with continuous partial derivatives. When, for example, both $M=0$ and two or more

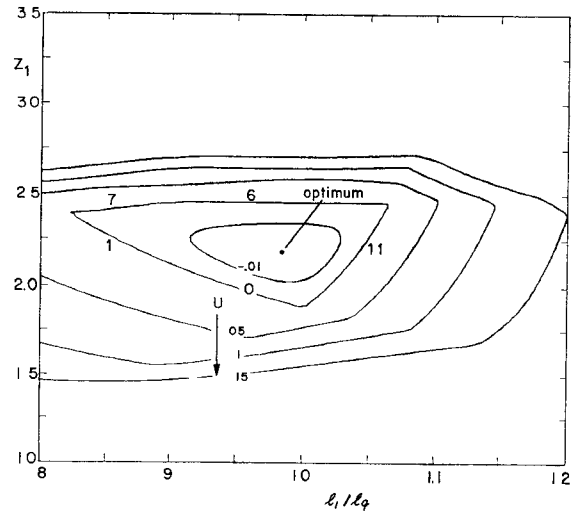


Fig. 1. Contours of U for $p=1$ (see footnote 2).

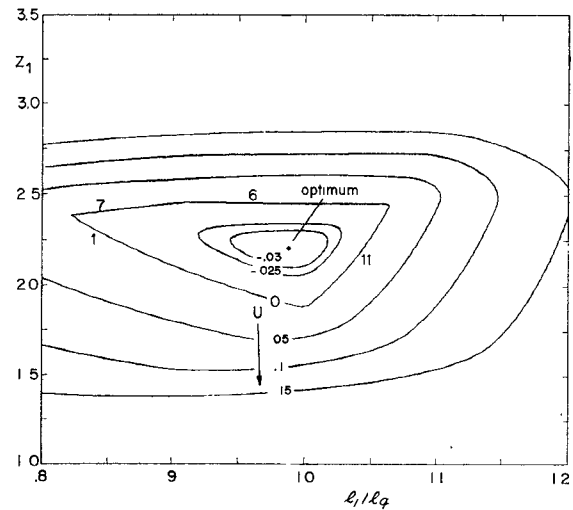


Fig. 2. Contours of U for $p=2$ (see footnote 2).

maxima are equal, the function is continuous but the derivatives are discontinuous. On the rare occasions when this situation causes a gradient minimization algorithm to stall, one can simply change the value of ξ and restart the optimization process.

Examples

It is convenient to illustrate some of these theoretical ideas on a simple problem which has already received attention from the optimization point of view [10]–[12]. In particular, we will consider a 2-section 10:1 transmission-line transformer over a 100-percent relative bandwidth [13]. The section lengths l_1 and l_2 and the characteristic impedances Z_1 and Z_2 may be considered as design variables.

Let

$$e_{ui}' = |\rho_i| - 0.5 \quad (26)$$

where ρ is the reflection coefficient. We have no lower specification, and we consider 11 uniformly spaced fre-

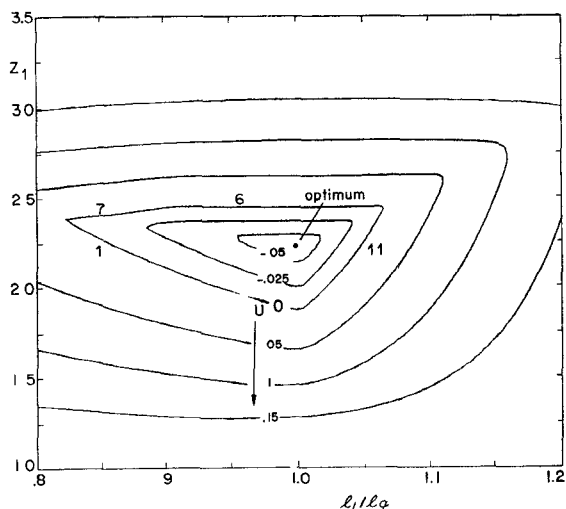


Fig. 3. Contours of U for $p=10$ (see footnote 2).

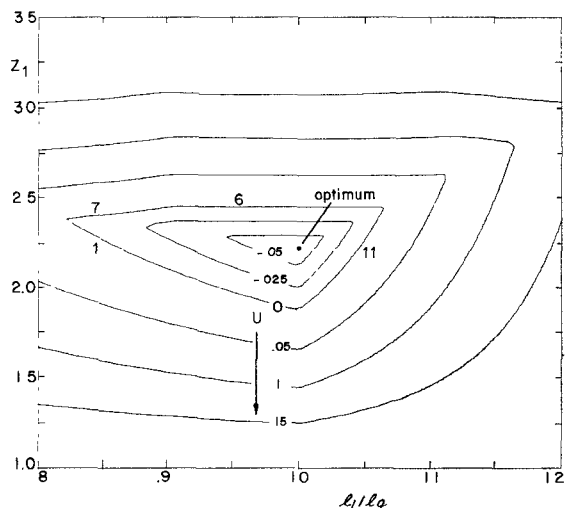


Fig. 4. Contours of U for $p=\infty$ (see footnote 2).

quencies in the band 0.5–1.5 GHz. The weighting is 1, $\xi=0.5$, while the upper specification is taken as 0.

Figs. 1–4 show contours of the function (21) with $K_l = \emptyset$ and e_{ui}' as in (26) for different values of p , namely, 1, 2, 10, and ∞ . l_2 is held fixed at the optimum quarter-wave value l_q , and Z_2 is held at the minimax optimum normalized value of 4.4721. The contours are plotted with respect to l_1/l_q and Z_1 .

Note that the contour $M=0$, i.e., $\max_i |\rho_i| = 0.5$, is common to all the figures, and the presence of discontinuous derivatives (sharp points) on this contour. Sample points generating particular sections of this contour are noted.² For $p=1$ with $M>0$, and for $p=\infty$ for all M , discontinuous derivatives are observed elsewhere. These situations have been deliberately avoided in the foregoing formulation.

² The numbers 1, 6, 7, and 11 indicated in Figs. 1–4 refer to sections of the contour $M=0$ generated by the first, sixth, seventh, and eleventh sample points of the 11 uniformly spaced points on the band from 0.5 to 1.5 GHz.

III. EXAMPLES

Example 1

To compare the performance of two efficient optimization methods, the Fletcher–Powell method [8] and the Fletcher method [9], for widely different values of p we will consider the design of 2- and 3-section 10:1 transmission-line transformers [13], as previously, over a 100-percent band. In this case let

$$e_{ui}' = |\rho_i| \tag{27}$$

and the 11 sample points are, in gigahertz,

$$\{0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5\}$$

for the 2-section case (as previously) and

$$\{0.5, 0.6, 0.7, 0.77, 0.9, 1.0, 1.1, 1.23, 1.3, 1.4, 1.5\}$$

for convenience, for the 3-section case. The weighting is 1, and both $\xi=0$ and the upper specification is 0. Appropriate gradient vectors with respect to length and characteristic impedance of transmission lines are calculated by the adjoint network method [4].

The progress of the two algorithms used on a CDC 6400 computer from indicated starting points is summarized in Tables I and II and Figs. 5–9. For convenience we show results for different values of n where $p=10^n$. The figures are plots of M of (20) against N , the number of function evaluations at the beginning of an iteration. One function evaluation includes the evaluation of appropriate gradients.

Example 2

For the circuit shown in Fig. 10, suppose it is desired to have solutions to the following two problems.

- 1) An insertion loss in the passband, 0–1 GHz, of no more than 0.01 dB while maximizing the minimum stop-band insertion loss at 5 GHz.
- 2) An insertion loss in the passband, 0–1 GHz, of no more than 0.01 dB while maximizing the minimum stop-band insertion loss over the range 2.5–10 GHz.

The characteristic impedances are to be set fixed at the values [14]

$$Z_1 = Z_3 = Z_5 = 0.2$$

$$Z_2 = Z_4 = 5$$

and the section lengths used as variables. It was decided initially to select 21 uniformly spaced sample points from 0–1 GHz, letting

$$e_{ui}' = |\rho_i| - (0.047960 + \xi), \quad i = 1, 2, \dots, 21 \tag{28}$$

where 0.047960 is the passband reflection coefficient corresponding to 0.01 dB, and a lower error function at 5 GHz for problem 1, and at 16 uniformly spaced sample points from 2.5–10 GHz for problem 2, i.e.,

$$e_{ui}' = |\rho_i| - (1 - \xi) \begin{cases} i=22 \text{ for problem 1} \\ i=22, 23, \dots, 37 \text{ for problem 2.} \end{cases} \tag{29}$$

In problem 1 we set $\xi=0.00204$ and in problem 2 we set

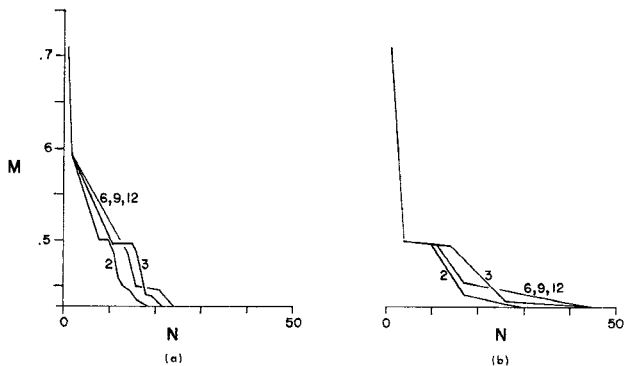


Fig. 5. Optimization from $Z_1=1.0, Z_2=3.0$. (a) Fletcher. (b) Fletcher-Powell.

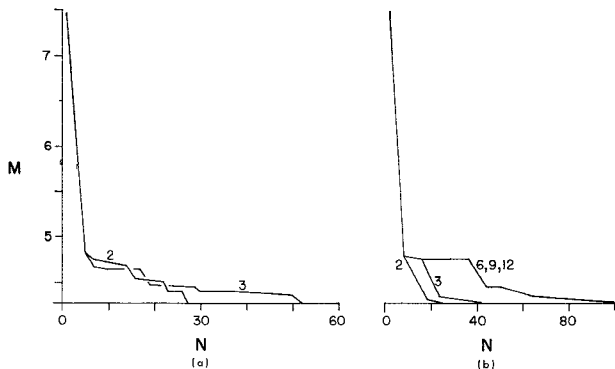


Fig. 6. Optimization from $Z_1=1.0, Z_2=6.0$. (a) Fletcher. (b) Fletcher-Powell.

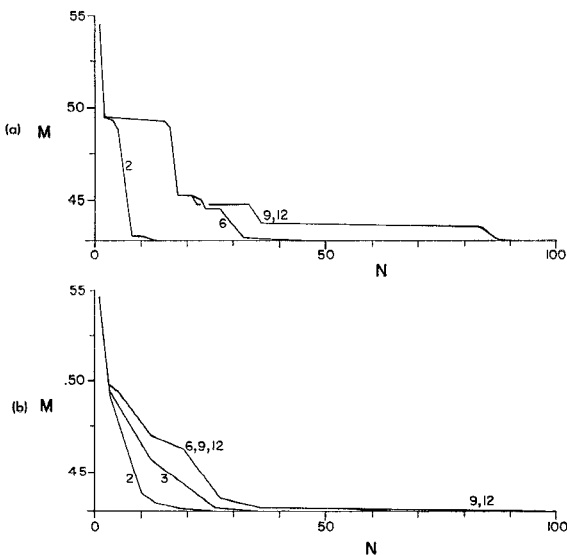


Fig. 7. Optimization from $Z_1=3.5, Z_2=6.0$. (a) Fletcher. (b) Fletcher-Powell.

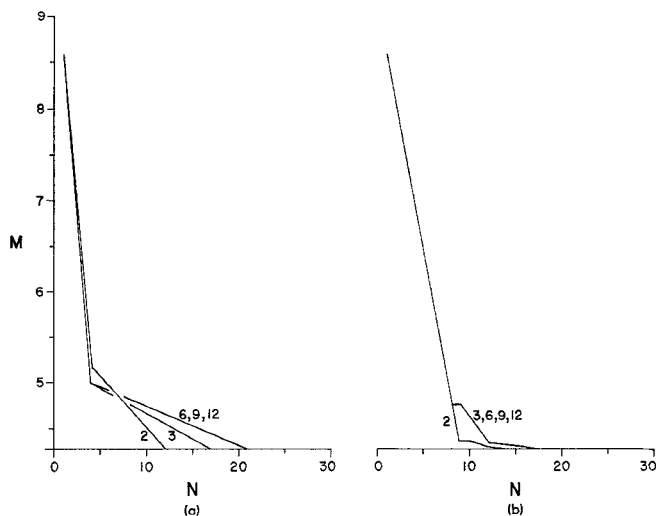


Fig. 8. Optimization from $Z_1=3.5, Z_2=3.0$. (a) Fletcher. (b) Fletcher-Powell.

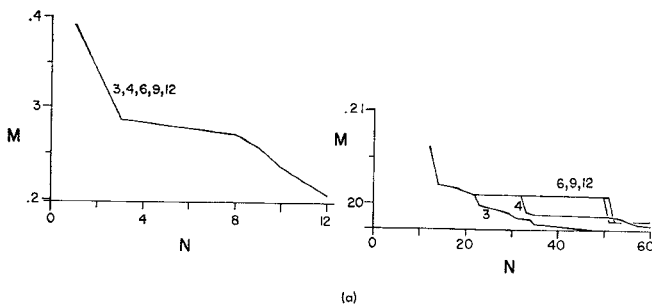


Fig. 9. Optimization from $Z_1=1.5, Z_2=3.0, Z_3=6.0, l_1/l_q=0.8, l_2/l_q=1.2, l_3/l_q=0.8$. (a) Fletcher. (b) Fletcher-Powell.

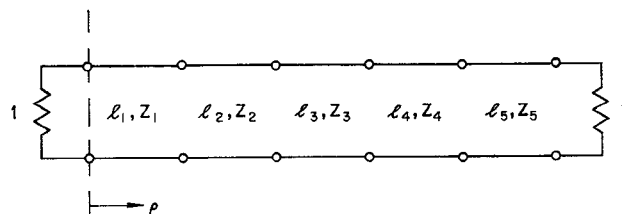


Fig. 10. 5-section transmission-line low-pass filter.

$\xi=0$. Since the response at zero frequency is independent of the parameters, and to avoid numerical difficulties, the frequency point 0.02 replaced 0 in problem 1.

Optimization using the Fletcher method in accordance

with the foregoing ideas with $p=1000$ gave the results shown in Table III, where l_q is the quarter-wave value at 1 GHz.

The responses are depicted in Figs. 11-13. The final

TABLE I

OPTIMIZATION OF A 2-SECTION 10:1 QUARTER-WAVE TRANSFORMER OVER 100-PERCENT BANDWIDTH WITH VARIABLE CHARACTERISTIC IMPEDANCES Z_1 AND Z_2

Figure	Starting Point		n where $p=10^n$	Number of Function Evaluations N^a	
	Z_1	Z_2		Fletcher [9]	Fletcher-Powell [8]
5	1.0	3.0	2	22	31
			3	28	49
			6	33	56
			9	33	56
			12	33	56
6	1.0	6.0	2	30	26
			3	58	50
			6	^b	133
			9	^b	172
			12	^b	198
7	3.5	6.0	2	15	23
			3	^b	41
			6	44	101
			9	102	118
			12	102	118
8	3.5	3.0	2	14	16
			3	19	56
			6	21	75
			9	21	85
			12	21	310

^a The number of N listed are those required to bring M within 0.01 percent of the known optimum value, namely, 0.42857.

^b Missing entries are due to parameters becoming negative—constraints were not imposed during optimization.

TABLE II

OPTIMIZATION OF A 3-SECTION 10:1 TRANSFORMER OVER 100-PERCENT BANDWIDTH WITH VARIABLE LENGTHS AND CHARACTERISTIC IMPEDANCES

Starting Point: $Z_1=1.5$, $Z_2=3.0$, $Z_3=6.0$, $l_1/l_q=0.8$, $l_2/l_q=1.2$, $l_3/l_q=0.8$, where l_q is the quarter wavelength at center frequency

Number of Function Evaluations N to Reach the Value of M Shown in Brackets^a; the Optimum Value of M is 0.19729

n where $p=10^n$	Fletcher [9]	Fletcher-Powell [8]
3	57 (0.19734)	115 (0.19733)
4	86 (0.19730)	378 (0.19729)
6	418 (0.19729)	702 (0.19740)
9	634 (0.19730)	661 (0.19740)
12	668 (0.19736)	645 (0.19851)

^a A time limit of 64 s/run was imposed, at which time the optimum for large p had still not been reached.

TABLE III

OPTIMIZATION OF THE CIRCUIT SHOWN IN FIG. 10 USING VARIABLE LENGTHS

Parameters	Starting Point	Problem 1	Problem 2
$\frac{l_1}{l_q} = \frac{l_3}{l_q}$	0.07	0.09593	0.09098
$\frac{l_2}{l_q}$	0.15	0.16278	0.18928
$\frac{l_2}{l_q} = \frac{l_4}{l_q}$	0.15	0.19798	0.15821

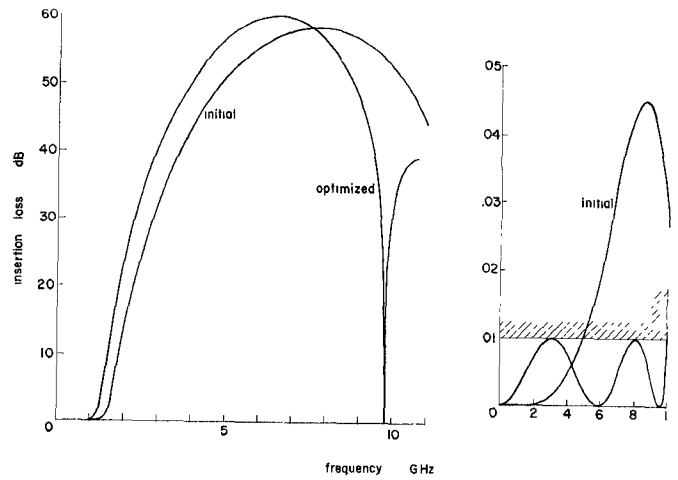


Fig. 11. Optimized response of the circuit of Fig. 10 subject to the constraints imposed for problem 1.

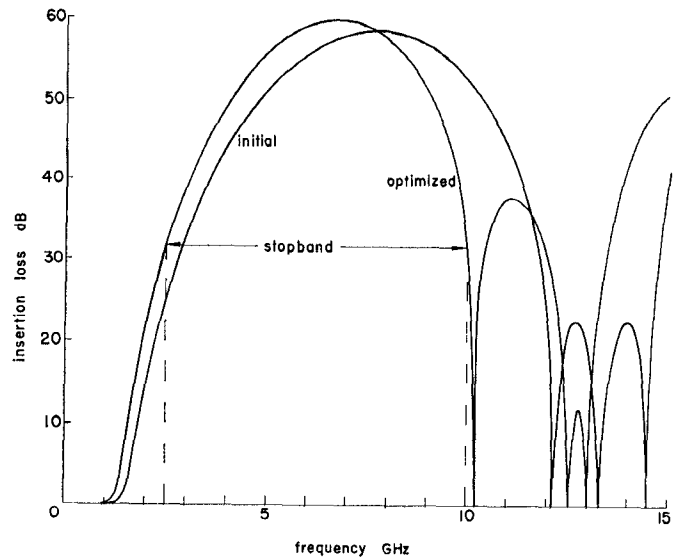


Fig. 12. Optimized response of the circuit of Fig. 10 subject to the constraints imposed for problem 2.

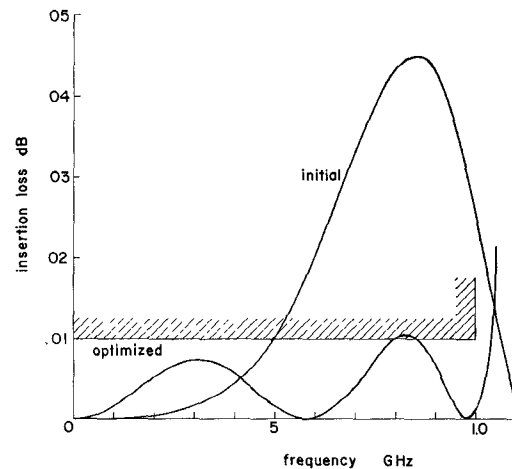


Fig. 13. Passband details of the optimized response shown in Fig. 12.

results are the same whether or not symmetry is assumed. Problem 2, for example, was solved using the Fletcher-Powell method without the symmetry assumption.

Three comments on the responses shown in Figs. 11-13 are in order. The first is that perfectly equal-ripple responses should not be expected in general nonlinear approximation problems, with or without constraints. The second is that, unless interpolation methods [3], [11] are used, actual extrema in the response errors will usually lie between adjacent discrete sample points. The third is that, in the present examples, slight deviations from the passband specification are to be expected, since the stopband specification is unattainable in practice.

IV. DISCUSSION

From a minimax point of view ($p = \infty$), the value of the parameter ξ does not affect the location of the optimum. For finite values of p , however, it can play an extremely important role. It can be chosen, if desired, so that the M of (20) is always positive or, alternatively, always negative. In the first case an economy in gradient computation may be realized since only sample points satisfying the conditions in (10) and (11) are considered. This is a subset of all the possible sample points. In the second case all the sample points would generally have to be considered, but in our experience convergence to a good solution is usually faster. In this case, of course, we avoid the mild possible hazards mentioned in Section II encountered in the transition region when $M = 0$.

Theoretically, if ξ is chosen such that $M = 0$ is optimal, then a finite value of p will yield the minimax solution! In practice, a good estimate for ξ may allow relatively low values of p to yield results much closer to the minimax solution than a bad estimate. As Figs. 5-9 indicate the lower the value of p the faster M is reduced in the early stages of optimization. This is not unexpected since the minimization algorithms used are based on quadratic models. As p increases the objective function will generally deviate further and further from a quadratic form so that the algorithms will progressively slow down.

It is, incidentally, always good practice to monitor the current minimum value of M and the associated ϕ while U is being minimized, since a lower, and hence presumably preferable, value of M may be realized on the way than might prevail at the minimizing point for U .

V. CONCLUSIONS

An approach to computer-aided minimax design of microwave circuits employing highly efficient optimization methods has been presented. Typically, less than 1 min of CDC 6400 computer time is sufficient to optimize the type of examples given in this paper to a high degree of accuracy.

Other recent work on least p th approximation using very large p is the work by Fletcher *et al.* [15] on linear

approximation problems, and the work by Bandler *et al.* [16] on optimum system modeling problems in the time domain. The latter paper, in particular, compares the grazor search method [12] with the present approach.

No attempts at modifying the minimization methods to improve convergence for extremely large values of p nor a detailed study of other possible effects of numerical ill-conditioning have as yet been carried out. But, if the success we have had is widely repeatable, then far-reaching consequences are foreseen, not only in nonlinear approximation, but in the closely related field of nonlinear programming [17].

ACKNOWLEDGMENT

The authors wish to thank Dr. R. Fletcher, Atomic Energy Research Establishment, Harwell, England, who made available his computer program [9]. They also wish to thank Dr. M. D. Hebden, formerly at McMaster University, now also at the Atomic Energy Research Establishment, Harwell, England, for useful suggestions on least p th approximation.

REFERENCES

- [1] G. C. Temes and D. Y. F. Zai, "Least p th approximation," *IEEE Trans. Circuit Theory* (Corresp.), vol. CT-16, pp. 235-237, May 1969.
- [2] 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.
- [3] J. W. Bandler, "Optimization methods for computer-aided design," *IEEE Trans. Microwave Theory Tech. (Special Issue on Computer-Oriented Microwave Practices)*, vol. MTT-17, pp. 533-552, Aug. 1969.
- [4] J. W. Bandler and R. E. Seviara, "Current trends in network optimization," *IEEE Trans. Microwave Theory Tech. (1970 Symposium Issue)*, vol. MTT-18, pp. 1159-1170, Dec. 1970.
- [5] J. W. Bandler and C. Charalambous, "Theory of generalized least p th approximation," *IEEE Trans. Circuit Theory* (Corresp.), vol. CT-19, pp. 287-289, May 1972.
- [6] —, "On conditions for optimality in least p th approximation with $p \rightarrow \infty$," *J. Optimiz. Theory Appl.*, to be published.
- [7] J. W. Bandler, "Conditions for a minimax optimum," *IEEE Trans. Circuit Theory* (Corresp.), vol. CT-18, pp. 476-479, July 1971.
- [8] R. Fletcher and M. J. D. Powell, "A rapidly convergent descent method for minimization," *Comput. J.*, vol. 6, pp. 163-168, June 1963.
- [9] R. Fletcher, "A new approach to variable metric algorithms," *Comput. J.*, vol. 13, pp. 317-322, Aug. 1970.
- [10] 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.
- [11] —, "Optimization of microwave networks by razor search," *IEEE Trans. Microwave Theory Tech. (Special Issue on Computer-Oriented Microwave Practices)*, vol. MTT-17, pp. 552-562, Aug. 1969.
- [12] J. W. Bandler, T. V. Srinivasan, and C. Charalambous, "Minimax optimization of networks by grazor search," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-20, pp. 596-604, Sept. 1972.
- [13] G. L. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance Matching Networks and Coupling Structures*. New York: McGraw-Hill, 1964, ch. 6.
- [14] C. Brancher, F. Maffioli, and A. Premoli, "Computer optimization of cascaded noncommensurable-line lowpass filters," *Electron. Lett.*, vol. 6, pp. 513-515, Aug. 1970.
- [15] R. Fletcher, J. A. Grant, and M. D. Hebden, "Linear minimax approximation as the limit of best L_p approximation," Atomic Energy Res. Estab., Harwell, England, TP470, Dec. 1971.
- [16] J. W. Bandler, N. D. Marketos, and T. V. Srinivasan, "A comparison of recent minimax techniques on optimum system modeling," in *Proc. 6th Princeton Conf. Information Sciences and Systems* (Princeton, N. J., Mar. 1972), pp. 540-544.
- [17] J. W. Bandler and C. Charalambous, "A new approach to nonlinear programming," in *Proc. 5th Hawaii Int. Conf. System Sciences* (Honolulu, Hawaii, Jan. 1972), pp. 127-129.