INTERNAL REPORTS IN

SIMULATION, OPTIMIZATION AND CONTROL

No. SOC-80

NEW RESULTS IN THE LEAST PTH APPROACH TO MINIMAX DESIGN

J.W. Bandler, C. Charalambous, J.H.K. Chen and W.Y. Chu

March 1975

FACULTY OF ENGINEERING
McMASTER UNIVERSITY
HAMILTON, ONTARIO, CANADA

NEW RESULTS IN THE LEAST PTH APPROACH TO MINIMAX DESIGN

J.W. Bandler, C. Charalambous, J.H.K. Chen and W.Y. Chu

Abstract We present two general approaches for obtaining minimax designs through a sequence of least pth approximations yielding significant improvement in computational efficiency over previous least pth algorithms as well as highly accurate solutions. One utilizes a method for estimating a lower bound on the minimum (which has important design implications in itself) and the other extrapolation of least pth solutions to $p = \infty$. Documented computer programs are available for both methods. A practical feature is the successive and automatic reduction in sample points used in the optimization process allowing minimax solutions to be reached with only moderately more effort than required by a single least pth approximation. The application of these new techniques is illustrated by microwave transformer and filter examples.

This work was supported by the National Research Council of Canada under Grant A7239, by a Post doctorate Fellowship to C. Charalambous and a scholarship to W.Y. Chu.

J.W. Bandler and W.Y. Chu are with the Group on Simulation, Optimization and Control and Department of Electrical Engineering, McMaster University, Hamilton, Canada.

C. Charalambous is with the Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Canada.

J.H.K. Chen is with Bell-Northern Research, Ottawa, Canada.

I. INTRODUCTION

This paper is directed towards improving the efficiency of computing minimax optimum solutions to network design problems. Bandler and Charalambous have shown [1] how near minimax solutions can be obtained using least pth approximation with very large values of p. Here, we consider accelerating convergence to minimax solutions by extrapolating on a sequence of least pth solutions with geometrically increasing values of p. Another approach uses work by Charalambous and Bandler [2-4], in which a sequence of least pth solutions with finite, usually low, values of p are obtained in an effort to reach a minimax solution. The new feature here is based on recently derived results by Charalambous [5] using duality theory for nonlinear programming. Both approaches lead to efficient algorithms, as our results indicate.

Potentially inactive sample points are dropped from the optimization process as one proceeds with the computations. In particular, the approach based on the work of Charalambous permits the estimation of a true lower bound on the optimal minimax error function at any least pth optimum. We feel that the result is a very powerful one and has important design implications.

In this paper we compare the performance of the new methods with the results of Charalambous and Bandler [2] on a three-section quarter wave transformer. We also investigate the design of a seven-section microwave filter to illustrate the flexibility of the ideas presented.

II. THE ALGORITHMS

General Considerations

We briefly summarize the two approaches as follows. Basically, we minimize w.r.t. $\underset{\sim}{\phi}$ for given ξ and $p > 1$ the function

$$U(\underset{\sim}{\phi},\xi,p) \triangleq \begin{cases} M(\underset{\sim}{\phi},\xi)\left[\sum_{i\in K}\left(\dfrac{f_i(\underset{\sim}{\phi}) - \xi}{M(\underset{\sim}{\phi},\xi)}\right)^q\right]^{\frac{1}{q}} & \text{for } M(\underset{\sim}{\phi},\xi) \neq 0 \\ 0 \quad \text{for } M(\underset{\sim}{\phi},\xi) = 0 & \end{cases} \qquad (1)$$

where

$$M_f(\underset{\sim}{\phi}) \triangleq \max_{i\in I} f_i(\underset{\sim}{\phi}) \qquad (2)$$

$$M(\underset{\sim}{\phi},\xi) \triangleq M_f(\underset{\sim}{\phi}) - \xi \qquad (3)$$

$$K = \begin{cases} I \subset \{1,2,\ldots,m\} & \text{for } M(\underset{\sim}{\phi},\xi) < 0 \\ J \triangleq \{i \mid f_i(\underset{\sim}{\phi}) - \xi \geq 0,\ i \in I\} & \text{for } M(\underset{\sim}{\phi},\xi) > 0 \end{cases} \qquad (4)$$

$$q \triangleq p \text{ sgn } M(\underset{\sim}{\phi},\xi) \qquad (5)$$

and where $\underset{\sim}{\phi} \triangleq [\phi_1\ \phi_2\ \ldots\ \phi_k]^T$ is the design parameter vector, and $f_1(\underset{\sim}{\phi}), f_2(\underset{\sim}{\phi}), \ldots, f_m(\underset{\sim}{\phi})$ are m linear or nonlinear functions directly related to the response error functions such that if

$$M_f(\underset{\sim}{\phi}) \begin{cases} > 0 & \text{the specifications are violated} \\ = 0 & \text{the specifications are just met} \\ < 0 & \text{the specifications are satisfied} \end{cases}$$

Lower Bound for Minimax Solution

Charalambous has shown [5] that if we have $\underset{\sim}{u}$ and $\underset{\sim}{\phi}$ such that

$$\begin{aligned} &\sum_{i=1}^{m} u_i \underset{\sim}{\nabla} f_i(\underset{\sim}{\phi}) = \underset{\sim}{0} \\ &\sum_{i=1}^{m} u_i = 1 \end{aligned} \qquad (6)$$

$$u_i \geq 0, \; i = 1,2,\ldots,m$$

then

$$\sum_{i=1}^{m} u_i f_i(\underset{\sim}{\phi}) \leq M_f(\check{\underset{\sim}{\phi}}) \leq M_f(\underset{\sim}{\phi}) \tag{7}$$

where $\check{\underset{\sim}{\phi}}$ is the minimax optimum which is being sought, $\underset{\sim}{u} \triangleq [u_1 \; u_2 \; \ldots \; u_m]^T$ and $\underset{\sim}{\nabla} \triangleq [\partial/\partial\phi_1 \quad \partial/\partial\phi_2 \; \ldots \; \partial/\partial\phi_k]^T$. Furthermore, it is readily shown that the conditions (6) are satisfied at each optimum point $\check{\underset{\sim}{\phi}}(p,\xi)$ for a least pth objective function, yielding

$$\sum_{i=1}^{m} u_i f_i(\check{\underset{\sim}{\phi}}(p,\xi)) \leq M_f(\check{\underset{\sim}{\phi}}) \leq M_f(\check{\underset{\sim}{\phi}}(p,\xi)) \tag{8}$$

where, assuming K contains all critical sample points,

$$u_i = \frac{v_i}{\sum\limits_{i \varepsilon K} v_i} \tag{9}$$

$$v_i = \begin{cases} \left(\dfrac{f_i(\check{\underset{\sim}{\phi}}(p,\xi)) - \xi}{M(\check{\underset{\sim}{\phi}},\xi)} \right)^{q-1} & \text{for } i \; \varepsilon \; K \\ 0 \quad \text{for } i \not\varepsilon \; K \end{cases} \tag{10}$$

The first term of the inequality (8) is a lower bound on $M_f(\check{\underset{\sim}{\phi}})$. Together with the result that any single least pth solution will indicate whether the design specifications can ever be satisfied [2], this provides at any least pth solution an optimistic indication of the ultimate minimax error to be expected for a particular design.

The ξ-Algorithm

An algorithm has been developed and extensively tested based on setting the ξ for the next or (r+1)th optimization equal to the lower bound estimated using the solution to the rth optimization. p is kept constant.

Our algorithm begins at ϕ^0 by setting ξ^1 to an initial guess or estimate of a lower bound and $\xi^0 < \xi^1$ to a margin such that we have the option immediately to discard all functions f_i for which $f_i(\phi^0) < \xi^0$, $i = 1,2,\ldots,m$, which are considered a priori unlikely to be active at $\check{\phi}$. Subsequently, we obtain ξ^{r+1} for $r = 1,2,\ldots$ as indicated previously, and ξ^r is used as a level for discarding functions, so that $I^{r+1} = \{i | f_i(\check{\phi}^r) \geq \xi^r\}$. Thus, we can afford to start approximation with a large number of sampling points in order to reduce the possibility of missing some crucial points. The functions specifically required for the least pth objective and its gradients are successively reduced as the optimum is approached enabling a saving of effort in gradient computations. Features are built into our implementation to safeguard against our assumptions not being satisfied. The program, called MINOPT [6], is written in such a way that the optimization can be restarted efficiently from any point instead of having to repeat the entire process.

Extrapolation of Least pth Solutions

Under certain assumptions [7] we may use the same least pth objective, keep ξ constant throughout and minimize with respect to ϕ for geometrically increasing values of p. Suppose we have the (unique) minima corresponding to n_t values of p. Then we can develop the parameter vector ϕ as a polynomial function of $p' \triangleq \frac{1}{p}$, namely,

$$\phi(p_r') = \sum_{j=0}^{n-1} a_j (p_r')^j \quad , \quad r=1,\ldots,n_t \tag{11}$$

where the a_j are k-component vectors. This yields an approximation of $\phi(p')$ on $[0,p_1']$ and $\phi(0) = \check{\phi}$ is approximated by a_0. With $p_{r+1}' = p_r'/c$, where $c > 1$, we can obtain an extrapolation procedure of order $n(\leq n_t-1)$

based on the Richardson-Romberg principle [8] to estimate $\check{\phi}$. An algorithm which incorporates the extrapolation procedure and a scheme for dropping inactive error functions is described below. Theoretical justification of the extrapolation procedure has been developed [7].

The p-Algorithm

1. $\xi \leftarrow$ constant, $r \leftarrow 1$, $I^1 = \{1,2,\ldots,m\}$.

2. Minimize (1) w.r.t. ϕ for $p=p_r$.

3. $n_0 \leftarrow \min[r-1,n]$, where n is the highest order of extrapolation.

4. Compute estimates of the minimax solution using the following extrapolation formula

$$\phi_0^r \leftarrow \phi(p_r')$$

and for $r > 1$,

$$\phi_j^r \leftarrow \frac{c^j \phi_{j-1}^r - \phi_{j-1}^{r-1}}{c^j - 1}, \quad j=1,\ldots,n_0,$$

where ϕ_j^r signifies the jth order estimate of $\phi(0)$ after r minima have been obtained, and c is the multiplying factor for increasing the value of p_r.

5. $\phi(0) \leftarrow \phi_{n_0}^r$.

If $r = n_t$, stop.

Let $\phi_{n_0}^{r+1} \leftarrow \phi_{n_0}^r$ and estimate the next minimum by solving for $n_0 \geq 1$

$$\phi_{j-1}^{r+1} \leftarrow \frac{(c^j-1)\phi_j^{r+1} + \phi_{j-1}^r}{c^j}$$

for $j = n_0, \ldots, 1$. The starting point for the next optimization is

$$\phi \leftarrow \phi_0^{r+1}.$$

6. $p_{r+1} \leftarrow p_r * c$.

If $r = 1$, $p_{r+1} \leftarrow p_{r+1}/c$.

$$\mu_i \leftarrow \frac{v_i}{\sum_{i \in K} v_i}$$

where

$$v_i \leftarrow \begin{cases} \left(\dfrac{f_i(\phi) - \xi}{M(\phi, \xi)} \right)^{q_{r+1}} & \text{for } i \in K \\ 0 \quad \text{for } i \notin K \end{cases}$$

and K is chosen according to (4) with I set to I^r.

If $r = 1$, $p_{r+1} \leftarrow p_{r+1}{}^* c$.

$I^{r+1} = \{i | \mu_i > \eta \, , \, i \in I^r\}$.

7. $r \leftarrow r + 1$ and go to 2.

In Step 6 of the p-algorithm p is not increased when $r = 1$, i.e., after the first minimization, in determining the multipliers μ_i. No extrapolation can be made at this stage so the starting point for the next minimization is the solution to the first one. It was felt that these multipliers should, more conservatively, be based on the first value of p, in that case.

In the program, called FLOPT2[9], past solutions may be retained for future runs permitting extrapolation to be implemented immediately.

III. EXAMPLES

The unconstrained minimization method throughout was a recent quasi-Newton method [10], and the computer used was a CDC 6400.

Three-Section Transformer

The algorithms are first compared on a well-known test problem. We consider two starting points for the optimization in the minimax sense of a three-section 100 percent relative bandwidth 10:1 transmission-line transformer [2] as shown in Table I. We use the same sample points as Charalambous and Bandler [2] and compare the effort required to reach or exceed a reflection coefficient of 0.19729 (optimal to 5 figures) letting f_i be the modulus of the reflection coefficient. All 6 parameters are varied, lengths and characteristic impedances. The results, indicating about 1/3 to 1/2 the response evaluations used by the Charalambous-Bandler algorithms [2], are shown in Table II. Increasing p in the ξ-algorithm gave poorer results. The extrapolation approach (p-algorithm) appears relatively insensitive to the sequence of p used.

Table III shows details of the progress of the p-algorithm on Problem 1 and Table IV the corresponding progress of the ξ-algorithm.

Seven-Section Filter

A seven-section resistively terminated bandpass filter consisting of two unit elements and five stubs [11] as shown in Fig. 1 is considered next. Specifications of 0.1 dB from 1.0875 to 3.2625 GHz (passband) and 50 dB at 0.6 and 3.75 GHz is desired. The initial normalized characteristic impedances were taken as 0.63, 0.33, 1.27, 0.26, 1.27, 0.33, 0.63. The lengths were fixed at optimal normalized values of 1. Consider the difference in dB between the response and specifications. The functions f_i were set to $\pm$ the difference, the positive sign corresponding to the passband, the negative sign to the stopband. Twenty-one uniformly spaced passband points were initially considered but, due to symmetry, only the first 10 were actually used.

Table V shows the effort required by the 3rd order extrapolation approach. The final extrapolated solution gives characteristic impedances of 0.606458, 0.303062, 0.722085, 0.235612, 0.722085, 0.303062, 0.606458. (Symmetry was not enforced here but was obtained, as expected.) The 8 passband ripples as evaluated at the 21 points were all in the range 0.06530 - 0.06531 dB and the stopband responses were 50.0347 dB. CPU time on the CDC 6400 was 5 sec. About half this time would be expected if characteristic impedance symmetry were exploited. The response is plotted in Fig. 2. We note that, because of the uniformly spaced sample points, the response is not exactly equal ripple but meets the design specifications.

To illustrate the usefulness of the lower bound estimation, the ξ-algorithm was used to generate the results of Table VI employing sample points corresponding to the ripple maxima for the above example as shown in Table VII, for 4 sets of specifications. The ripple maxima in the passband were found by quadratic approximations based on sets of three adjacent sample points taken from 101 uniformly spaced candidates.

We note from Table VI that the process works whether the specifications are violated or satisfied, the results yielding an immediate indication of how good a design in the minimax sense one can expect from the results of only one least squares approximation. The weighting of the differences between response and specifications are uniform throughout this example. Different weights or values of p, however, cannot change the nature of the result, but only the amounts by which specifications are violated on the one hand or satisfied on the other.

Fig. 3 shows responses corresponding to the 4 sets of specifications after only two optimizations with $p = 2$ of the ξ-algorithm. The maximum errors are indicated in the 4th column of Table VI. The responses are essentially equal ripple for engineering purposes. Table VII summarizes the final solution for the 50 dB specification.

IV. CONCLUSIONS

Two new algorithms and related results for the least pth approach to minimax design have been presented. Documented computer programs, namely, MINOPT [6] and FLOPT2 [9] are available from the first author at nominal charge. The mathematical background has been ommitted, but is also available [5], [7]. Although impedance symmetry in the seven-section filter example and well-known corresponding assumptions for the three-section transformer example could easily have been made to simplify the computations (with appropriate reduction in running times) we felt that a demonstration of the power of the algorithms in readily forcing or maintaining these properties was worthwhile for testing purposes.

REFERENCES

[1] J.W. Bandler and C. Charalambous, "Practical least pth optimization of networks", IEEE Trans. Microwave Theory Tech., vol. MTT-20, Dec. 1972, pp. 834-840.

[2] C. Charalambous and J.W. Bandler, "New algorithms for network optimization", IEEE Trans. Microwave Theory Tech., vol. MTT-21, Dec. 1973, pp. 815-818.

[3] C. Charalambous, "Minimax design of recursive digital filters", Computer Aided Design, vol. 6, April 1974, pp. 73-82.

[4] C. Charalambous and J.W. Bandler, "Nonlinear minimax optimization as a sequence of least pth optimization with finite values of p", Int. J. Systems Science, to be published.

[5] C. Charalambous, "Minimax optimization of recursive digital filters using recent minimax results", IEEE Trans. Acoust. Speech and Signal Processing, to be published.

[6] J.W. Bandler, C. Charalambous and J.H.K. Chen, "MINOPT - an optimization program based on recent minimax results", Faculty of Engineering, McMaster University, Hamilton, Canada, Report SOC-70, Dec. 1974.

[7] W.Y. Chu, "Extrapolation in least pth approximation and nonlinear

programming", Faculty of Engineering, McMaster University, Hamilton, Canada, Report SOC-71, Dec. 1974.

[8] A.V. Fiacco and G.P. McCormick, Nonlinear Programming: Sequential Unconstrained Minimization Techniques. New York: Wiley, 1968.

[9] J.W. Bandler and W.Y. Chu, "FLOPT2- a program for least pth optimization with extrapolation to minimax solutions", Faculty of Engineering, McMaster University, Hamilton, Canada, Internal Report in Simulation, Optimization and Control, to appear.

[10] R. Fletcher, "FORTRAN subroutines for minimization by quasi-Newton methods", Atomic Energy Research Establishment, Harwell, Berkshire, England, Report AERE-R7125, 1972.

[11] M.C. Horton and R.J. Wenzel, "General theory and design of optimum quarter-wave TEM filters", IEEE Trans. Microwave Theory Tech., vol. MTT-13, May 1965, pp. 316-327.

TABLE I

THE STARTING AND SAMPLE POINTS IN THE OPTIMIZATION OF A THREE-SECTION 10:1 TRANSFORMER OVER 100-PERCENT RELATIVE BANDWIDTH

Parameters ϕ_i	Problem 1	Problem 2
ℓ_1/ℓ_q	1.0	0.8
Z_1	1.0	1.5
ℓ_2/ℓ_q	1.0	1.2
Z_2	3.16228	3.0
ℓ_3/ℓ_q	1.0	0.8
Z_3	10.0	6.0
Maximum reflection coefficient	0.70930	0.38813

Sample points (Normalized frequencies)

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

TABLE II

OPTIMIZATION OF A THREE-SECTION 10:1
TRANSFORMER OVER 100-PERCENT RELATIVE BANDWIDTH

Effort* required to reach or exceed a
reflection coefficient of 0.19729 (optimal to 5 figures)

Method		Problem 1			Problem 2		
	Parameter ξ or p	Function Evaluations	Sample Points	Response Evaluations	Function Evaluations	Sample Points	Response Evaluations
ξ-algorithm	0.1	28	11	308	19	11	209
(lower bound)	0.18846	16	7	112	16	7	112
p = 2	0.19730	45	4	180	53	4	212
$\xi^0 = 0$	0.19729	89		600	88		533
p-algorithm	8	39	11	429	29	11	319
3rd order	48	17	8	136	18	8	144
extrapolation	288	15	4	60	14	4	56
$\xi = 0$	1728	12	4	48	11	4	44
$\eta = 0.001$		83		673	72		563
Charalambous -	Alg.1	165	11	1815	105	11	1155
Bandler [2]	Alg.2	155	11	1705	95	11	1045

*Does not include response evaluations to determine sample
points to be used

TABLE III

PROGRESS OF THE p-ALGORITHM ON PROBLEM 1

r	Value of p_r	rth optimum	Extrapolated Solution	Max. reflection coefficient: rth optimum	Max. reflection coefficient: at extrapolated solution
1	8	.98828 1.62868 1.00004 3.16228 .98828 6.13993	same	.21017	.21017
2	48	.99833 1.63478 .99991 3.16228 .99833 6.11703	1.00035 1.63600 .99988 3.16228 1.00035 6.11246	.19838	.19863
3	288	.99973 1.63472 .99999 3.16228 .99973 6.11726	1.00000 1.63467 1.00000 3.16228 1.00000 6.11744	.19747	.19732
4	1728	.99995 1.63471 1.00000 3.16228 .99995 6.11730	1.00000 1.63471 1.00000 3.16228 1.00000 6.11730	.19732	.19729

TABLE IV

PROGRESS OF THE ξ-ALGORITHM ON PROBLEM 1

r	Value of ξ^r	rth optimum	Max. reflection coefficient
1	.1	.97238 1.59720 .98791 3.16228 .97238 6.26097	.25530
2	.18846	.99709 1.63451 1.00013 3.16228 .99709 6.11804	.19929
3	.19730	1.00000 1.63471 1.00000 3.16228 1.00000 6.11730	.19729

TABLE V

OPTIMIZATION OF THE SEVEN-SECTION FILTER BY THE p-ALGORITHM

Method	Parameter p	Function Evaluations	Sample Points	Response Evaluations
3rd order	2	73	11	803
extrapolation	12	16	11	176
$\xi = 0$	72	14	7	98
$\eta = 0.0001$	432	12	6	72
		115		1149

TABLE VI

LOWER BOUNDS FOR THE SEVEN-SECTION FILTER

Passband Specification 0.1 dB
Value of p_0 2
Value of ξ 0

Stopband Specification (dB)	First Maximum Error (dB)	Predicted Lower Bound (dB)	Next Maximum Error (dB)
50	-0.0256	-0.0283	-0.0282
55	0.1430	0.1154	0.1160
60	0.6211	0.4954	0.4986
65	1.5486	1.3148	1.3195

TABLE VII

FINAL SAMPLE POINTS AND SOLUTION FOR THE SEVEN-SECTION FILTER

Solution†: $Z_1 = Z_7 = 0.606595$

$Z_2 = Z_6 = 0.303547$

$Z_3 = Z_5 = 0.722287$

$Z_4 = 0.235183$

Sample Points (GHz)	Sample Points (Normalized w.r.t. 2.175 GHz)	Insertion Loss (dB)	Errors†† (dB)
0.6	-	50.028245	-0.028245
1.0875	0.5000	0.071755	-0.028245
	0.5395		
	0.6636		
	0.8741		
	1.1259		
	1.3364		
	1.4605		
3.2625	1.5000		
3.75	-	50.028245	-0.028245

† Symmetrical to at least the accuracy of the CDC 6400.
†† Equal to at least 5 figures.

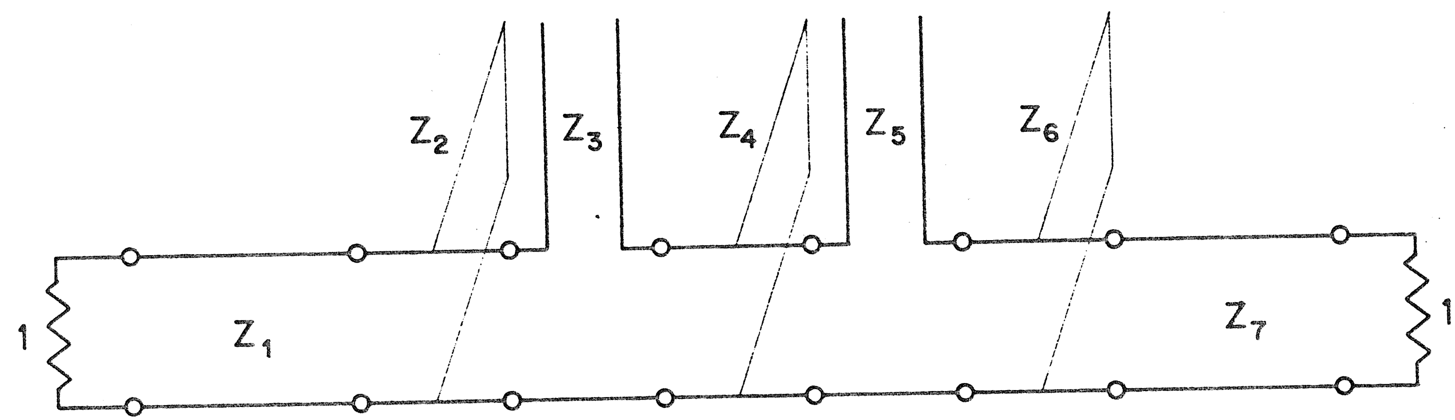

Fig. 1 The seven section filter.

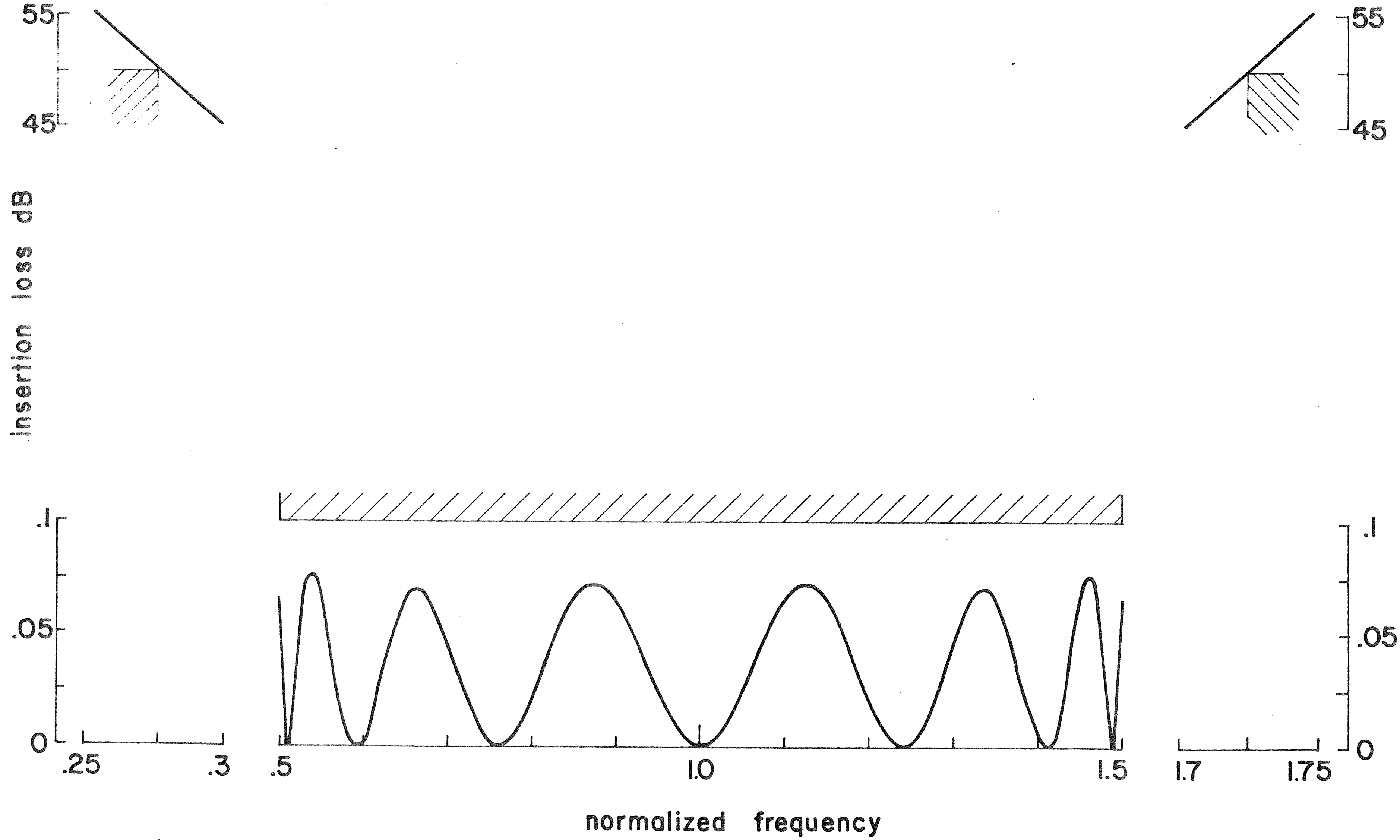

Fig. 2 Response of the filter optimized by the p-algorithm with 21 uniformly spaced passband points (only 10 were actually employed because of symmetry).

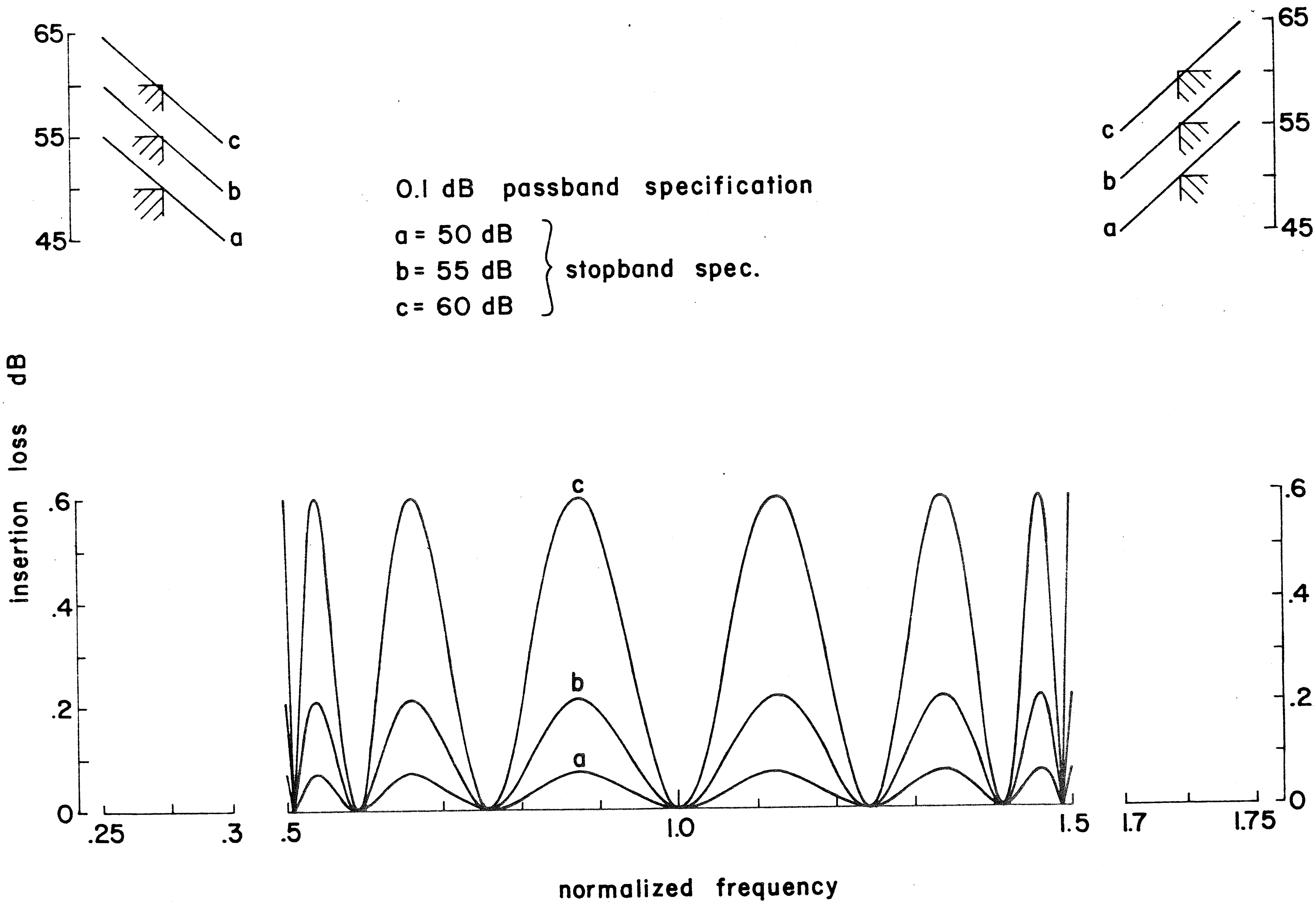

Fig. 3 Responses of the filter after two optimizations with the ξ-algorithm for p=2 using the ripple maxima from Fig. 2. Four sets of specifications were considered. The original starting point was taken in all cases.

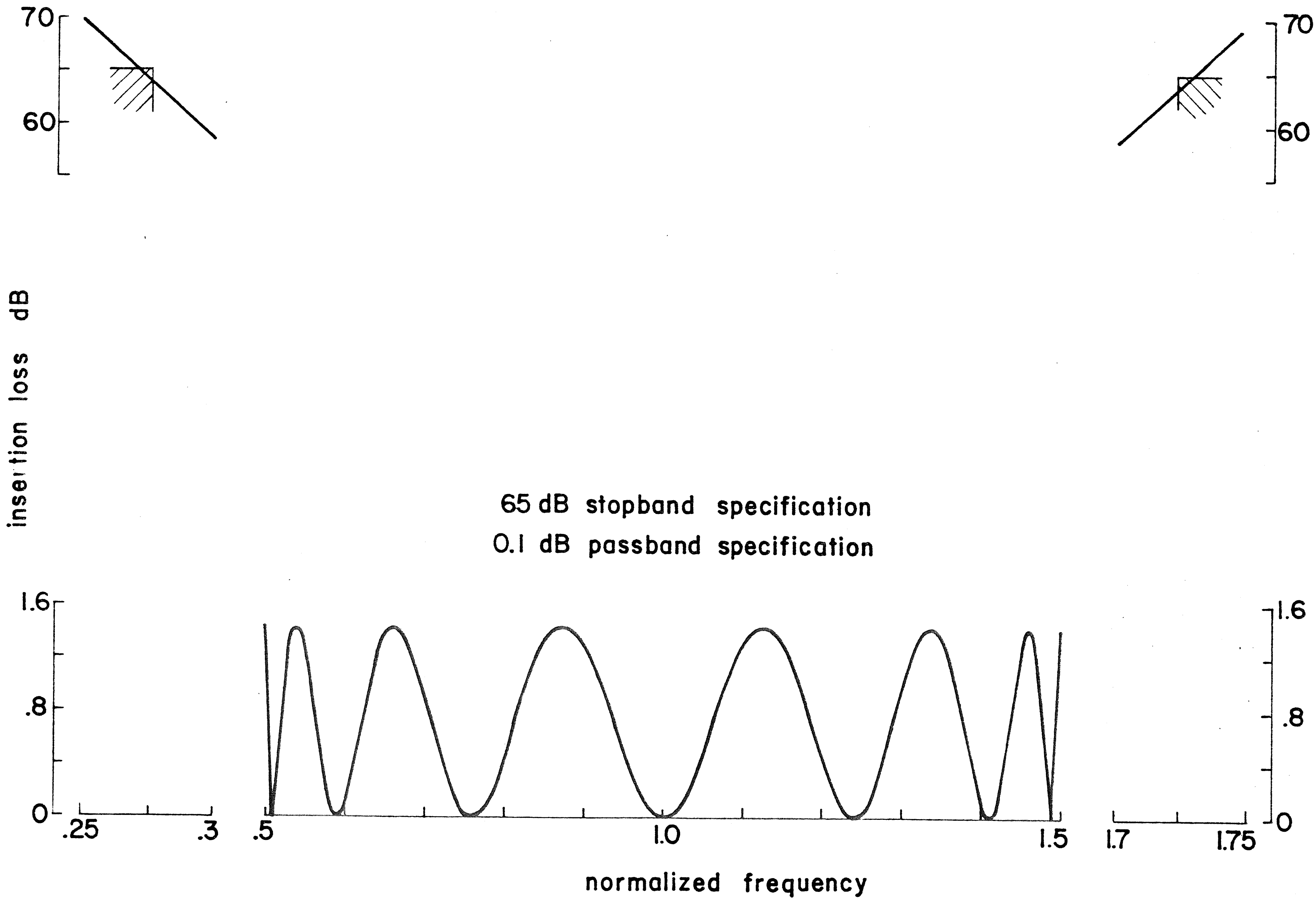

Fig. 3 [continued].