INTERNAL REPORTS IN

SIMULATION, OPTIMIZATION AND CONTROL

No. SOC-13

NEW ALGORITHMS FOR NETWORK OPTIMIZATION

C. Charalambous and J.W. Bandler

June 1973

FACULTY OF ENGINEERING
McMASTER UNIVERSITY
HAMILTON, ONTARIO, CANADA

NEW ALGORITHMS FOR NETWORK OPTIMIZATION

CHRISTAKIS CHARALAMBOUS and JOHN W. BANDLER

Abstract   Two new algorithms suitable for computer-aided optimization of networks are presented. They are both based on the nonlinear least pth approximation approach which has been successfully applied by the authors to microwave network design problems requiring minimax or near minimax solutions. A basic difference here is that, instead of requiring very large values of p, any finite value of p greater than one can be used to produce extremely accurate minimax solutions. The paper discusses a six variable transformer example where values of p equal to 2, 4, 6, 10, 100, 1,000, and 10,000 have all been used separately to obtain substantially the same solution. Both the adjoint network method for gradient evaluation and the Fletcher method are employed for greater efficiency. Comparisons with the razor search and grazor search methods are made. Some far-reaching observations concerning minimax design are also made.

This work was supported by the National Research Council of Canada under Grants A7239 and C154 and through a scholarship to C. Charalambous. This paper was presented at the 1973 IEEE International Microwave Symposium, Boulder, Colo., June 4-6, 1973.

C. Charalambous was with the Department of Electrical Engineering, McMaster University, Hamilton, Ont., Canada. He is now with the Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ont., Canada.

J.W. Bandler is with the Department of Electrical Engineering, McMaster University, Hamilton, Ontario, Canada.

I. INTRODUCTION

The authors have already presented a justification of the use of least pth approximation techniques with large values of p for computer-aided network design [1]. They showed that the use of a fairly well-conditioned objective function with efficient gradient minimization methods such as the method by Fletcher [2] and the adjoint network method for gradient evaluation [3], yields very near minimax designs with little computational effort.

The present paper exploits all the advantages of that approach in presenting two new algorithms for practical minimax approximation. A basic difference in these algorithms is that, instead of requiring very large values of p, any finite value of p greater than one can be used to produce minimax solutions.

The paper discusses a six variable example (namely, a three-section transmission-line transformer) where values of p equal to 2, 4, 6, 10, 100, 1,000, and 10,000 have all been used separately to obtain substantially the same solution. A comparison with other methods already known to microwave engineers is made. The Fletcher minimization method is used throughout. The advantage of the new algorithms is a combination of efficiency and flexibility which, it is believed, has not been previously enjoyed by computer-aided circuit designers.

II. THE NEW ALGORITHMS

Most of the notation to be used here is the same as in a previous publication by the authors [1].

We denote by $F(\underset{\sim}{\phi},\psi)$ the approximating function or network response, as indicated in Fig. 1. $\underset{\sim}{\phi}$ is a vector containing the k independent variables and ψ is an independent variable, for example, frequencey. $S_u(\psi)$ and $S_\ell(\psi)$ are upper and lower response specifications. ξ, as indicated in Fig. 1, is an artifical margin which plays an important role in the algorithms to be described. It is used to shift the error functions $e_u(\underset{\sim}{\phi},\psi) \triangleq w_u(\psi) \ (F(\underset{\sim}{\phi},\psi) - S_u(\psi))$ and $e_\ell(\underset{\sim}{\phi},\psi) \triangleq w_\ell(\psi) \ (F(\underset{\sim}{\phi},\psi) - S_\ell(\psi))$ to obtain $e_u'(\underset{\sim}{\phi},\psi,\xi) \triangleq e_u(\underset{\sim}{\phi},\psi) - \xi$ and $e_\ell'(\underset{\sim}{\phi},\psi,\xi) \triangleq e_\ell(\underset{\sim}{\phi},\psi) + \xi$, respectively, where $w_u(\psi)$ and $w_\ell(\psi)$ are appropriate weighting functions.

As usual, we will sample the various functions to obtain, for example, $e_{ui}(\underset{\sim}{\phi}) \triangleq e_u(\underset{\sim}{\phi},\psi_i)$ for $i\epsilon I_u$ and $e_{\ell i}(\underset{\sim}{\phi}) \triangleq e_\ell(\underset{\sim}{\phi},\psi_i)$ for $i\epsilon I_\ell$, where the sets I_u and I_ℓ correspond to appropriate sample points. We also assume that these error functions are differentiable and that an optimum exists.

<u>Algorithm 1</u>

(1) $\xi^1 \leftarrow \min[0, M_e(\underset{\sim}{\phi}^0) + \varepsilon]$,

where $\underset{\sim}{\phi}^0$ is the starting point, ε is a small positive number, and

$$M_e(\underset{\sim}{\phi}) \triangleq \max_{i,j} \ [e_{ui}(\underset{\sim}{\phi}), \ - e_{\ell j}(\underset{\sim}{\phi})] \ , \ i\epsilon I_u, j\epsilon I_\ell \tag{1}$$

(2) $r \leftarrow 1$

(3) Minimize with respect to $\underset{\sim}{\phi}$, the function

$$U(\underset{\sim}{\phi},\xi^r) \leftarrow M(\underset{\sim}{\phi},\xi^r)\left(\sum_{i\varepsilon K_u} \left(\frac{e_{ui}(\underset{\sim}{\phi})-\xi^r}{M(\underset{\sim}{\phi},\xi^r)}\right)^q + \sum_{i\varepsilon K_\ell} \left(\frac{-e_{\ell i}(\underset{\sim}{\phi})-\xi^r}{M(\underset{\sim}{\phi},\xi^r)}\right)^q \right)^{\frac{1}{q}}$$

$$\text{for } M(\underset{\sim}{\phi},\xi^r) \neq 0 \qquad (2)$$

$$U(\underset{\sim}{\phi},\xi^r) \leftarrow 0 \text{ for } M(\underset{\sim}{\phi},\xi^r) = 0$$

where

$$M(\underset{\sim}{\phi},\xi^r) \leftarrow M_e(\underset{\sim}{\phi}) - \xi^r \qquad (3)$$

$$q \leftarrow p \text{ sign } M(\underset{\sim}{\phi},\xi^r) \qquad (4)$$

and

$$\text{if } M(\underset{\sim}{\phi},\xi^r) \begin{cases} >0 \text{ then } 1 < p < \infty,\ K_u \leftarrow \{ i \mid e_{ui}(\underset{\sim}{\phi}) \geq \xi^r,\ i\varepsilon I_u \} \\ \qquad\qquad \text{and } K_\ell \leftarrow \{ i \mid -e_{\ell i}(\underset{\sim}{\phi}) \geq \xi^r,\ i\varepsilon I_\ell \} \\ <0 \text{ then } 1 \leq p < \infty,\ K_u \leftarrow I_u,\ K_\ell \leftarrow I_\ell \end{cases} \qquad (5)$$

(4) $\xi^{r+1} \leftarrow M_e(\underset{\sim}{\check{\phi}}^r) + \varepsilon$

where $\underset{\sim}{\check{\phi}}^r$ is the solution vector for step (3).

(5) If $|\xi^{r+1} - \xi^r| < \eta$ stop. Otherwise $r \leftarrow r+1$. The quantity η is a suitable small positive number.

(6) Go to step (3).

Comments on Algorithm 1

If $M_e(\underset{\sim}{\phi}^0)$ is positive, $\xi^1 = 0$ and $q = p$, which is positive. If $M_e(\underset{\sim}{\phi}^0)$ is negative and for subsequent optimizations with $r = 2,3,\ldots$ we have $q = -p$. We note that the reason for inclusion of ε in steps (1) and (4) is to avoid having $M = 0$. In this case, when two or more maxima in the error functions are equal, the objective function's first derivatives are discontinuous.

It can be shown [4] that the objective function decreases as ξ increases keeping the parameter vector $\underset{\sim}{\phi}$ constant. As the sequence of optimizations proceeds for $r \geq 2$, the modulus of the objective function at each optimum in the sequence, namely $|U(\underset{\sim}{\check{\phi}}^r,\xi^r)|$, decreases as r increases. Furthermore, as $r \to \infty$, $|U(\underset{\sim}{\check{\phi}}^r,\xi^r)| \to 0$ and $M_e(\underset{\sim}{\check{\phi}}^r) \to M_e(\underset{\sim}{\check{\phi}})$, where $\underset{\sim}{\check{\phi}}$ is the minimax solution vector desired [4].

Algorithm 2

(1)-(3) As in Algorithm 1

(4) If $M(\underset{\sim}{\check{\phi}}^r,\xi^r) < 0$ we remain with Algorithm 1. Otherwise

$$\xi^{r+1} \leftarrow \xi^r + \lambda^r M(\underset{\sim}{\check{\phi}}^r,\xi^r) = (1 - \lambda^r)\, \xi^r + \lambda^r M_e(\underset{\sim}{\check{\phi}}^r) \qquad (6)$$

where

$$0 < \lambda^r < 1 \qquad (7)$$

(5)-(6) As in Algorithm 1

Comments on Algorithm 2

Algorithm 2 differs from Algorithm 1 only when $M_e(\check{\underset{\sim}{\phi}})$ is positive. Under this condition q = p, which is positive. Furthermore, assuming we remain with Algorithm 2, the objective function at each optimum, namely, $U(\check{\underset{\sim}{\phi}}^r, \xi^r)$, decreases as r increases. As $r \to \infty$, $U(\check{\underset{\sim}{\phi}}^r, \xi^r) \to 0$ and $M_e(\check{\underset{\sim}{\phi}}^r) \to M_e(\check{\underset{\sim}{\phi}})$.

Algorithm 2 differs from Algorithm 1 in that Algorithm 2 attempts to predict $M_e(\check{\underset{\sim}{\phi}})$ by increasing the level ξ^r, whereas Algorithm 1 attempts to push the maximum away from the level ξ^r. To use Algorithm 2 when $M_e(\check{\underset{\sim}{\phi}})$ is negative we simply add a suitably large positive constant value to all the error functions.

III. EXAMPLES

The algorithms have been applied to a wide range of design problems. Here, we will compare their performance for selected values of p using the Fletcher method [2] on a 3-section, 100% relative bandwidth, 10:1 transmission-line transformer problem, which has already received attention in the literature [5], [6]. See Table I for a statement of the problem and the starting points used in the optimization. Specifically, we let e_{ui} be the modulus of the reflection coefficient sampled at the 11 normalized frequencies

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

Gradient vectors with respect to section lengths and characteristic impedances are obtained using the adjoint network method [3].

The number of function evaluations (one function evaluation comprising an evaluation of the objective function (2) and its derivatives) needed by Algorithms 1 and 2 to reach a reflection coefficient magnitude of 0.19729, which is optimal to five figures, is shown in Tables II-IV. All six parameters were varied. Seven values of p were tried ranging from 2 to 10,000. It is important to remember that the value of p is fixed throughout each optimization so that the tables represent the effort required to reach substantially the same minimax solution for the particular value indicated.

Table II compares the present approach with previous results on the same problem. The paper by Bandler etal. [5] briefly discusses the method by Osborne and Watson [7]. The value of ε was 10^{-8}.

Figs. 2 and 3 show how the maximum error M_e defined in (1) is reduced at the end of each step in the sequence of optimizations.

As shown in Tables III and IV, different values of λ, the parameter given in (7), were used in testing Algorithm 2. In Table IV the quantity M_e^1 is the maximum error at the end of the first optimization for the particular value of p indicated and m_e^1 is the corresponding value of the smallest ripple. On average, the choice of λ shown in Table IV produced the best results. Note that M_e^1 decreases as p increases, as would be expected.

Figs. 4 and 5 illustrate ideas similar to those in Figs. 2 and 3. In addition, the rise of ξ (relevant only in Algorithm 2) is also depicted.

IV. DISCUSSION

An important result of the investigation is that if the maximum error corresponding to the solution of the first optimization with any value of p greater than one is positive (specification violated) then it remains positive for all permissible values of p. Similarly, if it is negative (specification satisfied) it remains negative for all permissible values of p. The same is true of the objective function in (2). Thus, if we are investigating whether a particular structure will satisfy design specifications in the minimax sense, any single suitable least pth optimization will reveal this! We do not need to carry out the full sequence of optimizations to obtain this information.

In practice, as the results indicate, only two or three optimizations with low values of p will result in a good minimax design. For p = 10, two optimizations usually suffice for practical purposes.

V. CONCLUSIONS

Two new algorithms for computer-aided minimax optimizations have been presented and discussed. Unlike previous work by the authors [1], where a single optimization is carried out with large p in an effort to reach near minimax results, these algorithms can be described as true minimax algorithms. We believe that by adopting them the designer will have an extremely flexible and efficient way of achieving minimax designs. The mathematical background with proofs of convergence, and so on, is rather involved and lengthly. It is available elsewhere [4].

REFERENCES

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

[2] R. Fletcher, "A new approach to variable metric algorithms", Computer J., vol. 13, pp. 317-322, August 1970.

[3] J.W. Bandler and R.E. Seviora, "Current trends in network optimization", IEEE Trans. Microwave Theory and Techniques, vol. MTT-18, pp. 1159-1170, December 1970.

[4] C. Charalambous and J.W. Bandler, "Nonlinear minimax optimization as a sequence of least pth optimization with finite values of p", J. Optimization Theory and Applications, submitted.

[5] J.W. Bandler, T.V. Srinivasan and C. Charalambous, "Minimax optimization of networks by grazor search", IEEE Trans. Microwave Theory and Techniques, vol. MTT-20, pp. 596-604, September 1972.

[6] J.W. Bandler and P.A. Macdonald, "Optimization of microwave networks by razor search", IEEE Trans. Microwave Theory and Techniques, vol. MTT-17, pp. 552-562, August 1969.

[7] M.R. Osborne and G.A. Watson, "An algorithm for minimax approximation in the non-linear case", Computer J., vol. 12, pp. 63-68, February 1969.

TABLE I

THE STARTING POINTS IN THE OPTIMIZATION OF A 3-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

TABLE II

OPTIMIZATION OF A 3-SECTION 10:1 TRANSFORMER

OVER 100 PERCENT RELATIVE BANDWIDTH USING

ALGORITHM 1

Value of p	Number of function evaluations to reach a reflection coefficient of 0.19729	
	Problem 1	Problem 2
2	178	160
4	143	128
6	142	116
10	112	89
100	136	69
1000	193	66
10000	249	104
Average number of function evaluations	165	105
Grazor search [5]	696	498
Osborne* & Watson [5]	860(0.20831)	237(0.19788)
Razor* search [6]	1300(0.19733)	1250(0.19731)

* Number of function evaluations to reach the value shown in brackets.

TABLE III

OPTIMIZATION OF A 3-SECTION 10:1 TRANSFORMER

OVER 100 PERCENT RELATIVE BANDWIDTH USING

ALGORITHM 2

Value of p	Number of function evaluations to reach a reflection coefficient of 0.19729					
	Problem 1			Problem 2		
	$\lambda=0.5$	$\lambda=0.6$	$\lambda=0.7$	$\lambda=0.5$	$\lambda=0.6$	$\lambda=0.7$
2	183	146	151	165	128	133
4	193	162	122	151	149	109
6	199	182	138	150	135	112
10	191	159	146	168	136	114
100	185	171	165	126	104	99
1000	211	211	202	83	91	83
10000	248	248	248	103	103	103
Average number of function evaluations	202	182	168	135	121	108

TABLE IV

OPTIMIZATION OF A 3-SECTION 10:1 TRANSFORMER

OVER 100 PERCENT RELATIVE BANDWIDTH USING

ALGORITHM 2 WITH $\lambda = \frac{M_e^1 + m_e^1}{2M_e^1}$

Value of p	M_e^1	m_e^1	Number of function evaluations to reach a reflection coefficient of 0.19729	
			Problem 1	Problem 2
2	0.30819	0.11626	161	141
4	0.23680	0.16278	135	123
6	0.21759	0.17612	115	89
10	0.20636	0.18280	95	72
100	0.19781	0.19562	131	65
1000	0.19734	0.19712	193	66
10000	0.19730	0.19727	249	104
Average number of function evaluations			155	95

Figure Captions

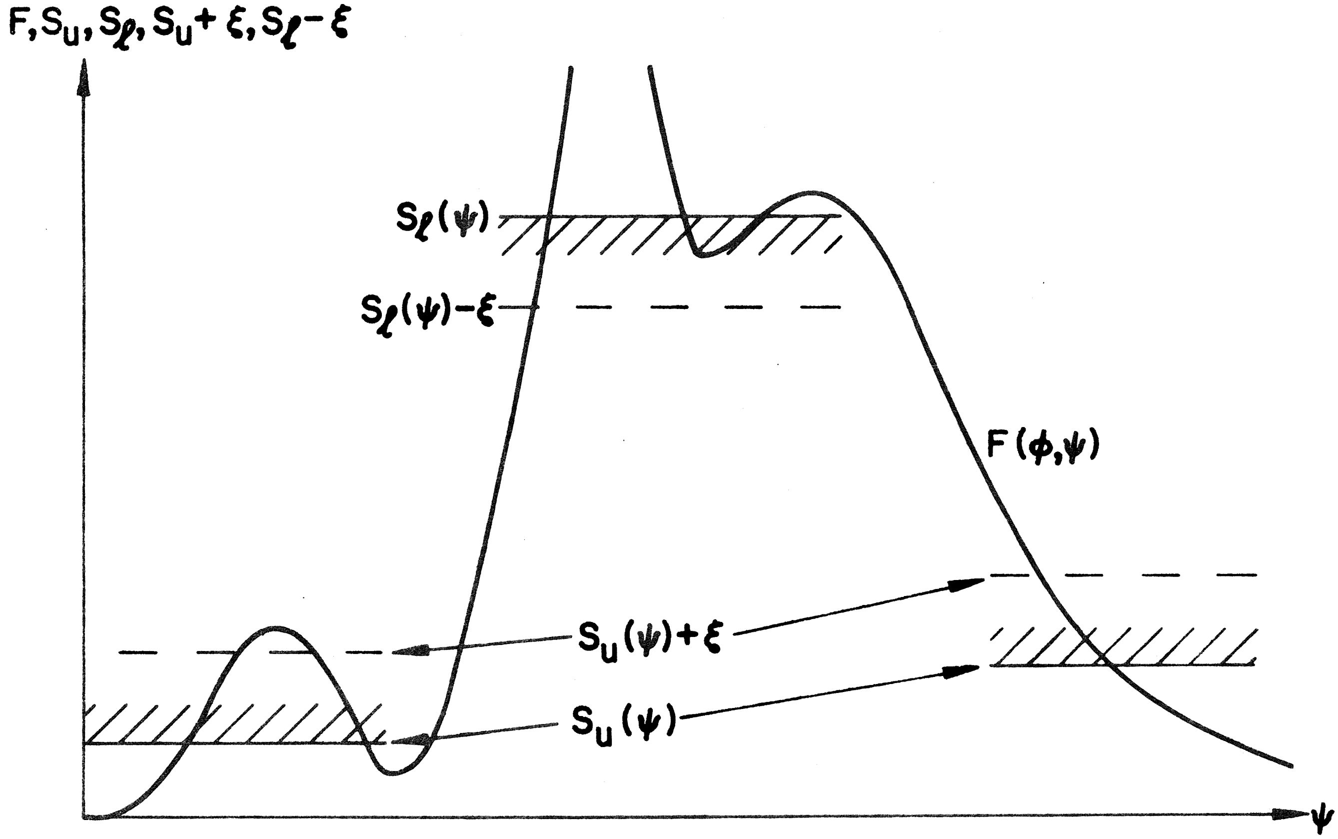
Fig. 1. An approximation problem where the response specifications are violated and the weighting factors are unity.

Fig. 2. Maximum error against number of function evaluations using Algorithm 1 on Problem 1.

Fig. 3. Maximum error against number of function evaluations using Algorithm 1 on Problem 2.

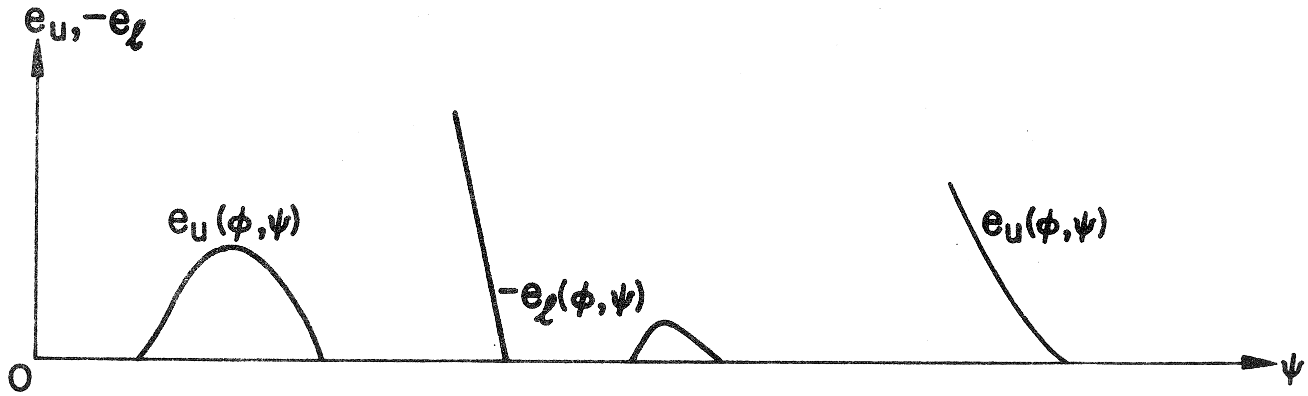
Fig. 4. Maximum error and ξ against number of function evaluations using Algorithm 2 on Problem 1 with $\lambda=0.5$.

Fig. 5. Maximum error and ξ against number of function evaluations using Algorithm 2 on Problem 2 with $\lambda=0.5$.

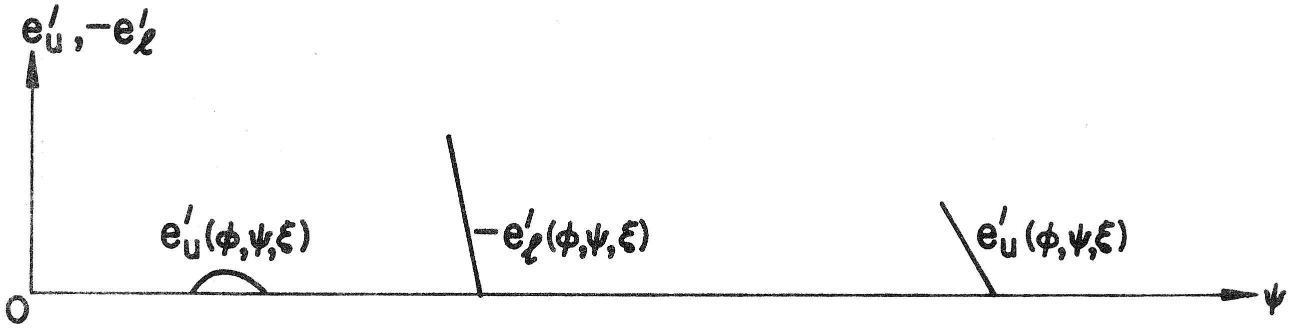


$F, S_u, S_\ell, S_u+\xi, S_\ell-\xi$
$S_\ell(\psi)$
$S_\ell(\psi)-\xi$
$F(\phi,\psi)$
$S_u(\psi)+\xi$
$S_u(\psi)$
ψ

Fig. 1



$e_u, -e_\ell$
$e_u(\phi,\psi)$
$-e_\ell(\phi,\psi)$
$e_u(\phi,\psi)$
O
ψ



$e'_u, -e'_\ell$
$e'_u(\phi,\psi,\xi)$
$-e'_\ell(\phi,\psi,\xi)$
$e'_u(\phi,\psi,\xi)$
O
ψ

Fig. 1 [cont.]

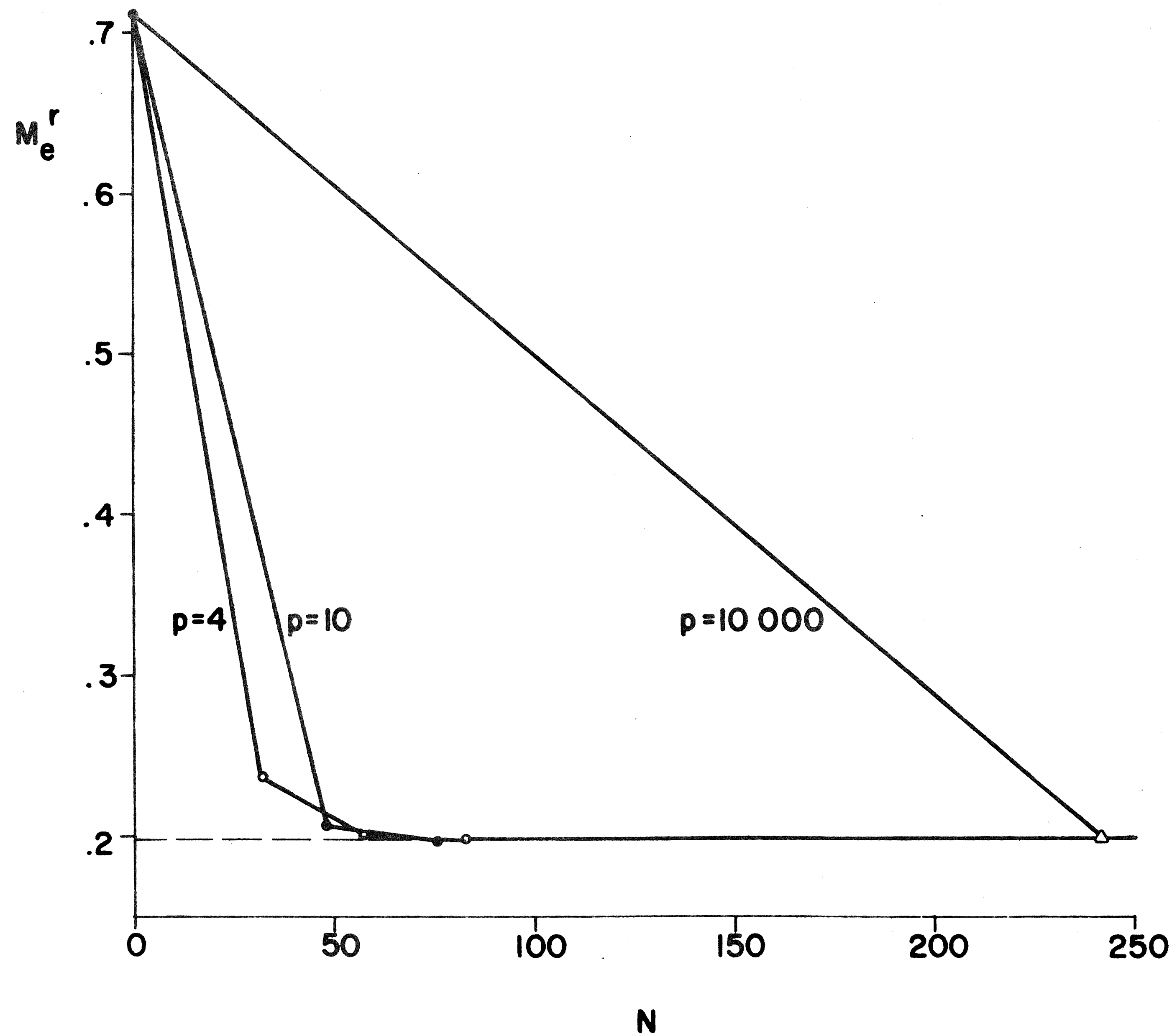


Fig. 2

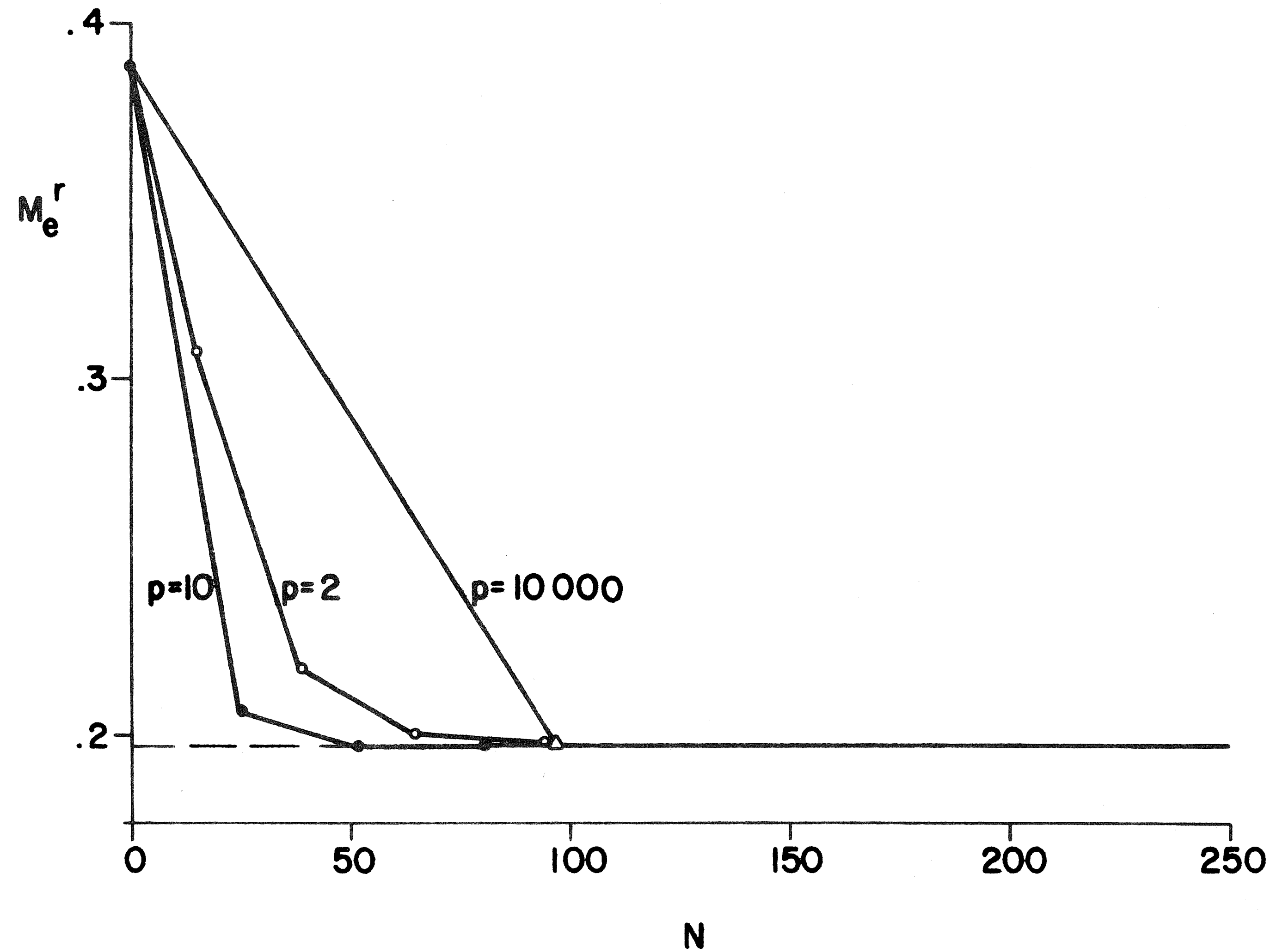
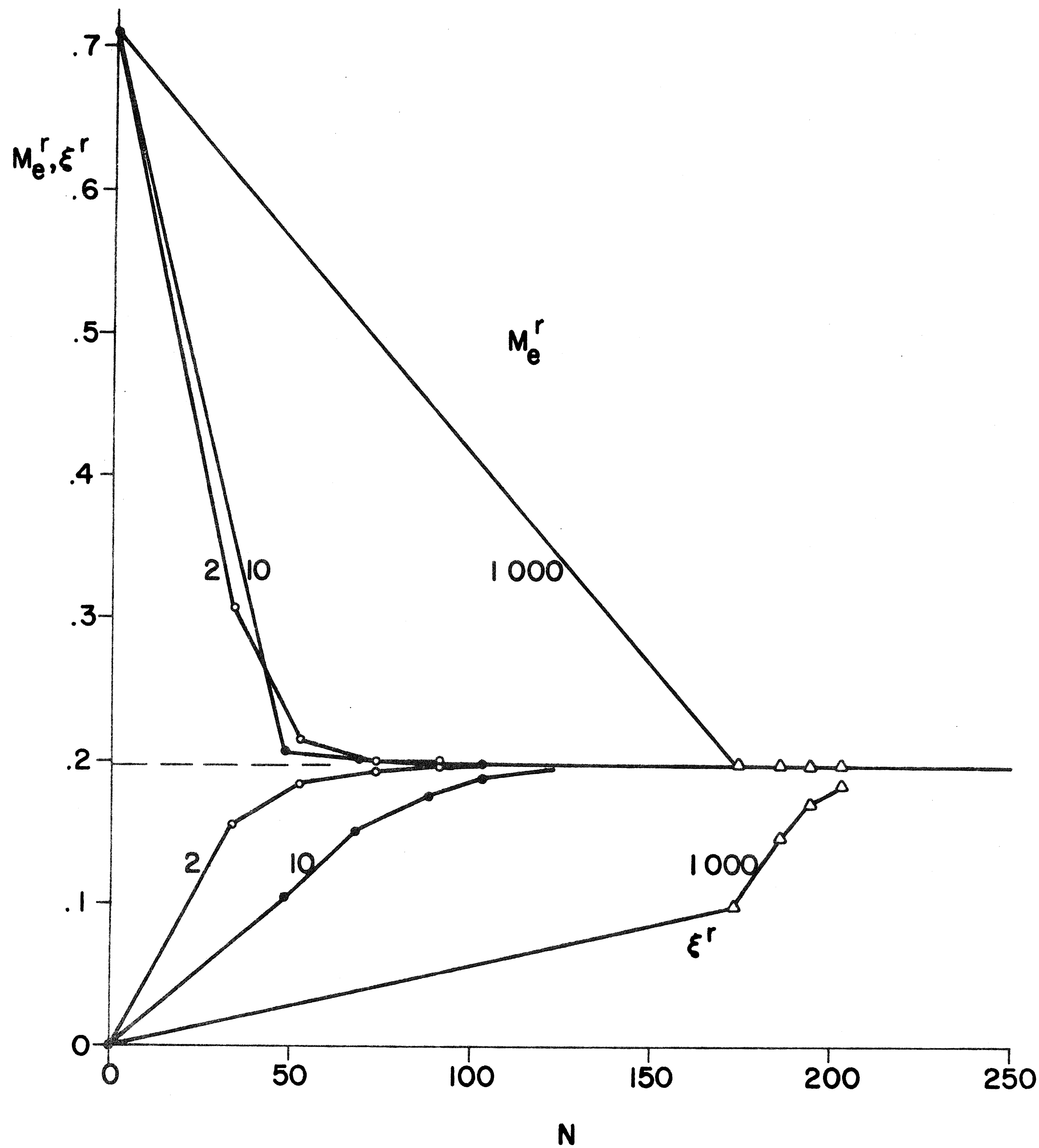


Fig. 3



M_e^r, ξ^r
.7
.6
.5
.4
.3
.2
.1
0
M_e^r
2
10
1 000
2
10
1000
ξ^r
0
50
100
150
200
250
N

Fig. 4

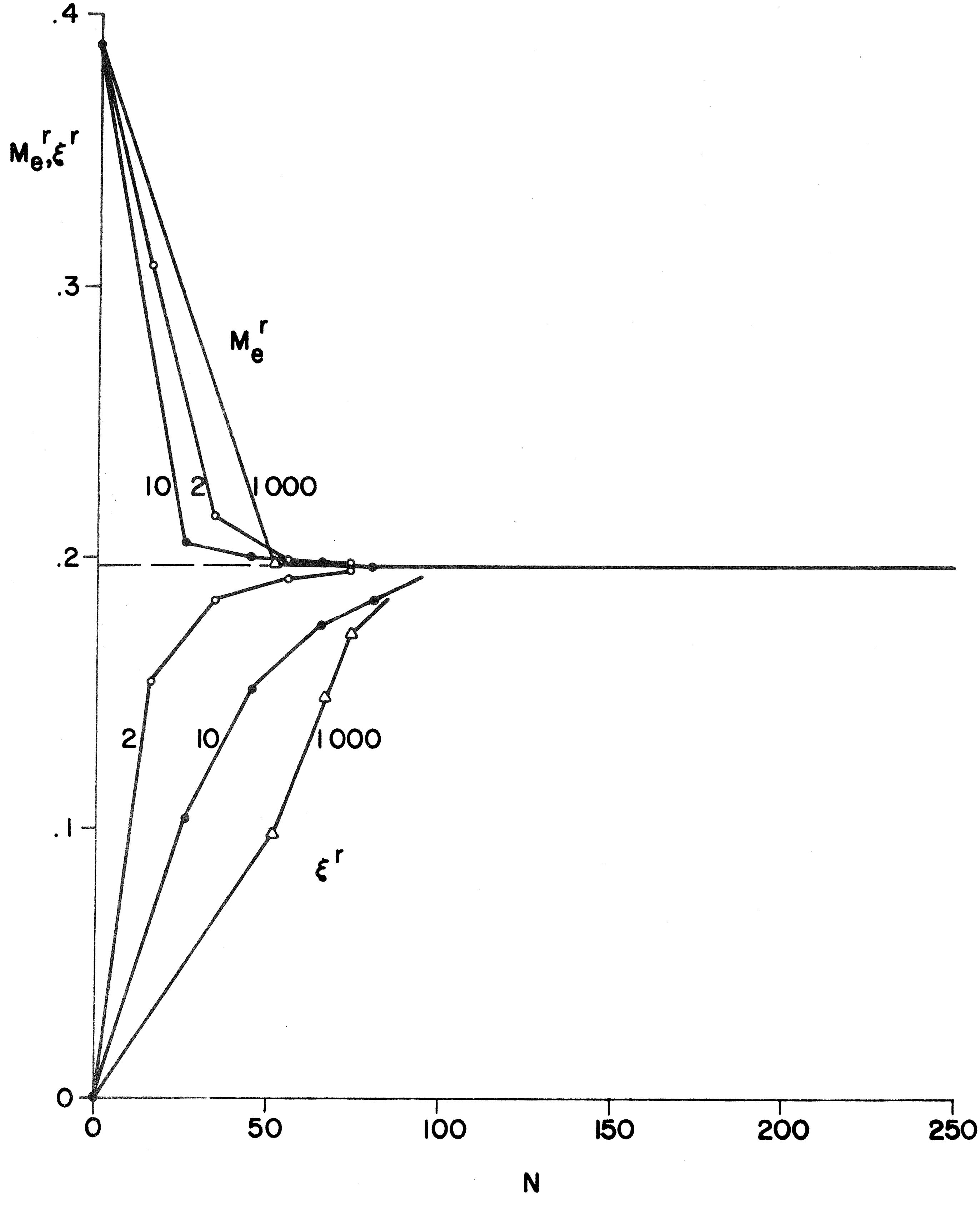


Fig. 5