AN EFFICIENT GRADIENT BASED MINIMAX ALGORITHM REQUIRING ONLY CIRCUIT RESPONSE EVALUATIONS

J.W. Bandler, S.H. Chen, S. Daijavad and K. Madsen

SOS-85-17-R

December 1985

AN EFFICIENT GRADIENT BASED MINIMAX ALGORITHM REQUIRING ONLY CIRCUIT RESPONSE EVALUATIONS

J.W. Bandler*, S.H. Chen*, S. Daijavad* and K. Madsen**

* Simulation Optimization Systems Research Laboratory
and Department of Electrical and Computer Engineering
McMaster University, Hamilton, Canada L8S 4L7
(416) 525-9140 Ext. 4305

** Institute for Numerical Analysis
Technical University of Denmark
DK-2800 Lyngby, Denmark

Abstract

A flexible and powerful minimax algorithm not requiring user-provided gradients is presented. It extends the practical applicability of a recent algorithm of Bandler, Kellermann and Madsen by integrating it with efficient gradient approximations. The applications are illustrated by practical design of a 3-channel multiplexer involving 45 nonlinear variables and worst-case tolerance design of a microwave amplifier.

SUMMARY

Many microwave circuit design, centering and tuning problems can be formulated as minimax optimization problems. Recently, a highly efficient nonlinear minimax algorithm was presented by Bandler, Kellermann and Madsen [1]. The main obstacle to the extensive applications of such an algorithm is the requirement of exact gradients. Deriving explicit sensitivity expressions is usually very difficult or impossible. Using numerical differentiations at every iteration becomes prohibitively expensive when attempting to solve large scale problems.

In this paper we present a new algorithm for linearly constrained minimax optimization integrated with effective and efficient gradient approximations. This algorithm embodies the recent results of Bandler et al. [1], namely a 2-stage combined LP and quasi-Newton method for nonlinear minimax optimization. It also utilizes a hybrid approach to gradient approximation which consists of numerical differentiation, the Broyden rank-one update [2] and the special iterations of Powell [3].

Compared to an earlier minimax algorithm not requiring derivatives by Madsen et al. [4], our new algorithm is more powerful and flexible for the following reasons. The quasi-Newton method (Stage 2) of the new algorithm effectively solves singular problems where the earlier algorithm suffers from a slow convergence rate. Furthermore, for large scale problems with a large number of variables, approximation using only a rank-one update, as implemented in the earlier algorithm, may fail to converge. In our case, the gradient approximation can be corrected by numerical differentiation with some prescribed regularity. Hence, both the computational efficiency and convergence can be achieved. This is proved through practical design of a 3-channel multiplexer involving 45 nonlinear variables. We also describe the worst-case tolerance design of a microwave amplifier as a further illustration of other possible practical applications. It is worth mentioning that many existing circuit simulation programs can now be easily equipped with a powerful minimax optimizer.

Another feature of our method for gradient approximation is that any possible sparsity existing in the Jacobians can be explored to improve the accuracy of approximation. Also, linear constraints on the optimization variables are easily handled. These features can be utilized to great advantage in computer-assisted tuning of microwave circuits.

Review of the 2-Stage Algorithm

We consider the minimax optimization problem

$$\underset{\mathbf{x}}{\text{minimize}}\ F(\mathbf{x}) \triangleq \max_{j} \{f_j(\mathbf{x})\}$$

subject to a set of linear constraints

$$\begin{aligned} \mathbf{a}_i^T \mathbf{x} + b_i &= 0\,, \qquad i = 1, \ldots, \ell_{eq}\,, \\ \mathbf{a}_i^T \mathbf{x} + b_i &\geq 0\,, \qquad i = \ell_{eq}+1, \ldots, \ell\,, \end{aligned} \tag{1}$$

where

$$f_j(\mathbf{x}) = f_j(x_1, \ldots, x_n)\,, \qquad j = 1, \ldots, m\,,$$

are a set of nonlinear, continuously differentiable functions, $\mathbf{x} \triangleq [x_1 \ldots x_n]^T$ is the variable vector, $\mathbf{a}_i$ and b_i are constants. The algorithm presented by Bandler et al. [1] consists of two stages. The Stage 1 is a first-order method which solves a locally linearized subproblem using linear programming technique. In Stage 2, a quasi-Newton method is employed to solve a set of nonlinear equations arising from the optimality conditions that should hold at the solution. Details can be found in [1]. Our algorithm embodies essentially the same optimizer except that the rules for revising local bounds (see [1]) have been modified because approximate gradients are used.

Approximation of the Derivatives

A hybrid method is used for gradient approximations. Numerical differentiations

$$g_{ji} = \frac{\partial f_j(\mathbf{x})}{\partial x_i} \approx \frac{f_j(\mathbf{x} + t\,\mathbf{e}_i) - f_j(\mathbf{x})}{t} \tag{2}$$

are used to obtain an initial approximation of the Jacobian

$$G^0 \approx \left[\frac{\partial \mathbf{f}^T(\mathbf{x})}{\partial \mathbf{x}} \right]^T \tag{3}$$

and to provide corrections to the subsequent approximations. At each step of optimization, a modified Broyden rank-one formula is used to update the approximate derivatives

$$\mathbf{g}_j^{k+1} = \mathbf{g}_j^k + \frac{f_j(\mathbf{x}^{k+1}) - f_j(\mathbf{x}^k) - (\mathbf{g}_j^k)^T \mathbf{h}^k}{(\mathbf{q}_j^k)^T \mathbf{h}^k} \, \mathbf{q}_j^k \ , \tag{4}$$

where $\mathbf{g}_j{}^k$ is the approximation to $\partial f_j(\mathbf{x})/\partial \mathbf{x}$ at $\mathbf{x}^k$, $\mathbf{h}^k = \mathbf{x}^{k+1} - \mathbf{x}^k$ is the incremental change and

$$\mathbf{q}_j^k \triangleq [w_{j1} h_1^k \ldots w_{jn} h_n^k]^T \ . \tag{5}$$

The weights w_{ji} are introduced to represent any possible linearity of some functions in some variables, or, in other words, any possible constants existing in the Jacobian. To prevent a degenerate approximation, the incremental direction $\mathbf{h}^k$ should satisfy some linear independence condition. The special iterations of Powell [3] are introduced for this purpose. The approach is very similar to the one adopted by Madsen [4] except for the ideas of regular correction by numerical differentiations and weights in the Broyden update. For Stage 2, however, an effective solution of a singular problem by quasi-Newton method requires accurate first-order derivatives and approximate second-order information. We found that in this case it is usually more desirable to calculate the gradient by numerical differentiations and approximate the second-order derivatives using a BFGS update. With a fast final convergence rate, Stage 2 usually terminates in a few steps. Therefore, the computational effort involved in the numerical differentiations is reasonable and well justified.

We have compared the performance of our algorithm with the earlier algorithm of Madsen et al. [4] for the classical problems of two-section and three-section transmission line transformers [5]. For the two-section transformer, the characteristic impedances Z_1 and Z_2 are optimized. Starting from 4 different points, namely (1,3), (3.5,6), (1,6) and (3.5,3), our algorithm requires respectively 21, 21, 23 and 28 function evaluations to reach the solution. The corresponding results reported in [4] are 28, 25, 23 and 29 function evaluations. Three tests are made for the three section transformer: fixed lengths and optimizing the characteristic impedances; optimizing the lengths and impedances from two different starting points. In terms of the number of function evaluations required to reach the solution, our results are 107, 129 and 272, and the respective results in [4] are 251, 540 and 824. The quasi-Newton iteration employed in the new algorithm has significantly accelerated the convergence rate when solving singular problems.

Worst-Case Design of an Amplifier

Worst-case network design is considered here as a fixed tolerance problem [6], [7]. It consists of design data of a nominal point $\mathbf{x}^0$ and a set of associated tolerances $\boldsymbol{\varepsilon} \triangleq [\varepsilon_1 \ldots \varepsilon_n]^T$. The tolerance region R_ε is defined as $R_\varepsilon \triangleq \{\mathbf{x} \mid x_i^0 - \varepsilon_i \leq x_i \leq x_i^0 + \varepsilon_i, i=1, \ldots, n\}$. The aim is to center the nominal point $\mathbf{x}^0$ such that the specifications are best satisfied for any outcome in R_ε. We assume that the worst-case occurs at some vertices of R_ε.

A microwave amplifier consisting of an NEC 70000 FET and five transmission lines is considered (Fig. 1). The NEC 70000 FET is characterized by tabulated scattering parameters. The design specification is given as 7.25 dB $\leq$ 20 log $|S_{21}| \leq$ 8 dB between 6 GHz and 18 GHz. A nominal design is first obtained by taking ℓ_1, ℓ_2, ℓ_3, ℓ_4, ℓ_5 and Z as variables, where ℓ_1 to ℓ_5 are the lengths of the transmission lines. Thirteen sample points uniformly spaced between 6 and 18 GHz are taken, resulting in 26 error functions corresponding to the upper and lower specifications. Using the algorithm described in this paper, a nominal solution is obtained as $[\ell_1^0 \ \ell_2^0 \ \ell_3^0 \ \ell_4^0 \ \ell_5^0 \ Z^0] = [52.962 \ 148.13 \ 26.798 \ 24.010 \ 46.627 \ 81.266]$. This requires 65 function evaluations. The circuit response at the solution is shown in Fig. 2.

For the worst-case design, a 5 percent tolerance is associated with each length ℓ_i and Z is assumed tolerance-free, i.e., $\boldsymbol{\varepsilon} = 0.05\,[\ell_1 \ \ell_2 \ \ell_3 \ \ell_4 \ \ell_5 \ 0]^T$. The worst-case outcomes give extreme values of 20 log $|S_{21}|$ of 8.3 dB and 6.8 dB. At the starting point, 26 worst-case vertices (corresponding to the upper and lower specifications at 13 frequency sample points) are selected. These vertices are predicted using the gradient at $\mathbf{x}^0$ which is obtained by numerical differentiations. Using these vertices we obtained a solution at which 10 new worst-case vertices are detected. These new worst-case vertices are added to give a total of 36 error functions. In general such a procedure is repeated until the set of selected vertices is complete. For the case considered here, the second solution is found to be final. The optimally centered design is given by [69.01 152.01 18.48 5.095 36.49 126.39]. It reduces the extremes of 20 log $|S_{21}|$ to 8.1 dB and 7.1 dB.

The total number of function evaluations required is 280. We have also solved the same problem with derivatives being calculated entirely by numerical differentiations, which required 585 function evaluations.

<u>Practical Design of a 3-channel Multiplexer</u>

Design of multiplexers consisting of multi-cavity filters distributed along a waveguide manifold is a large scale problem of significant interest [8], [9]. A general multiplexer optimization procedure using exact network sensitivities has been reported by Bandler et al. [1], [9]. The minimax error functions are created using specifications on common port return loss and individual channel insertion losses, simulated multiplexer responses and weighting factors. Since our new algorithm does not require sensitivities, the size and complexity of the simulation program are greatly reduced.

Here, we consider a 12 GHz, 3-channel multiplexer [9]. A total of 45 network parameters are optimized, including spacings, transformer ratios, cavity resonances and coupling coefficients. The channel filters are assumed lossy and dispersive, and waveguide junctions are assumed nonideal. The network responses at the starting point and at the solution are shown in Figs. 3 and 4, respectively.

To reach the solution, 476 response evaluations are performed. The solution reported in [9] was obtained in more than 70 iterations, which would need more than 3000 response evaluations to provide the required exact gradient by numerical differentiations. The new algorithm is significantly more efficient.

References

[1] J.W. Bandler, W. Kellermann and K. Madsen, "A superlinearly convergent minimax algorithm for microwave circuit design", IEEE Trans. Microwave Theory Tech., vol. MTT-33, 1985.

[2] C.G. Broyden, "A class of methods for solving non-linear simultaneous equations", Mathematics of Computation, vol. 19, 1965, pp. 577-593.

[3] M.J.D. Powell, "A Fortran subroutine for unconstrained minimization, requiring first derivatives of the objective functions", AERE, Harwell, Oxon., England, Report R.6469, 1970, pp. 20-27.

[4] K. Madsen, O. Nielsen, H. Schjaer-Jacobsen and L. Thrane, "Efficient minimax design of networks without using derivatives", IEEE Trans. Microwave Theory and Tech., vol. MTT-23, 1975, pp. 803-809.

[5] J.W. Bandler and P.A. Macdonald, "Optimization of microwave networks by razor search", IEEE Trans. Microwave Theory Tech., vol. MTT-17, 1969, pp. 522-562.

[6] J.W. Bandler, P.C. Liu and H. Tromp, "A nonlinear programming approach to optimal design centering, tolerancing and tuning", IEEE Trans. Circuits and Systems, vol. CAS-23, 1976, pp. 155-165.

[7] J.W. Bandler, P.C. Liu and J.H.K. Chen, "Worst case network tolerance optimization", IEEE Trans. Microwave Theory Tech., vol. MTT-23, 1975, pp. 630-641.

[8] A.E. Atia, "Computer-aided design of waveguide multiplexers", IEEE Trans. Microwave Theory Tech., vol. MTT-22, 1974, pp. 332-336.

[9] J.W. Bandler, S.H. Chen, S. Daijavad and W. Kellermann, "Optimal design of multi-cavity filters and contiguous-band multiplexers", Proc. 14th European Microwave Conference (Liege, Belgium, 1984), pp. 863-868.

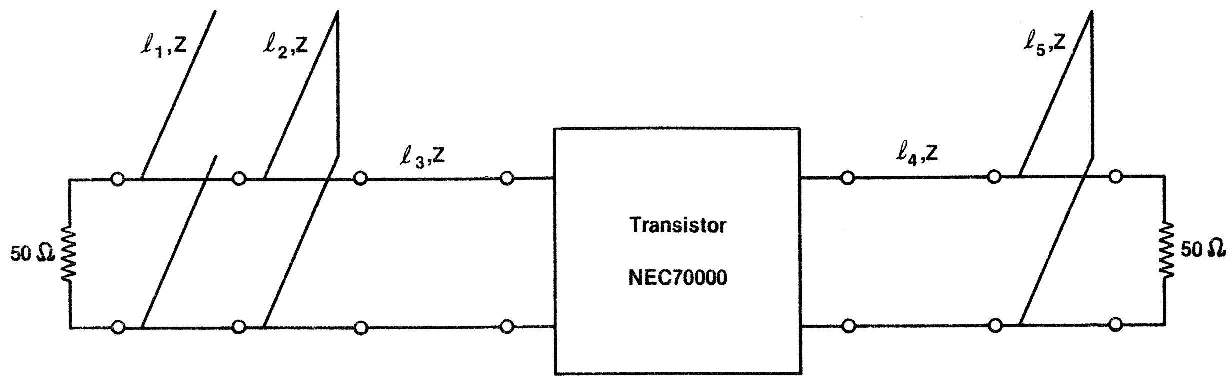

Fig. 1 A microwave amplifier consisting of a NEC70000 FET and transmission lines.

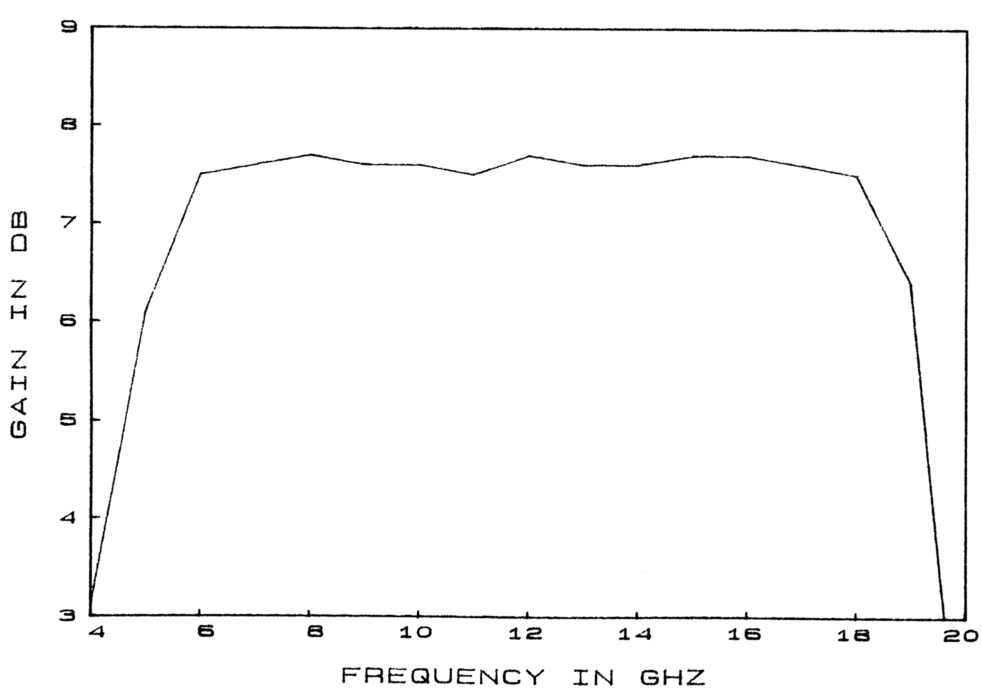

Fig. 2 The nominal response of the amplifier in Fig. 1.

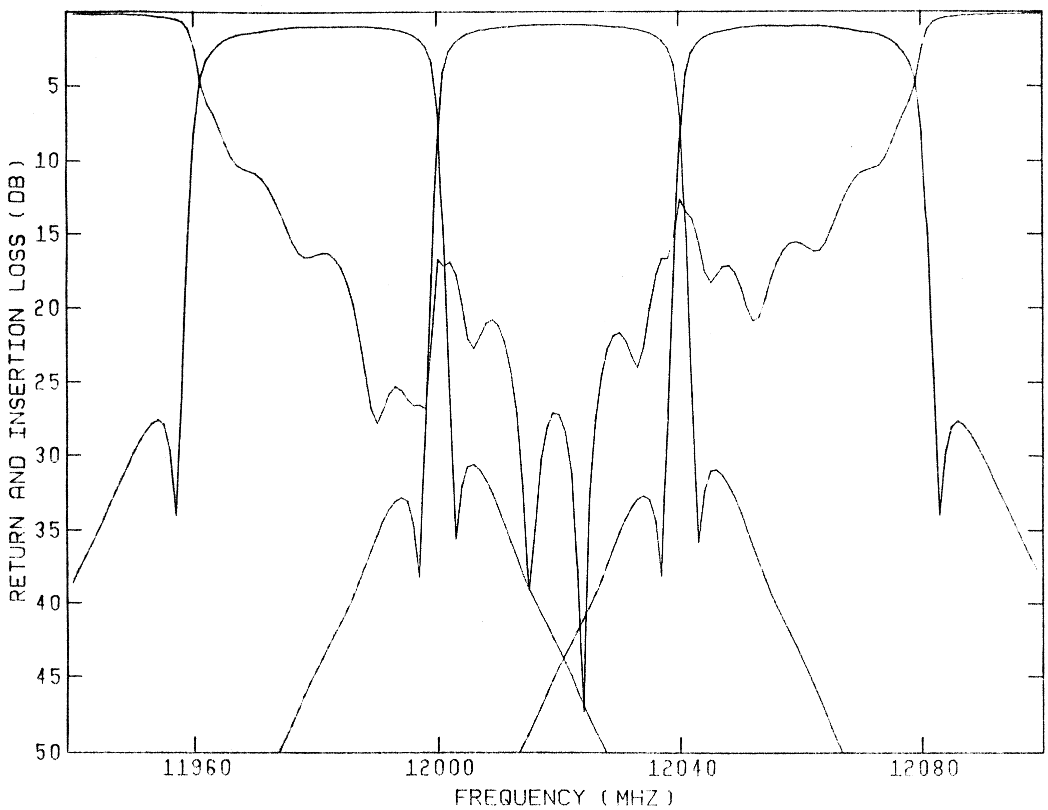

Fig. 3 Responses of the 3-channel multiplexer at the starting point.

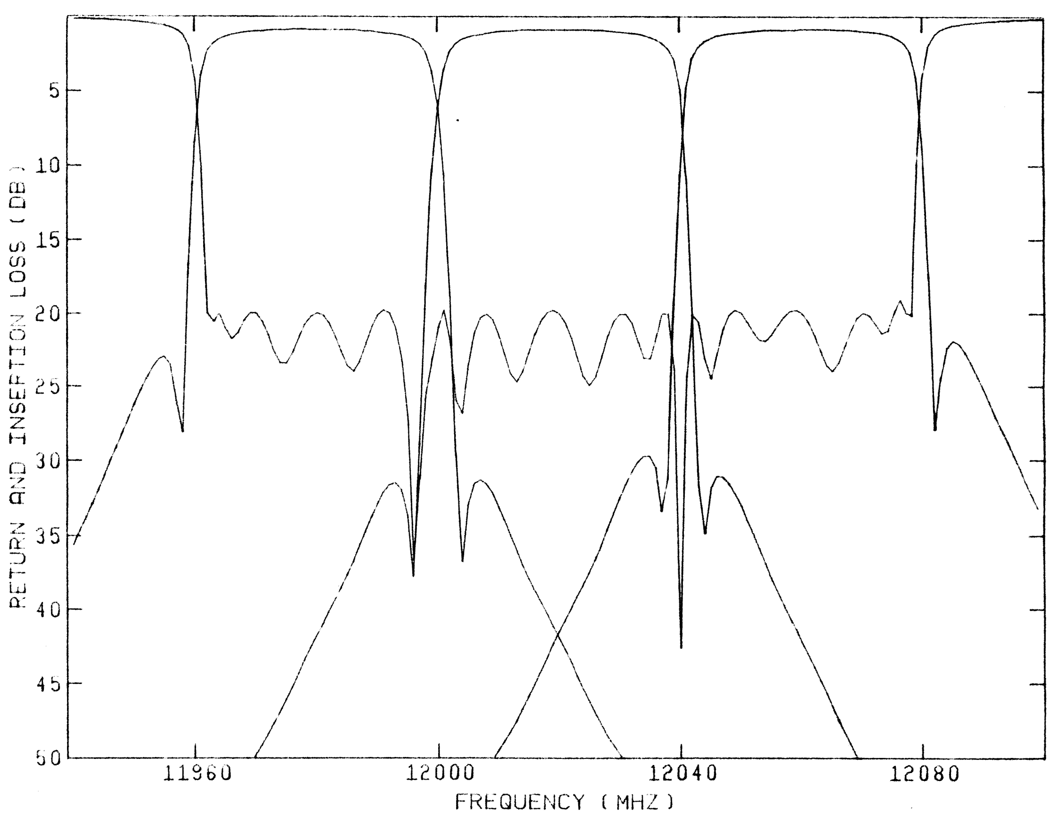

Fig. 4 Responses of the 3-channel multiplexer with 45 optimized parameters.