An Automatic Decomposition Approach to Optimization of Large Microwave Systems

JOHN W. BANDLER, FELLOW, IEEE, AND QI-JUN ZHANG, STUDENT MEMBER, IEEE

Abstract—**We present a novel and general technique applicable to the optimization of large microwave systems. Using sensitivity information obtained from a suitable Monte Carlo analysis, we extract possible decomposition properties which could otherwise be deduced only through a detailed physical and topological investigation. The overall problem is automatically separated into a sequence of subproblems, each being characterized by the optimization of a subset of circuit functions w.r.t. variables which are sensitive to the selected responses. A heuristic algorithm for automatic decomposition is developed. The decomposition patterns are dynamically updated until a satisfactory solution is reached. The partitioning approach proposed by Kondoh for FET modeling problems is verified. The technique was successfully tested on large-scale optimizations of microwave multiplexers involving 16 channels, 399 nonlinear functions, and 240 variables.**

I. Introduction

A SERIOUS CHALLENGE to researchers in microwave CAD areas is due to the size of practical microwave systems. Existing CAD techniques, mature enough to handle systems of ordinary size, generally balk at large circuits. The reasons for their failure include prohibitive computer storage and CPU times required. A frequent frustration with large-scale optimization is the increased likelihood of stopping at an undesired local optimum. Other difficulties, especially in prototype and production tuning, are due to human inability to cope with problems involving large numbers of independent variables to be adjusted simultaneously to meet a specified response pattern over a wide frequency range.

Recently, FET modeling [1] and manifold multiplexer design [2] problems were solved using appropriate decomposition schemes. The success of these efforts motivated us to pursue the generalization and automation of decomposition approaches for microwave optimization problems.

The concept of decomposition has been a traditional, mathematically based vehicle for approaching large-scale problems, e.g., in mathematical programming [3]–[6]; in circuit analysis [7]–[13], design [3], and fault diagnosis [14]; and in optimal power flow [15], [16], state estimation [17], and real and reactive power optimization [18].

Manuscript received March 17, 1987; revised August 4, 1987. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Grants A7239 and G1135 and in part by Optimization Systems Associates, Inc.

The authors are with the Simulation Optimization Systems Research Laboratory and the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada L8S 4L7. J. W. Bandler is also with Optimization Systems Associates Inc., Dundas, Ontario, Canada L9H 6L1.

IEEE Log Number 8717104.

Microwave engineers have their own special concerns with decomposition, as exposed in [1], [2] and [19]. Thorough laboratory experimentation has to be performed before using certain function structures assumed in mathematical programming theory. They do not take advantage of topological analysis often exploited in the areas of circuits and systems since microwave device models are oriented more to physical than topological analysis. Unlike power systems, most microwave responses are much more complicated and highly nonlinear. It is often difficult for microwave engineers to analytically indicate possible decomposition patterns. To our knowledge, there does not exist a general and abstract theory describing a decomposition approach to microwave circuit optimization not requiring particular physical or topological knowledge of the system.

In this paper, we present a novel technique applicable to the optimization of large microwave systems. Using sensitivity information obtained from a suitable Monte Carlo analysis, we extract possible decomposition properties which could otherwise be deduced only through a physical and topological investigation. The overall problem is automatically separated into a sequence of subproblems, each being characterized by the optimization of a subset of circuit functions w.r.t. variables which are sensitive to the selected responses. Our suggested technique has been successfully tested on microwave multiplexers involving up to 16 channels and 240 variables.

In Section II, we describe the basic concepts of decomposition for circuit optimization problems. Using these concepts, the partitioning approach for FET modeling problems suggested by Kondoh [1] is verified. Section III illustrates the automatic determination of suboptimization problems. An automated decomposition algorithm for large-scale microwave optimization is presented in Section IV. In Section V, the method is applied to the optimization of microwave multiplexers. Interesting results demonstrating the procedure of automated decomposition for a five-channel multiplexer are depicted in illustrative graphs. The results of optimizing a 16-channel multiplexer using our approach are provided.

II. The Decomposition Approach

A. Circuit Optimization Problems

Let $\phi = [\phi_1\ \phi_2 \cdots \phi_n]^T$ represent the system parameters. The circuit responses, denoted as $F_k(\phi, \omega)$, $k = 1, 2, \cdots, n_F$, are functions of variables ϕ and frequency ω. In an

0018-9480/87/1200-1231$01.00 ©1987 IEEE

optimization problem for circuit design, the objective function usually involves a set of nonlinear error functions $f_j(\phi)$, $j=1,2,\cdots,m$. Typically, the error functions represent the weighted differences between circuit responses and given specifications in the form

$$\begin{aligned} & w_{Uk}(\omega)(F_k(\phi,\omega)-S_{Uk}(\omega)) \\ & -w_{Lk}(\omega)(F_k(\phi,\omega)-S_{Lk}(\omega)) \\ & k\in\{1,2,\cdots,n_F\} \end{aligned} \tag{1}$$

where S_{Uk} and S_{Lk} are upper and lower specifications, respectively, and w_{Uk} and w_{Lk} are weighting factors.

Suppose sets I and J are defined as

$$I \triangleq \{1,2,\cdots,n\} \tag{2}$$

$$J \triangleq \{1,2,\cdots,m\}. \tag{3}$$

The overall optimization problem, e.g., a minimax optimization, is

$$\underset{\phi_i, i\in I}{\text{minimize}} \max_{j\in J} f_j(\phi). \tag{4}$$

B. Description of the Decomposition Approach

In a decomposition approach, one attempts to reach the overall solution by solving a sequence of subproblems. A typical subproblem is characterized by

$$\underset{\phi_i, i\in I^s}{\text{minimize}} \max_{j\in J^s} f_j(\phi) \tag{5}$$

where I^s and J^s are subsets of I and J, respectively.

The basic idea for decomposition is to decouple a variable ϕ_i from a function f_j if the interaction between them is weak. A subproblem contains only the sensitively related variables and functions. A proper arrangement of the sequence of different subproblems to be solved is often important to ensure convergence and efficiency.

C. Sensitivity Analysis

We perform sensitivity analysis at a set of randomly chosen points ϕ^l, $l=1,2,\cdots$. A measure of the interaction between ϕ_i and f_j is defined as

$$S_{ij} \triangleq \sum_l \left(\frac{\partial f_j(\phi^l)}{\partial \phi_i} \frac{\phi_i^0}{f_j^0} \right)^2 \tag{6}$$

where ϕ_i^0 and f_j^0 are used for scaling. All the S_{ij}, $i=1,2,\cdots,n$ and $j=1,2,\cdots,m$ constitute an $n\times m$ sensitivity matrix $\boldsymbol{S}$. It is reasonable to conclude that ϕ_i and f_j can be decoupled if S_{ij} is very small.

D. Grouping of Variables and Functions

The examination of various interaction patterns between ϕ_i, $i\in I$, and f_j, $j\in J$, results in the breakdown of all variables ϕ into p groups identified by index sets $I_1, I_2,\cdots, I_p$, and all functions $\boldsymbol{f}$ into q groups identified by sets $J_1, J_2,\cdots, J_q$. We have

$$I = I_1 \cup I_2 \cup \cdots \cup I_p \tag{7}$$

and

$$J = J_1 \cup J_2 \cup \cdots \cup J_q. \tag{8}$$

The partitioning of ϕ or $\boldsymbol{f}$ can be achieved either manually or automatically. The manual procedure corresponds to the manual determinations of variable groups and function groups using *a priori* knowledge. Such knowledge is typically obtained through extensive laboratory experiment and an excellent understanding of the particular device. The automatic procedure corresponds to the computerized partitioning of ϕ or $\boldsymbol{f}$ based upon the sensitivity matrix $\boldsymbol{S}$.

As an example for manual partitioning of $\boldsymbol{f}$, we consider an N-channel multiplexer. The common port return loss and channel insertion loss responses associated with the same channel can be grouped together since their behavior is similarly affected by variables ϕ. Therefore, we have N groups of functions, i.e., $q=N$. J_l contains indices of error functions related to channel l, $l=1,2,\cdots,N$.

E. Automatic Partitioning of Variables ϕ

Suppose the function groups have been determined, i.e., J has been decomposed into J_l, $l=1,2,\cdots,q$. We define an $n\times q$ matrix $\boldsymbol{C}$ whose (i,l)th component is

$$C_{il} \triangleq \sum_{j\in J_l} S_{ij}. \tag{9}$$

A very small value of an entry in the $\boldsymbol{C}$ matrix, say C_{il}, implies that the ith variable and the lth function group are weakly interconnected.

Let C_{ave} represent the average value of all components in the $\boldsymbol{C}$ matrix. For a given factor λ, $\lambda\geq 0$, the matrix is made sparse such that C_{il} is set to zero if it is less than λC_{ave}. By making $\boldsymbol{C}$ sparse, insensitive variables are eliminated and weak interactions between variables and function groups are decoupled.

Two variables ϕ_i and ϕ_j belong to the same group if they interact only with the same groups of functions, i.e., if the ith and the jth rows of $\boldsymbol{C}$ have the same zero/nonzero pattern. A thorough computerized checking of the $\boldsymbol{C}$ matrix results in the automatic determination of index sets I_k, $k=1,2,\cdots,p$.

F. Example of Matrix $\boldsymbol{C}$

Consider the fictitious relations between variables and function groups shown in Fig. 1(a). The functions $\boldsymbol{f}$ have been arranged into five groups. The $\boldsymbol{C}$ matrix (already made sparse) is

$$\begin{pmatrix} 22. & 100. & 32. & 0. & 0. \\ 0. & 100. & 0. & 0. & 0. \\ 0. & 100. & 0. & 0. & 0. \\ 0. & 0. & 83. & 100. & 0. \\ 0. & 0. & 0. & 0. & 100. \\ 0. & 0. & 100. & 86. & 0. \\ 0. & 0. & 100. & 0. & 0. \\ 0. & 78. & 100. & 55. & 0. \\ 100. & 0. & 0. & 0. & 0. \end{pmatrix} \tag{10}$$

As seen from Fig. 1(a), ϕ_2 and ϕ_3 both affect only the second function group. In the $\boldsymbol{C}$ matrix, rows 2 and 3 both have only one nonzero located at the second column.

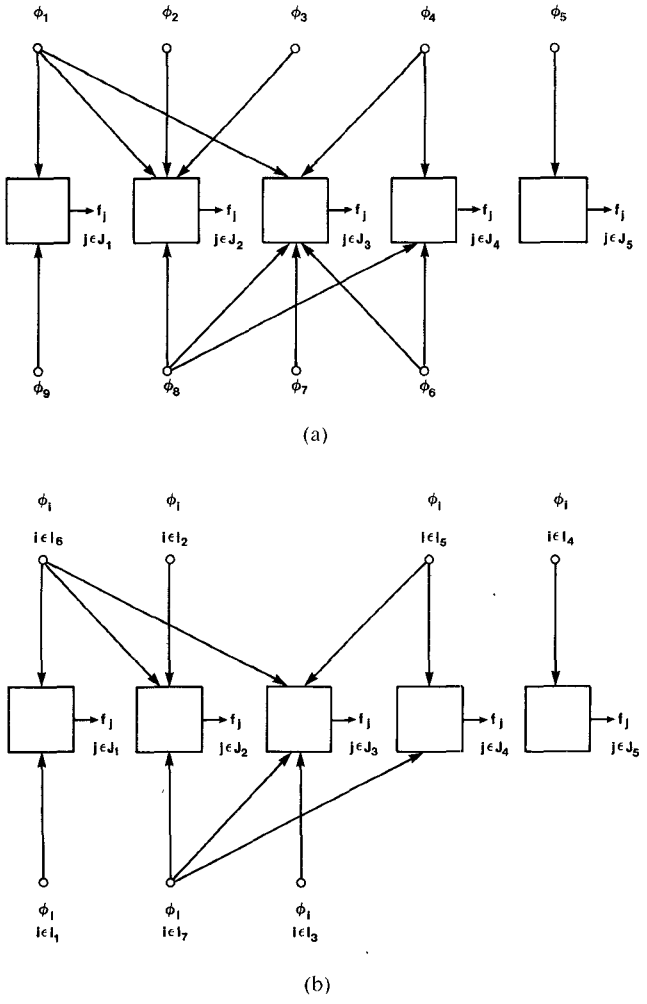


Fig. 1. A fictitous example showing only the strong interconnections between variables and function groups (a) System configuration corresponding to matrix C. (b) System configuration corresponding to matrix D.

Therefore, variables ϕ_2 and ϕ_3 are grouped together. Similarly, variables ϕ_4 and ϕ_6 belong to the same group. The resulting index sets for variable groups are $I_1 = \{9\}$, $I_2 = \{2, 3\}$, $I_3 = \{7\}$, $I_4 = \{5\}$, $I_5 = \{4,6\}$, $I_6 = \{1\}$, and $I_7 = \{8\}$. The index sets have been ordered such that the kth variable group correlates with no more function groups than the $(k+1)$th variable group does, $k = 1,2,\cdots,6$. Such an arrangement is made to keep subsequent description simple.

G. Decomposition Dictionary

To manipulate directly with groups of variables and groups of functions, we construct a $p \times q$ dictionary decomposition matrix D. Define the (k,l)th component of D as

$$D_{kl} \triangleq \sum_{i \in I_k} \sum_{j \in J_l} S_{ij} = \sum_{i \in I_k} C_{il}. \tag{11}$$

If D_{kl} is zero, variables in the kth group are decoupled from functions in the lth group. Otherwise if $D_{kl} \neq 0$, we say that ϕ_i, $i \in I_k$, and f_j, $j \in J_l$, are correlated. The decomposition dictionary gives a clear picture of the correlation patterns between groups of variables and functions, facilitating the automatic determination of suboptimization problems. The ideal dictionary is a diagonal matrix where a subproblem simply corresponds to a diagonal element. In this case, only one variable group and one function group are involved in a subproblem. If a diagonal dictionary can be obtained without artificially making C sparse (i.e., using sparse factor $\lambda = 0$), then the system is completely decomposable [20]. For a completely decomposable system, different subproblems can be calculated in parallel.

H. Example of the Decomposition Dictionary

Consider the previous example with the resulting C matrix defined in (10). According to the index sets I_k, $k = 1,2,\cdots,7$, the decomposition dictionary D can be obtained from C by adding rows 2 and 3, and adding rows 4 and 6, respectively. The relations between groups of variables and functions are shown in Fig. 1(b). The resulting dictionary is

$$\begin{pmatrix} 100. & 0. & 0. & 0. & 0. \\ 0. & 200. & 0. & 0. & 0. \\ 0. & 0. & 100. & 0. & 0. \\ 0. & 0. & 0. & 0. & 100. \\ 0. & 0. & 180. & 180. & 0. \\ 20. & 100. & 30. & 0. & 0. \\ 0. & 70. & 100. & 50. & 0. \end{pmatrix} \tag{12}$$

where each entry has been rounded to multiples of 10.

I. Decomposition for FET Device Models

Through extensive experiment on practical FET devices, Kondoh [1] summarized eight suboptimization problems which can be repeatedly solved to yield a FET model with improved accuracy. The equivalent circuit is shown in Fig. 2. We perform sensitivity analysis at ten randomly chosen parameter points in the 10 percent neighborhood of ϕ^0, ϕ^0 representing the true value listed in [1]. The function f_j used in (6) is defined as the weighted difference between the calculated and the measured values of the modulus or the phase of a particular S parameter. The entire frequency band for calculating S is [1.5,26.5] GHz. Functions associated with the same S parameter are grouped together. Table I shows the C matrix of (9) before being made sparse, indicating strong as well as weak interconnections between each individual parameter and different groups of functions. In the table, each row has been scaled. Table II provides an example of the decomposition dictionary calculated and normalized from Table I. Table II yields eight subproblems which agree with and further verify the decomposition scheme proposed in [1]. When the C matrix is made sparse, certain entries, whose values are only slightly less than the dominant ones, are also set to zero. Therefore, as mentioned in [1], repeated cycling and careful ordering of the eight suboptimizations are necessary.

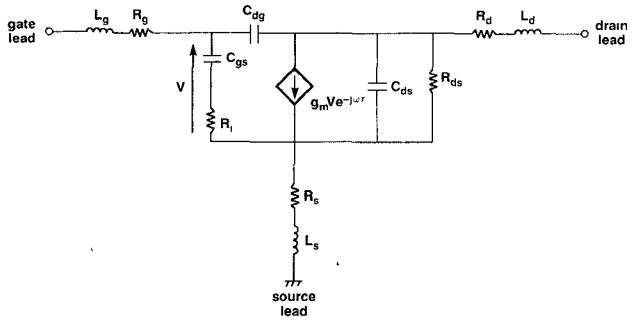


Fig. 2. A FET equivalent circuit.

TABLE I
THE *C* MATRIX FOR THE FET MODEL

Frequency Band	Variables	Function Groups S_{11}	S_{21}	S_{12}	S_{22}
	g_m	18 55	100.00	87.55	68 33
	C_{gs}	100.00	89.74	67.98	62.25
entire	C_{ds}	4.88	67.74	45.73	100.00
band	C_{dg}	4.24	48.88	100.00	81.27
	R_s	35.53	37.14	100.00	5.88
	R_{ds}	17 44	97.68	70.51	100.00
	τ	31.91	100.00	36.61	59.31
	R_g	100 00	50.67	24.87	29.89
upper	R_d	34.65	74.31	85.85	100.00
half	R_i	100.00	65.63	88 43	39.53
band	L_g	100.00	87.85	57.16	37.44
	L_d	9.99	97.88	61.78	100.00
	L_s	62.94	31.31	100 00	21.99

TABLE II
NORMALIZED DECOMPOSITION DICTIONARY *D*

Frequency Band	Variable Groups	Function Groups S_{11}	S_{21}	S_{12}	S_{22}
	R_{ds}, C_{ds}	0.00	0.00	0 00	1.00
entire	C_{gs}	1.00	0.00	0 00	0.00
band	C_{dg}, R_s	0.00	0.00	1 00	0.00
	g_m	0.00	1.00	0 00	0.00
upper	R_d, L_d	0.00	0.00	0 00	1.00
half	R_g, R_i, L_g	1.00	0.00	0 00	0.00
band	L_s	0.00	0.00	1.00	0.00
	τ	0.00	1.00	0 00	0.00

The feasibility of computerized automatic decomposition is demonstrated by this example.

III. AUTOMATIC DETERMINATION OF SUBOPTIMIZATION PROBLEMS

A. Reference Function Group and Candidate Variable Group

Usually, the decomposition dictionary is not diagonal. A suboptimization often involves several function groups and several variable groups. Among the function groups involved, there is a key group which we call the reference group. Such a group typically contains the worst error function. The reference function group is used to initiate a subproblem as described in the subsequent text.

Suppose the index set J_l indicates the reference function group. The candidate groups of variables to be used for the suboptimization are those which affect f_j, $j \in J_l$.

In the decomposition dictionary, the lth column associates with the reference function group. Rows having a nonzero in the lth column are candidate rows, each corresponding to a candidate variable group. Take Fig. 1(b) as an example. Suppose that the function group J_2 is the reference group, i.e., $l = 2$. The candidate groups of variables are I_2, I_6, and I_7 since they correlate with the reference function group. Correspondingly, in the ***D*** matrix of (12), rows 2, 6, and 7 are candidate rows since they all have a nonzero in the second column.

B. Determination of a Suboptimization Problem

An automatic procedure for the determination of I^s and J^s for the suboptimization of (5) has been developed. Suppose J_l indicates the reference function group. For a selected candidate variable group, e.g., the one corresponding to set I_k, the index set J^s indicates the union of all function groups which correlate with variable group k. I^s identifies variables in the kth group, as well as all other variables which correlate with functions only within f_j, $j \in J^s$. Also, I^s excludes variables not correlating with any active functions in f_j, $j \in J^s$. A function f is said to be active if

$$f > 0.8M_f \quad \text{when } M_f > 0$$
$$f > 1.25M_f \quad \text{when } M_f < 0 \tag{13}$$

where

$$M_f \triangleq \max_{j \in J^s} f_j. \tag{14}$$

C. Priority of Candidate Variable Groups

It can be seen that a pair of (I^s, J^s) associate with a pair of (I_k, J_l). For a selected reference function group, each candidate variable group leads to a subproblem. The sequence of subproblems used to penalize f_j, $j \in J_l$, is determined by the priority of all resulting candidates.

Since each candidate determines the function set J^s for a suboptimization, the priority of the candidate is based upon the pattern of error functions it will affect, i.e., patterns of f_j, $j \in J^s$. Firstly, the fewer the number of function groups in J^s, the higher the priority. Secondly, the worse the overall error functions in J^s, the higher the priority. The overall error functions in J^s are ranked by an appropriate measure, e.g., the generalized least pth function (GLP) [21].

The priority of candidate variable groups can be similarly determined in the decomposition dictionary. The fewer the number of nonzeros that exist in a candidate row, the higher the priority. For two candidate rows containing an equal number of nonzeros, a higher priority is

given to the candidate having a larger value in its generalized least pth function.

D. Example

For the example of Fig. 1, suppose that the maximum error functions within each of the five function groups are [3.8 4. 1. −1. 2.]. Suppose that we choose the worst group, i.e., group 2, as the reference function group. According to our previous discussions, the candidate variable groups are I_2, I_6, and I_7. I_2 has the highest priority since it affects fewer (i.e., only one) function groups than I_6 or I_7 does. The functions affected by variables in I_6 (or I_7) are f_j, $j \in J^s = J_1 \cup J_2 \cup J_3$ (or $J^s = J_2 \cup J_3 \cup J_4$). I_6 has a higher priority than I_7 since the overall error functions in $J_1 \cup J_2 \cup J_3$ are worse than in $J_2 \cup J_3 \cup J_4$.

Correspondingly, in the decomposition dictionary of (12), rows 2, 6, and 7 are candidates. Row 2 has the highest priority since it contains fewer nonzeros than others. Row 6 has the second highest priority since its GLP value is larger than the GLP value for row 7.

To formulate a suboptimization problem, i.e., to decide on I^s and J^s, we choose a pair of (I_k, J_l), e.g., candidate variable group I_6 and reference function group J_2. The index set $J^s = J_1 \cup J_2 \cup J_3$. The variable index set I^s includes I_6 (indicating the candidate variable group), as well as I_1, I_2, and I_3 (indicating all other variables affecting functions only within J^s). Further, I_3 can be excluded from I^s since variables in I_3 do not affect active functions in J^s. Therefore, we have $I^s = I_6 \cup I_1 \cup I_2$.

E. Circuit Responses and Sample Frequencies

When a subset of error functions $f_j(\phi)$, $j \in J^s$, are included in a subproblem, the necessary circuit response functions $F_a(\phi, \omega_b)$, $a \in \{1, 2, \cdots, n_F\}$ and frequency points ω_b, $b \in \{1, 2, \cdots, n_\omega\}$, should be selected for circuit simulation programs. This is accomplished using a coding scheme representing the one-to-one correspondence between j and (a, b). We define weighting factor matrices $\boldsymbol{W}_U$ (for upper specification) and $\boldsymbol{W}_L$ (for lower specification). Both matrices are n_F by n_ω. The (a, b)th component of $\boldsymbol{W}_U$ and $\boldsymbol{W}_L$ are the weighting factors $w_{Ua}(\omega_b)$ and $w_{La}(\omega_b)$, respectively, as defined in (1). The quantity $w_{Ua}(\omega_b)$ or $w_{La}(\omega_b)$ is zero if no upper or lower specification is imposed on $F_a(\phi, \omega_b)$. The coding scheme relating the index of f_j to the indices of nonzeros in $\boldsymbol{W}_U$ and $\boldsymbol{W}_L$ is constructed by systematically scanning though $\boldsymbol{W}_U$ and then $\boldsymbol{W}_L$, respectively.

IV. An Automatic Decomposition Algorithm

An automatic decomposition algorithm for optimization of microwave systems has been developed and implemented. The algorithm heuristically decides when to update the sensitivity matrix and the decomposition dictionary. The formulation and the sequence of suboptimization problems are dynamically determined. The degree of decomposition is reduced as the system converges to its overall solution. As a special case, if all variables interact with all functions, our approach solves only one subproblem, this being identical to the original overall optimization.

Step 1: Initialize sparse factor λ. Calculate the sensitivity matrix $\boldsymbol{S}$ and the decomposition dictionary $\boldsymbol{D}$. Calculate $\boldsymbol{f}$.

Comment: The initial sensitivity matrix can be obtained from a suitable Monte Carlo sensitivity analysis performed off-line. All error functions are calculated in this step.

Step 2: Define l such that

$$f_{\text{worst}} = \max_{j \in J_l} f_j = \max_{j \in J} f_j.$$

Comment: The lth function group contains the worst response. Such a function group will be frequently chosen as the reference group to be penalized.

Step 3: For the given l, determine the sequence of candidate rows in $\boldsymbol{D}$. Rank the candidates in decreasing priority. Set $k = 0$.

Comment: The lth function group is the reference group to be penalized. All variable groups correlating with the lth function group are considered as candidates.

Step 4: If $k = 0$ then set k to the row index of the first candidate; otherwise set k to the row index of the next candidate. If such a candidate does not exist, then go to step 8.

Comment: The candidate groups of variables are sequentially selected. Each entry into this step results in a selection of a candidate with a lower priority than the current one.

Step 5: Define I^s and J^s using the current k, l. If I^s and J^s are identical with their previous values, then go to step 4. Solve the suboptimization problem of (5). Terminate the optimization if

$$\max_{j \notin J^s} f_j > \lambda' f_{\text{worst}}.$$

Comment: A subproblem is formulated and solved in this step. By checking the functions not covered in the present suboptimization, any significant deterioration in the overall objective function is prevented. The factor λ′ can be, e.g., 1.2.

Step 6: If $I^s = I$ and $J^s = J$, then stop.

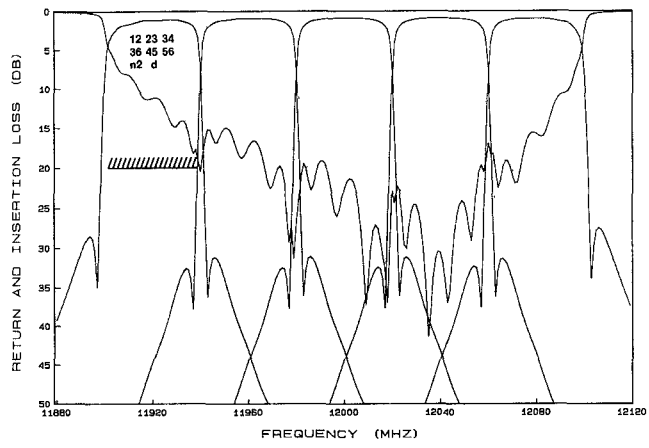
Comment: The program terminates following the completion of an overall optimization which is considered as the last subproblem.

Step 7: Calculate $\boldsymbol{f}$. Calculate

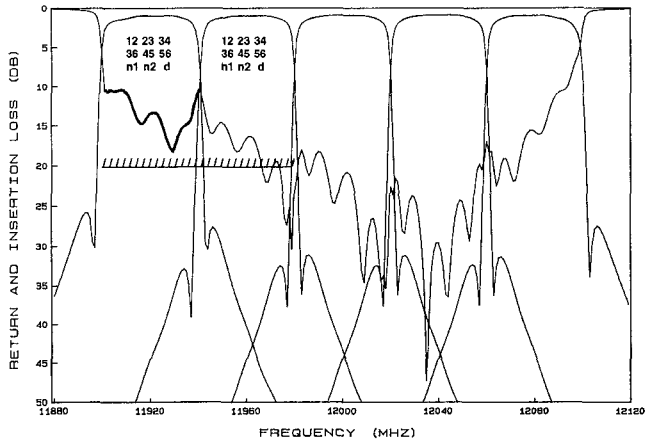
$$f_{\text{worst}} = \max_{j \in J} f_j.$$

Go to step 5.

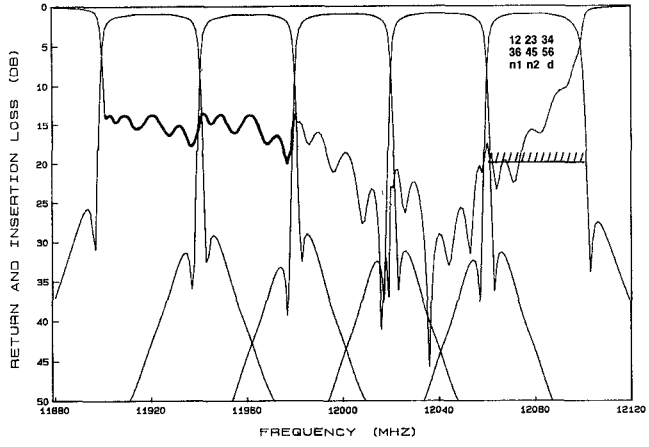
Comment: An overall simulation is performed. By going to step 5, the current reference function group can be continuously penalized in the next



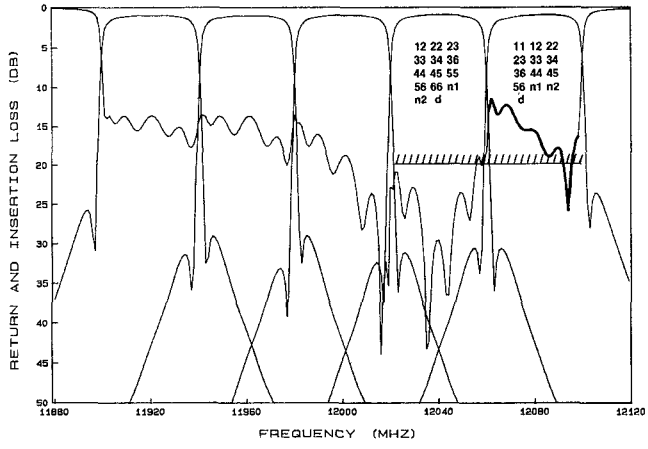
(a)



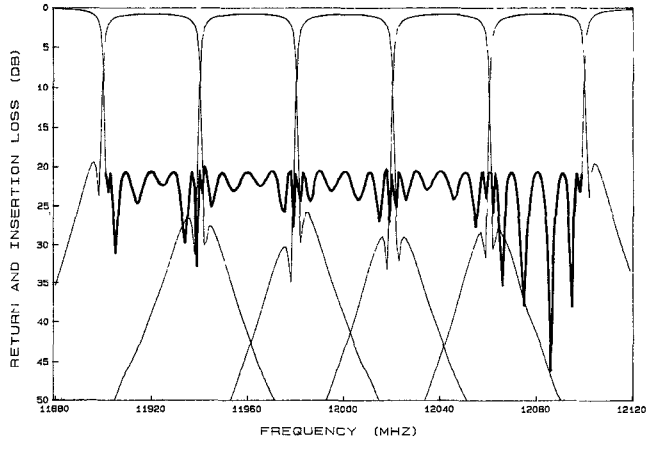
(b)



(c)



(d)



(e)

Fig. 3. Return and insertion loss responses of the five-channel multiplexer for each suboptimization. The 20-dB specification line indicates which channel(s) is to be optimized in the next subproblem. The variables to be selected are indicated, e.g., 35 representing coupling M_{35}, d representing the distance of the corresponding channel filter from the short circuit main cascade termination. The previously optimized channels are highlighted by thick response curves. (a) Responses at the starting point. (b) Responses after the first suboptimization. (c) Responses after the second suboptimization. (d) Responses after the third suboptimization. (e) Responses at the optimum solution.

subproblem even if this group does not include the worst error functions.

Step 8: If

$$\max_{j \in J^s} f_j < \max_{j \in J} f_j$$

then go to step 2. If $\lambda \approx 0$, then stop; otherwise, update $\boldsymbol{S}$, reduce λ, update dictionary $\boldsymbol{D}$, and go to step 3.

Comment: When the selection of a candidate fails, a new sequence of candidates will be defined by going to step 2 or 3. By reducing the sparse factor λ, the degree of decomposition is reduced as the overall solution is being approached. The reference function group will be readjusted if the existing one does not contain the maximum error function. For completely decomposable problems, the terminating conditions in step 6 will not be satisfied and the program will exit from step 8.

While the theory in the previous sections is applicable to general optimization problems such as the least pth optimization, the algorithm described in this section is particularly suitable for the minimax optimization defined by (4). Both variables and functions are allowed to overlap between different subproblems. Convergence of the algorithm is not theoretically guaranteed.

V. Large-Scale Optimization of Multiplexers

The automatic decomposition technique was tested on the optimization of microwave multiplexers used in satellite communications. Specifications were imposed on the common port return loss and individual channel insertion loss functions. Each suboptimization was solved using a recent minimax algorithm [22]. Until our recent paper on multiplexers [2], the reported design and manufacturing of these devices was limited to 12 channels [23]–[27].

A contiguous band five-channel multiplexer was specifically optimized to illustrate the novel process of automatic decomposition, as shown in Fig. 3. Functions associated with the same channel are grouped together. Variables for each channel include six coupling parameters, six cavity resonances, input and output transformer ratios (n_1 and n_2), and the distance measure from the channel filter to the short circuit main cascade termination. The overall problem involved 75 variables and 124 nonlinear functions. Fig. 3(a)–(d) shows the multiplexer responses for the first three suboptimizations. Eleven suboptimizations were used reaching the optimal solution shown in Fig. 3(e). The final subproblem was the overall optimization.

We also tested our approach on a 16-channel multiplexer involving 240 variables and 399 nonlinear functions. The responses at the starting point are shown in Fig. 4. Only ten suboptimizations were performed before reaching the response of Fig. 5. Then a full optimization was activated, resulting in all responses satisfying their specifications as shown in Fig. 6. A comparison between the optimal design with and without decomposition is provided in Table III. When used to obtain a good starting point for subsequent optimization, the decomposition approach offers considerable reductions in both CPU time and storage. The feasibility of obtaining a near optimum

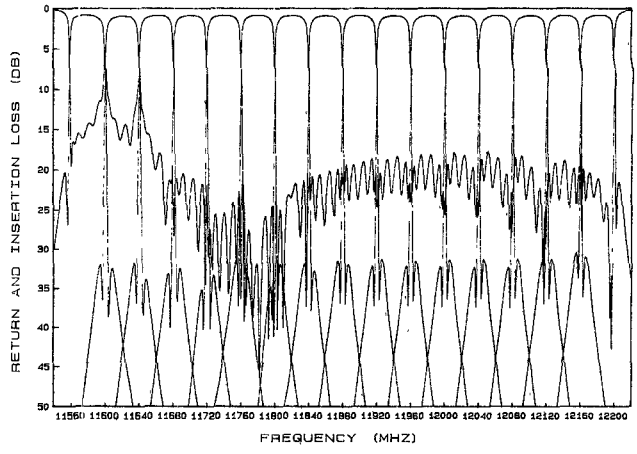


Fig. 4. Return and insertion loss responses of the 16-channel multiplexer before optimization.

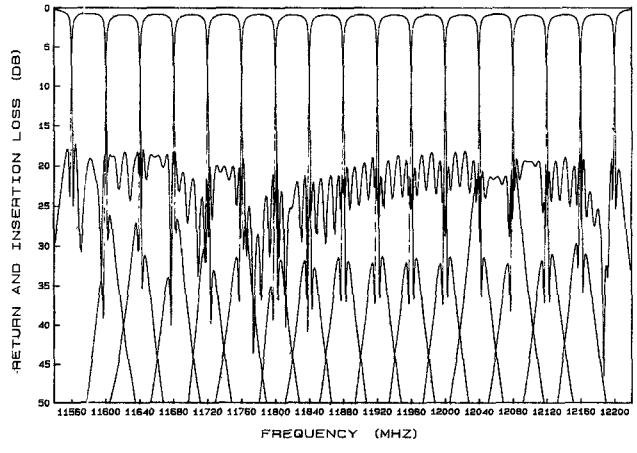


Fig. 5. Return and insertion loss responses of the 16-channel multiplexer after ten suboptimizations. Each of the ten suboptimizations involved responses associated with only one channel and no more than 15 variables.

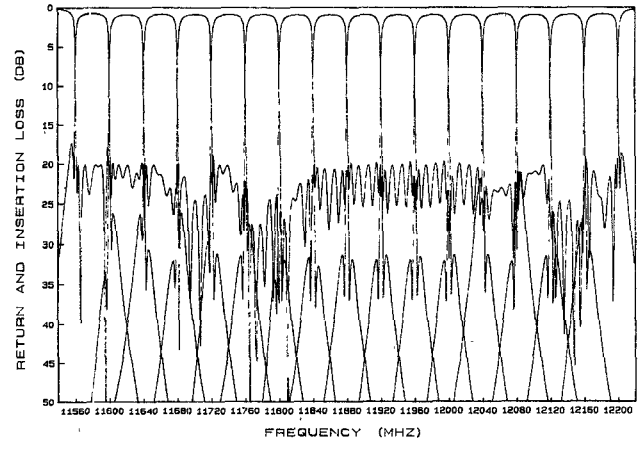


Fig. 6. Return and insertion loss responses of the 16-channel multiplexer at the overall solution. All design specifications are satisfied.

TABLE III
COMPARISON OF 16-CHANNEL MULTIPLEXER OPTIMIZATION WITH AND WITHOUT DECOMPOSITION

Purpose of Optimization+	Reduction in Objective Function	Criteria for Comparison	With Decomp	Without Decomp
to provide a good starting point for further optimization	from 13.46 to 2.4	CPU time *	99	250
		working space needed†	2,197	483,036
to obtain a near optimum solution	from 13.46 to 0.32	CPU time *	651	553
		working space needed†	73,972	483,036
to obtain optimum solution	from 13.46 to −0.09	CPU time *	1045	1289
		working space needed†	483,036	483,036

+ Different sparse factors λ have been used to control the degree of decomposition for the three different purposes.
*Seconds on the FPS-264 mainframe.
† Machine memory units (one unit per real number) required by the minimax optimization package [22]

for large problems using computers with memory limitations is observed from the table. Such a near optimum is obtained at the cost of increased CPU time. When close to the desired solution, the sizes of the subproblems may approach that of the overall problem. In this case, the performance of optimization does not differ significantly with or without decomposition, unless the original problem is almost completely decomposable.

VI. CONCLUSIONS

We have presented an automated decomposition approach for optimization of large microwave systems. Compared with the existing decomposition methods, the novelty of our approach lies in its generality in terms of device independency and its automation. Advantages of the approach are 1) a very significant saving of CPU time and/or computer storage and 2) efficient decomposition by automation. By partitioning the overall problem into smaller ones, the approach promises to provide a basis for computer-assisted tuning. It contributes positively towards future general computer software for large-scale optimization of microwave systems.

ACKNOWLEDGMENT

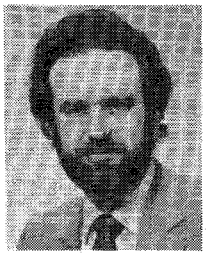
The authors acknowledge fruitful interactions with Dr. W. Kellermann, now with Ontario Hydro, Toronto, Canada, and Dr. S. Daijavad, now with the University of California, Berkeley. Dr. S. H. Chen, now with Optimization Systems Associates Inc., Dundas, Ontario, Canada, and Ms. M. L. Renault of McMaster University, Hamilton, Canada, provided stimulating suggestions and cooperation throughout this project. Technical discussions with Dr. M. H. Chen of TRW, Redondo Beach, CA. and Dr. K. Madsen of the Technical University of Denmark, Lyngby, are also appreciated.

REFERENCES

[1] H. Kondoh, "An accurate FET modelling from measured S-parameters," in *IEEE Int. Microwave Sym. Dig.* (Baltimore, MD), 1986, pp. 377–380.
[2] J. W Bandler, S. H. Chen, S Daijavad, W. Kellermann, M. Renault, and Q. J. Zhang, "Large scale minimax optimization of microwave multiplexers," in *Proc European Microwave Conf.* (Dublin, Ireland), 1986, pp. 435–440.
[3] D. M Himmelblau, Ed., *Decomposition of Large-Scale Problems*. Amsterdam: North-Holland, 1973.
[4] A. M. Geoffrion, "Elements of large-scale mathematical programming," *Management Sci.*, vol. 16, pp. 652–691, 1970.
[5] L. S. Lasdon, *Optimization Theory for Large Systems*. New York: MacMillan, 1970.
[6] H. P. L Luna, "A survey on informational decentralization and mathematical programming decomposition," in *Mathematical Programming*, R. W Cottle, M L. Kelmanson, and B. Korte, Eds. Amsterdam: Elsevier, 1984.
[7] F. Luccio and M Sami, "On the decomposition of networks in minimally interconnected subnetworks," *IEEE Trans. Circuit Theory*, vol CT-16, pp. 184–188, 1969.
[8] L O. Chua and L. K. Chen, "Diakoptic and generalized hybrid analysis," *IEEE Trans. Circuits Syst.*, vol. CAS-23, pp. 694–705, 1976.
[9] F. F Wu, "Solution of large-scale networks by tearing," *IEEE Trans Circuits Syst*, vol. CAS-23, pp. 706–713, 1976.
[10] M. D Tong and W K. Chen, "Hybrid analysis of a large-scale network by node-tearing," in *Proc. IEEE Int. Symp. Circuits Syst.* (San Jose, CA), 1986, pp. 178–181.
[11] H Asai, M. Urano, and M. Tanaka, "A new simulation technique for a large scale transistor circuit based on relaxation method using network tearing," in *Proc. IEEE Int. Symp Circuits Syst* (San Jose, CA), 1986, pp 161–164.
[12] J Pospisil, "Linear network analysis using scattering matrix decomposition," in *Proc. IEEE Int. Symp Circuits Syst* (San Jose, CA), 1986, pp. 193–196.
[13] A Sangiovanni-Vincentelli, L. K Chen, and L. O. Chua, "An efficient heuristic cluster algorithm for tearing large-scale networks," *IEEE Trans. Circuit. Syst.*, vol. CAS-24, pp. 709–717, 1977.
[14] A. E. Salama, J. A. Starzyk, and J. W. Bandler, "A unified decomposition approach for fault location in large analog circuits," *IEEE Trans Circuits Syst.*, vol. CAS-31, pp 609–622, 1984.
[15] S. N. Talukdar, T. C. Giras, and V. K. Kalyan, "Decompositions for optimal power flows," *IEEE Trans. Power App. Syst.*, vol. PAS-102, pp. 3877–3884, 1983.
[16] G C. Contaxis, C. Delkis, and G. Korres, "Decoupled optimal load flow using linear or quadratic programming," *IEEE Trans. Power Syst.*, vol. PWRS-1, pp 1–7, 1986
[17] K. L. Lo and Y. M. Mahmoud, "A decoupled linear programming technique for power system state estimation," *IEEE Trans. Power Syst.*, vol. PWRS-1, pp. 154–160, 1986.
[18] R. Billinton and S. S. Sachdeva, "Real and reactive power optimization by suboptimum techniques," *IEEE Trans Power App Syst.*, vol. PAS-92, pp. 950–956, 1973.
[19] C. Tsironis and R. Meierer, "Microwave wide-band model of GaAs dual gate MESFET's," *IEEE Trans Microwave Theory Tech.*, vol. MTT-30, pp. 243–251, 1982.
[20] P. J Courtois, *Decomposability*. New York: Academic Press, 1977
[21] J W Bandler and M. R. M. Rizk, "Optimization of electrical circuits," *Mathematical Programming Study on Engineering Optimization*, vol. 11, pp. 1–64, 1979
[22] 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, pp. 1519–1530, 1985
[23] M H. Chen, "Current state-of-the-art technology on contiguous band multiplexer," in *Proc. IEEE Int Symp Circuits Syst* (Kyoto, Japan), 1985, pp. 1583–1586.
[24] R G. Egri, A E. Williams, and A E. Atia, "A contiguous-band multiplexer design," in *IEEE Int. Microwave Symp Dig* (Boston, MA), 1983, pp. 86–88
[25] R. Tong and D. Smith "A 12-channel contiguous-band multiplexer for satellite application," in *IEEE Int Microwave Symp. Dig.* (San Francisco, CA), 1984, pp. 297–298.
[26] S C Holme, "A 12 GHz 12 channel multiplexer for satellite applications," in *IEEE Int. Microwave Symp. Dig.* (San Francisco, CA), 1984, pp 295–296

[27] J. W. Bandler, S. Daijavad, and Q. J. Zhang, "Exact simulation and sensitivity analysis of multiplexing networks," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-34, pp. 93–102, 1986.

✠

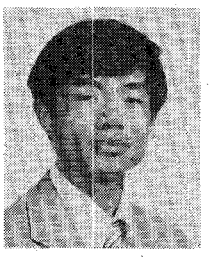


John W. Bandler (S'66–M'66–SM'74–F'78) was born in Jerusalem, Palestine, on November 9, 1941. He studied at Imperial College of Science and Technology, London, England, from 1960 to 1966. He received the B.Sc. (Eng.), Ph.D., and D.Sc. (Eng.) degrees from the University of London, London, England, in 1963, 1967 and 1976, respectively.

He joined Mullard Research Laboratories, Redhill, Surrey, England, in 1966. From 1967 to 1969 he was a Postdoctorate Fellow and Sessional Lecturer at the University of Manitoba, Winnipeg, Canada. He joined McMaster University, Hamilton, Canada, in 1969, where he is currently a Professor of Electrical and Computer Engineering. He has served as Chairman of the Department of Electrical Engineering and Dean of the Faculty of Engineering. He currently directs research in the Simulation Optimization Systems Research Laboratory. Dr. Bandler is President of Optimization Systems Associates Inc., which he established in 1983. OSA currently provides consulting services and software to numerous microwave clients in CAE of microwave integrated circuits.

Dr. Bandler is a contributor to *Modern Filter Theory and Design* (Wiley-Interscience, 1973) and to the forthcoming *Analog Circuits: Computer-Aided Analysis and Diagnosis* (Marcel Dekker). He has more than 220 publications, four of which appear in *Computer-Aided Filter Design* (IEEE Press, 1973), one in *Microwave Integrated Circuits* (Artech House, 1975), one in *Low-Noise Microwave Transistors and Amplifiers* (IEEE Press, 1981), one in *Microwave Integrated Circuits* (2nd ed., Artech House, 1985), one to be published in *Statistical Design of Integrated Circuits* (IEEE Press), and one to be published in *Analog Fault Diagnosis* (IEEE Press). Dr. Bandler was an Associate Editor of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES (1969–1974). He was Guest Editor of the Special Issue of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES on Computer-Oriented Microwave Practices (March 1974). Dr. Bandler is a Fellow of the Royal Society of Canada and of the Institution of Electrical Engineers (Great Britain). He is a member of the Association of Professional Engineers of the Province of Ontario (Canada).

✠



Qi-jun Zhang (S'84) was born in Xiangyan, Shanxi, China, on October 8, 1959. He received the B.Eng. degree from the East China Engineering Institute, Nanjing, China, in 1982.

From 1982 to 1983, he was with the Research Institute of Systems Engineering, Tianjin University, Tianjin, China, as a research assistant. He was awarded a Chinese Government Scholarship and joined the Simulation Optimization Systems Research Laboratory, McMaster University, Hamilton, Canada, in 1983. He is currently completing the Ph.D. degree in electrical engineering at McMaster University. He has held an Ontario Graduate Scholarship and a Teaching Assistantship during his doctoral studies. His research interests are in CAD techniques for large-scale simulation and in sensitivity analysis and optimization with applications in circuit design, modeling and diagnosis.