

A Hybrid Aggressive Space-Mapping Algorithm for EM Optimization

Mohamed H. Bakr, *Student Member, IEEE*, John W. Bandler, *Fellow, IEEE*,
Natalia Georgieva, *Member, IEEE*, and Kaj Madsen

Abstract— We propose a novel hybrid aggressive space-mapping (HASM) optimization algorithm. HASM exploits both the trust-region aggressive space-mapping (TRASM) strategy and direct optimization. Severe differences between the coarse and fine models and nonuniqueness of the parameter extraction procedure may cause the TRASM algorithm to be trapped in local minima. The HASM algorithm is based on a novel lemma that enables smooth switching from the TRASM optimization to direct optimization if the TRASM algorithm is not converging. It also enables switching back from direct optimization to the TRASM algorithm in a smooth way. The uniqueness of the extraction step is improved by utilizing a good starting point. The algorithm does not assume that the final space-mapped design is the true optimal design and is robust against severe misalignment between the coarse and fine models. The examples include a seven-section waveguide transformer, the design of an H -plane waveguide filter, and a double-folded stub filter.

Index Terms— CAD, design automation, electromagnetic simulation, iterative methods, microstrip filters, optimization methods, space mapping, waveguide filters.

I. INTRODUCTION

IN THIS PAPER, a hybrid aggressive space-mapping (HASM) algorithm is presented. Space-mapping (SM) optimization [1]–[4] assumes that the circuit under consideration can be simulated using two models: a fine model and a coarse model. The fine model is accurate, but is computationally intensive, e.g., a full-wave electromagnetic (EM) simulator. The coarse model is assumed to be fast, but not very accurate. SM optimization directs most of the optimization computational effort toward the coarse model while maintaining the accuracy of the fine model. The overall computational effort needed is much smaller than that for direct optimization.

Manuscript received March 26, 1999; revised July 12, 1999. This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada under Grant OGP0007239 and under Grant STP0201832, by Com Dev, by Remec Nanowave, and by the Micronet Network of Centres of Excellence. The work of N. Georgieva was supported by an NSERC Postdoctorate Fellowship. The work of M. H. Bakr was supported under an Ontario Graduate Scholarship.

M. H. Bakr and N. Georgieva are with the Simulation Optimization Systems Research Laboratory, Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ont., Canada L8S 4K1.

J. W. Bandler is with the Simulation Optimization Systems Research Laboratory, Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ont., Canada L8S 4K1, and is also with the Bandler Corporation, Dundas, Ont., Canada L9H 5E7.

K. Madsen is with the Department of Mathematical Modeling, Technical University of Denmark, DK-2800 Lyngby, Denmark.

Publisher Item Identifier S 0018-9480(99)08440-9.

Parameter extraction is crucial to the aggressive space mapping (ASM) technique [4]. Here, a coarse-model point whose response matches a given fine-model response is obtained. This is essentially an optimization problem. The nonuniqueness of the extracted parameters may lead to divergence or oscillation of the iterations [2]. To alleviate this problem the trust-region aggressive space-mapping (TRASM) algorithm was introduced [3]. TRASM integrates a trust region methodology [5] with the ASM technique. Also, it utilizes a recursive multi-point parameter extraction in order to improve the uniqueness of the extraction step.

In this paper, we address the convergence behavior of the TRASM algorithm. We discuss the effect of a severely misaligned coarse model. We show that, in this case, TRASM optimization may be trapped in local minima.

The design obtained by pure SM optimization in most cases is very near optimal. However, the optimality of the final design cannot be guaranteed. This is because the final space-mapped response matches the optimal coarse-model response, which may be different from the optimal fine-model response obtained by solving the original design problem in the fine-model space.

Our HASM algorithm is designed to overcome these limitations. The algorithm switches smoothly between SM optimization and direct optimization and vice versa. A novel lemma enables such a switch. The algorithm also integrates a novel prediction of the starting point for the parameter-extraction problem to enhance the uniqueness.

II. SM OPTIMIZATION: A BRIEF REVIEW

We refer to the vectors of “fine”-model parameters and “coarse”-model parameters as \mathbf{x}_{em} and \mathbf{x}_{os} , respectively. The first step in any SM algorithm is to obtain the optimal design of the coarse model \mathbf{x}_{os}^* . The corresponding response is denoted by \mathbf{R}_{os}^* . ASM aims at establishing a mapping P between the two spaces [4]

$$\mathbf{x}_{os} = P(\mathbf{x}_{em}) \quad (1)$$

such that

$$\|\mathbf{R}_{em}(\mathbf{x}_{em}) - \mathbf{R}_{os}(\mathbf{x}_{os})\| \leq \varepsilon \quad (2)$$

where \mathbf{R}_{em} is the vector of fine-model response, \mathbf{R}_{os} is the vector of coarse mode response, and $\|\cdot\|$ is a suitable norm. We define the error function

$$\mathbf{f} = P(\mathbf{x}_{em}) - \mathbf{x}_{os}^*. \quad (3)$$

The final fine-model design is obtained and the mapping is established by solving the nonlinear system

$$\mathbf{f}(\mathbf{x}_{em}) = \mathbf{0}. \quad (4)$$

Let $\mathbf{x}_{em}^{(i)}$ be the i th iterate in the solution of (4). In the ASM technique, the next iterate $\mathbf{x}_{em}^{(i+1)}$ is found by a quasi-Newton iteration

$$\mathbf{x}_{em}^{(i+1)} = \mathbf{x}_{em}^{(i)} + \mathbf{h}^{(i)} \quad (5)$$

where $\mathbf{h}^{(i)}$ is obtained from

$$\mathbf{B}^{(i)}\mathbf{h}^{(i)} = -\mathbf{f}(\mathbf{x}_{em}^{(i)}) \quad (6)$$

and $\mathbf{B}^{(i)}$ is an approximation to the Jacobian of the vector \mathbf{f} with respect to \mathbf{x}_{em} at the i th iteration. The matrix \mathbf{B} is updated at each iteration using Broyden's update [6].

Vector \mathbf{f} is obtained by evaluating $\mathbf{P}(\mathbf{x}_{em})$, which is done indirectly through parameter extraction. The nonuniqueness of this problem may lead to divergence or oscillation of the ASM technique. The TRASM algorithm [3] was designed to overcome this problem. At the i th iteration, the residual vector $\mathbf{f}^{(i)} = \mathbf{P}(\mathbf{x}_{em}^{(i)}) - \mathbf{x}_{os}^*$ defines the difference between the vector of extracted coarse-model parameters $\mathbf{x}_{os}^{(i)} = \mathbf{P}(\mathbf{x}_{em}^{(i)})$ and the optimal coarse-model design. The value $\|\mathbf{f}^{(i)}\|$ serves as a measure of the misalignment between the two spaces in the i th iteration. The TRASM algorithm aims at minimizing $\|\mathbf{f}^{(i)}\|$. The i th iteration of the algorithm is given by

$$(\mathbf{B}^{(i)T}\mathbf{B}^{(i)} + \lambda\mathbf{I})\mathbf{h}^{(i)} = -\mathbf{B}^{(i)T}\mathbf{f}^{(i)}. \quad (7)$$

Parameter λ is selected such that the step obtained satisfies $\|\mathbf{h}^{(i)}\| \leq \delta$, where δ is the size of the trust region.

III. SOME PROPERTIES OF SM

ASM and TRASM are efficient algorithms. The number of fine-model simulations needed to obtain the space-mapped design is of the order of the problem dimensionality. However, both algorithms depend on the existence of a coarse model that is fast and has enough accuracy.

If the coarse model is bad (i.e., very different from the fine model) SM might not work. To illustrate this, we consider the Rosenbrock function [7]. We form an artificial problem in which the ‘‘coarse’’ model is given by

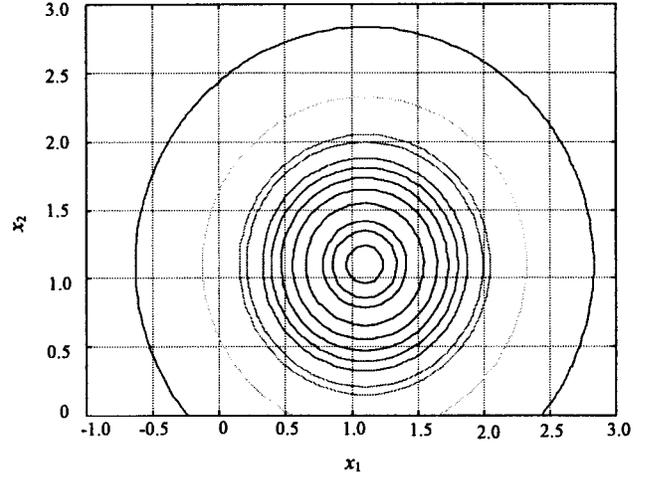
$$R_{os} = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \quad (8)$$

and the ‘‘fine’’ model by

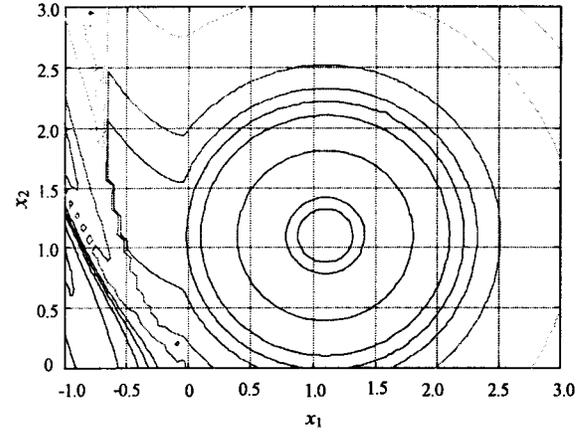
$$R_{em} = 100((x_2 + \alpha_2) - (x_1 + \alpha_1)^2)^2 + (1 - (x_1 + \alpha_1))^2 \quad (9)$$

where α_1 and α_2 are constant shifts. The target of the direct optimization problem is to minimize R_{em} . Considering (8) and (9), we notice that $\mathbf{x}_{os}^* = [1.0 \ 1.0]^T$ and $\mathbf{x}_{em}^* = (\mathbf{x}_{os}^* - \boldsymbol{\alpha})$ where $\boldsymbol{\alpha} = [\alpha_1 \ \alpha_2]^T$. The misalignment between the two models is thus given by the two shifts α_1 and α_2 .

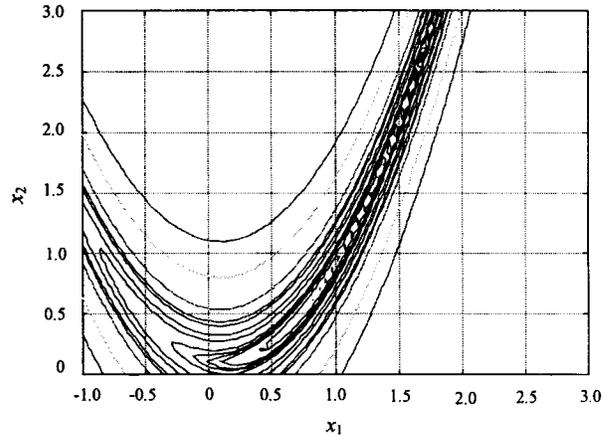
We discuss two sets of values for the shifts. First, we consider $\boldsymbol{\alpha} = [-0.1 \ -0.1]^T$. Using (8) and (9), we notice that the coarse-model point whose response matches the fine response at a point \mathbf{x}_{em} is $\mathbf{x}_{os} = (\mathbf{x}_{em} + \boldsymbol{\alpha})$. It follows that the mapping between the two spaces is given by $\mathbf{P}(\mathbf{x}_{em}) =$



(a)



(b)



(c)

Fig. 1. Different contour plots for the Rosenbrock problem for the case $\boldsymbol{\alpha} = [-0.1 \ -0.1]^T$. (a) The contour plot of $\|\mathbf{x}_{em} + \boldsymbol{\alpha} - \mathbf{x}_{os}^*\|_2^2$. (b) The contour plot of $\|\mathbf{P}(\mathbf{x}_{em}) - \mathbf{x}_{os}^*\|_2^2$ obtained through parameter extraction. (c) Contours of the fine-model Rosenbrock function.

$\mathbf{x}_{em} + \boldsymbol{\alpha}$. The contours of $\|\mathbf{x}_{em} + \boldsymbol{\alpha} - \mathbf{x}_{os}^*\|_2^2$ are shown in Fig. 1(a). The mapping $\mathbf{P}(\mathbf{x}_{em})$ is then approximated through multipoint parameter extraction [3]. Only one perturbed fine-model point is utilized. The contours of $\|\mathbf{P}(\mathbf{x}_{em}) - \mathbf{x}_{os}^*\|_2^2$ obtained in this manner are shown in Fig. 1(b). Fig. 1(a) and

(b) shows that $\|\mathbf{f}\|_2^2$ has a single minimum, which is the solution that would have been obtained by direct optimization. The differences between the two plots are attributed to the nonuniqueness of the parameter-extraction process. Taking the point \mathbf{x}_{os}^* as the initial solution of the fine model, the TRASM algorithm is expected to converge to \mathbf{x}_{em}^* . The corresponding contours of R_{em} are shown in Fig. 1(c), whose minimum value is at $[1.1 \ 1.1]^T$, as expected.

The same steps are repeated for the case $\boldsymbol{\alpha} = [-1.5 \ -1.5]^T$. The contour plot of $\|\mathbf{x}_{em} + \boldsymbol{\alpha} - \mathbf{x}_{os}^*\|_2^2$ is shown in Fig. 2(a). The contour plot of $\|\mathbf{P}(\mathbf{x}_{em}) - \mathbf{x}_{os}^*\|_2^2$ obtained using parameter extraction is shown in Fig. 2(b). Fig. 2(b) illustrates the existence of a minimum of $\|\mathbf{P}(\mathbf{x}_{em}) - \mathbf{x}_{os}^*\|_2^2$ other than \mathbf{x}_{em}^* , which is closer to the starting point of the TRASM algorithm \mathbf{x}_{os}^* . It follows that the TRASM algorithm is unlikely to converge to \mathbf{x}_{em}^* . The corresponding contours of R_{em} for this case are shown in Fig. 2(c), whose minimum value is at $[2.5 \ 2.5]^T$.

The solution obtained using SM optimization is very near optimal if \mathbf{R}_{os}^* is similar to the optimal fine-model response \mathbf{R}_{em}^* . However, this cannot be guaranteed. For example, consider

$$R_{em} = 100(x_2 - x_1^2)^2 + (1 - x_1)^2. \quad (10)$$

Assume also that

$$R_{os} = R_{em} + \varepsilon \quad (11)$$

where $\varepsilon > 0$. It is clear that R_{os}^* is equal to ε while R_{em}^* is zero. It follows that SM may converge to a solution other than \mathbf{x}_{em}^* .

IV. SM AND DIRECT OPTIMIZATION

The properties of SM suggest that a hybrid algorithm be used. This algorithm exploits the efficiency of SM and defaults to direct optimization when SM fails. Our HASM algorithm utilizes a novel lemma that enables smooth switching between the TRASM algorithm and direct optimization and vice versa.

Lemma: Assume that \mathbf{x}_{os} corresponds to \mathbf{x}_{em} through a parameter-extraction process. The Jacobian \mathbf{J}_{em} of the fine-model response at \mathbf{x}_{em} and the Jacobian \mathbf{J}_{os} of the coarse-model response at \mathbf{x}_{os} are then related by

$$\mathbf{J}_{em} = \mathbf{J}_{os}\mathbf{B} \quad (12)$$

where \mathbf{B} is a valid mapping between the two spaces at \mathbf{x}_{os} and \mathbf{x}_{em} .

Proof: As the points \mathbf{x}_{em} and \mathbf{x}_{os} correspond to each other, it follows that their responses match, i.e.,

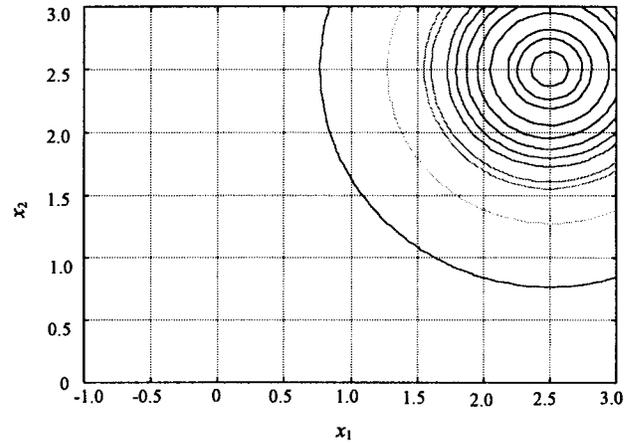
$$\mathbf{R}_{em}(\mathbf{x}_{em}) = \mathbf{R}_{os}(\mathbf{x}_{os}). \quad (13)$$

Now define a new fine-model point $\mathbf{x}_n = \mathbf{x}_{em} + \Delta\mathbf{x}_{em}$ where $\Delta\mathbf{x}_{em}$ is a small perturbation. The response at this new point is perturbed from the response at the point \mathbf{x}_{em} by

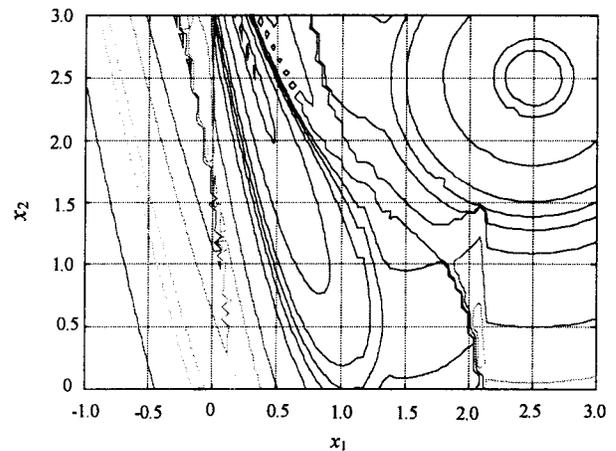
$$\Delta\mathbf{R} = \mathbf{J}_{em}\Delta\mathbf{x}_{em}. \quad (14)$$

The point \mathbf{x}_n corresponds to a coarse-model point $\mathbf{x}_{os} + \Delta\mathbf{x}_{os}$ that satisfies

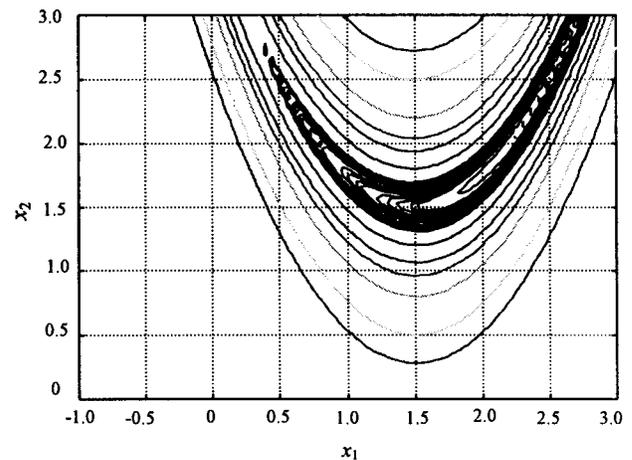
$$\Delta\mathbf{R} = \mathbf{J}_{em}\Delta\mathbf{x}_{em} = \mathbf{J}_{os}\Delta\mathbf{x}_{os}. \quad (15)$$



(a)



(b)



(c)

Fig. 2. Different contour plots for the Rosenbrock problem for the case $\boldsymbol{\alpha} = [-1.5 \ -1.5]^T$. (a) The contour plot of $\|\mathbf{x}_{em} + \boldsymbol{\alpha} - \mathbf{x}_{os}^*\|_2^2$. (b) The contour plot of $\|\mathbf{P}(\mathbf{x}_{em}) - \mathbf{x}_{os}^*\|_2^2$ obtained through parameter extraction. (c) Contours of the fine-model Rosenbrock function.

Also, by definition of the mapping \mathbf{B} , the two perturbations $\Delta\mathbf{x}_{em}$ and $\Delta\mathbf{x}_{os}$ are related by

$$\mathbf{B}\Delta\mathbf{x}_{em} = \Delta\mathbf{x}_{os}. \quad (16)$$

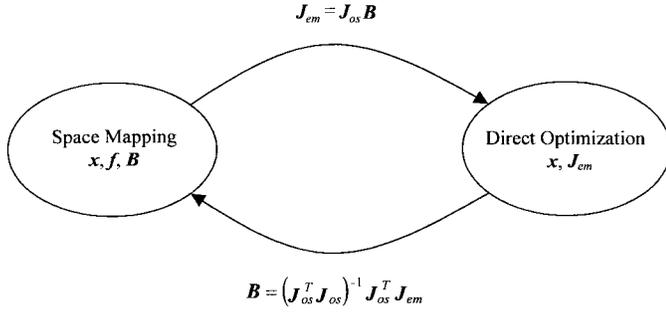


Fig. 3. Illustration of the connection between SM optimization and direct optimization.

Multiplying both sides of (16) by \mathbf{J}_{os} , we get

$$\mathbf{J}_{os}\mathbf{B}\Delta\mathbf{x}_{em} = \mathbf{J}_{os}\Delta\mathbf{x}_{os}. \quad (17)$$

Comparing (17) with (15), we conclude that

$$\mathbf{J}_{em} = \mathbf{J}_{os}\mathbf{B}. \quad (18)$$

Relation (18) is interesting. It shows that by having the matrix \mathbf{B} and the coarse-model Jacobian \mathbf{J}_{os} , we are able to obtain a good estimate of the Jacobian of the fine-model response without any further fine-model simulations. It follows that when SM optimization is not converging, we can switch to direct optimization and supply it with the available first-order derivatives given by (18).

Another relationship that can be easily obtained from (18) is

$$\mathbf{B} = (\mathbf{J}_{os}^T \mathbf{J}_{os})^{-1} \mathbf{J}_{os}^T \mathbf{J}_{em}. \quad (19)$$

Relation (19) assumes that \mathbf{J}_{os} is full rank and $m \geq n$, where n is the number of parameters and m is the dimensionality of both \mathbf{R}_{em} and \mathbf{R}_{os} . It is used for switching back from direct optimization to SM optimization. Fig. 3 illustrates the switching between SM optimization and direct optimization.

We illustrate the lemma as follows. Consider

$$R_{em} = (0.9x_1 + 0.1x_2)^2 + (0.1x_1 + 0.9x_2)^2 \quad (20)$$

and

$$R_{os} = x_1^2 + x_2^2. \quad (21)$$

Take $\mathbf{x}_{em} = [2.0 \ 1.0]^T$. Here, $R_{em} = 4.82$. The solution for the parameter-extraction problem is $\mathbf{x}_{os} = [1.90 \ 1.10]^T$. The Jacobian \mathbf{J}_{os} at \mathbf{x}_{os} is

$$\mathbf{J}_{os} = [3.8 \ 2.2]. \quad (22)$$

From (20) and (21), it is seen that

$$\mathbf{B} = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}. \quad (23)$$

It follows that \mathbf{J}_{em} at \mathbf{x}_{em} is estimated by

$$\mathbf{J}_{em} = [3.8 \ 2.2] \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix} = [3.64 \ 2.36] \quad (24)$$

which is the exact Jacobian of the fine-model response.

V. SELECTION OF THE STARTING POINT

The discussion in Section III reveals how the nonuniqueness of the parameter-extraction process can affect the convergence of SM optimization. The uniqueness of this procedure can be improved by utilizing a good starting point. In the first iteration of the algorithm, there is no available information about the mapping between the two spaces. A reasonable assumption is to take \mathbf{x}_{os}^* as the starting point for the parameter-extraction optimization problem. As the algorithm proceeds, the matrix $\mathbf{B}^{(i)}$ approximates the mapping between the two spaces. A prediction of the extracted parameters in the i th iteration is given by

$$\mathbf{x}_{os}^{(p)} = \mathbf{x}_{os}^{(i)} + \mathbf{B}^{(i)} (\mathbf{x}_{em}^{(i+1)} - \mathbf{x}_{em}^{(i)}). \quad (25)$$

This predicted point is then taken as a starting point for the parameter-extraction optimization problem. It supplies a good starting point provided that $\mathbf{x}_{os}^{(i)}$ is a valid solution to the parameter extraction in the previous iteration and $\mathbf{B}^{(i)}$ approximates the mapping between the two spaces.

VI. THE HASM ALGORITHM

The HASM algorithm exploits SM when effective, otherwise it defaults to direct optimization. Two objective functions are utilized by the algorithm. The first objective function is

$$\|\mathbf{f}\|_2^2 = \|\mathbf{P}(\mathbf{x}_{em}) - \mathbf{x}_{os}^*\|_2^2 \quad (26)$$

while the second function is

$$\|\mathbf{g}\|_2^2 = \|\mathbf{R}_{em}(\mathbf{x}_{em}) - \mathbf{R}_{os}(\mathbf{x}_{os}^*)\|_2^2. \quad (27)$$

While (26) aims at matching the extracted coarse-model parameters to \mathbf{x}_{os}^* in the parameter space, (27) aims at matching the same points mapped through the appropriate responses in the response space.

The HASM algorithm consists of two phases: the first phase follows the TRASM strategy, while the second phase exploits direct optimization. It utilizes (18) and (19) for switching between phases as dictated by the smoothness of convergence.

In the i th iteration, we assume the existence of trusted extracted coarse-model parameters $\mathbf{x}_{os}^{(i)} = \mathbf{P}(\mathbf{x}_{em}^{(i)})$. The step taken in this iteration is given by (7) where $\mathbf{x}_{em}^{(i+1)} = \mathbf{x}_{em}^{(i)} + \mathbf{h}^{(i)}$. Single-point parameter extraction is then applied at the point $\mathbf{x}_{em}^{(i+1)}$ to get $\mathbf{f}^{(i+1)} = \mathbf{P}(\mathbf{x}_{em}^{(i+1)}) - \mathbf{x}_{os}^*$.

The first phase utilizes two success criteria related to the reduction in (26) and (27). The SM success criterion is defined as

$$\left(\|\mathbf{f}^{(i)}\| - \|\mathbf{f}^{(i+1)}\| \right) > 0.01 \left(\|\mathbf{f}^{(i)}\| - \|\mathbf{f}^{(i)} + \mathbf{B}^{(i)} \mathbf{h}^{(i)}\| \right) \quad (28)$$

which indicates that the actual reduction in the objective function (26) should be greater than a certain fraction of the predicted reduction. The direct optimization success criterion is

$$\|\mathbf{g}^{(i+1)}\| < \|\mathbf{g}^{(i)}\| \quad (29)$$

which implies that the new iterate $\mathbf{x}_{em}^{(i+1)}$ is a descent iterate of (27).

The new point $\mathbf{x}_{em}^{(i+1)}$ is accepted and the first phase resumes if this point satisfies both (28) and (29). $\mathbf{B}^{(i)}$ is then updated. The vector $\mathbf{f}^{(i+1)}$ satisfying (28) is either obtained through single-point extraction or through recursive multipoint extraction [3] that approaches a limit satisfying (28). We denote by \mathbf{x}'_{em} and \mathbf{R}'_{em} the solution obtained at the end of the first phase and the corresponding fine-model response, respectively.

Switching to the second phase takes place in two cases. The first case is that (29) is not satisfied, which means that we have to reject the new point $\mathbf{x}_{em}^{(i+1)}$. The Jacobian of the fine-model response at the point $\mathbf{x}_{em}^{(i)}$ is then evaluated. This is done by first evaluating the Jacobian of the coarse-model response $\mathbf{J}_{os}^{(i)}$ at the previously extracted coarse-model point $\mathbf{x}_{os}^{(i)} = \mathbf{P}(\mathbf{x}_{em}^{(i)})$. $\mathbf{J}_{em}^{(i)}$ is then approximated using (18). The second phase is then supplied by $\mathbf{x}_{em}^{(i)}$, $\mathbf{J}_{em}^{(i)}$, and $\mathbf{f}^{(i)}$.

The second case occurs when the new point $\mathbf{x}_{em}^{(i+1)}$ satisfies (29), but does not satisfy (28). In this case, the point $\mathbf{x}_{em}^{(i+1)}$ is better than the previous point $\mathbf{x}_{em}^{(i)}$ and is accepted. As the vector of extracted parameters does not satisfy (28), the vector $\mathbf{f}^{(i+1)}$ can not be trusted. In order to trust this vector, recursive multipoint parameter extraction is applied at the point $\mathbf{x}_{em}^{(i+1)}$ until either $\mathbf{f}^{(i+1)}$ approaches a limiting value or the number of additional points used for multipoint parameter extraction reaches n . If $\mathbf{f}^{(i+1)}$ approaches a limit that does not satisfy (28), $\mathbf{B}^{(i)}$ is updated to $\mathbf{B}^{(i+1)}$, $\mathbf{J}_{os}^{(i+1)}$ at the extracted coarse-model point $\mathbf{x}_{os}^{(i+1)} = \mathbf{P}(\mathbf{x}_{em}^{(i+1)})$ is evaluated and $\mathbf{J}_{em}^{(i+1)}$ is then approximated using (18). Otherwise, $\mathbf{J}_{em}^{(i+1)}$ is approximated using the $n + 1$ fine-model points used for multipoint parameter extraction. The second phase is then supplied by the point $\mathbf{x}_{em}^{(i+1)}$, $\mathbf{f}^{(i+1)}$, and the Jacobian estimate $\mathbf{J}_{em}^{(i+1)}$, which is either calculated using (18) or through finite differences.

The second phase utilizes the first-order derivatives supplied by SM to carry out a number of successful iterations. By a successful iteration, we mean an iteration that satisfies the success criterion

$$\left(\left\| \mathbf{g}^{(k)} \right\| - \left\| \mathbf{g}^{(k+1)} \right\| \right) > 0.01 \left(\left\| \mathbf{g}^{(k)} \right\| - \left\| \mathbf{g}^{(k)} + \mathbf{J}_{em}^{(k)} \mathbf{h}^{(k)} \right\| \right) \tag{30}$$

which indicates that the actual reduction in the objective function (27) should be greater than a certain fraction of the predicted reduction. Notice that the superscript k is used as an index for the successful iterates of the direct optimization phase. At the end of each successful iteration, parameter extraction is applied at the new iterate $\mathbf{x}_{em}^{(k)}$ and is used to check whether a switch back to the first phase can take place. The criterion for such a switch is

$$\left\| \mathbf{f}^{(k+1)} \right\| < \left\| \mathbf{f}^{(k)} \right\|. \tag{31}$$

If it is satisfied, $\mathbf{J}_{os}^{(k+1)}$ is evaluated at the point $\mathbf{x}_{os}^{(k+1)} = \mathbf{P}(\mathbf{x}_{em}^{(k+1)})$, \mathbf{B} is reevaluated using (19), and the algorithm switches back to the first phase. Otherwise, the second phase continues. We denote by \mathbf{x}''_{em} and \mathbf{R}''_{em} the solution obtained at

the end of the second phase and the corresponding fine-model response, respectively.

For any iteration $i \geq 0$, the two phases are given by the following steps.

Phase 1:

Step 0: Given $\mathbf{x}_{em}^{(i)}$, $\mathbf{f}^{(i)}$, $\mathbf{B}^{(i)}$, and $\delta^{(i)}$. Set $\delta^{(i+1)} = \delta^{(i)}$.

Comment: $\delta^{(i)}$ is the utilized trust region size.

Step 1: Obtain $\mathbf{h}^{(i)}$ by solving (7) with $\delta = \delta^{(i+1)}$. Let $\delta^{(i+1)} = \|\mathbf{h}^{(i)}\|_2$.

Step 2: Evaluate $\mathbf{x}_{em}^{(i+1)}$ using (5) and set $V = \{\mathbf{x}_{em}^{(i+1)}\}$.

Comment: V is the set of fine-model points utilized in the multipoint extraction.

Step 3: Apply multipoint parameter extraction using the points in the set V to obtain $\mathbf{f}^{(i+1)}$.

Comment: The prediction given in (25) is used as a starting point for the multipoint parameter extraction.

Step 4: If both (28) and (29) are satisfied, update the matrix $\mathbf{B}^{(i)}$ to $\mathbf{B}^{(i+1)}$ using Broyden's formula [6] and update δ . Go to Step 10.

Comment: The trust region size δ is updated based on how the predicted reduction in $\|\mathbf{f}\|_2$ agrees with the actual reduction [3].

Step 5: If (29) is not satisfied, obtain $\mathbf{J}_{os}^{(i)}$ and evaluate $\mathbf{J}_{em}^{(i)} = \mathbf{J}_{os}^{(i)} \mathbf{B}^{(i)}$. Switch to the second phase.

Comment: The second phase takes as arguments $\mathbf{x}_{em}^{(i)}$, $\mathbf{f}^{(i)}$, and $\mathbf{J}_{em}^{(i)}$ and returns $\mathbf{x}_{em}^{(k)}$, $\mathbf{B}^{(k)}$, and $\mathbf{f}^{(k)}$. It should be clear that several iterations might be executed in the second phase before switching back to the first phase at Step 10.

Step 6: If $|V|$ is equal to one, go to Step 9.

Comment: $|V|$ denotes the cardinality of the set V .

Step 7: Compare $\mathbf{f}^{(i+1)}$ obtained using $|V|$ fine-model points with that previously obtained using $|V| - 1$ fine-model points. If $\mathbf{f}^{(i+1)}$ is approaching a limit, update the matrix $\mathbf{B}^{(i)}$ to get $\mathbf{B}^{(i+1)}$, obtain $\mathbf{J}_{os}^{(i+1)}$, evaluate $\mathbf{J}_{em}^{(i+1)} = \mathbf{J}_{os}^{(i+1)} \mathbf{B}^{(i+1)}$ and switch to the second phase.

Step 8: If $|V|$ is equal to $n+1$, obtain the matrix $\mathbf{J}_{em}^{(i+1)}$ by finite differences using the set V . Switch to the second phase.

Step 9: Obtain a temporary point $\mathbf{x}_t = \mathbf{x}_{em}^{(i+1)} + \mathbf{h}_t$, where

$$\left(\mathbf{B}^{(i)T} \mathbf{B}^{(i)} + \lambda \mathbf{I} \right) \mathbf{h}_t = -\mathbf{B}^{(i)T} \mathbf{f}^{(i+1)}$$

and $\|\mathbf{h}_t\| \leq \delta^{(i+1)}$. Add this point to the set V and go to Step 3.

Step 10: Let $i = i + 1$. Go to Step 0.

The second phase can be summarized in the following steps.

Phase 2:

Step 0: Given the current iterate of the SM technique $\mathbf{x}_{em}^{(k)}$, the corresponding Jacobian matrix $\mathbf{J}_{em}^{(k)}$ and $\mathbf{f}^{(k)}$.

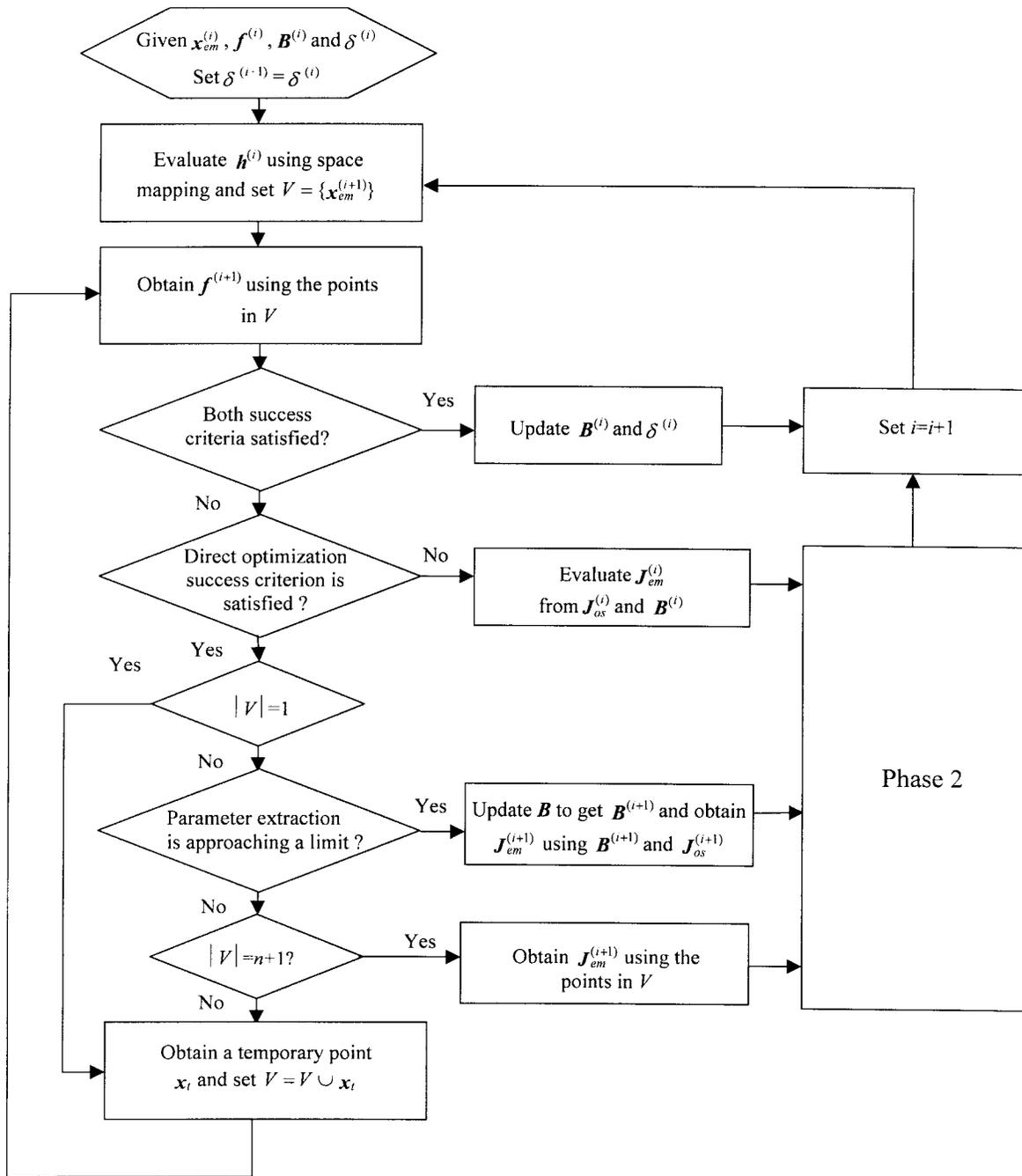


Fig. 4. A flowchart of the first phase of the proposed algorithm.

Step 1: Obtain a successful iterate $\mathbf{x}_{em}^{(k+1)}$ by solving

$$\left(\mathbf{J}_{em}^{(k)T} \mathbf{J}_{em}^{(k)} + \lambda \mathbf{I} \right) \Delta \mathbf{x} = -\mathbf{J}_{em}^{(k)T} \mathbf{g}^{(k)}$$

for a suitable value of λ that satisfies the direct optimization success criterion.

Step 2: Update $\mathbf{J}_{em}^{(k)}$ to $\mathbf{J}_{em}^{(k+1)}$.

Step 3: Apply parameter extraction at $\mathbf{x}_{em}^{(k+1)}$ to get $\mathbf{f}^{(k+1)}$.

Step 4: If (31) is satisfied, obtain $\mathbf{J}_{os}^{(k+1)}$ at the point $\mathbf{x}_{os}^{(k+1)} = \mathbf{P}(\mathbf{x}_{em}^{(k+1)})$, evaluate the matrix $\mathbf{B} = (\mathbf{J}_{os}^{(k+1)T} \mathbf{J}_{os}^{(k+1)} - 1) \mathbf{J}_{os}^{(k+1)T} \mathbf{J}_{em}^{(k+1)}$ and switch to the first phase.

Step 5: If the termination condition is satisfied invoke the minimax optimizer else set $k = k + 1$ and go to Step 1.

A flowchart of the first phase of the HASM algorithm is shown in Fig. 4.

To ensure optimality, direct optimization is used to solve the original design problem using a minimax optimizer [8] starting from \mathbf{x}_{em}'' . The current implementation of the HASM algorithm is in MATLAB.¹

¹MATLAB Version 5.0, The Math Works Inc., Natick, MA, 1997.

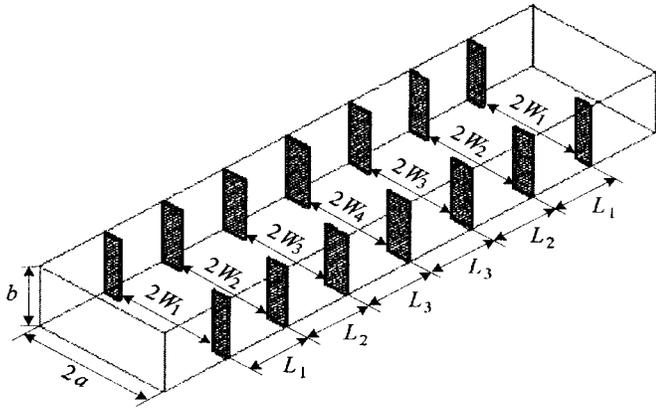


Fig. 5. The fine-model of the six-section H -plane waveguide filter [9], [10].

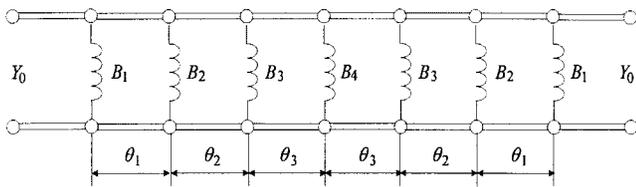


Fig. 6. The coarse model of the six-section H -plane waveguide filter [11].

VII. EXAMPLES

A. Six-Section H -Plane Waveguide Filter

We consider a six-section H -plane waveguide filter [9], [10]. Design specifications are taken as

$$|S_{11}| \leq 0.16, \quad \text{for } 5.4 \text{ GHz} \leq f \leq 9.0 \text{ GHz} \quad (32)$$

$$|S_{11}| \geq 0.85, \quad \text{for } f \leq 5.2 \text{ GHz and } |S_{11}| \geq 0.5 \\ \text{for } 9.5 \text{ GHz} \leq f. \quad (33)$$

A waveguide with a cross section of 1.372 by 0.622 in (3.485 by 1.58 cm) is used. The six sections are separated by seven H -plane septa, which have a finite thickness of 0.02 in (0.508 mm). The filter is shown in Fig. 5.

The optimizable parameters are the four septa widths W_1 , W_2 , W_3 , and W_4 and the three waveguide-section lengths L_1 , L_2 , and L_3 . The coarse model consists of lumped inductances and dispersive transmission-line sections. It is simulated using OSA90/hope.² There are various approaches to calculate the equivalent inductive susceptance corresponding to an H -plane septum. We utilize a simplified version of a formula due to Marcuvitz [11] in evaluating the inductances. The coarse model is shown in Fig. 6. The fine model exploits HP HFSS³ through HP Empipe3D.⁴

The fine-model response at the starting point x_{os}^* is shown in Fig. 7. The first phase required four iterations to reach the design x_{em}' . A total of five fine-model simulations were needed. The second phase did not carry out any successful iteration. The response R_{em}'' is shown in Fig. 8.

²OSA90/hope Version 4.0, formerly Optimization Systems Associates Inc., Dundas, Ont., Canada, now HP EEsof Division, Santa Rosa, CA.

³HP HFSS Version 5.2, HP EEsof Division, Santa Rosa, CA, 1998.

⁴HP Empipe3D Version 5.2, HP EEsof Division, Santa Rosa, CA, 1998.

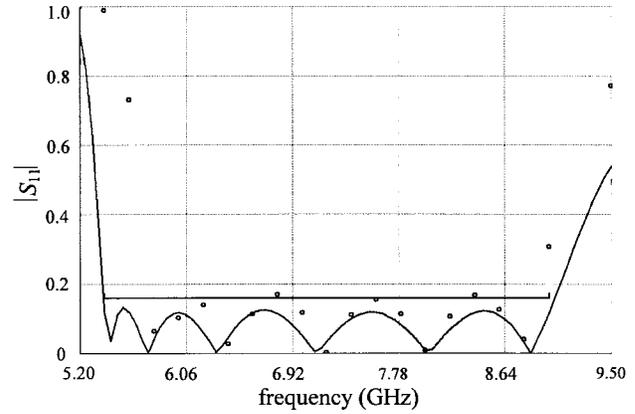


Fig. 7. The coarse response R_{os}^* (—) and the fine response $R_{em}(x_{os}^*)$ (o) for the six-section H -plane waveguide filter.

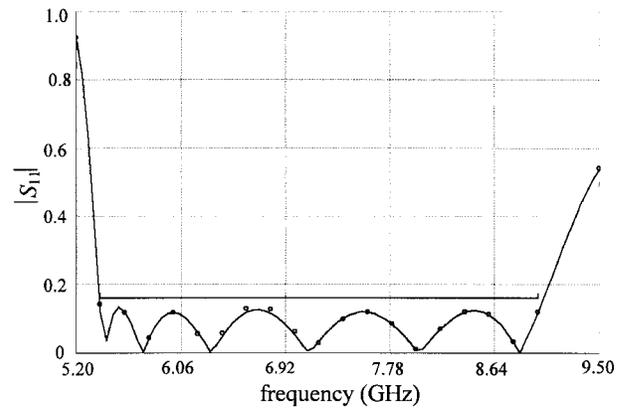


Fig. 8. The coarse response R_{os}^* (—) and the fine response R_{em}'' (o) for the six-section H -plane waveguide filter.

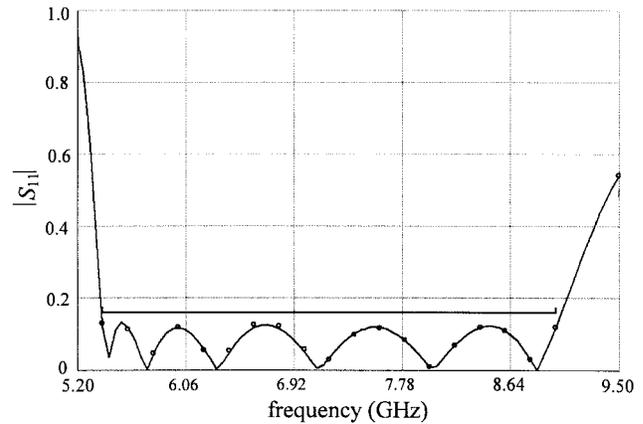


Fig. 9. The coarse response R_{os}^* (—) and the fine response R_{em}^* (o) for the six-section H -plane waveguide filter.

The response R_{em}^* is obtained through direct minimax optimization (see Fig. 9). The different fine-model designs are given in Table I. It is clear that the convergence of the first phase is smooth as $R_{em}' \approx R_{em}^* \approx R_{os}^*$.

B. Seven-Section Waveguide Transformer

The design of a seven-section waveguide transformer is also considered. The transformer is shown in Fig. 10. This example is a classical microwave circuit problem [12]. The fine

TABLE I
DESIGNS OBTAINED DURING DIFFERENT PHASES FOR THE SIX-SECTION
H-PLANE WAVEGUIDE FILTER

Parameter	x_{os}^*	x'_{em}, x''_{em}	x_{em}^*
W_1	0.48583	0.51326	0.51344
W_2	0.43494	0.47379	0.47396
W_3	0.40433	0.45091	0.45100
W_4	0.39796	0.44675	0.44664
L_1	0.65585	0.63701	0.63695
L_2	0.65923	0.63954	0.63977
L_3	0.67666	0.65704	0.65694

All values are in inches

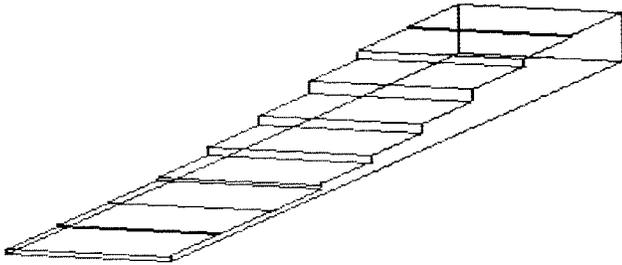


Fig. 10. The seven-section waveguide transformer [12].

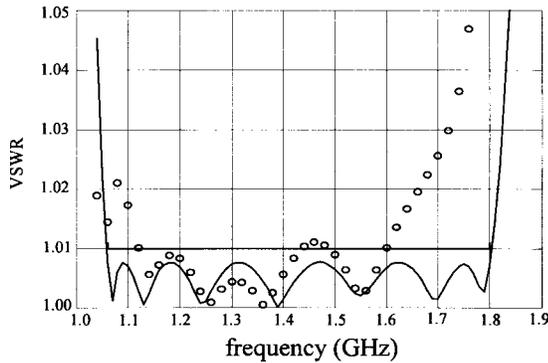


Fig. 11. The coarse response R_{os}^* (—) and the fine response $R_{em}(x_{os}^*)$ (o) for the seven-section waveguide transformer.

model is simulated using HP HFSS through HP Empire3D. The coarse model is an analytical model that neglects the junction discontinuity [12]. The design specifications are taken as

$$VSWR \leq 1.01, \quad \text{for } 1.06 \text{ GHz} \leq f \leq 1.8 \text{ GHz}. \quad (34)$$

The designable parameters for this problem are the height and length of each waveguide section. The fine-model response at x_{os}^* is shown in Fig. 11. The first phase executed three successful iterations that required six fine-model simulations. The response R'_{em} is shown in Fig. 12. The second phase executed four iterations (see Fig. 13). The response R^*_{em} is shown in Fig. 14. Table II shows the different designs.

C. Double-Folded Stub Filter

We consider the design of the double-folded stub (DFS) microstrip structure [1]. Folding the stubs reduces the filter area with respect to the conventional double stub structure

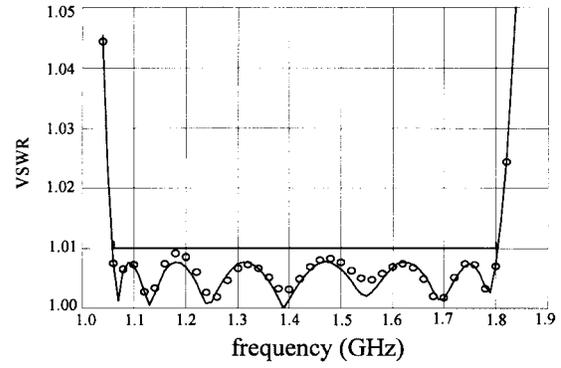


Fig. 12. The coarse response R_{os}^* (—) and the fine response R'_{em} (o) for the seven-section waveguide transformer.

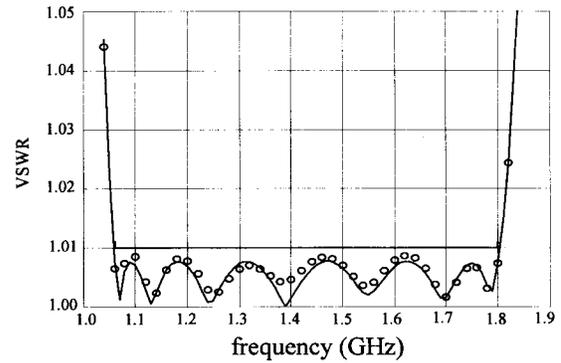


Fig. 13. The coarse response R_{os}^* (—) and the fine response R''_{em} (o) for the seven-section waveguide transformer.

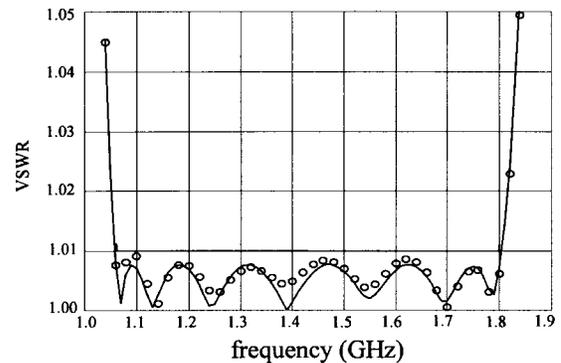


Fig. 14. The coarse response R_{os}^* (—) and the fine response R^*_{em} (o) for the seven-section waveguide transformer.

[13]. The filter is characterized by five parameters: $W_1, W_2, S, L_1,$ and L_2 (see Fig. 15). $L_1, L_2,$ and S are chosen as optimization variables. W_1 and W_2 are fixed at 4.8 mil. The design specifications are

$$|S_{21}| \geq -3 \text{ dB}, \quad \text{for } f \leq 9.5 \text{ GHz and } 16.5 \text{ GHz} \leq f \quad (35)$$

$$|S_{21}| \leq -30 \text{ dB}, \quad \text{for } 12 \text{ GHz} \leq f \leq 14 \text{ GHz}. \quad (36)$$

The fine model is the structure simulated by HP HFSS through HP Empire3D. The coarse model exploits the microstrip line and microstrip T-junction models available in OSA90/hope. The coupling between the folded stubs and

TABLE II
DESIGNS OBTAINED DURING DIFFERENT PHASES FOR THE SEVEN-SECTION
WAVEGUIDE TRANSFORMER

Parameter	x_{os}^*	x'_{em}	x''_{em}	x^*_{em}
b_1	7.86732	7.84126	7.84321	7.84319
b_2	6.61888	6.56661	6.56753	6.56746
b_3	4.68540	4.63369	4.63275	4.63267
b_4	2.91987	2.88266	2.88266	2.88268
b_5	1.81412	1.79307	1.79273	1.79272
b_6	1.27658	1.26697	1.26721	1.26723
b_7	1.06847	1.06475	1.06477	1.06474
L_1	7.10588	7.27059	7.27141	7.27145
L_2	7.12201	7.03866	7.04043	7.04047
L_3	7.11760	6.89568	6.89549	6.89552
L_4	7.12331	6.89253	6.89192	6.89189
L_5	7.12815	6.98273	6.97985	6.98000
L_6	7.12154	7.03160	7.03020	7.03023
L_7	7.12945	7.02606	7.02503	7.02509

All values are in cm

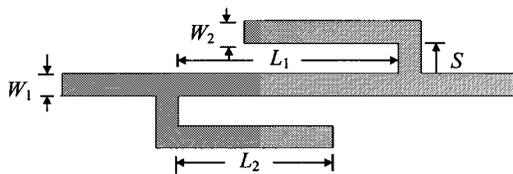


Fig. 15. The DFS filter [1], [13].

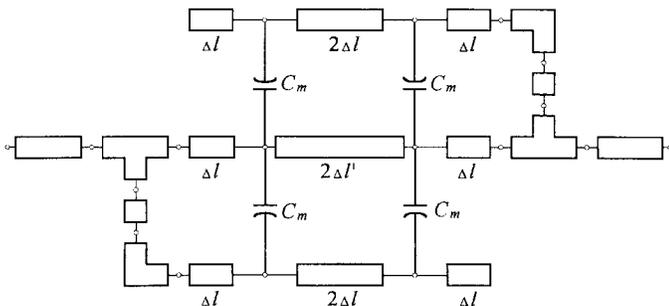


Fig. 16. The coarse model of the DFS filter.

the microstrip line is simulated using equivalent capacitors. The values of these capacitors are determined using Walker's formulas [14]. Jansen's microstrip bend model [15] is used to model the folding effect of the stub. The coarse model is shown in Fig. 16.

The fine-model response at x_{os}^* is shown in Fig. 17. This figure shows a big shift between the optimal coarse response and the initial fine response. This signals considerable misalignment between the two models.

The first phase successfully carried out eight iterations that required 12 fine-model simulations. The response R'_{em} is shown in Fig. 18. The mapping established in the first phase is utilized to get a good estimate of J_{em} and a switch to the second phase is carried out. The response R''_{em} (Fig. 19) shows

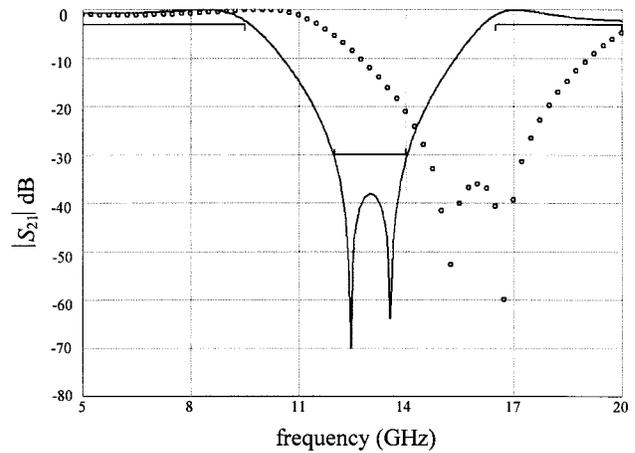


Fig. 17. The coarse response R_{os}^* (—) and the fine response $R_{em}(x_{os}^*)$ (o) for the DFS filter.

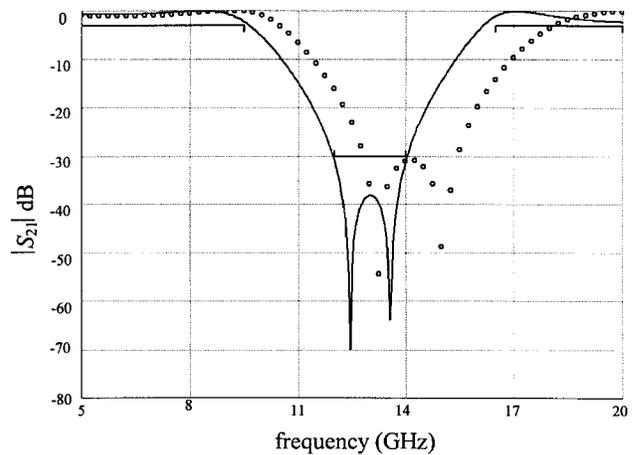


Fig. 18. The coarse response R_{os}^* (—) and the fine response R'_{em} (o) for the DFS filter.

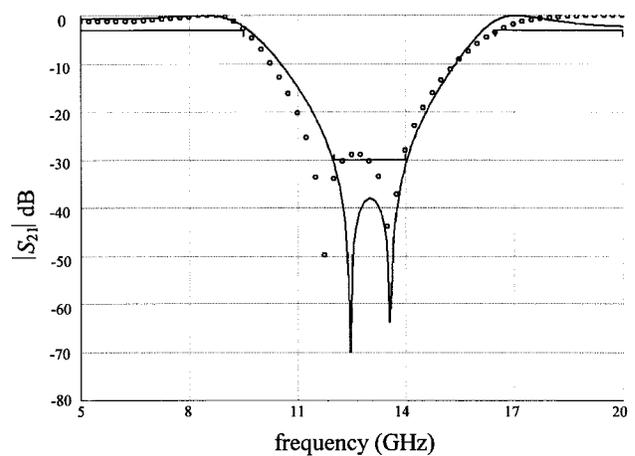


Fig. 19. The coarse response R_{os}^* (—) and the fine response R''_{em} (o) for the DFS filter.

a significant improvement in the response. The design x''_{em} is then taken as the starting point for the minimax optimizer. The response R^*_{em} is shown in Fig. 20. The designs are given in Table III.

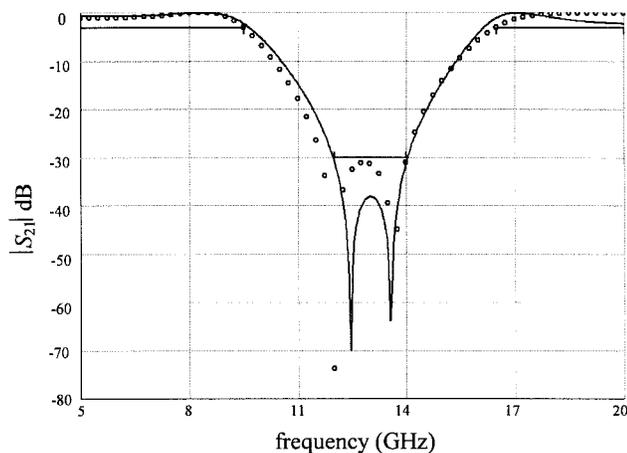


Fig. 20. The coarse response R_{os}^* (—) and the fine response R_{em}^* (o) for the DFS filter.

TABLE III
DESIGNS OBTAINED DURING DIFFERENT PHASES FOR THE DFS FILTER

Parameter	x_{os}^*	x'_{em}	x''_{em}	x_{em}^*
L_1	66.727	72.454	73.869	78.964
L_2	60.228	72.728	82.939	81.210
S	9.592	7.621	8.170	7.901

All values are in mil

VIII. CONCLUSIONS

We present a novel HASM optimization algorithm. The algorithm is designed to handle severely misaligned cases. It enables smooth switching from SM optimization to direct optimization if SM fails. The direct optimization phase utilizes all the available information accumulated by SM optimization about the mapping between the coarse and fine-model spaces. The algorithm also enables switching back from direct optimization to SM if SM is potentially convergent. The connection between SM and direct optimization is based on a novel lemma. An original approach for the prediction of the starting point of the parameter-extraction optimization problem is also utilized. This approach improves the uniqueness of the extraction step and, consequently, enhances the convergence of the algorithm. The algorithm is successfully demonstrated through the design of waveguide transformers and filters.

ACKNOWLEDGMENT

The authors would like to thank HP EEsof, Santa Rosa, CA, for making HP HFSS and HP Empipe3D available for this work.

REFERENCES

- [1] J. W. Bandler, R. M. Biernacki, S. H. Chen, P. A. Grobely, and R. H. Hemmers, "Space mapping technique for electromagnetic optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 42, pp. 2536–2544, Dec. 1994.
- [2] J. W. Bandler, R. M. Biernacki, and S. H. Chen, "Fully automated space mapping optimization of 3D structures," in *IEEE MTT-S Int. Microwave Symp. Dig.*, San Francisco, CA, 1996, pp. 753–756.

- [3] M. H. Bakr, J. W. Bandler, R. M. Biernacki, S. H. Chen, and K. Madsen, "A trust region aggressive space mapping algorithm for EM optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 46, pp. 2412–2425, Dec. 1998.
- [4] J. W. Bandler, R. M. Biernacki, S. H. Chen, R. H. Hemmers, and K. Madsen, "Electromagnetic optimization exploiting aggressive space mapping," *IEEE Trans. Microwave Theory Tech.*, vol. 43, pp. 2874–2882, Dec. 1995.
- [5] J. J. Moré and D. C. Sorenson, "Computing a trust region step," *SIAM J. Sci. Stat. Comput.*, vol. 4, pp. 553–572, 1983.
- [6] C. G. Broyden, "A class of methods for solving nonlinear simultaneous equations," *Math. Comput.*, vol. 19, pp. 577–593, 1965.
- [7] R. Fletcher, *Practical Methods of Optimization*, 2nd ed. New York: Wiley, 1987.
- [8] 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, Dec. 1985.
- [9] L. Young and B. M. Schiffman, "A useful high-pass filter design," *Microwave J.*, vol. 6, pp. 78–80, 1963.
- [10] G. L. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance-Matching Network and Coupling Structures*, 1st ed. New York: McGraw-Hill, 1964.
- [11] N. Marcuvitz, *Waveguide Handbook*, 1st ed. New York: McGraw-Hill, 1951, p. 221.
- [12] J. W. Bandler, "Computer optimization of inhomogeneous waveguide transformers," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 563–571, Aug. 1969.
- [13] J. C. Rautio, private communication, 1992.
- [14] C. S. Walker, *Capacitance, Inductance and Crosstalk Analysis*. Norwood, MA: Artech House, 1990.
- [15] M. Kirschning, R. Jansen, and N. Koster, "Measurement and computer-aided modeling of microstrip discontinuities by an improved resonator method," in *IEEE MTT-S Int. Microwave Symp. Dig.*, Boston, MA, 1983, pp. 495–497.

Mohamed H. Bakr (S'98), for photograph and biography, see this issue, p. 2439.

John W. Bandler (S'66–M'66–SM'74–F'78), for photograph and biography, see this issue, p. 2426.

Natalia Georgieva (S'93–M'97), for photograph and biography, see this issue, p. 2439.



Kaj Madsen was born in Denmark, in 1943. He received the cand.scient. degree in mathematics from the University of Aarhus, Aarhus, Denmark, in 1968, and the dr.techn. degree from the Technical University of Denmark, Lyngby, Denmark, in 1986.

From 1968 to 1988, he was a Lecturer in Numerical Analysis, apart from the academic year 1973 to 1974, when he was with AERE Harwell, Didcot, U.K. Most of his career has been spent with the Institute for Numerical Analysis, Technical University of Denmark, but from 1981 to 1983, he was with the Computer Science Department, Copenhagen University. During the summer of 1978, he visited McMaster University, Hamilton, Ont., Canada. In 1988, he became a Professor with the Institute for Numerical Analysis, Technical University of Denmark. In 1993, he joined the Department of Mathematical Modeling, Technical University of Denmark. From 1988 to 1993, and since 1995, he has been Head of Department. His fields of interest in teaching and research are parallel algorithms, optimization, and interval analysis.