AN AGGRESSIVE APPROACH TO PARAMETER EXTRACTION

M.H. Bakr, J.W. Bandler and N. Georgieva

SOS-99-10-R

March 1999

© M.H. Bakr, J.W. Bandler and N. Georgieva 1999

No part of this document may be copied, translated, transcribed or entered in any form into any machine without written permission. Address inquiries in this regard to Dr. J.W. Bandler. Excerpts may be quoted for scholarly purposes with full acknowledgment of source. This document may not be lent or circulated without this title page and its original cover.

AN AGGRESSIVE APPROACH TO PARAMETER EXTRACTION

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

Abstract A novel Aggressive Parameter Extraction (APE) algorithm is presented. Our APE algorithm addresses the optimal selection of parameter perturbations used to increase trust in parameter extraction uniqueness. The uniqueness of the parameter extraction problem is crucial especially in the Space Mapping (SM) approach to circuit design. We establish an appropriate criterion for the generation of these perturbations. The APE algorithm classifies possible solutions for the parameter extraction problem. Two different approaches for obtaining subsequent perturbations are utilized based on a classification of the extracted parameters. The algorithm is demonstrated through parameter extraction of microwave filters and transformers. The examples include the parameter extraction of a decomposed electromagnetic model of an HTS filter. Also, the parameter extraction of an empirical model of a DFS filter is carried out.

I. INTRODUCTION

Parameter extraction is important in device modeling and characterization. It also plays a crucial role in Space Mapping (SM) technology [1–3]. Optimization approaches to parameter extraction often yield nonunique solutions. In SM optimization this nonuniqueness may lead to divergence or oscillatory behavior.

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Grants OGP0007239, STP0201832 and through the Micronet Network of Centres of Excellence. N. Georgieva is supported by an NSERC Postdoctorate Fellowship and M.H. Bakr by an Ontario Graduate Scholarship.

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

J.W. Bandler is also with Bandler Corporation, P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7.

We present an "aggressive" approach to parameter extraction. While generally applicable, the new algorithm is discussed here in the context of SM technology. We assume the existence of a "fine" model that generates the target response and a "coarse" model whose parameters are to be extracted.

Several authors have addressed nonuniqueness in parameter extraction. For example, Bandler et *al.* [4] proposed the idea of making unknown perturbations to a certain system whose parameters are to be extracted. Later Bandler *et al.* [5] suggested that Multi-Point Extraction (MPE) be used to match the first-order derivatives of the two models to ensure a global minimum. The perturbations used in that approach are predefined and arbitrary. The optimality of the selection of those perturbations was not addressed. Recently, a recursive MPE technique was suggested by Bakr *et al.* [3]. This approach employs a mapping between the two models to enhance uniqueness.

Our algorithm aims at minimizing the number of perturbations used in the MPE process by utilizing perturbations that significantly improve the uniqueness in each iteration. Consequently, we designate this as an Aggressive Parameter Extraction (APE) algorithm. Each perturbation requires an additional fine model simulation which could be very CPU intensive. We classify the different solutions returned by the MPE process and, based on this classification, a new perturbation that is likely to sharpen the result is suggested.

II. PARAMETER EXTRACTION

The objective of parameter extraction is to find a set of parameters of a model whose response matches a given set of measurements. It can be formulated as

$$\mathbf{x}_{os}^{e} = \arg \left\{ \min_{\mathbf{x}_{os}} \left\| \mathbf{R}_{m} - \mathbf{R}_{os}(\mathbf{x}_{os}) \right\| \right\}$$
(1)

where \mathbf{R}_m is the vector of given measurements, \mathbf{R}_{os} is the vector of circuit response and \mathbf{x}_{os}^e is the vector of extracted parameters. In the context of SM the fine model response \mathbf{R}_{em} , typically from an

electromagnetic simulator, at a certain point x_{em} supplies the target response R_m . Fig. 1 illustrates the Single Point Extraction (SPE) for the two dimensional case. An MPE procedure [5] was suggested to improve the uniqueness of the step. The vector of extracted coarse model parameters x_{os}^{e} should satisfy

$$\mathbf{x}_{os}^{e} = \arg \left\{ \min_{\mathbf{x}_{os}} \left\| \begin{bmatrix} \mathbf{e}_{0}^{T} & \mathbf{e}_{1}^{T} & \cdots & \mathbf{e}_{N_{p}}^{T} \end{bmatrix}^{T} \right\| \right\}$$
(2)

where

$$\boldsymbol{e}_{0} = \boldsymbol{R}_{os}(\boldsymbol{x}_{os}) - \boldsymbol{R}_{em}(\boldsymbol{x}_{em})$$
(3)

and

$$\boldsymbol{e}_{i} = \boldsymbol{R}_{os}(\boldsymbol{x}_{os} + \Delta \boldsymbol{x}_{os}^{(i)}) - \boldsymbol{R}_{em}(\boldsymbol{x}_{em} + \Delta \boldsymbol{x}_{em}^{(i)})$$

$$\tag{4}$$

The set of perturbations in the coarse model space is represented by $\Delta \mathbf{x}_{os}^{(i)} \in V_p$, where $i=1, 2, ..., N_p$ and $|V_p| = N_p$. $\Delta \mathbf{x}_{em}^{(i)}$ is the corresponding perturbation in the fine model space. The perturbations $\Delta \mathbf{x}_{os}^{(i)}$ and $\Delta \mathbf{x}_{em}^{(i)}$ in this MPE procedure are related by

$$\Delta \boldsymbol{x}_{os}^{(i)} = \Delta \boldsymbol{x}_{em}^{(i)} \tag{5}$$

It follows that its solution simultaneously matches the responses of a set of corresponding points in both spaces. The number of fine model points needed for this process is arbitrary. There is no clear way of how to select the set of perturbations. Also, the available information about the mapping between the two spaces was not utilized.

Bakr et al. [3] suggested that the perturbations utilized in (2)-(4) should satisfy

$$\Delta \boldsymbol{x}_{os}^{(i)} = \boldsymbol{B} \Delta \boldsymbol{x}_{em}^{(i)} \tag{6}$$

The matrix B approximates the mapping between the two spaces. In the context of space mapping (6) is superior to (5) as it integrates the available mapping B into the MPE procedure. It is also suggested in [3] that the parameter extraction step terminates if the vector of extracted parameters approaches a limit. The algorithm generates the perturbations used for the MPE process. The perturbations used in [3] are not guaranteed to result in significant improvement in the uniqueness of the extracted parameters. A large number of additional fine model simulations may be needed to ensure the uniqueness of the step.

For both (5) and (6) the set V of fine model points utilized in MPE is

$$V = \left\{ \boldsymbol{x}_{em} \right\} \cup \left\{ \boldsymbol{x}_{em} + \Delta \boldsymbol{x}_{em}^{(i)} \middle| \forall \Delta \boldsymbol{x}_{os}^{(i)} \in V_p \right\}$$
(7)

Fig. 2 illustrates the MPE procedure.

III. THE SELECTION OF PERTURBATIONS

The vector of coarse model responses \boldsymbol{R} used to match the two models is given by

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{R}_{os}(\boldsymbol{x}_{os}) \\ \boldsymbol{R}_{os}(\boldsymbol{x}_{os} + \Delta \boldsymbol{x}_{os}^{(1)}) \\ \vdots \\ \boldsymbol{R}_{os}(\boldsymbol{x}_{os} + \Delta \boldsymbol{x}_{os}^{(Np)}) \end{bmatrix}$$
(8)

The dimensionality of \mathbf{R} is m_p , where $m_p = (N_p + 1)m$ and m is the dimensionality of both \mathbf{R}_{os} and \mathbf{R}_{em} . Vector \mathbf{x}_{os}^e is labeled locally unique [6] if there exists an open neighborhood of \mathbf{x}_{os}^e containing no other point \mathbf{x}_{os} such that $\mathbf{R}(\mathbf{x}_{os}) = \mathbf{R}(\mathbf{x}_{os}^e)$. Otherwise, it is labeled locally nonunique. It was shown in [6] that the local uniqueness condition is equivalent to the condition that the Jacobian of \mathbf{R} has rank n where n is the number of parameters.

To achieve local uniqueness, it was suggested in the context of system identification [4] that increasing the number of perturbations enhances the possibility that the Jacobian matrix J of R becomes full rank. The perturbations suggested by Daijavad [4] were unidentified perturbations and thus result in an increase in the number of the optimizable parameters. However, it was pointed out that the improvement in the rank of J outweighs the increase in the number of parameters.

In a later work, the idea of using known perturbations to achieve global uniqueness of parameter extraction was introduced [5]. By global uniqueness we mean that there exists only one minimum \mathbf{x}_{os}^{e}

for the MPE problem. The existence of more than one locally unique solution for the parameter extraction procedure may result in the divergence or oscillation of the SM optimization algorithm. It was also pointed out in [5] that using MPE is equivalent to matching the first-order derivatives of the coarse and fine models.

We suggest two different types of perturbation depending on whether the solution of the MPE is locally unique or locally nonunique. If the solution obtained is locally nonunique we choose a perturbation that is likely to make the new extracted parameters locally unique.

Assume that a locally nonunique minimum \mathbf{x}_{os}^{e} is obtained using the current set of coarse model perturbations V_{p} . Here, the rank of the Jacobian J of \mathbf{R} is k where k < n and n is the number of parameters. We suggest a perturbation $\Delta \mathbf{x}$ that attempts to increase the rank of the Jacobian of the responses corresponding to the augmented set $\{V_p \cup \Delta \mathbf{x}\}$ at \mathbf{x}_{os}^{e} by at least one. This is achieved by imposing the condition that the gradients of n-k of the coarse model responses in the new response vector $\mathbf{R}_{os}(\mathbf{x}_{os}^{e} + \Delta \mathbf{x})$ be normal to a linearly independent set of gradients of cardinality k of the responses in the vector \mathbf{R} at the point \mathbf{x}_{os}^{e} . We denote the set of linearly independent gradients by S where

$$S = \left\{ \boldsymbol{g}^{(1)}, \dots, \boldsymbol{g}^{(k)} \right\}$$
(9)

We denote the set of the gradients of the newly selected *n*-*k* responses in $\mathbf{R}_{os}(\mathbf{x}_{os}^{e} + \Delta \mathbf{x})$ by S_{a} , where

$$S_a = \left\{ \boldsymbol{g}_a^{(k+1)}, \dots, \boldsymbol{g}_a^n \right\}$$
(10)

Each of these gradients is approximated by

$$g_a^{(i)} = g^{(i)} + G^{(i)} \Delta x, \quad i = k+1, \dots, n$$
 (11)

where $g^{(i)}$ is the gradient of the *i*th response at the point x_{os}^{e} and $G^{(i)}$ is the corresponding Hessian. The imposed condition on the perturbation is that

$$\boldsymbol{g}_{a}^{(i)T} \boldsymbol{g}^{(j)} = 0 \quad \forall \, \boldsymbol{g}^{(j)} \in S \text{ and } \forall \, \boldsymbol{g}_{a}^{(i)} \in S_{a}$$
 (12)

Using (11) and (12), the perturbation Δx is obtained by solving the system of linear equations

$$\boldsymbol{A}^{T}\Delta\boldsymbol{x} = -\boldsymbol{c} \tag{13}$$

where the matrix A is given by

$$A = [G^{(k+1)}g^{(1)} \cdots G^{(n)}g^{(1)} \cdots G^{(n)}g^{(k)}]$$
(14)

and the vector \boldsymbol{c} is given by

$$\boldsymbol{c} = \begin{bmatrix} \boldsymbol{g}^{(k+1)T} \, \boldsymbol{g}^{(1)} \\ \vdots \\ \boldsymbol{g}^{(n)T} \, \boldsymbol{g}^{(1)} \\ \vdots \\ \boldsymbol{g}^{(n)T} \, \boldsymbol{g}^{(k)} \end{bmatrix}$$
(15)

A complete derivation of (13) is given in Appendix A. It should be noted that the system of linear equations (13) may be an over-determined, under-determined or well-determined system of equations depending on *k* and *n*. The pseudoinverse of the matrix A^T obtains the solution with minimum ℓ_2 norm in all cases. The fact that this solution is a minimum length solution is of importance since (13) is based on a linear approximation of the gradients which can only be trusted within a certain trust region. If the perturbation $\Delta \mathbf{x}$ is outside this trust region, it is rescaled.

If the minimum obtained by the MPE is locally unique we still have to ensure that this is the true solution to the extraction problem. The following lemma leads to a robust way to weaken any other existing locally unique minimum.

Lemma

Assume that there exist two locally unique minima $\mathbf{x}_{os}^{e,1}$ and $\mathbf{x}_{os}^{e,2}$ for the MPE problem obtained using least squares optimization and a set of perturbations V_p . A possible perturbation $\Delta \mathbf{x}$ that can be added to the set V_p and can be used to weaken one of these minima as a solution for the MPE is in the direction of an eigenvector for the matrix $\mathbf{H}_1 - \mathbf{H}_2$ where \mathbf{H}_1 and \mathbf{H}_2 are the Hessian matrices for the ℓ_2 objective function at the points $\mathbf{x}_{os}^{e,1}$ and $\mathbf{x}_{os}^{e,2}$, respectively. Proof

We denote by $Q(\mathbf{x}, V)$ the value of the ℓ_2 objective function of the MPE problem at a coarse model point \mathbf{x} using a set of fine model points V, where V is given by (7). The quadratic approximations of $Q(\mathbf{x}, V)$ in a neighborhood centered at the two locally unique minima $\mathbf{x}_{os}^{e,1}$ and $\mathbf{x}_{os}^{e,2}$, respectively are given by

$$q_1(\Delta \mathbf{x}, V) = Q(\mathbf{x}_{os}^{e,1}, V) + 0.5 \Delta \mathbf{x}^T \mathbf{H}_1 \Delta \mathbf{x}$$
(16)

$$q_2(\Delta \mathbf{x}, V) = Q(\mathbf{x}_{os}^{e,2}, V) + 0.5 \Delta \mathbf{x}^T \mathbf{H}_2 \Delta \mathbf{x}$$
(17)

The perturbation $\Delta \mathbf{x}$ that results in the maximum difference between the two quadratic models (16) and (17) for a specific trust region \mathbf{d} is obtained by formulating the Lagrangian

$$L(\boldsymbol{x},\boldsymbol{l}) = (Q(\boldsymbol{x}_{os}^{e,2},V) - Q(\boldsymbol{x}_{os}^{e,1},V)) + 0.5\Delta\boldsymbol{x}^{T}(\boldsymbol{H}_{2} - \boldsymbol{H}_{1})\Delta\boldsymbol{x} + \boldsymbol{l}(\Delta\boldsymbol{x}^{T}\Delta\boldsymbol{x} - \boldsymbol{d}^{2})$$
(18)

Taking the derivative with respect to Δx gives

$$(\boldsymbol{H}_1 - \boldsymbol{H}_2)\Delta \boldsymbol{x} = 2\boldsymbol{I}\Delta \boldsymbol{x} \tag{19}$$

It follows that Δx is an eigenvector of the matrix $H_1 - H_2$. Δx provides a direction that maximizes the difference between the quadratic models. In other words, it provides a perturbation that maximizes the contrast between the changes of the coarse model responses at these two minima. It follows that the true minimum is the one whose response changes match better the changes of the fine model responses obtained using the perturbation Δx .

A similar result to that obtained in (19) can be obtained using a different approach. Assume that a perturbation of $\Delta \mathbf{x}$ is sought. This perturbation results in a perturbation of the coarse model responses at the two minima by

$$\Delta \boldsymbol{R}_{1} = \boldsymbol{J}_{os} \left(\boldsymbol{x}_{os}^{e,1} \right) \Delta \boldsymbol{x} \tag{20}$$

and

$$\Delta \boldsymbol{R}_2 = \boldsymbol{J}_{os} \left(\boldsymbol{x}_{os}^{e,2} \right) \Delta \boldsymbol{x} \tag{21}$$

Where $J_{os}(\mathbf{x}_{os})$ is the Jacobian of the coarse model response \mathbf{R}_{os} . We impose the condition that the difference between the ℓ_2 norms of these two response perturbations be maximized subject to certain trust region size. Therefore, the following Lagrangian can be formed

$$L(\Delta \mathbf{x}, \mathbf{l}) = \Delta \mathbf{x}^T \mathbf{J}_{os}(\mathbf{x}_{os}^{e,1})^T \mathbf{J}_{os}(\mathbf{x}_{os}^{e,1}) \Delta \mathbf{x} - \Delta \mathbf{x}^T \mathbf{J}_{os}(\mathbf{x}_{os}^{e,2})^T \mathbf{J}_{os}(\mathbf{x}_{os}^{e,2}) \Delta \mathbf{x} + \mathbf{l}(\Delta \mathbf{x}^T \Delta \mathbf{x} - \mathbf{d}^2)$$
(22)

Using a similar approach to that used in deriving (19) it can be shown that the perturbation Δx is obtained by solving the eigenvalue problem

$$(\boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,1})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,1}) - \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,2})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,2}))\Delta \boldsymbol{x} = \boldsymbol{I}\Delta \boldsymbol{x}$$
(23)

The two perturbations (19) and (23) can be shown to be almost identical by writing the Hessian matrix of $Q(\mathbf{x}, V)$ in terms of the Jacobian of the coarse model responses [7]. However, the perturbation calculated in (23) is more computionally efficient than that of (19).

There is one substantial difficulty in the exact evaluation of the perturbation given by (23). Once a locally unique minimum is reached the Hessian of Q at this point can be obtained while no information is available about the Hessian at the other locally unique minima that may exist. In such a case, a reasonable assumption is to take $H_2=I$, the identity matrix or alternatively take $J_{os}(x_{os}^{e,2})^T J_{os}(x_{os}^{e,2})$ as the identity matrix in (23). This assumption implies that no information is available about the curvature of the objective function at the other minima. It follows that Δx is an eigenvector of the matrix $J_{os}(x_{os}^{e,1})^T J_{os}(x_{os}^{e,1})$.

The perturbation given by (23) is a suggested perturbation in the coarse model space. The new fine model point that should be added to the set V is $\mathbf{x}_{em} + \Delta \mathbf{x}_{em}$ where $\Delta \mathbf{x}_{em}$ obtained by solving the system of linear equations

$$\Delta \boldsymbol{x} = \boldsymbol{B} \Delta \boldsymbol{x}_{em} \tag{24}$$

The relation (24) is used if some information is available about the mapping between the two spaces. However, in most cases we make the assumption that B=I. The scheme that we utilized for the selection of points in (23) is as follows. First, the eigenvalue problem is solved. The eigenvector $v^{(1)}$ with the largest eigenvalue in modulus is initially selected as the candidate eigenvector. The suggested perturbation in this case is

$$\Delta \boldsymbol{x}_{os} = \frac{\boldsymbol{d}}{\|\boldsymbol{v}^{(1)}\|} \boldsymbol{v}^{(1)}$$
(25)

where d is the current size of the trust region. This perturbation is tested to see whether it belongs to the current set of perturbations. It follows that $\Delta \mathbf{x}_{os}$ is rejected if the condition

$$\frac{\Delta \boldsymbol{x}_{os}^{T} \Delta \boldsymbol{x}_{os}^{(i)}}{\left\|\Delta \boldsymbol{x}_{os}\right\|^{2}} > (1 - \boldsymbol{e})$$
(26)

is satisfied for a perturbation $\Delta \mathbf{x}_{os}^{(i)} \in V_p$, where $\mathbf{e} > 0$ is a small number. In this case the alternative perturbation

$$\Delta \boldsymbol{x}_{os} = \frac{-\boldsymbol{d}}{\|\boldsymbol{v}^{(1)}\|} \boldsymbol{v}^{(1)}$$
(27)

is tested against the condition (26). If it also fails, we switch to the eigenvector with second largest eigenvalue in modulus and repeat steps (25) – (27). This is repeated until either a perturbation is found such that condition (26) is not satisfied or all the eigenvectors are exhausted for perturbations of length d In this case the trust region size d is scaled by a where a > 1.0. The perturbation is then taken in the direction of eigenvector with largest eigenvalue in modulus.

IV. THE AGGRESSIVE PARAMETER EXTRACTION (APE) ALGORITHM

In this section we present the APE algorithm for the MPE process. This algorithm is based on the two methods discussed in the previous section. It is given by the following steps.

Step 0 Given \mathbf{x}_{em} , \mathbf{d} and n. Initialize $V^{(1)} = \{\mathbf{x}_{em}^{(1)}\}$, where $\mathbf{x}_{em}^{(1)} = \mathbf{x}_{em}$ and set i=1.

Comment The set $V^{(i)}$ contains the points used for the MPE in the *i*th iteration. The index *i* is equal to $|V^{(i)}|$, the cardinality of $V^{(i)}$.

Step 1 Apply MPE using the set $V^{(i)}$ to get $\mathbf{x}_{os}^{e(i)}$.

Comment The point $\mathbf{x}_{os}^{e(i)}$ is the solution to the MPE problem obtained using the set $V^{(i)}$.

Step 2 If the Jacobian of **R** at $\mathbf{x}_{os}^{e(i)}$ has full rank, go to Step 4.

Step 3 Obtain a new perturbation $\Delta \mathbf{x}$ using (13), use (24) to get $\Delta \mathbf{x}_{em}$ and let $V^{(i+1)} = V^{(i)} \cup \{\mathbf{x}_{em}^{(i+1)}\}$, where $\mathbf{x}_{em}^{(i+1)} = \mathbf{x}_{em} + \Delta \mathbf{x}_{em}$. Set i=i+1 and go to Step 1.

Comment The perturbation Δx is rescaled to satisfy the trust region condition $\|\Delta x\| = d$.

Step 4 If
$$|V^{(i)}|$$
 is equal to one, go to Step 6

Step 5 If $\mathbf{x}_{os}^{e(i)}$ is approaching a limit, stop.

- Step 6 Obtain a new perturbation $\Delta \mathbf{x}$ using (23) and use (24) to get $\Delta \mathbf{x}_{em}$. Update \mathbf{d} and let $V^{(i+1)} = V^{(i)} \cup \{\mathbf{x}_{em}^{(i+1)}\}$, where $\mathbf{x}_{em}^{(i+1)} = \mathbf{x}_{em} + \Delta \mathbf{x}_{em}$. Set i=i+1 and go to Step 1.
- *Comment* In Step 6 the eigenvalue problem is solved and the perturbation Δx is selected according to the scheme discussed in the previous section. This scheme may result in updating the trust region size. The algorithm terminates if the vector of extracted coarse model parameters obtained using *i* fine model points is close enough in terms of some norm to the vector of extracted parameters obtained using *i*-1 fine model points.

Fig. 3 illustrates the relationship between the generated sets $V^{(i)}$, the fine model points $\mathbf{x}_{em}^{(i)}$ and the extracted coarse model points $\mathbf{x}_{os}^{e(i)}$. A flowchart of the APE algorithm is shown in Fig. 4. The current implementation of the algorithm is in MATLAB [8].

V. EXAMPLES

The Rosenbrock Function

The first example utilizes the famous Rosenbrock function [7]. The coarse model for this problem is given by

$$R_c = 100 \left(u_2 - u_1^2\right)^2 + \left(1 - u_1\right)^2 \tag{28}$$

The fine model is another "Rosenbrock" function but with a shift applied to the parameters

$$R_f = 100 \left(\left(u_2 + 0.2 \right) - \left(u_1 - 0.2 \right)^2 \right)^2 + \left(1 - \left(u_1 - 0.2 \right) \right)^2$$
(29)

It is required to extract the coarse model parameters corresponding to the fine model point $[1.0 \ 1.0]^T$. The result of the SPE at this point is $\mathbf{x}_{os}^{e(1)} = [1.21541 \ 0.91728]^T$. The contours of $Q(\mathbf{x}, V^{(1)})$ are shown in Fig. 5. It is clear from the contour plot that the minimum obtained is a locally nonunique minimum. The algorithm detects this and generates a perturbation that attempts to improve the rank of the Jacobian of \mathbf{R} in the Double Parameter Extraction (DPE) using (13). Utilizing a trust region size of 0.25 the set $V^{(2)}$ is given by

$$V^{(2)} = \left\{ \begin{bmatrix} 1.0 & 1.0 \end{bmatrix}^T, \begin{bmatrix} 0.7643 & 1.0833 \end{bmatrix}^T \right\}$$
(30)

The contours of $Q(\mathbf{x}, V^{(2)})$ are shown in Fig. 6. It is clear that by using only one additional fine model simulation the uniqueness of the problem has improved dramatically. Actually, the only existing minimum is a global unique minimum. To ensure uniqueness a third point is generated by solving the eigenvalue problem (23). Thus we have

$$V^{(3)} = \left\{ \begin{bmatrix} 1.0 & 1.0 \end{bmatrix}^T, \begin{bmatrix} 0.7643 & 1.0833 \end{bmatrix}^T, \begin{bmatrix} 0.8185 & 0.8281 \end{bmatrix}^T \right\}$$
(31)

The contours of $Q(\mathbf{x}, V^{(3)})$ are shown in Fig. 7. The algorithm then terminates as it detects that the extracted parameters are approaching a limit. It returns the last set of extracted parameters as the solution for the MPE problem. The variation of the extracted parameters obtained using the ℓ_2 optimizer with the number of fine model points used is shown in Table I.

A 10:1 Impedance Transformer

The second example is the well-known 10:1 impedance transformer [9]. The parameters for this problem are the characteristic impedance of the two transmission lines Z_1 and Z_2 while the two lengths of the transmission lines are kept fixed at their optimal values (quarter wave length). The coarse model utilizes nonscaled parameters while a "fine" model scales each of the two impedances by a factor of 1.6.

It is required in this synthetic problem to extract the coarse model parameters whose response matches the fine model response at the point $[2.2628 \quad 4.5259]^T$. This point is the optimal coarse model design according to the specifications in [9].

The two models are matched using the reflection coefficients at eleven equally spaced frequencies in the frequency range 0.5 GHz $\leq f \leq 1.5$ GHz. The fine model response at $\mathbf{x}_{em}^{(1)}$ and the coarse model response at the point $\mathbf{x}_{os}^{e(1)}$ are shown in Fig. 8. The contours of $Q(\mathbf{x}, V^{(1)})$ are shown in Fig. 9. It is clear from this figure that there exist three locally unique minima for the extraction problem. The algorithm then generates a second perturbation using (23). The set $V^{(2)}$ is given by

$$V^{(2)} = \left\{ \begin{bmatrix} 2.26277 & 4.52592 \end{bmatrix}^T, \begin{bmatrix} 1.49975 & 4.76634 \end{bmatrix}^T \right\}$$
(32)

The fine model response for every point in $V^{(2)}$ and the coarse response at the corresponding extracted coarse model point are shown in Fig. 10. The corresponding contours of $Q(\mathbf{x}, V^{(2)})$ are shown in Fig. 11. It is clear that there still exist two locally unique minima. Using (23) we have

$$V^{(3)} = \left\{ \begin{bmatrix} 2.26277 & 4.52592 \end{bmatrix}^T, \begin{bmatrix} 1.49975 & 4.76634 \end{bmatrix}^T, \begin{bmatrix} 3.02024 & 4.26855 \end{bmatrix}^T \right\}$$
(33)

The fine model response for each point in the set $V^{(3)}$ and the coarse model response at the corresponding extracted coarse model point are shown in Fig. 12. The contours of $Q(\mathbf{x}, V^{(3)})$ are shown in Fig. 13. The algorithm terminates as the termination condition is satisfied. The variation of the extracted coarse model point with $|V^{(i)}|$ is given in Table II.

The fine model for the HTS filter (Fig. 14) is simulated as a whole using Sonnet's *em* [11]. The "coarse" model is a decomposed Sonnet version of the fine model. This model exploits a coarser grid than that used for the fine model. The physical parameters of the coarse and fine models are given in Table III.

It is required to extract the coarse model parameters corresponding to the fine model parameters given in Table IV. The values in this table are the optimal coarse model design obtained using the minimax optimizer in OSA90/hope [12] according to specifications given in [10]. We utilize the responses at 15 discrete frequencies in the range [3.967 GHz, 4.099 GHz] in the parameter extraction process.

The algorithm first started by applying SPE where $V^{(1)}$ contains only the point given in Table IV. The point $\mathbf{x}_{os}^{e(1)}$ is given in Table VI. Fig. 15 shows the fine model response at $\mathbf{x}_{em}^{(1)}$ and the coarse model response at $\mathbf{x}_{os}^{e(1)}$.

The algorithm detected that this extracted point is a locally unique minimum. A new fine model point is then generated by solving the eigenvalue problem (23). A DPE step is then carried out. The set $V^{(2)}$ includes the points given in the second and third columns of Table V. The point $x_{os}^{e(2)}$ is given in Table VI. Fig. 16 shows the fine model responses at the two utilized fine model points and the responses at the corresponding extracted coarse model points, respectively. Again the algorithm detected that the extracted point is locally unique and a new fine model point is generated and added to the set of points. The same steps were then repeated for three-point and four point parameter extraction. The points utilized are given in Table V. The results are shown in the fourth and fifth columns of Table VI. It is clear that the extracted parameters are approaching a limit. The fine model responses and the responses at the corresponding extracted coarse model points for the last two iterations are shown in Figs. 17 and

18, respectively. Fig. 18(a) demonstrates that a good match between the responses of both models over a wider range of frequencies than that used for parameter extraction is achieved.

Double-folded Stub Filter [1]

We consider the design of the double-folded stub (DFS) microstrip structure shown in Fig. 19 (Bandler *et al.* [1]). Folding the stubs reduces the filter area w.r.t. the conventional double stub structure (Rautio [13]). The filter is characterized by five parameters : W_1 , W_2 , S, L_1 and L_2 . L_1 , L_2 and S are chosen as optimization variables. W_1 and W_2 are fixed at 4.8 mil. The fine model is simulated by HP HFSS [14] through HP Empipe3D [15]. The coarse model exploits the microstrip line and microstrip T-junction models available in OSA90/hope [12]. The coupling between the folded stubs and the microstrip line is simulated using equivalent capacitors. The values of these capacitors is determined using Walker's formulas [16]. Jansen's microstrip bend model [17] is used to model the folding effect of the stub. The coarse model is shown in Fig. 20.

It is required in this example to extract the coarse model parameters corresponding to the fine model parameters given in Table VII. This vector is the optimal design of the coarse model obtained by minimax optimization. The optimal coarse model response and the fine model response at the optimal coarse model design are shown in Fig. 21. This figure shows clearly the large misalignment between the two models which implies nonuniqueness of the extracted parameters.

The algorithm started by applying SPE using the fine model point given in Table VII. Fig. 22 shows the fine model response at $\mathbf{x}_{em}^{(1)}$ and the coarse model response at the point $\mathbf{x}_{os}^{e(1)}$. The algorithm detected that the extracted parameters are locally unique. A new fine model point is generated using (23) and added to the set of fine model points used for the MPE. The algorithm needed nine iterations to trust the extracted coarse model parameters. The fine model points utilized are given in Table VIII and the extracted coarse model points are given in Table IX. Fig. 23 shows the fine model response at $\mathbf{x}_{em}^{(1)}$ and the coarse model response at the point $\mathbf{x}_{os}^{e(9)}$.

Table IX shows the large relative change in parameter values between the first set of extracted parameters $\mathbf{x}_{os}^{e(1)}$ and the trusted set of parameters $\mathbf{x}_{os}^{e(9)}$. If the step taken by any SM optimization algorithm utilizes $\mathbf{x}_{os}^{e(1)}$, the algorithm would have probably failed. The contours of $Q(\mathbf{x}, V^{(1)})$ in the neighborhood of $\mathbf{x}_{os}^{e(9)}$ are shown in Fig. 24 (a) and (b), respectively. Also, The contours of $Q(\mathbf{x}, V^{(9)})$ in the neighborhood of $\mathbf{x}_{os}^{e(1)}$ and in the neighborhood of $\mathbf{x}_{os}^{e(1)}$ has been weakened as a possible solution for the parameter extraction problem.

Fig. 25 shows the change of $Q(\mathbf{x}_{os}^{e(1)}, V^{(i)})$ and of $Q(\mathbf{x}_{os}^{e(9)}, V^{(i)})$ with $|V^{(i)}|$. The value of $Q(\mathbf{x}_{os}^{e(9)}, V^{(i)})$ remains almost constant and small in value. On the other hand the value of $Q(\mathbf{x}_{os}^{e(1)}, V^{(i)})$ increases significantly with each new point added to the set of utilized fine model points signaling a false minimum.

VI. CONCLUSIONS

An Aggressive Parameter Extraction (APE) algorithm is proposed. Our APE algorithm addresses the optimal selection of parameter perturbations used to improve the uniqueness of a multi-point parameter extraction procedure. The nonuniqueness of the parameter extraction problem may lead to the divergence or oscillation of the SM approach to circuit design. New parameter perturbations are generated based on the nature of the minimum reached in the previous iteration. We consider possibly locally unique and locally nonunique minima. The suggested perturbations in each of these two cases are obtained either by solving a system of linear equations or by solving an eigenvalue problem. The APE algorithm continues until the extracted coarse model parameters can be trusted. The algorithm is successfully demonstrated through the parameter extraction of microwave filters and impedance transformers.

ACKNOWLEDGEMENT

The authors thank Sonnet Software, Inc., Liverpool, NY for making *em* available for this work.

They also thank HP EEsof, Santa Rosa, CA, for making HP HFSS and HP Empipe3D available.

REFERENCES

- [1] J.W. Bandler, R.M. Biernacki, S.H. Chen, P.A. Grobelny and R.H. Hemmers, "Space mapping technique for electromagnetic optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 42, 1994, pp. 2536-2544.
- [2] 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, 1995, pp. 2874-2882.
- [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, 1998, pp. 2412-2425.
- [4] J.W. Bandler, S.H. Chen and S. Daijavad, "Microwave device modeling using efficient ℓ_1 optimization: a novel approach," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-34, 1986, pp. 1282-1293.
- [5] J.W. Bandler, R.M. Biernacki and S.H. Chen, "Fully automated space mapping optimization of 3D structures," *IEEE MTT-S Int. Microwave Symp. Dig.* (San Francisco, CA), 1996, pp. 753-756.
- [6] J.W. Bandler and A.E. Salama, "Fault diagnosis of analog circuits," *Proc. IEEE*, vol. 73, pp. 1279-1325, 1985.
- [7] R. Fletcher, *Practical Methods of Optimization*. New York: Wiley, Second Edition, 1987.
- [8] MATLAB[®] Version 5.0, The Math. Works, Inc., 24 Prime Park Way, Natick, MA 01760, 1997.
- [9] J.W. Bandler, "Optimization methods for computer-aided design," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, 1969, pp. 533-552.
- [10] J.W. Bandler, R.M. Biernacki, S.H. Chen, W.J. Gestinger, P.A. Grobelny, C. Moskowitz and S.H. Talisa, "Electromagnetic design of high-temperature superconducting filters," *Int. J. Microwave and Millimeter-Wave Computer-Aided Engineering*, vol. 5, 1995, pp. 331-343.
- [11] $em^{\mathbb{M}}$, Sonnet Software, Inc., 1020 Seventh North Street, Suite 210, Liverpool, NY 13088, 1997.
- [12] OSA90/hope[™] Version 4.0, formerly Optimization Systems Associates Inc., P.O. Box 8083, Dundas, ON, Canada, L9H 5E7, 1997, now HP EEsof Division, 1400 Fountaingrove Parkway, Santa Rosa, CA 95403-1799.

- [13] J.C. Rautio, Sonnet Software, Inc., 1020 Seventh North Street, Suite 210, Liverpool, NY 13088, Private Communication, 1992.
- [14] HP HFSS[™] Version 5.2, HP EEsof Division, 1400 Fountaingrove Parkway, Santa Rosa, CA 95403-1799, 1998.
- [15] HP Empipe3D[™] Version 5.2, HP EEsof Division, 1400 Fountaingrove Parkway, Santa Rosa, CA 95403-1799, 1998.
- [16] C.S. Walker, Capacitance, Inductance and Crosstalk Analysis, Norwood, MA: Artech House, 1990.
- [17] M. Kirschning, R. Jansen and N. Koster, "Measurement and computer-aided modeling of microstrip discontinuities by an improved resonator method," *IEEE MTT-S Int. Microwave Symp. Dig.* (Boston, MA), 1983, pp. 495-497.

TABLE I THE VARIATION OF THE EXTRACTED PARAMETERS FOR THE ROSENBROCK FUNCTION WITH THE NUMBER OF POINTS USED FOR EXTRACTION

Parameter	$\boldsymbol{x}_{os}^{e(1)}$	$\boldsymbol{x}_{os}^{e(2)}$	$\boldsymbol{x}_{os}^{e(3)}$
u_1	1.21541	0.80008	0.80008
u_2	0.91728	1.20012	1.20014

TABLE II THE VARIATION OF THE EXTRACTED PARAMETERS FOR THE 10:1 IMPEDANCE TRANSFORMER WITH THE NUMBER OF POINTS USED FOR EXTRACTION USING

Parameter	$\boldsymbol{x}_{os}^{e(1)}$	$\boldsymbol{x}_{os}^{e(2)}$	$\boldsymbol{x}_{os}^{e(3)}$
$egin{array}{c} Z_1 \ Z_2 \end{array}$	3.62043	3.47160	3.60357
	7.24147	7.43214	7.35052

TABLE III
MATERIAL AND PHYSICAL PARAMETERS
FOR THE COARSE AND FINE MODELS OF THE HTS FILTER

Model Parameter	Coarse Model	Fine Model
substrate dielectric constant substrate thickness (mil) shielding cover height (mil) conducting material thickness substrate dielectric loss tangent resistivity of metal (Ω m) magnetic loss tangent	23.425 19.9516 100 0 0 0	$23.425 \\19.9516 \\250 \\0 \\0 \\0 \\0 \\0 \\0 \\0 \\0 \\0 \\0 \\0 \\0 \\0$
surface reactance (Ω /sq) x-grid cell size (mil) y-grid cell size (mil)	0 2.00 1.75	0 1.00 1.75

TOR III		
Parameter	Value	
L_1	181.00	
L_2	201.59	
L_3	180.97	
S_1	20.12	
S_2	67.89	
S_3	66.85	

TABLE IV THE OPTIMAL COARSE MODEL DESIGN FOR THE HTS FILTER

all values are in mils

TABLE V THE FINE MODEL POINTS USED IN THE APE ALGORITHM FOR THE HTS FILTER

Parameter	$oldsymbol{x}_{em}^{(1)}$	$oldsymbol{x}_{em}^{(2)}$	$\boldsymbol{x}_{em}^{(3)}$	$oldsymbol{x}_{em}^{(4)}$
L_1	181.00	182.55	181.34	179.86
L_2	201.59	205.64	205.38	197.74
L_3^2	180.97	183.36	184.20	178.08
S_1	20.12	20.05	20.07	20.46
S_2	67.89	68.40	68.08	67.35
S_3	66.85	67.25	66.98	66.46

all values are in mils

TABLE VI THE VARIATION IN THE EXTRACTED PARAMETERS FOR THE HTS FILTER WITH THE NUMBER OF FINE MODEL POINTS

Parameter	$\boldsymbol{x}_{os}^{e(1)}$	$\boldsymbol{x}_{os}^{e(2)}$	$\boldsymbol{x}_{os}^{e(3)}$	$\boldsymbol{x}_{os}^{e(4)}$
L_1	188.31	179.99	176.67	178.50
L_2	197.69	204.52	208.52	206.78
L_3	189.72	181.23	178.00	179.09
S_1	19.34	17.13	17.21	18.99
S_2	52.67	63.44	56.52	57.99
S_2	52.67	63.44	56.52	57.99
S_3	52.06	53.18	53.47	56.77
	all va	lues are in mils		

Parameter	Value
L_1	66.73
L_2	60.23
S	9.59
all valu	es are in mils

TABLE VII THE OPTIMAL COARSE MODEL DESIGN

TABLE VIII
THE FINE MODEL POINTS USED IN THE APE
ALGORITHM FOR THE DFS FILTER

Parameter	$oldsymbol{x}_{em}^{(1)}$	$\boldsymbol{x}_{em}^{(2)}$	$\boldsymbol{x}_{em}^{(3)}$	$\boldsymbol{x}_{em}^{(4)}$	$x_{em}^{(5)}$	$\boldsymbol{x}_{em}^{(6)}$	$x_{em}^{(7)}$	$\boldsymbol{x}_{em}^{(8)}$	$x_{em}^{(9)}$
L_1	66.73	67.72	67.32	66.15	70.60	67.66	62.82	65.80	66.57
L_2	60.23	63.58	64.13	56.33	59.48	64.10	60.88	56.36	59.85
S	9.59	9.27	9.48	9.71	9.71	9.66	9.50	9.52	10.26
			all val	ues are in	mils				

	THE VA FOR	RIATION	N IN THE S FILTER	EXTRA	CTED PA	ARAMET	ERS		
	_		FINE MC	DEL PO	INTS				
Parameter	$\boldsymbol{x}_{os}^{e(1)}$	$\boldsymbol{x}_{os}^{e(2)}$	$\boldsymbol{x}_{os}^{e(3)}$	$\boldsymbol{x}_{os}^{e(4)}$	$\boldsymbol{x}_{os}^{e(5)}$	$\boldsymbol{x}_{os}^{e(6)}$	$\boldsymbol{x}_{os}^{e(7)}$	$x_{os}^{e(8)}$	$\boldsymbol{x}_{os}^{e(9)}$
L_1	58.01	67.05	66.11	64.36	56.46	66.10	56.50	56.39	56.5
L_2	38.40	40.47	40.40	43.28	42.94	42.02	42.81	43.00	43.0
S	3.24	6.86	6.64	8.83	18.10	7.99	18.25	17.93	17.8

TABLE IX
THE VARIATION IN THE EXTRACTED PARAMETERS
FOR THE DFS FILTER WITH THE NUMBER OF
ENE MODEL DONTS

APPENDIX A

Let the two sets S and S_a be defined as in (9) and (10), respectively. We impose the condition

that every gradient on the set S_a should be orthogonal to all gradients in the set S. It follows that

$$\boldsymbol{g}_{a}^{(i)T} \boldsymbol{g}^{(j)} = 0 \quad \forall \, \boldsymbol{g}^{(j)} \in S \text{ and } \forall \, \boldsymbol{g}_{a}^{(i)} \in S_{a}$$
 (34)

But each gradient in the set S_a can be approximated by

$$g_a^{(i)} = g^{(i)} + G^{(i)} \Delta x, \quad i = k+1, \dots, n$$
 (35)

where $g^{(i)}$ is the gradient of the *i*th response at the point x_{os}^{e} and $G^{(i)}$ is the corresponding Hessian. It follows that the condition (34) can be restated as

$$g^{(j)T}g^{(i)} = -g^{(j)T}G^{(i)}\Delta x, \quad j=1,...,k \text{ and } i=k+1,...,n$$
 (36)

Equation (36) is a linear equation in *n* unknowns (the components of Δx). There are k(n-k) such linear equations. Putting these equations into a matrix form we have

$$B\Delta x = -c \tag{37}$$

where the *m*th row of the matrix **B** is

$$\boldsymbol{B}^{(m)T} = \boldsymbol{g}^{(j)T} \boldsymbol{G}^{(i)}, \quad i = k+1, \dots, n \text{ and } j = 1, \dots, k$$
(38)

where m=(i-1)(n-k)+j. Similarly the *m*th component of the vector *c* is

$$c_{m,1} = g^{(j)T} g^{(i)}, \quad i = k+1, \dots, n \text{ and } j = 1, \dots, k$$
 (39)

Thus (13) follows.



Fig. 1. Illustration of the SPE procedure.



Fig. 2. Illustration of the MPE procedure.



Fig. 3. Illustration of the relationship between the generated sets $V^{(i)}$, the fine model points $\mathbf{x}_{em}^{(i)}$ and the extracted coarse model points $\mathbf{x}_{os}^{e(i)}$ generated by the APE algorithm.



Fig. 4. The flow chart of the APE algorithm.



Fig. 5. The contours of $Q(\mathbf{x}, V^{(1)})$ for the Rosenbrock function.



Fig. 6. The contours of $Q(\mathbf{x}, V^{(2)})$ for the Rosenbrock function.



Fig. 7. The contours of $Q(\mathbf{x}, V^{(3)})$ for the Rosenbrock function.



Fig. 8. The responses of the given fine model point (o) and the coarse model response (–) at the point $\mathbf{x}_{os}^{e(1)}$ for the 10:1 impedance transformer.



Fig. 9. The contours of $Q(\mathbf{x}, V^{(1)})$ for the 10:1 impedance transformer.



Fig. 10. The fine model response (o) and the corresponding coarse model response (-); (a) at the first point and (b) at the second point utilized in the DPE for the 10:1 impedance transformer.



Fig. 11. The contours of $Q(\mathbf{x}, V^{(2)})$ for the 10:1 impedance transformer.



Fig. 12. The fine model response (o) and the corresponding coarse model response (-); (a) at the first point, (b) at the second point and (c) at the third point utilized in the three-point parameter extraction for the 10:1 impedance transformer.



Fig. 13. The contours of $Q(\mathbf{x}, V^{(3)})$ for the 10:1 impedance transformer.



Fig. 14. The HTS filter [10].



Fig. 15. The fine model response (0) and the corresponding coarse model response (-) at the point utilized in the SPE for the HTS filter. Note that only points in the range 3.967 GHz to 4.099 GHz were actually used.



Fig. 16. The fine model response (o) and the corresponding coarse model response (–), (a) at the first point, and (b) at the second point utilized in the DPE for the HTS filter. Note that only points in the range 3.967 GHz to 4.099 GHz were actually used.



Fig. 17. The fine model response (o) and the corresponding coarse model response (–), (a) at the first point, (b) at the second point, and (c) at the third point utilized in the three-point parameter extraction for the HTS filter. Note that only points in the range 3.967 GHz to 4.099 GHz were actually used.



Fig. 18. The fine model response (o) and the corresponding coarse model response (–), (a) at the first point, (b) at the second point, (c) at the third point, and (d) at the fourth point utilized in the four-point parameter extraction for the HTS filter. Note that only points in the range 3.967 GHz to 4.099 GHz were actually used.



Fig. 19. The DFS filter [1].



Fig. 20. The coarse model of the DFS filter.



Fig. 21. The optimal coarse model response (—) and the fine model response (o) at the optimal coarse model design for the DFS filter.



Fig. 22. The fine model response (o) and the corresponding coarse model response (-) at the point $\mathbf{x}_{os}^{e(1)}$ for the DFS filter.



Fig. 23. The fine model response (o) and the corresponding coarse model response (-) at the point $\mathbf{x}_{os}^{e^{(9)}}$ for the DFS filter.



Fig. 24. Contour plots of the MPE problem for the DFS filter; (a) the contours of $Q(\mathbf{x}, V^{(1)})$ in the neighborhood of $\mathbf{x}_{os}^{e(1)}$, (b) the contours of $Q(\mathbf{x}, V^{(1)})$ in the neighborhood of $\mathbf{x}_{os}^{e(9)}$, (c) the contours of $Q(\mathbf{x}, V^{(9)})$ in the neighborhood of $\mathbf{x}_{os}^{e(1)}$ and (d) the contours of $Q(\mathbf{x}, V^{(9)})$ in the neighborhood of $\mathbf{x}_{os}^{e(9)}$.



Fig. 25. The variation of $Q(\mathbf{x}, V^{(i)})$ for the DFS filter at the point $\mathbf{x}_{os}^{e(1)}$ (— * —) and at the point $\mathbf{x}_{os}^{e(9)}$ (— o —) with the number of points utilized for parameter extraction.