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 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 examples include the parameter extraction of a decomposed electromagnetic model of a high-temperature superconducting filter. The parameter extraction of an empirical model of a double-folded stub filter is also carried out.**

***Index Terms*—Design automation, electromagnetic simulation, microstrip filters, optimization methods, parameter extraction, space mapping, waveguide filters.**

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.

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. Bandler *et al.* [5] later suggested that multipoint 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 that could be very central processing unit (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.

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, and by the Micronet Network of Centres of Excellence. The work of M. H. Bakr was supported under an Ontario Graduate Scholarship. The work of N. Georgieva was supported under an NSERC Postdoctorate Fellowship.

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.

Publisher Item Identifier S 0018-9480(99)08439-2.

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

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

where $\boldsymbol{R}_m$ is the vector of given measurements, $\boldsymbol{R}_{os}$ is the vector of circuit response, and $\boldsymbol{x}_{os}^{e}$ is the vector of extracted parameters. In the context of SM, the fine-model response $\boldsymbol{R}_{em}$, typically from an electromagnetic simulator, at a certain point $\boldsymbol{x}_{em}$ supplies the target response $\boldsymbol{R}_m$. 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 $\boldsymbol{x}_{os}^{e}$ should satisfy

$$\boldsymbol{x}_{os}^{e} = \arg\left\{\min_{\boldsymbol{x}_{os}} \left\| \begin{bmatrix} \boldsymbol{e}_0^T & \boldsymbol{e}_1^T & \cdots & \boldsymbol{e}_{N_p}^T \end{bmatrix}^T \right\|\right\} \tag{2}$$

where

$$\boldsymbol{e}_0 = \boldsymbol{R}_{os}(\boldsymbol{x}_{os}) - \boldsymbol{R}_{em}(\boldsymbol{x}_{em}) \tag{3}$$

and

$$\boldsymbol{e}_i = \boldsymbol{R}_{os}\left(\boldsymbol{x}_{os} + \Delta\boldsymbol{x}_{os}^{(i)}\right) - \boldsymbol{R}_{em}\left(\boldsymbol{x}_{em} + \Delta\boldsymbol{x}_{em}^{(i)}\right). \tag{4}$$

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

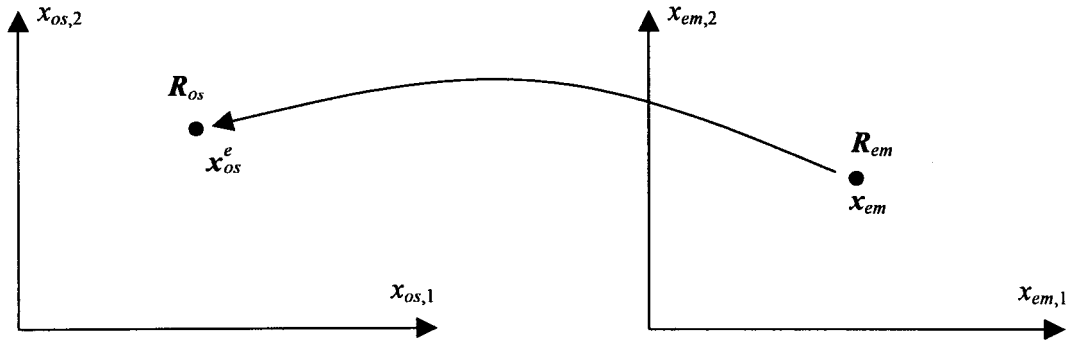

Fig. 1. Illustration of the SPE procedure.

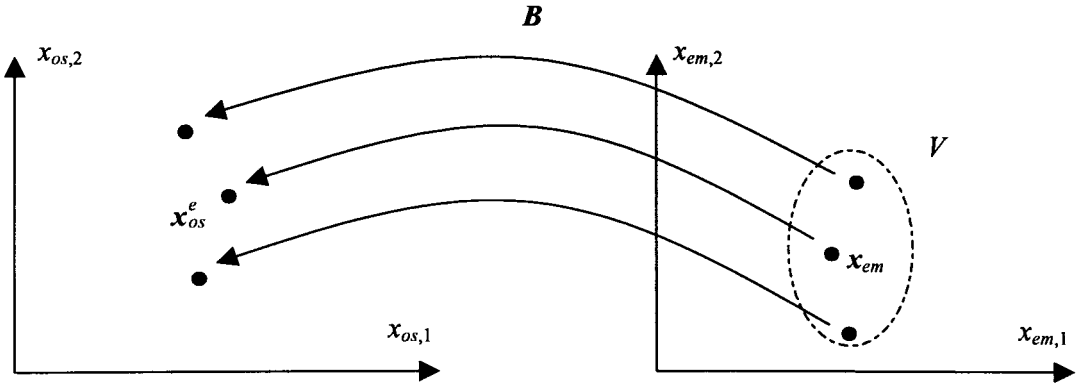

Fig. 2. Illustration of the MPE procedure.

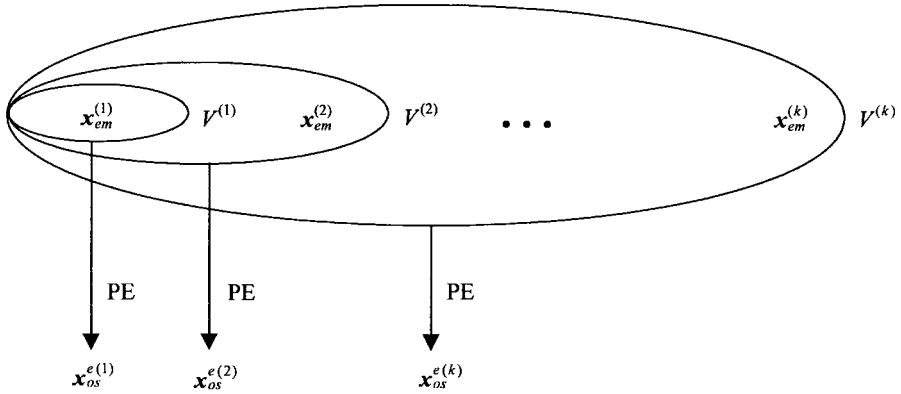

Fig. 3. Illustration of the relationship between the generated sets $V^{(i)}$, the fine-model points $\boldsymbol{x}^{(i)}_{em}$, and the extracted coarse-model points $\boldsymbol{x}^{e(i)}_{os}$ generated by the APE algorithm.

MPE procedure are related by

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

It follows that the solution of (2) simultaneously matches the responses of a set of corresponding points in both spaces.

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

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

The matrix $\boldsymbol{B}$ approximates the mapping between the two spaces. In the context of SM, (6) is superior to (5), as it integrates the available mapping $\boldsymbol{B}$ 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 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

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

Fig. 2 illustrates the MPE procedure.

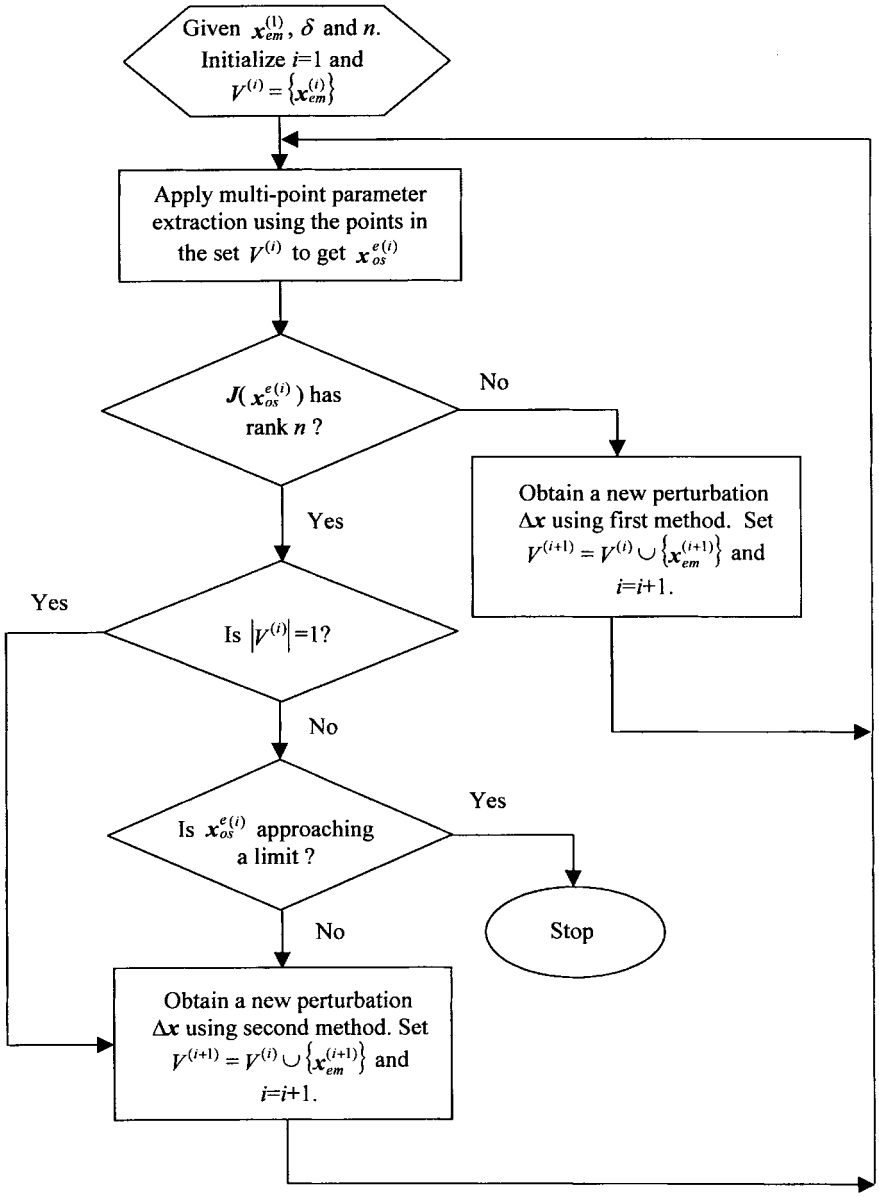

Fig. 4. The flowchart of the APE algorithm.

III. The Selection of Perturbations

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

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{R}_{os}(\boldsymbol{x}_{os}) \\ \boldsymbol{R}_{os}\left(\boldsymbol{x}_{os} + \Delta\boldsymbol{x}_{os}^{(1)}\right) \\ \vdots \\ \boldsymbol{R}_{os}\left(\boldsymbol{x}_{os} + \Delta\boldsymbol{x}_{os}^{(N_p)}\right) \end{bmatrix}. \tag{8}$$

The dimensionality of $\boldsymbol{R}$ is m_p, where $m_p = (N_p + 1)m$ and m is the dimensionality of both $\boldsymbol{R}_{os}$ and $\boldsymbol{R}_{em}$. Vector $\boldsymbol{x}_{os}^{e}$ is labeled locally unique [6] if there exists an open neighborhood of $\boldsymbol{x}_{os}^{e}$ containing no other point $\boldsymbol{x}_{os}$ such that $\boldsymbol{R}(\boldsymbol{x}_{os}) = \boldsymbol{R}(\boldsymbol{x}_{os}^{e})$. 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 $\boldsymbol{R}$ 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 $\boldsymbol{J}$ of $\boldsymbol{R}$ becomes full rank. The perturbations suggested in [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 $\boldsymbol{J}$ 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 $\boldsymbol{x}_{os}^{e}$ for the MPE problem.

Assume that a locally nonunique minimum $\boldsymbol{x}_{os}^{e}$ is obtained using the current set of coarse-model perturbations V_p. Here, the rank of the Jacobian $\boldsymbol{J}$ of $\boldsymbol{R}$ is k, where $k < n$ and n is the number of parameters. We suggest a perturbation $\Delta\boldsymbol{x}$ that attempts to increase the rank of the Jacobian of the responses

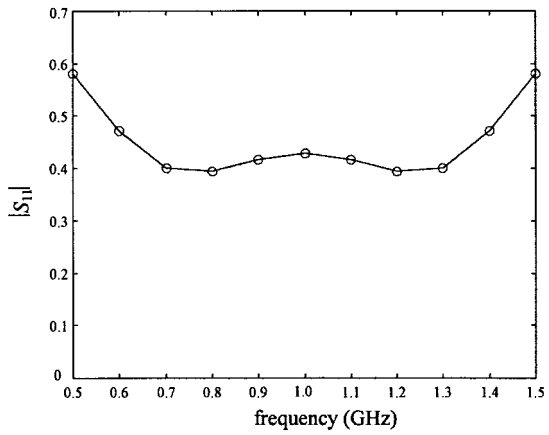

Fig. 5. The responses of the given fine-model point (o) and the coarse-model response (—) at the point $\boldsymbol{x}_{os}^{e(1)}$ for the 10 : 1 impedance transformer.

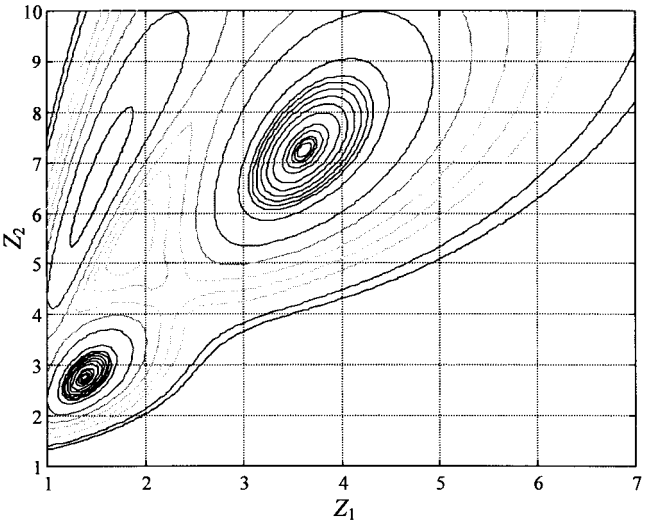

Fig. 6. The contours of $Q(\boldsymbol{x}, V^{(1)})$ for the 10 : 1 impedance transformer.

corresponding to the augmented set $\{V_p \cup \Delta\boldsymbol{x}\}$ at $\boldsymbol{x}_{os}^{e}$ 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 $\boldsymbol{R}_{os}(\boldsymbol{x}_{os}^{e}+\Delta\boldsymbol{x})$ be normal to a linearly independent set of gradients of cardinality k of the responses in the vector $\boldsymbol{R}$ at the point $\boldsymbol{x}_{os}^{e}$. We denote the set of linearly independent gradients by S where

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

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

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

Each of these gradients is approximated by

$$\boldsymbol{g}_a^{(i)}=\boldsymbol{g}^{(i)}+\boldsymbol{G}^{(i)}\Delta\boldsymbol{x},\qquad i=k+1,\cdots,n \tag{11}$$

where $\boldsymbol{g}^{(i)}$ is the gradient of the ith response at the point $\boldsymbol{x}_{os}^{e}$ and $\boldsymbol{G}^{(i)}$ is the corresponding Hessian. The imposed condition on the perturbation is that

$$\boldsymbol{g}_a^{(i)T}\boldsymbol{g}^{(j)}=0\qquad \forall\,\boldsymbol{g}^{(j)}\in S \text{ and } \forall\,\boldsymbol{g}_a^{(i)}\in S_a. \tag{12}$$

Using (11) and (12), the perturbation $\Delta\boldsymbol{x}$ is obtained by solving the system of linear equations

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

where the matrix $\boldsymbol{A}$ is given by

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

and the vector $\boldsymbol{c}$ is given by

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

A complete derivation of (13) is given in the Appendix. 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 $\boldsymbol{A}^T$ 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 $\Delta\boldsymbol{x}$ 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 $\boldsymbol{x}_{os}^{e,1}$ and $\boldsymbol{x}_{os}^{e,2}$ for the MPE problem obtained using least squares optimization and a set of perturbations V_p. A possible perturbation $\Delta\boldsymbol{x}$ 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 $\boldsymbol{H}_1-\boldsymbol{H}_2$ where $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are the Hessian matrices for the ℓ_2 objective function at the points $\boldsymbol{x}_{os}^{e,1}$ and $\boldsymbol{x}_{os}^{e,2}$, respectively.

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

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

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

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

$$L(\boldsymbol{x},\lambda)=\left(Q(\boldsymbol{x}_{os}^{e,2}, V)-Q(\boldsymbol{x}_{os}^{e,1}, V)\right)+0.5\Delta\boldsymbol{x}^T(\boldsymbol{H}_2-\boldsymbol{H}_1)\Delta\boldsymbol{x}+\lambda(\Delta\boldsymbol{x}^T\Delta\boldsymbol{x}-\delta^2). \tag{18}$$

Taking the derivative with respect to $\Delta\boldsymbol{x}$ gives

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

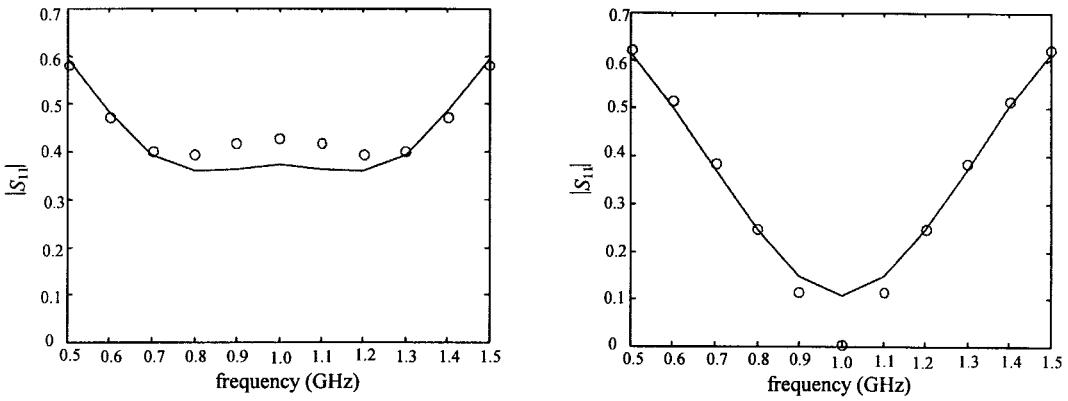

Fig. 7. The fine-model response (o) and the corresponding coarse-model response (—): (a) at the first point and (b) at the second point utilized in the DPE for the 10:1 impedance transformer.

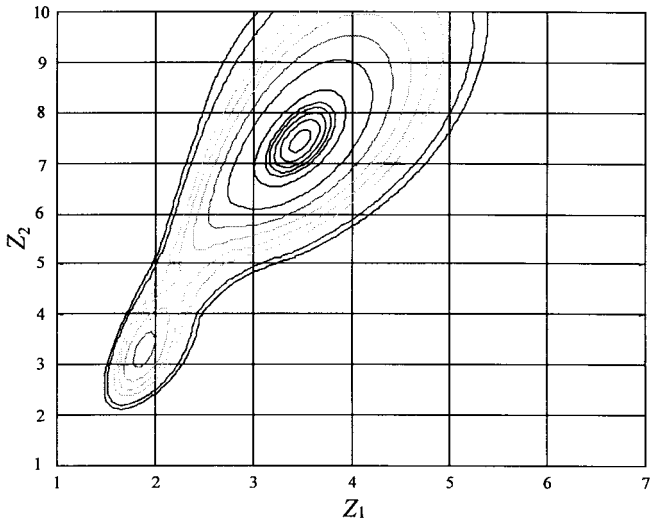

Fig. 8. The contours of $Q(\boldsymbol{x}, V^{(2)})$ for the 10:1 impedance transformer.

It follows that $\Delta\boldsymbol{x}$ is an eigenvector of the matrix $\boldsymbol{H}_1 - \boldsymbol{H}_2$. $\Delta\boldsymbol{x}$ 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 fine-model perturbation corresponding to $\Delta\boldsymbol{x}$.

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

$$\Delta\boldsymbol{R}_1 = \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,1})\Delta\boldsymbol{x} \tag{20}$$

and

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

where $\boldsymbol{J}_{os}(\boldsymbol{x}_{os})$ is the Jacobian of the coarse-model response $\boldsymbol{R}_{os}$. 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:

$$\begin{aligned} L(\Delta\boldsymbol{x}, \lambda) = {} & \Delta\boldsymbol{x}^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,1})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,1})\Delta\boldsymbol{x} \\ & - \Delta\boldsymbol{x}^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,2})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,2})\Delta\boldsymbol{x} \\ & + \lambda(\Delta\boldsymbol{x}^T\Delta\boldsymbol{x} - \delta^2). \end{aligned} \tag{22}$$

Using a similar approach to that used in deriving (19), it can be shown that the perturbation $\Delta\boldsymbol{x}$ is obtained by solving the eigenvalue problem

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

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

There is one substantial difficulty in the exact evaluation of the perturbation given by (19). 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 $\boldsymbol{H}_2 = \boldsymbol{I}$, the identity matrix or alternatively take $\boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,2})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,2})$ 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 $\Delta\boldsymbol{x}$ is an eigenvector of the matrix $\boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,1})^T \boldsymbol{J}_{os}(\boldsymbol{x}_{os}^{e,1})$.

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 $\boldsymbol{x}_{em} + \Delta\boldsymbol{x}_{em}$ where $\Delta\boldsymbol{x}_{em}$ is 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 $\boldsymbol{B} = \boldsymbol{I}$.

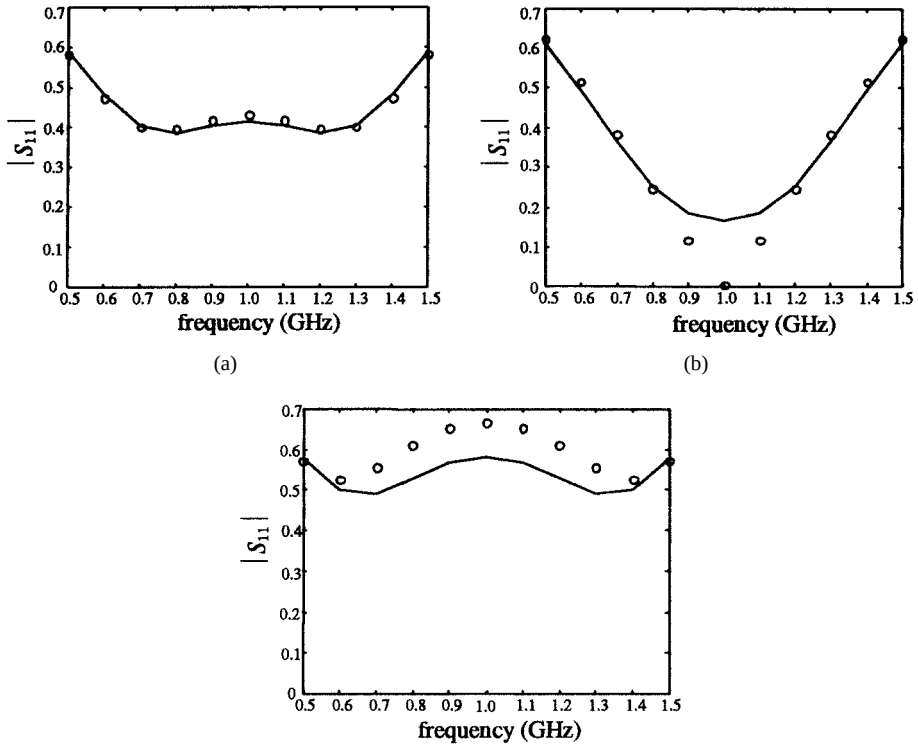

Fig. 9. The fine-model response (o) and the corresponding coarse-model response (—): (a) at the first point, (b) at the second point, and (c) at the third point utilized in the three-point parameter extraction for the 10 : 1 impedance transformer.

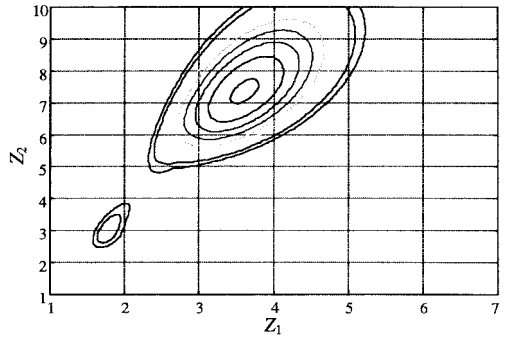

Fig. 10. The contours of $Q(\boldsymbol{x}, V^{(3)})$ for the 10 : 1 impedance transformer.

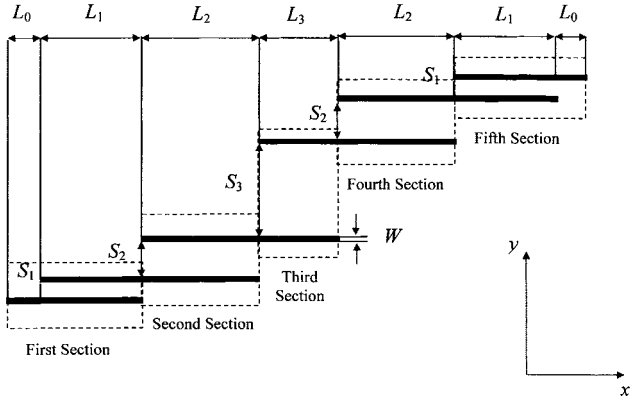

Fig. 11. The HTS filter [9].

The scheme that we utilized for the selection of points in (23) is as follows. First, the eigenvalue problem is solved. The eigenvector $\boldsymbol{v}^{(1)}$ with the largest eigenvalue in modulus is

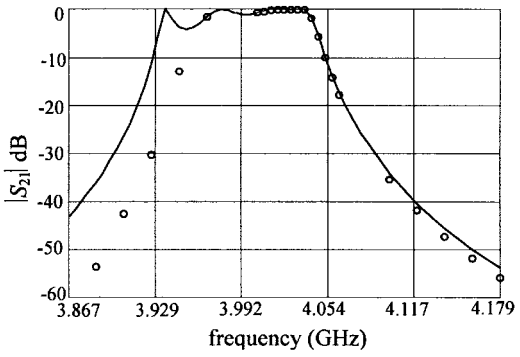

Fig. 12. The fine-model response (o) and the corresponding coarse-model response (—) at the point utilized in the SPE for the HTS filter. Note that only points in the range from 3.967 to 4.099 GHz were actually used.

initially selected as the candidate eigenvector. The suggested perturbation in this case is

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

where δ 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 $\Delta \boldsymbol{x}_{os}$ is rejected if the condition

$$\frac{\Delta \boldsymbol{x}_{os}^{T} \Delta \boldsymbol{x}_{os}^{(i)}}{\|\Delta \boldsymbol{x}_{os}\|^2} > (1-\varepsilon). \tag{26}$$

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

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

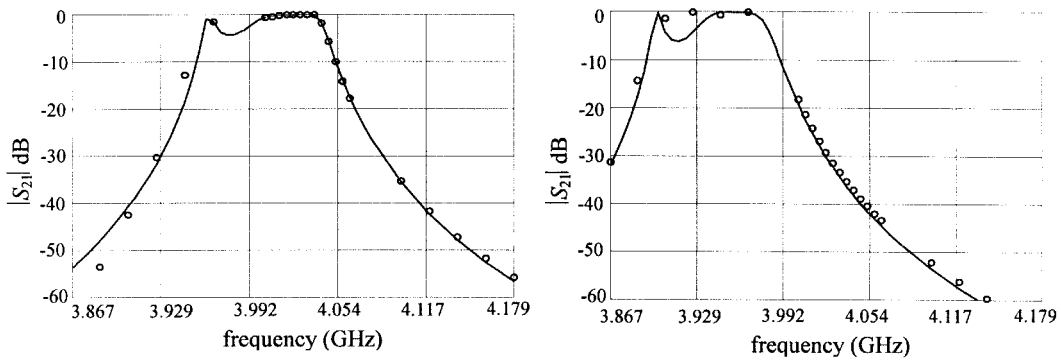

Fig. 13. The fine-model response (o) and the corresponding coarse-model response (—): (a) at the first point and (b) at the second point utilized in the DPE for the HTS filter. Note that only points in the range from 3.967 to 4.099 GHz were actually used.

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 (26) is not satisfied or all the eigenvectors are exhausted for perturbations of length δ. In this case, the trust region size δ is scaled by α where $\alpha > 1.0$. The perturbation is then taken in the direction of eigenvector with largest eigenvalue in modulus.

IV. The 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, and is given by the following steps.

1) Given $\boldsymbol{x}_{em}$, δ, and n. Initialize $V^{(1)} = \{\boldsymbol{x}_{em}^{(1)}\}$, where $\boldsymbol{x}_{em}^{(1)} = \boldsymbol{x}_{em}$ and set $i = 1$.
2) The set $V^{(i)}$ contains the points used for the MPE in the ith iteration. The index i is equal to $|V^{(i)}|$, the cardinality of $V^{(i)}$.
3) Apply MPE using the set $V^{(i)}$ to get $\boldsymbol{x}_{os}^{e(i)}$.
4) The point $\boldsymbol{x}_{os}^{e(i)}$ is the solution to the MPE problem obtained using the set $V^{(i)}$.
5) If the Jacobian of $\boldsymbol{R}$ at $\boldsymbol{x}_{os}^{e(i)}$ has full rank, go to Step 4.
6) Obtain a new perturbation $\Delta\boldsymbol{x}$ using (13), use (24) to get $\Delta\boldsymbol{x}_{em}$ and let $V^{(i+1)} = V^{(i)} \cup \{\boldsymbol{x}_{em}^{(i+1)}\}$, where $\boldsymbol{x}_{em}^{(i+1)} = \boldsymbol{x}_{em} + \Delta\boldsymbol{x}_{em}$. Set $i = i + 1$ and go to Step 1.
7) The perturbation $\Delta\boldsymbol{x}$ is rescaled to satisfy the trust region condition $\|\Delta\boldsymbol{x}\| = \delta$.
8) If $|V^{(i)}|$ is equal to one, go to Step 6.
9) If $\boldsymbol{x}_{os}^{e(i)}$ is approaching a limit, stop.
10) Obtain a new perturbation $\Delta\boldsymbol{x}$ using (23) and use (24) to get $\Delta\boldsymbol{x}_{em}$. Update δ and let $V^{(i+1)} = V^{(i)} \cup \{\boldsymbol{x}_{em}^{(i+1)}\}$, where $\boldsymbol{x}_{em}^{(i+1)} = \boldsymbol{x}_{em} + \Delta\boldsymbol{x}_{em}$. Set $i = i + 1$ and go to Step 1.
11) In Step 6, the eigenvalue problem is solved and the perturbation $\Delta\boldsymbol{x}$ 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.

TABLE I
The Variation of the Extracted Parameters for the 10 : 1 Impedance Transformer with the Number of Fine-Model Points

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

Fig. 3 illustrates the relationship between the generated sets $V^{(i)}$, the fine-model points $\boldsymbol{x}_{em}^{(i)}$, and the extracted coarse-model points $\boldsymbol{x}_{os}^{e(i)}$. A flowchart of the APE algorithm is shown in Fig. 4. The current implementation of the algorithm is in MATLAB.[1]

V. Examples

A. 10 : 1 Impedance Transformer

The first example is the well-known 10 : 1 impedance transformer [8]. 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-wavelength). 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 $[2.2628\ 4.5259]^T$. This point is the optimal coarse-model design according to the specifications in [8].

The two models are matched using the reflection coefficients at 11 equally spaced frequencies in the frequency range 0.5 GHz $\leq f \leq$ 1.5 GHz. The fine-model response at $\boldsymbol{x}_{em}^{(1)}$ and the coarse-model response at the point $\boldsymbol{x}_{os}^{e(1)}$ are shown in Fig. 5. The contours of $Q(\boldsymbol{x}, V^{(1)})$ are shown in Fig. 6. 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

$$V^{(2)} = \left\{[2.26277 \quad 4.52592]^T, \quad [1.49975 \quad 4.76634]^T\right\}. \tag{28}$$

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

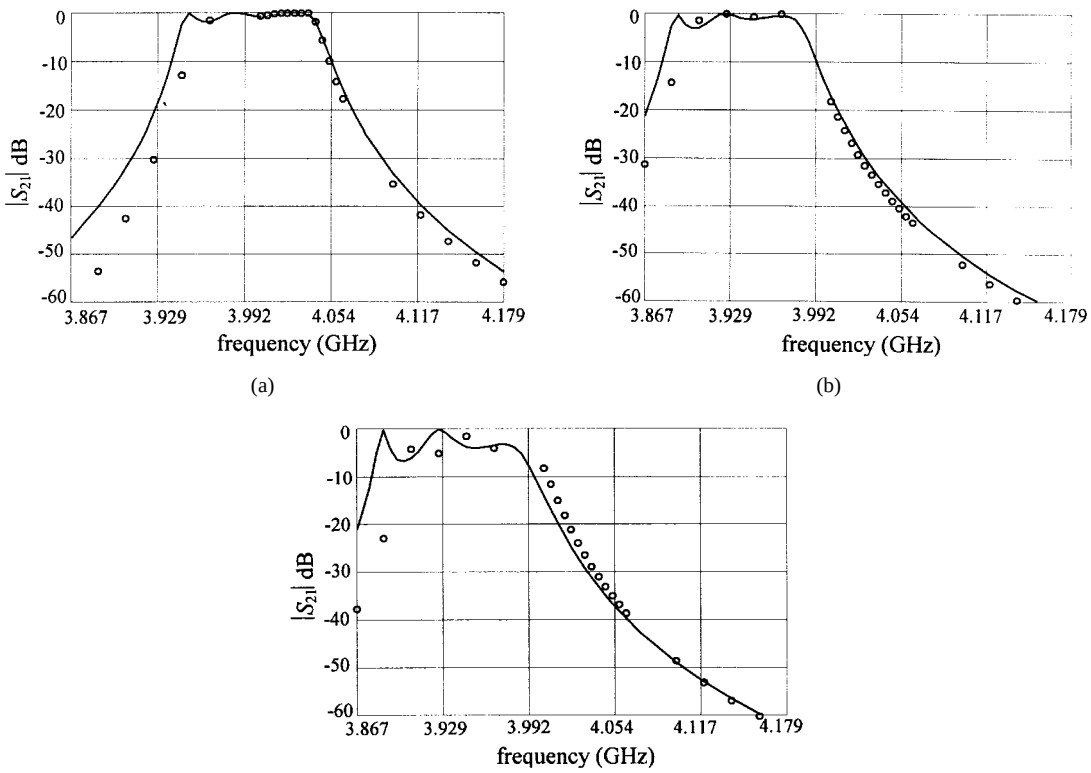

Fig. 14. The fine-model response (o) and the corresponding coarse-model response (—): (a) at the first point, (b) at the second point, and (c) at the third point utilized in the three-point parameter extraction for the HTS filter. Note that only points in the range from 3.967 to 4.099 GHz were actually used.

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

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

TABLE III
THE OPTIMAL COARSE-MODEL DESIGN FOR THE HTS FILTER

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
all values are in mils	

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. 7. The contours of this double-point extraction (DPE) are shown in Fig. 8. It is clear that there still exist two locally unique minima. Using (23), we have

$$V^{(3)} = \Big\{[2.26277 \quad 4.52592]^T, \quad [1.49975 \quad 4.76634]^T, \quad [3.02024 \quad 4.26855]^T\Big\}. \quad (29)$$

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. 9. The contours of $Q(\boldsymbol{x}, V^{(3)})$ are shown in Fig. 10. The algorithm terminates as the termination condition is satisfied. The variation of the extracted coarse-model point with $|V^{(i)}|$ is given in Table I.

B. The High-Temperature Superconducting Filter:

The fine model for the high-temperature superconducting (HTS) filter [9] (Fig. 11) is simulated as a whole using Sonnet's ***em***.[2] 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 II.

It is required to extract the coarse-model parameters corresponding to the fine-model parameters given in Table III. The values in this table are the optimal coarse-model design ob-

[2] ***em***, Sonnet Software Inc., Liverpool, NY, 1997.

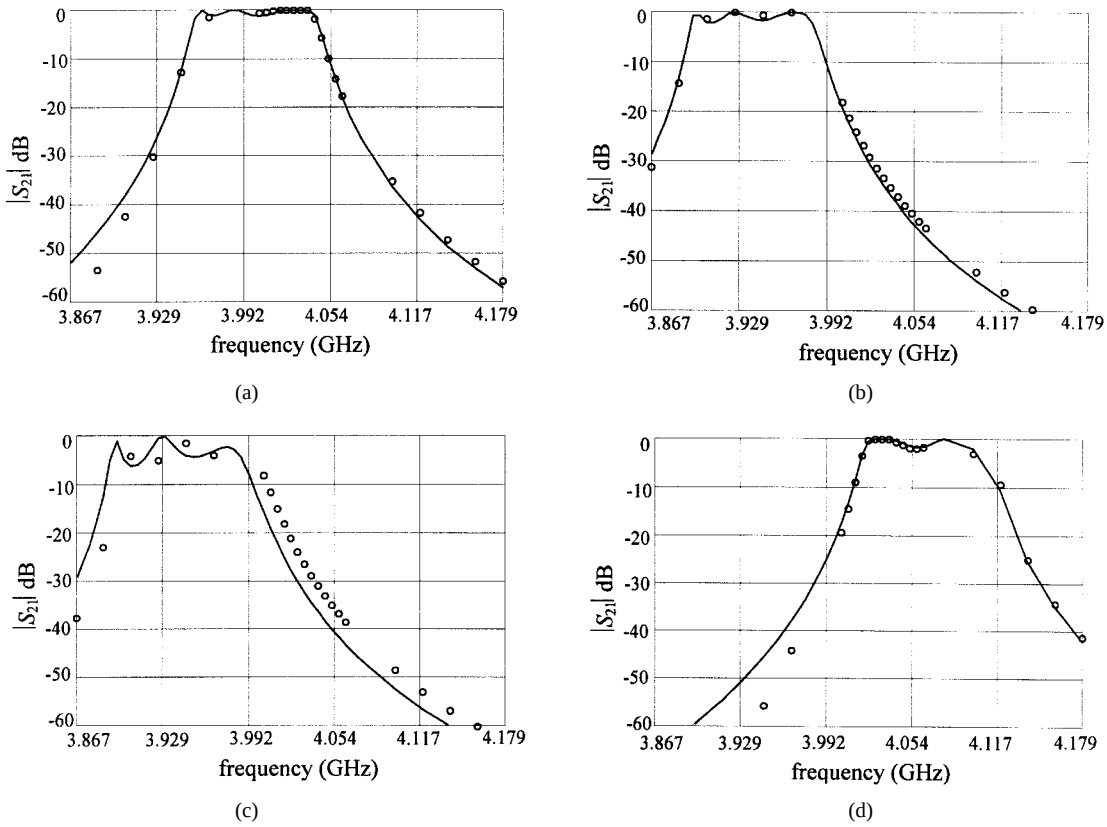

Fig. 15. The fine-model response (o) and the corresponding coarse-model response (—): (a) at the first point, (b) at the second point, (c) at the third point, and (d) at the fourth point utilized in the four-point parameter extraction for the HTS filter. Note that only points in the range from 3.967 to 4.099 GHz were actually used.

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

Parameter	$x_{em}^{(1)}$	$x_{em}^{(2)}$	$x_{em}^{(3)}$	$x_{em}^{(4)}$
L_1	181.00	182.55	181.34	179.86
L_2	201.59	205.64	205.38	197.74
L_3	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 V
THE VARIATION IN THE EXTRACTED PARAMETERS FOR THE HTS FILTER WITH THE NUMBER OF FINE-MODEL POINTS

Parameter	$x_{os}^{e(1)}$	$x_{os}^{e(2)}$	$x_{os}^{e(3)}$	$x_{os}^{e(4)}$
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_3	52.06	53.18	53.47	56.77

all values are in mils

TABLE VI
THE OPTIMAL COARSE-MODEL DESIGN FOR THE DFS FILTER

Parameter	Value
L_1	66.73
L_2	60.23
S	9.59

all values are in mils

tained using the minimax optimizer in OSA90/hope[3] according to specifications given in [9]. 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 the first column of Table IV. The point $\boldsymbol{x}_{os}^{e(1)}$ is given in Table V. Fig. 12 shows the fine-model response at $\boldsymbol{x}_{em}^{(1)}$ and the coarse-model response at $\boldsymbol{x}_{os}^{e(1)}$.

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 IV. The point $\boldsymbol{x}_{os}^{e(2)}$ is

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

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

Parameter	$x_{em}^{(1)}$	$x_{em}^{(2)}$	$x_{em}^{(3)}$	$x_{em}^{(4)}$	$x_{em}^{(5)}$	$x_{em}^{(6)}$	$x_{em}^{(7)}$	$x_{em}^{(8)}$	$x_{em}^{(9)}$
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 values are in mils

TABLE VIII
THE VARIATION IN THE EXTRACTED PARAMETERS FOR THE DFS FILTER WITH THE NUMBER OF FINE-MODEL POINTS

Parameter	$x_{os}^{e(1)}$	$x_{os}^{e(2)}$	$x_{os}^{e(3)}$	$x_{os}^{e(4)}$	$x_{os}^{e(5)}$	$x_{os}^{e(6)}$	$x_{os}^{e(7)}$	$x_{os}^{e(8)}$	$x_{os}^{e(9)}$
L_1	58.01	67.05	66.11	64.36	56.46	66.10	56.50	56.39	56.59
L_2	38.40	40.47	40.40	43.28	42.94	42.02	42.81	43.00	43.02
S	3.24	6.86	6.64	8.83	18.10	7.99	18.25	17.93	17.87

all values are in mils

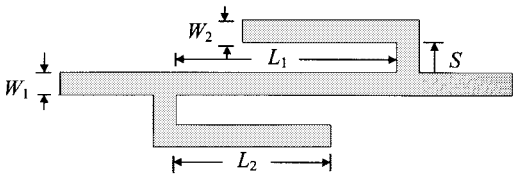

Fig. 16. The DFS filter [1].

given in Table V. Fig. 13 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- and four-point parameter extraction. The points utilized are given in Table IV. The results are shown in the fourth and fifth columns of Table V. 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. 14 and 15, respectively. Fig. 15(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.

C. Double-Folded Stub Filter

We consider the design of the double-folded stub (DFS) microstrip structure shown in Fig. 16 [1]. Folding the stubs reduces the filter area with respect to the conventional double-stub structure [10]. 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[4] through HP Empipe3D.[5]

[4]HP HFSS Version 5.2, HP EEsof Division, Santa Rosa, CA, 1998.

[5]HP Empipe3D Version 5.2, HP EEsof Division, Santa Rosa, CA, 1998.

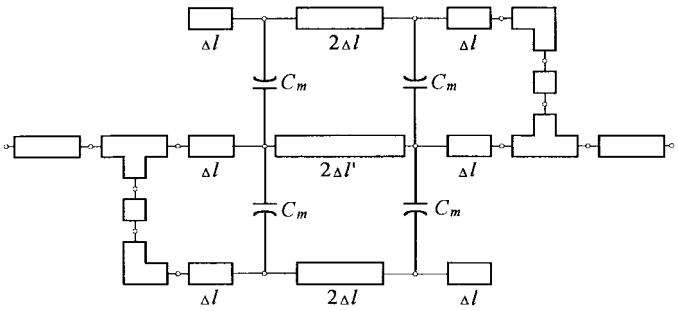

Fig. 17. The coarse model of the DFS filter.

The coarse model exploits the microstrip line and microstrip T-junction models available in OSA90/hope. 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 [11]. Jansen's microstrip bend model [12] is used to model the folding effect of the stub. The coarse model is shown in Fig. 17.

It is required in this example to extract the coarse-model parameters corresponding to the fine-model parameters given in Table VI. This vector is the optimal design of the coarse model obtained by minimax optimization.

The algorithm started by applying SPE using the fine-model point given in Table VI. Fig. 18 shows the fine-model response at $\boldsymbol{x}_{em}^{(1)}$ and the coarse-model response at the point $\boldsymbol{x}_{os}^{e(1)}$. 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 VII and the extracted coarse-model points are given in Table VIII. Fig. 19 shows the fine-model response at $\boldsymbol{x}_{em}^{(1)}$ and the coarse-model response at the point $\boldsymbol{x}_{os}^{e(9)}$.

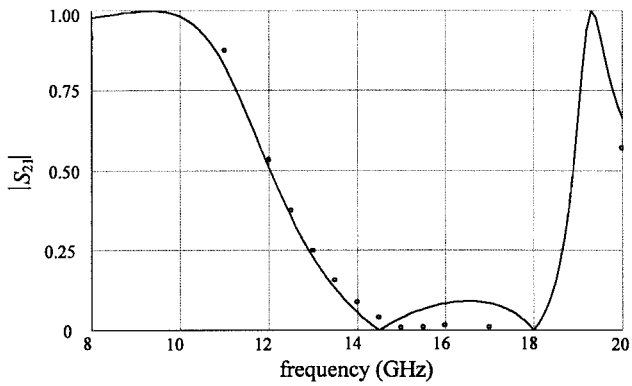

Fig. 18. The fine-model response (o) and the corresponding coarse-model response (—) at the point $\boldsymbol{x}_{os}^{e(1)}$ for the DFS filter.

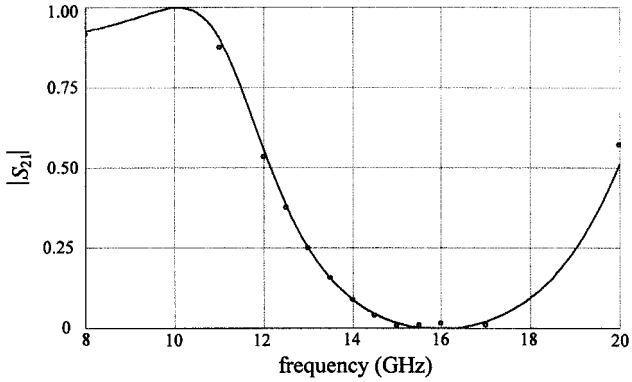

Fig. 19. The fine-model response (o) and the corresponding coarse-model response (—) at the point $\boldsymbol{x}_{os}^{e(9)}$ for the DFS filter.

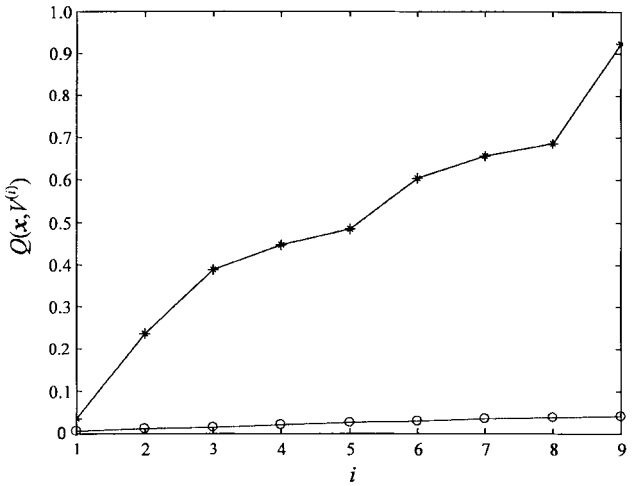

Fig. 20. The variation of $Q(\boldsymbol{x}, V^{(i)})$ for the DFS filter at the point $\boldsymbol{x}_{os}^{e(1)}$ (— $*$ —) and at the point $\boldsymbol{x}_{os}^{e(9)}$ (— o —) with the number of points utilized for parameter extraction.

Table VIII shows the large relative change in parameter values between the first set of extracted parameters $\boldsymbol{x}_{os}^{e(1)}$ and the trusted set of parameters $\boldsymbol{x}_{os}^{e(9)}$. If the step taken by any SM optimization algorithm utilizes $\boldsymbol{x}_{os}^{e(1)}$, the algorithm would have probably failed.

Fig. 20 shows the change of $Q(\boldsymbol{x}_{os}^{e(1)}, V^{(i)})$ and $Q(\boldsymbol{x}_{os}^{e(9)}, V^{(i)})$ with $|V^{(i)}|$. The value of $Q(\boldsymbol{x}_{os}^{e(9)}, V^{(i)})$ remains almost constant and small in value. On the other hand, the value of $Q(\boldsymbol{x}_{os}^{e(1)}, V^{(i)})$ increases significantly with each new point added to the set of utilized fine-model points signaling a false minimum.

VI. Conclusions

An APE algorithm has been proposed. Our APE algorithm addresses the optimal selection of parameter perturbations used to improve the uniqueness of a multipoint parameter-extraction procedure. 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.

Appendix

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

$$\boldsymbol{g}_a^{(i)T}\boldsymbol{g}^{(j)} = 0 \qquad \forall \boldsymbol{g}^{(j)} \in S \text{ and } \forall \boldsymbol{g}_a^{(i)} \in S_a. \tag{30}$$

However, each gradient in the set S_a can be approximated by

$$\boldsymbol{g}_a^{(i)} = \boldsymbol{g}^{(i)} + \boldsymbol{G}^{(i)}\Delta\boldsymbol{x}, \qquad i = k+1, \cdots, n \tag{31}$$

where $\boldsymbol{g}^{(i)}$ is the gradient of the ith response at the point $\boldsymbol{x}_{os}^{e}$ and $\boldsymbol{G}^{(i)}$ is the corresponding Hessian. It follows that the condition (30) can be restated as

$$\boldsymbol{g}^{(j)T}\boldsymbol{g}^{(i)} = -\boldsymbol{g}^{(j)T}\boldsymbol{G}^{(i)}\Delta\boldsymbol{x}, \qquad j = 1, \cdots, k; \quad i = k+1, \cdots, n. \tag{32}$$

Equation (32) is a linear equation in n unknowns (the components of $\Delta\boldsymbol{x}$). There are $k(n-k)$ such linear equations. Putting these equations into a matrix form, we have

$$\boldsymbol{B}\Delta\boldsymbol{x} = -\boldsymbol{c} \tag{33}$$

where the mth row of the matrix $\boldsymbol{B}$ is

$$\boldsymbol{B}^{(m)T} = \boldsymbol{g}^{(j)T}\boldsymbol{G}^{(i)}, \qquad i = k+1, \cdots, n; \quad j = 1, \cdots, k \tag{34}$$

where $m = (i-k-1)k + j$. Similarly, the mth component of the vector $\boldsymbol{c}$ is

$$\boldsymbol{c}_{m,1} = \boldsymbol{g}^{(j)T}\boldsymbol{g}^{(i)}, \qquad i = k+1, \cdots, n;\ j = 1, \cdots, k. \tag{35}$$

Thus, (13) follows.

ACKNOWLEDGMENT

The authors thank Sonnet Software Inc., Liverpool, NY, for making ***em*** available for this work. The authors 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, pp. 2536–2544, Dec. 1994.

[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, pp. 2874–2882, Dec. 1995.

[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, 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, pp. 1282–1293, Dec. 1986.

[5] 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.

[6] J. W. Bandler and A. E. Salama, "Fault diagnosis of analog circuits," *Proc. IEEE*, vol. 73, pp. 1279-1325, Aug. 1985.

[7] R. Fletcher, *Practical Methods of Optimization*, 2nd ed. New York: Wiley, 1987.

[8] J. W. Bandler, "Optimization methods for computer-aided design," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 533–552, Aug. 1969.

[9] 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 Millimeter-Wave Computer-Aided Eng.*, vol. 5, pp. 331–343, Sept. 1995.

[10] J. C. Rautio, private communication, 1992.

[11] C. S. Walker, *Capacitance, Inductance and Crosstalk Analysis.* Norwood, MA: Artech House, 1990.

[12] 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) was born on November 7, 1969. He received the B.Sc. degree (with distinction) in electronics and communications engineering and the Master's degree in engineering mathematics from Cairo University, Cairo, Egypt, in 1992 and 1996, respectively, and is currently working toward the Ph.D. degree at McMaster University, Hamilton, Ont., Canada.

In October 1992, he joined the Department of Engineering Mathematics and Physics, Faculty of Engineering, Cairo University. In September 1996, he joined the Department of Electrical and Computer Engineering, McMaster University. His research is performed in the Simulation Optimization Systems Research Laboratory. He is interested in optimization methods, computer-aided design and modeling of microwave circuits, and neural network applications.

Mr. Bakr has held the Ontario Graduate Scholarship (OGS) for two consecutive years.

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) received the Diploma in Engineering degree in electronics from the Technical University of Varna, Varna, Bulgaria, in 1989, and the Doctor of Engineering degree from the University of Electro-Communications, Tokyo, Japan, in 1997.

From 1997 to 1998, she was a Post-Doctoral Fellow in the Department of Electrical and Computer Engineering, Dalhousie University, Halifax, N.S., Canada, where she was involved in teaching and research projects concerning time-domain electromagnetic (EM) field simulation and modeling of RF/microwave multilayer passive structures, printed circuit boards, and printed antennas. In May 1998, she joined the Simulation Optimization Systems Research Laboratory, McMaster University, Hamilton, Ont., Canada. In 1999, she joined the Department of Electrical and Computer Engineering, McMaster University, as an Assistant Professor. Her research interests include numerical techniques for full-wave electromagnetic analysis and modeling as well as the implementation of efficient optimization algorithms to microwave design problems.