Accuracy Control in the Optimization of Microwave Devices by Finite-Element Methods

Minya M. Gavrilovic and Jon P. Webb, *Member, IEEE*

***Abstract*—Automatically optimizing the design of a microwave device can be prohibitively time-consuming when a numerical electromagnetic-field analysis is necessary at each iteration. However, the time taken for the field analysis depends on the accuracy required, and in the early stage of the optimization relatively inaccurate solutions are adequate. This idea is exploited in a scheme that combines a quasi-Newton constrained optimizer with a two-dimensional p-adaptive finite-element method for finding scattering parameters. The scheme has been tested on three H-plane rectangular waveguide devices: a T-junction, a miter bend with a dielectric column, and a two-cavity iris-coupled filter. Time savings of more than an order of magnitude were obtained, compared to the standard approach of requiring equally high accuracy throughout the optimization.**

***Index Terms*—Design automation, electromagnetic scattering, finite-element methods, microwave devices, optimization methods.**

I. Introduction

VARIOUS methods of computer-based electromagnetic-field analysis are now able to provide accurate predictions of performance of many microwave devices, even those that are irreducibly three-dimensional (3-D) and have no simplifying symmetries. However, in many cases, the computation times involved are great, i.e., hours rather than minutes or seconds on typical desktop computers. When these analysis techniques are embedded in an automatic optimizer, which tries to improve the performance of a design iteratively, the run times can become prohibitive. This has prompted the development of a variety of optimization techniques aimed at reducing to a bare minimum the number of times a full electromagnetic-field analysis has to be performed [1], [2].

One question that is rarely addressed in the optimization literature is the accuracy with which each individual field analysis is performed. The assumption is usually made that the field analysis is a "black box" that just gets the right answer. In reality, of course, any computational-field analysis method gives only an approximate answer. Moreover, there is always a tradeoff between accuracy and time: to get more accuracy, a longer computation time is required. A sensible approach to any field analysis, then, begins by asking the question: How accurate do I need the answers to be? There is no point in waiting an extra hour for an extra 0.1-dB accuracy when the quantity in question is only required (perhaps can only be measured) to within 1 dB. The idea behind this paper is that an optimizer too should be asking this question because, by doing so, it might greatly reduce the amount of computation time needed.

We shall consider specifically *direct* optimization, in which, from a single starting design, a path is followed from point to point in the design space, seeking a minimum of a cost function (generally some function of the scattering parameters of the device obtained by field analysis). Direct optimization is well suited to microwave device design, in which there is often a fairly good initial configuration available from approximate synthesis techniques and what is needed are relatively small alterations of the design parameters to bring the cost function to a nearby local minimum. Unlike stochastic methods, direct methods do not generally find a global minimum, but they take considerably fewer cost-function evaluations, particularly when gradient information is available.

In the early stages of a direct optimization, the cost function need not be very precise. It just has to be accurate enough to guide the optimizer in the right direction. As the local minimum is approached, the accuracy must be progressively improved. The goal is to obtain a final design that is as accurate as if the maximum precision had been required at every stage of the optimization, but more quickly.

In principle, this approach would work with any computational-field analysis technique, provided only that there is some mechanism for accuracy control, i.e., for increasing the accuracy of the results at the expense of longer run times. In the present study, we have used the frequency-domain finite-element method (FEM) because it is a well-developed technique for computing scattering parameters [3]–[5] and, particularly, because there is a body of literature on *adaptive* FEMs for electromagnetic-wave problems [6]–[10]. Adaptive methods involve a sequence of analyses of the same problem with increasing numbers of degrees of freedom (DOFs) and, therefore, increasing computational cost, until a pre-specified accuracy level is achieved. A further advantage of the FEM is that it can provide gradient information—the slope of the cost function with respect to each design parameter—at almost no extra cost [11]. This information is used by a gradient-based optimizer—specifically, a quasi-Newton constrained optimizer.

II. Adaptive FEM for Computing the Cost Function and Its Gradient

The cost function C is assumed to be a known function of the entries of the scattering matrix $\mathbf{S}$ of the device. In general, these are frequency-dependent; in this study, we assume that C is of the form $C = \sum_i w_i c_i$, where w_i are specified weight functions, c_i are known functions of the scattering parameters

Manuscript received April 2, 2001. This work was supported by the Natural Sciences and Research Council of Canada under a grant.

M. M. Gavrilovic is with EMS Technologies, Ste. Anne de Bellevue, QC, Canada H9X 3R2.

J. P. Webb is with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada H3A 2A7.

Publisher Item Identifier 10.1109/TMTT.2002.801329.

at frequency f_i, and the summation is over a number of discrete frequencies in the range of interest. In order to calculate c_i, we first need to calculate $\mathbf{S}$ at a single frequency.

For an N-port microwave device that is excited by the dominant mode at port j, the pth column of $\mathbf{S}$ can be calculated from the resulting electric field $\vec{E}^{(j)}$ inside the device using mode orthogonality, which leads to

$$S_{ij} = \gamma_0^{(i)}\left(\vec{E}^{(j)}\right) - \delta_{ij} \tag{1}$$

where δ_{ij} is the Kroenecker delta and $\gamma_l^{(i)}$ is a linear operator extracting the voltage of mode l at port i from the electric field [11]; subscript 0 denotes the dominant mode.

The FEM can be used to compute $\vec{E}^{(j)}$ by solving the modified weighted-residual equation

$$B\left(\vec{E}^{(j)}, \vec{w}\right) = R(\vec{w}) \tag{2}$$

for all weight functions $\vec{w}$, where B is the bilinear form

$$\begin{aligned} B\left(\vec{\boldsymbol{E}}^{(j)}, \vec{w}\right) &= \frac{1}{jk_o\eta_o}\int_\Omega \left\{\nabla\times\vec{\boldsymbol{E}}^{(j)}\cdot\frac{1}{\mu_r}\nabla\times\vec{w} - k_o^2\vec{\boldsymbol{E}}^{(j)}\cdot\varepsilon_r\vec{w}\right\}d\Omega \\ &\quad + \sum_{p=1}^{N}\sum_{l=0}^{\infty}\gamma_l^{(p)}\left(\vec{\boldsymbol{E}}^{(j)}\right)\gamma_l^{(p)}(\vec{w}) \end{aligned} \tag{3}$$

where Ω is the volume inside the device, ε_r is the relative permittivity, μ_r is the relative permeability, and k_o and η_o are the free-space wavenumber and intrinsic impedance, respectively. R is the linear function

$$R(\vec{w}) = 2\gamma_0^{(j)}(\vec{w}). \tag{4}$$

In order to solve the problem numerically, the unknown field must be represented by a finite number of DOF—in this case, with a mesh of nonoverlapping FEs filling the interior of the device. The basis functions that approximate the field distribution in each element allow for computation of $\vec{E}^{(j)}$ at any point in the problem domain. Representing $\vec{E}^{(j)}$ in this way, and setting the weight function to each basis function in turn in (2)–(4) results in a large, but sparse matrix equation for $\mathbf{e}^{(j)}$, a vector of unknown coefficients of electric field

$$\mathbf{K}\mathbf{e}^{(j)} = \mathbf{b} \tag{5}$$

where $\mathbf{K}$ is the global finite-element (FE) matrix and $\mathbf{b}$ is a known vector representing the sources of the problem and derived from R.

A point of interest with regards to optimization is that a derivative of a scattering parameter can then be efficiently calculated with the adjoint variable method [12]. It has been shown [11] that, for a scattering parameter S_{ij}, the adjoint variable conveniently becomes just the field solution when port i is excited and the derivative with respect to a design parameter g becomes

$$\frac{dS_{ij}}{dg} = -\frac{1}{2}\,\mathbf{e}^{(i)^{\mathrm{T}}}\,\frac{d\mathbf{K}}{dg}\,\mathbf{e}^{(j)}. \tag{6}$$

From (6), any derivative of a scattering parameter can be computed directly from the N field solutions needed to compute $\mathbf{S}$. No additional system of equations needs solving—thus providing cheap and efficient calculation of ∇S_{ij}. Furthermore, since any cost function is defined to be a function of S-parameters, computing ∇C becomes trivial.

Since any numerical method is an approximation to the actual solution, it will contain a certain error. The solution accuracy is a function of the number of DOFs in the problem. Increasing the number of DOFs reduces the error in the solution. Increasing DOFs in a FE problem is commonly achieved by subdividing elements or by approximating the field solution in each element with higher order basis functions. Regardless of the method, the computational benefits of concentrating DOFs in certain areas, adaptively, are well established. There are three key components to any adaptive algorithm: error estimation, error indication, and the refinement method.

An error estimator estimates the error in the quantity of interest in the problem for a particular mesh. Two common termination criteria are to adapt until the estimated field error in each element has been reduced to a certain tolerance or to adapt until the estimated error in a global quantity of interest has reached a certain tolerance.

An error indicator [13] assesses the relative error in each FE for use by the refinement algorithm in choosing where to add DOFs. An indicator does not have to provide an absolute estimate of the error. It only has to assess the error *relative* to the rest of the elements in the mesh. Note also that an indicator may be an assessment of the local contribution to a *global* quantity [14].

Refinement techniques are methods for adding DOFs to one or many elements in a mesh in order to increase the accuracy of the overall solution. An example of a refinement technique is to refine a fixed number of elements at each step (e.g., 25% of all elements). The elements refined are those with the highest errors, as predicted by the error indicator. As mentioned earlier, to increase the accuracy of the solution, one must increase the DOFs by either adding more elements or increasing the polynomial orders of existing elements in the mesh. The process of subdividing an element is referred to as h-type adaption [15]. P-type adaption (or p-adaption) is a method based on increasing the orders of elements to increase the accuracy of the solution [16], [17]. In order to allow mixing of the orders of elements in a mesh, basis functions in each element must be hierarchical. In this study, p-adaption is used with the hierarchical elements of [18].

The adaption is performed for a single frequency, the adaption frequency f_a. This is chosen to be the highest frequency in the range of interest in order to adapt with respect to the shortest wavelength—thus ensuring higher accuracy than adapting at lower frequencies. Once the estimate of error is below a certain tolerance, the adaptive process stops and a "post-adaptive" frequency sweep is then used (with the same distribution of DOFs as yielded by the adaptive procedure) to calculate the cost function over a range of frequencies. The frequency sweep requires additional FE solutions at a number of discrete frequency points. An alternative would be to use a technique that finds higher

derivatives of the field with respect to frequency, and builds the frequency response in this way [19]–[23].

For each adaptive iteration, 25% of all the elements are increased in order by one. Since accurate gradient information is necessary for the types of optimizers used in this study, the error estimator used for termination of the adaption is based on the gradient of the cost function. The derivative of a cost function with respect to a geometric parameter g can be expressed as a sum of derivatives at nodes on all boundaries and interfaces parameterized by g. A good estimate of the error in this derivative can be found from the rate of change in the cost function due to the perturbation of *internal* nodes of the mesh (internal nodes being those whose movement do not alter the geometry of the problem) [24]. For various cost functions (defined for a single frequency), good estimates of the errors in their gradients can be computed with little computational effort [24].

Error indication is based on refining c_a, the cost function at the adaption frequency. It has been shown that, in problems where a global quantity (such as a cost function based on S-parameters) is desired rather than the field itself, a targeted error indicator (TEI) is the better choice [14]. From the framework in [14], a TEI for microwave devices can be derived.

Quantities $I_1, \ldots, I_V$ are defined to be the real and imaginary parts of the complex scattering parameters such that

$$I_1, \ldots, I_V = S^r_{11}, S^i_{11}, S^r_{12}, S^i_{12}, \ldots, S^r_{NN}, S^i_{NN} \tag{7}$$

where $V = 2N^2$. The estimated errors in $I_1, \ldots, I_V$ are

$$\Delta I_1, \ldots, \Delta I_V = \Delta S^r_{11}, \Delta S^i_{11}, \Delta S^r_{12}, \Delta S^i_{12}, \ldots, \Delta S^r_{NN}, \Delta S^i_{NN} \tag{8}$$

where ΔS_{ij} is an estimate of the change in the complex S-parameter when the kth DOF is zeroed. An error indicator for the kth DOF, based on global quantity c_a, is then defined as follows:

$$\varepsilon^k = \left| \sum_i \frac{\partial c_a}{\partial I_i} \Delta I_i \right| = \left| \frac{\partial c_a}{\partial S^r_{11}} \Delta S^r_{11} + \frac{\partial c_a}{\partial S^i_{11}} \Delta S^i_{11} + \cdots + \frac{\partial c_a}{\partial S^i_{NN}} \Delta S^i_{NN} \right| \tag{9}$$

where the terms of the sum depend on the choice of cost function.

An error indicator for a pth-order element can now be defined as a sum of the contributions of ε^k for all DOFs k of that element that are of order p (i.e., not contained in the element of order $p-1$)

$$\varepsilon = \sum_{p\text{th order DOFs}} \varepsilon^k. \tag{10}$$

While it may seem computationally expensive to compute ΔS_{ij} a number of times for each element in order to evaluate the indicator of (10), the programming is quite simple and the computational costs are low. For calculation of S-parameters, (1) can be alternately written in terms of the bilinear form

$$S_{ij} = \frac{1}{2jk_o\eta_o} B\left(\vec{E}^{(i)}, \vec{E}^{(j)}\right) - \delta_{ij}. \tag{11}$$

Introducing basis functions reduces (11) to

$$S_{ij} = \frac{1}{2jk_o\eta_o} \mathbf{e}^{(i)^{\mathrm{T}}} \mathbf{K} \mathbf{e}^{(j)} - \delta_{ij}. \tag{12}$$

From (12), we see that the computational expense in calculating an S-parameter comes from the matrix calculation of the general form

$$F = \mathbf{u}^{\mathrm{T}} \mathbf{K} \mathbf{v} \tag{13}$$

where $\mathbf{u}$ and $\mathbf{v}$ are a pair of FE solutions and $\mathbf{K}$ is the $n \times n$ global matrix. Function F can be decomposed into two parts as follows:

$$F = F^z_k + \Delta F_k \tag{14}$$

where F^z_k is F with its kth DOF zeroed, i.e.,

$$F^z_k = \sum_{\substack{i=1 \\ i \neq k}}^{n} \sum_{\substack{j=1 \\ j \neq k}}^{n} u_i K_{ij} v_j \tag{15}$$

and ΔF_k can be computed by

$$\Delta F_k = \sum_{j=1}^{n} u_k K_{kj} v_j + \sum_{\substack{i=1 \\ i \neq k}}^{n} u_i K_{ik} v_k. \tag{16}$$

An efficient algorithm that uses (16) to calculate all changes, i.e., $\Delta F_1, \ldots, \Delta F_n$, is

$i = 1, \ldots, n$:
  $\Delta F_i = 0$

$i = 1, \ldots, n$:
  $j = 1, \ldots, n$:
    If $K_{ij} = 0$, skip (to account for the sparsity of K_{ij})
    $\Delta F_i = \Delta F_i + u_i K_{ij} v_j$
    If $i \neq j$:
      $\Delta F_j = \Delta F_j + u_i K_{ij} v_j$

This algorithm computes $\Delta F_1, \ldots, \Delta F_n$ with no more computational cost than computing $\mathbf{u}^{\mathrm{T}} \mathbf{K} \mathbf{v}$. Using this approach, the estimates of the changes of the complex S-parameters when each DOF is zeroed, i.e., $\Delta S_{11}, \ldots, \Delta S_{NN}$, need only be computed *once*, globally, and stored for subsequent use. Any estimate of ΔS_{ij}, needed in (9) for DOF k, is readily available from global quantities.

III. Optimization Technique

Let C be the cost function to be minimized and $\mathbf{g}$ be a column vector of d real-valued design variables. Since, in general, it is desirable to satisfy constraints on the design variables during the optimization, the problem to be solved is

$$\begin{aligned} &\underset{\mathbf{g} \in \Re^d}{\text{minimize}} \quad C(\mathbf{g}) \\ &\text{subject to:} \quad h_i(\mathbf{g}) = 0, \qquad i = 1, \ldots, m' \\ &\qquad\qquad\quad h_i(\mathbf{g}) \leq 0, \qquad i = m'+1, \ldots, m \end{aligned} \tag{17}$$

where m is the total number of constraints, m' of which are equality constraints. The solution $\hat{\mathbf{g}}$ of (17) can be shown to satisfy the Kuhn–Tucker equations

$$\begin{aligned} \nabla L\left(\hat{\mathbf{g}}, \hat{\boldsymbol{\lambda}}\right) &= 0 \\ \hat{\lambda}_i h_i(\hat{\mathbf{g}}) &= 0, \qquad i = 1, \ldots, m \\ \hat{\lambda}_i &\geq 0, \qquad i = m'+1, \ldots, m \end{aligned} \tag{18}$$

where L is the Lagrangian function

$$L(\mathbf{g}, \boldsymbol{\lambda}) = C(\mathbf{g}) + \sum_{i=1}^{m} \lambda_i h_i(\mathbf{g}) \tag{19}$$

and $\boldsymbol{\lambda}$ is an m-vector of Lagrange multipliers, taking the value $\hat{\boldsymbol{\lambda}}$ at the optimum. The approach taken by direct optimizers is to find a solution to (18) that is *locally* minimal; in general, such a point will not also be a global minimum. In this study, the MATLAB toolbox (function "constr.m") was used to do the constrained optimization. This solves (18) for a local minimum using sequential quadratic programming (SQP).

The details are in [25]. At each step of SQP, a new search direction is generated by solving a simpler quadratic-programming (QP) subproblem. To set up this subproblem requires the latest estimate of $\mathbf{g}$ and the gradient ∇C at this point. Once a new search direction is found, a line search determines the point along that direction that has the smallest value of C consistent with the constraints. Another QP subproblem is then set up.

There are a number of ways this iteration can be terminated. Perhaps the simplest method is to stop when the change in C or $\mathbf{g}$ from one step to the next is sufficiently small, but this can be unreliable because, occasionally, there will be a small change even a long way from the local optimum. The method used in this study is to terminate when the gradient of the Lagrangian is small enough since, according to (18), this gradient should be identically zero at the optimum. "Small enough" we take to mean a sufficient reduction for the initial value, i.e., stop when

$$\left\|\nabla L^{k+1}\right\|_2 < \frac{1}{R}\left\|\nabla L^{\text{initial}}\right\|_2 \tag{20}$$

where superscript k denotes the value at the end of the kth line search and R is the required reduction factor. A value of 1000 was used in this study.

IV. Controlling Accuracy During Optimization

Many optimization problems deal with cost functions that can be determined analytically. In such cases, the cost function is in effect "infinitely" accurate and the accuracy of cost-function evaluations (CFEs) is not a concern. In problems where a CFE is calculated using a numerical approximation, the accuracy of the CFE becomes an issue.

Two quantities that must be chosen in a design optimization are the accuracy of the optimization process itself (termination criteria) and the accuracy of the cost function at the optimum. One way of ensuring that an optimizer finds an accurate optimum is to perform every CFE as accurately as that of the final CFE. However, heavy computation can be avoided by using CFEs of lower accuracy for steps where $\mathbf{g}^k$ is far from the optimum. The early steps of an optimization scheme may not require CFEs that are very accurate. As the process continues and the parameters near an optimum, every CFE can become increasingly accurate. The design space of the cost function in such a scenario is dynamic in nature and is allowed to change (rather than stay fixed) throughout the optimization until it gets close to an optimum, where the changes in the design space become very small.

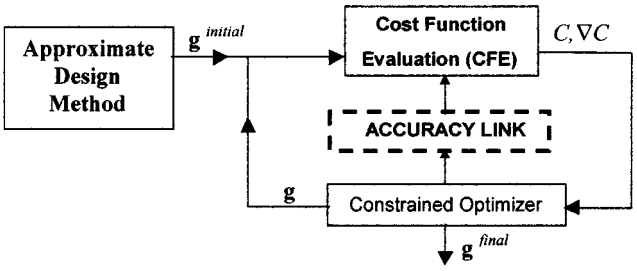

Fig. 1. Block diagram of an optimization scheme with an accuracy link.

An automated optimization scheme links the optimizer to each CFE in order to control the level of accuracy of the CFE throughout the optimization (see Fig. 1) instead of demanding a fixed accuracy.

In the present case, the CFE is an adaptive FE analysis whose accuracy is dependent on an error estimate. When the error estimated in the FE solution at an adaptive step is below a certain tolerance, the adaptive process stops. The accuracy link can increase the accuracy of the CFE by reducing the error tolerance.

A quantity that decreases throughout an optimization is the gradient of the cost function (or, rather, the gradient of the Lagrangian ∇L for constrained problems). At an optimum, theoretically, $\|\nabla L\|_2 = 0$. We then require of the CFE that, for a given positive number α

$$\left\|\text{estimated error in } \nabla L^k\right\|_2 < \alpha\left\|\nabla L^{k-1}\right\|_2. \tag{21}$$

As the optimization progresses, this holds the percentage error in $\|\nabla L\|_2$ fixed at a level determined by α. As $\|\nabla L^k\|_2$ reduces with an increase in k, this has the effect of increasing the accuracy of the FE solution as an optimum is approached. Of course, if α is set too small, it may not be possible for the CFE algorithm to satisfy (21). However, in any cases, a value that is too small will lead to unnecessarily long computation times, as demonstrated by the results below. A practical range for alpha is between 0.01–1; a value of 0.1 was found to work well for the examples tried to date.

For problems involving analysis over a range of frequencies, the cost function and its gradient (and, in turn, $\|\nabla L^k\|_2$) is calculated as a sum over a number of discrete frequency points. Since the adaptive process is performed for a single frequency, the estimate of the error in $\|\nabla L^k\|_2$ is available only at the adaption frequency f_a. Requirement (21) is approximated by

$$N_f\left\|\text{estimate of error in } \nabla L_a^k\right\|_2 < \alpha\left\|\nabla L^{k-1}\right\|_2 \tag{22}$$

where ∇L_a^k is the gradient of the Lagrangian of the single frequency cost function c_a and N_f is the number of frequency points sampled. Since ∇L_a^k is calculated by

$$\nabla L_a^k = \nabla c_a^k + \sum_{i=1}^{m} \lambda_i^k \nabla h_i^k \tag{23}$$

from (22) we have

$$N_f \left\| \text{est. error in } \left(\nabla c_a^k\right) + \text{est. error in } \left(\sum_{i=1}^{m} \lambda_i^k \nabla h_i^k\right) \right\|_2 < \alpha \left\| \nabla L^{k-1} \right\|_2 . \tag{24}$$

Since we do not have error estimates for the Lagrange multipliers, we drop this term and require simply that

$$N_f \left\| \text{estimate of error in } \nabla c_a^k \right\|_2 < \alpha \left\| \nabla L^{k-1} \right\|_2 . \tag{25}$$

A good estimate for the error in ∇c_a is available from the FE solution, as mentioned above and explained in [24]. In order to smooth out any potential discontinuities due to large changes in $\|\nabla L^k\|_2$, a weighted-average approach is used and (25) becomes

$$\Delta_{\text{est}}^k < \alpha \left(w \Delta_{\text{tol}}^{k-1} + (1-w) \Delta_{\text{tol}}^{k-2} \right) \tag{26}$$

where $\Delta_{\text{est}}^k = \|\text{estimate of error in } \nabla c_a^k\|_2$, $\Delta_{\text{tol}}^k = \|\nabla L^k\|_2 / N_f$ and the weight w was typically chosen to be 0.9.

The present CFE has two parts. The first part is the p-adaptive process, where the FE solution adapts at a single frequency until (26) is satisfied. (The error indicator used for the adaption is (10), targeted toward the single frequency cost function c_a.)

After the adaption is complete, the second part of the FE analysis is a post-adaptive frequency sweep over the N_f points performed with the same distribution of DOFs as the final step of the adaption. The adaptive frequency is available from the first part and does need to be recalculated, thus, the frequency sweep is actually performed over $N_f - 1$ points. The result of the frequency sweep is the cost function and its gradient—the desired quantities needed as input to the optimizer.

The adaption need not necessarily be performed over a single frequency. Alternatively, a frequency sweep could be performed at each adaptive step to calculate the full gradient ∇C^k. However, this would be very costly computationally and the approximation $\|\nabla c_a^k\|_2 \cong \|\nabla C^k\|_2 / N_f$ works adequately in practice.

A subtlety that was encountered in testing was the phenomenon of "diagonal flipping." For small changes in geometry, the structure of the mesh might change. While the change may be as small as a diagonal being flipped between two adjacent triangles (Fig. 2), this variation in the FE solution will make the numerical cost-function discontinuous (see Fig. 3).

Small discontinuities in C throughout the optimization are usually acceptable, except near termination when $\mathbf{g}^k$ is changing by small amounts and the gradients are small. Here, there can be an adverse effect on convergence. A certain level of continuity near local optima can be preserved by avoiding re-meshing; instead all nodes affected by that change are dragged in the appropriate direction by the amount $\Delta \mathbf{g}^{k+1}$. Forcing the same mesh configuration for large changes in geometry can create poorly shaped elements; thus, dragging the nodes replaces re-meshing only when the maximum change in a parameter is less than 10%. This was found to work well in practice.

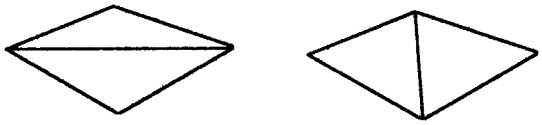

Fig. 2. Diagonal flipping between adjacent triangular elements.

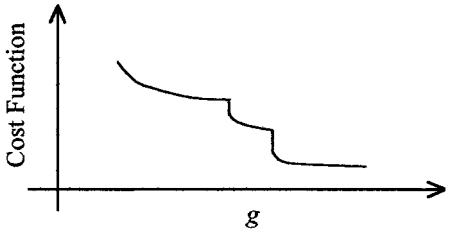

Fig. 3. Discontinuous cost function.

The accuracy of a CFE is only changed at main optimization steps. Line-search CFEs are kept at a constant level of accuracy as far as possible, i.e., using the same right-hand side in (26). Varying the accuracy of line-search CFEs changes the target of the bisection algorithm (the line-search method used). Attempts to reduce the cost function in such cases can lead to a large number of line-search steps and possible nonconvergence.

The optimization–adaption system is illustrated as a block diagram in Fig. 4. In order to allow the first CFE of the optimization to be adaptive, a value for $\|\nabla L^0\|_2$ is required. To provide a rough reference value, a pre-optimization CFE is performed nonadaptively at a uniform low order (chosen to be second order). The geometry for this pre-optimization step is the same as in the first optimization step. The approximation is made that $\|\nabla L^0\|_2$ is approximately equal to the pre-optimization computation of $\|\nabla L^1\|_2$. This quick uniform-order CFE adds little to the overall computational cost, yet allows adaptive solving for the first CFE of the optimization.

V. Results

We optimize three H-plane rectangular waveguide junction models: a T-junction with an inductive post, a mitered right-angled bend with a dielectric column, and a two-cavity iris-coupled filter. Three optimizations with different accuracy controls are compared for each problem: $\alpha = 0.1$, $\alpha = 0.01$ and a third benchmark case. The benchmark optimization requires each CFE to be of the same accuracy as the CFE at termination. The termination criterion (20) guarantees that the final gradient of the Lagrangian (when the optimization terminates) is $1/R$th of its initial value. Combining (20) and (25) gives the accuracy link for any CFE for the benchmark or "fixed accuracy" case

$$N_f \left\| \text{estimate of error in } \nabla c_a^k \right\|_2 < \frac{\alpha}{R} \left\| \nabla L^{\text{initial}} \right\|_2 \tag{27}$$

where $\alpha = 0.01$.

In comparing the costs of optimizing the three different devices using the system described by Fig. 4, cumulative floating-

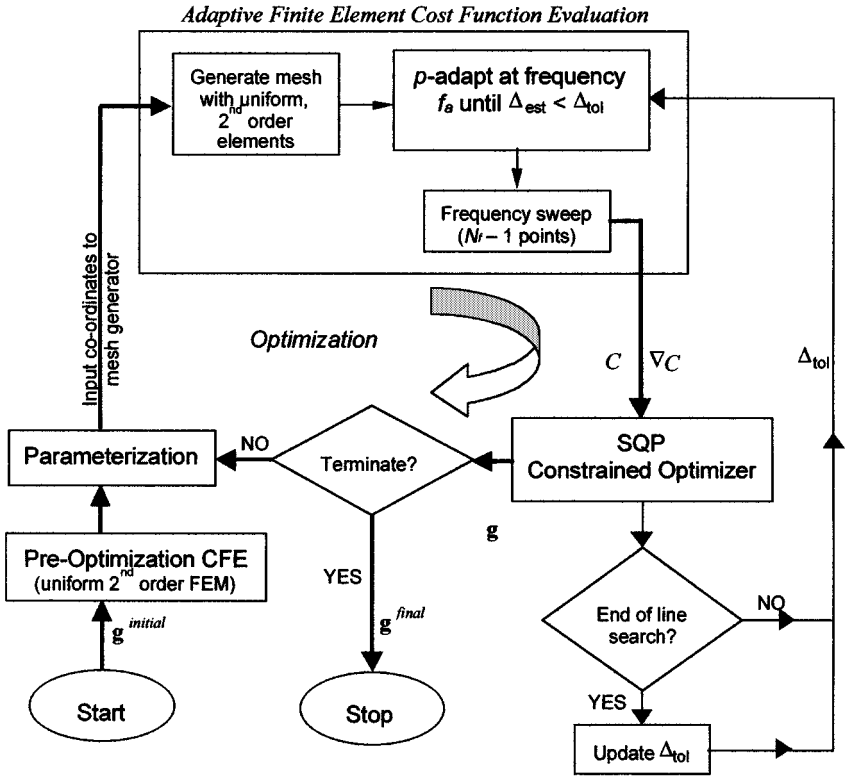

Fig. 4. Optimization–adaption block diagram.

point operation counts are used rather than CPU times. In a typical FEM, most of the computational effort lies in solving the matrix equation. The size of the matrix equation depends on the total number of DOFs (n) for that solution. In two dimensions, the computational cost of solving a matrix equation by an efficient sparse direct method, such as the frontal method [26], is roughly proportional to n^2 for large n. Since the computational cost of the optimization program, parameterization, and meshing is negligible compared to the cost of the matrix equations solved during a CFE, $O(n^2)$ is taken as approximately the total computational cost of a single optimization step.

The computational cost of any CFE then is based on the cumulative cost of solving a number of matrix equations. An N-port device requires N FE solutions per frequency. In the p-adaptive FE solution, if N_a adaptive steps are taken to converge the solution at a single frequency and the subsequent frequency sweep requires an additional $N_f - 1$ solutions, the cost of one CFE is

$$N\left(\sum_{i=1}^{N_a}\left(n[i]\right)^2 + (N_f - 1)\left(n[N_a]\right)^2\right) \tag{28}$$

where $n[i]$ is the number of DOFs of the ith adaptive step. The overall cost of an optimization is taken to be the sum of the computational costs of the CFEs.

A. Waveguide T-Junction With Inductive Post

In this example (Fig. 5), the return loss at the input (port 1) is optimized with the use of a symmetrical inductive septum (or post). Reflection can be minimized by varying the dimensions

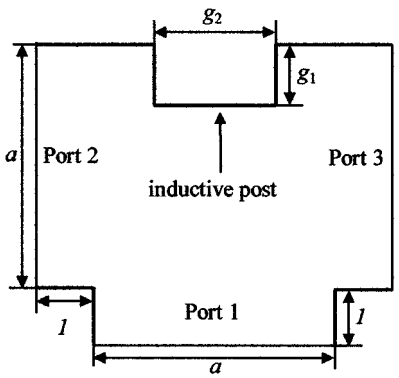

Fig. 5. Initial geometry and dimensions of the waveguide T-junction. $a = 2$ cm, $l = 0.1\lambda = 0.23$ cm, $g_1 = 0.5$ cm, $g_2 = 1$ cm.

of the inductive post to help the incident wave split and change direction by 90° (i.e., maximizing the transmission to ports 2 and 3) [27], [28]. The initial geometry of the model and its geometric parameters are given in Fig. 5.

The cost function for the problem is defined as

$$C(\mathbf{g}) = \sum_{i=1}^{N_f} c_i \tag{29}$$

where the single-frequency cost function is

$$c_i = |S_{11}(f_i)|^2. \tag{30}$$

The cost function calculation samples $N_f = 5$ discrete points between 8–12 GHz. The frequency of adaption is taken to be the center frequency $f_a = 10$ GHz.

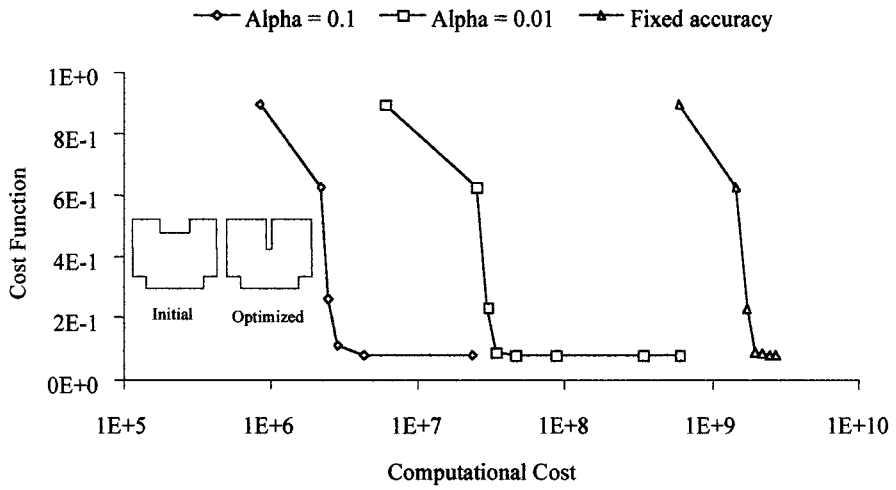

Fig. 6. "True" values of the cost function and the corresponding computational costs for different accuracy links in the optimization of the T-junctions.

For any problem, it is important to choose constraints that limit the geometric parameters to a range where any geometry of the device within that range is physically possible. A poorly constrained optimization problem (or an unconstrained one) may allow edges or boundaries to overlap—making it physically impossible to build and a useless design. A well-constrained optimization problem will converge more quickly than a poorly constrained or unconstrained problem. Although the constraints for the inductive post are quite simple, the choice of upper and lower limits must be made. An additional concern in the optimization of a device is meshing. While a particular geometry may be feasible (in a physical sense), meshing difficulties can easily arise when modeling the device. Edges too close together or too small in themselves may force the generation of many small elements—which may not be necessary. To prevent this, a minimum "cushion" of space was left between any two boundaries—thus affecting the choice of bounds for the parameters. For the T-junction meshes, a cushion of 0.1 cm worked well. The lower limits for both parameters are taken to be 0.1 cm. The upper limit for g_1 (1.9 cm) was chosen because at $g_1 = 1.9$ cm, the bottom edge of the post is 0.1 cm shorter than the port width of 2 cm. The upper limit of g_2 was chosen because when g_2 is at its upper limit, the left and right edges of the post are 0.1 cm away from ports 2 and 3, respectively. Four inequality constraints of the form $h_i(\mathbf{g}) \leq \mathbf{0}$ are

$$\begin{aligned} h_1(\mathbf{g}) &= g_1 - 1.9 \\ h_2(\mathbf{g}) &= 0.1 - g_1 \\ h_3(\mathbf{g}) &= g_2 - 2.70712 \\ h_4(\mathbf{g}) &= 0.1 - g_2. \end{aligned} \tag{31}$$

Table I gives the final computational cost for each optimization and a speed-up factor in each case. The speed-up factor is relative to the cost of the benchmark fixed-accuracy optimization. The results in Table I show the impressive speed-up that can be attained by controlling accuracy using α. In the case $\alpha = 0.1$, the design process finds the same optimum as the most accurate case, but at 117 times the speed. However, Table I gives information only about the final costs. Fig. 6 shows the values of the cost function versus computational cost throughout the optimization. In order to provide a fair comparison of cost-function values, each cost function must be of the same accuracy level. Since tenth-order elements (for the work in this paper) achieve

TABLE I
COMPUTATIONAL COSTS AND SPEED-UP FACTORS FOR DIFFERENT ACCURACY-LINKS IN THE T-JUNCTION OPTIMIZATION

Accuracy Link	Computational Cost (28)	Speed-up Factor
Fixed accuracy	2.72E+09	1
$\alpha = 0.01$	2.32E+07	4.6
$\alpha = 0.1$	5.92E+08	117

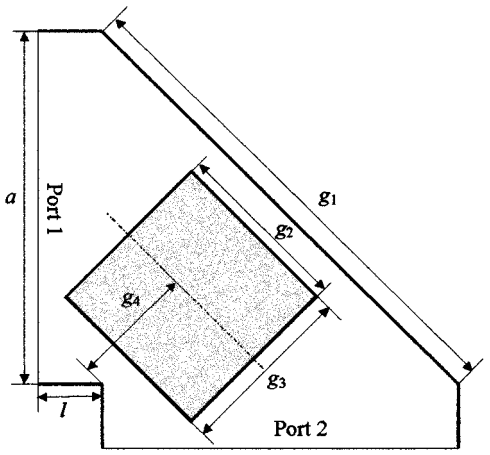

Fig. 7. Initial geometry and dimensions of the miter bend. $a = 2$ cm, $l = 0.1\lambda = 0.454$ cm, $g_1 = 2\sqrt{2}$ cm, $g_2 = 1$ cm, $g_3 = 1$ cm, $g_4 = 2\sqrt{2}$ cm.

TABLE II
COMPUTATIONAL COSTS AND SPEED-UP FACTORS FOR DIFFERENT ACCURACY LINKS IN THE MITER-BEND PROBLEM

Accuracy Link	Computational Cost	Speed-up Factor
Fixed accuracy	6.95E+09	1
$\alpha = 0.01$	2.54E+09	2.7
$\alpha = 0.1$	3.19E+08	22

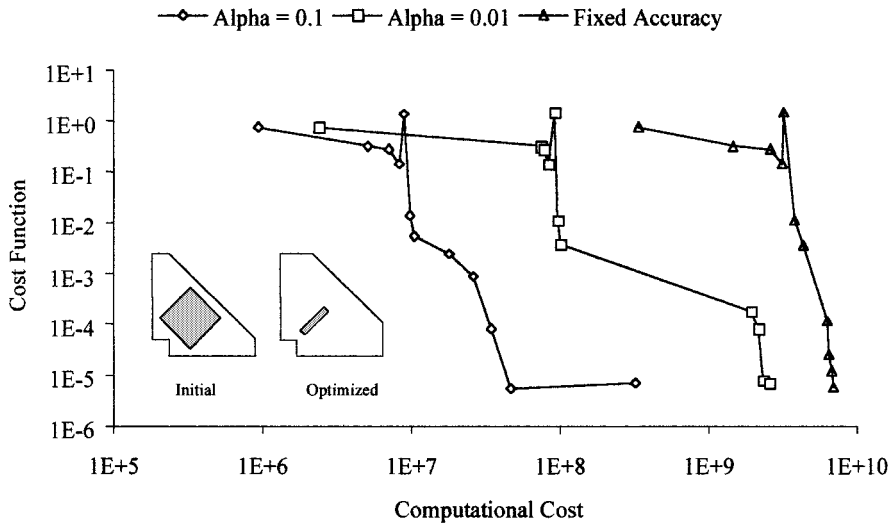

Fig. 8. "True" values of the cost function and the corresponding computational costs for different accuracy links in the optimization of the miter bend.

the highest accuracy (for a fixed mesh size), each point in Fig. 6 is the tenth-order cost-function value at the geometry of that particular optimization step. In other words, it is a plot of the "true" value of the cost function obtained using the different accuracy links.

Fig. 6 shows that a great deal of the computational costs in the optimizations lie in the final few steps, which is expected because the CFEs require higher accuracy levels as an optimum is approached. In addition, Fig. 6 shows that throughout an optimization (not only at the end), controlling accuracy can drastically reduce computation time without sacrificing the validity of the cost function. In other words, despite the fact that CFEs are of lower accuracy at the initial stages of an optimization (for $\alpha = 0.1$ and $\alpha = 0.01$), the cost function does genuinely reduce by the same amount as the fixed accuracy optimization—at a much lower computational cost.

B. Miter Bend With Dielectric Column

In this example (Fig. 7), the return loss is optimized in two ways: by varying the length of the chamfer at the 90° bend and varying the dimensions and position of the dielectric block [29]. The definitions of the cost function and frequency range are identical to those given in the previous problem. The optimization parameters and initial geometry for the problem are shown in Fig. 7. The constraints that were derived for this problem are

$$\begin{aligned}
h_1(\mathbf{g}) &= g_1 - \sqrt{2}(1+l) + \delta \\
h_2(\mathbf{g}) &= \delta - g_1 \\
h_3(\mathbf{g}) &= \delta - g_2 \\
h_4(\mathbf{g}) &= \delta - g_3 \\
h_5(\mathbf{g}) &= 2\delta - g_4 \\
h_6(\mathbf{g}) &= g_1 + g_3 + 2g_4 + 2\left(\delta - \sqrt{2}\right) \\
h_7(\mathbf{g}) &= g_3 - 2g_4 + 2\delta \\
h_8(\mathbf{g}) &= g_2 + g_3 - 2g_4 + 2\sqrt{2}(-l + \delta) \\
h_9(\mathbf{g}) &= g_2 + g_3 + 2g_4 - 2\sqrt{2}(1 - \delta)
\end{aligned} \tag{32}$$

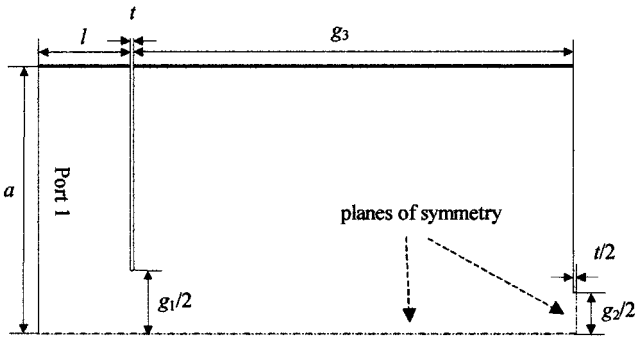

Fig. 9. Initial geometry and dimensions of the model of one-quarter of the two-cavity filter. $a = 19.05$ mm, $l = 0.1\lambda = 3.31$ mm, $t = 0.1$ mm, $g_1 = 4.5$ mm, $g_2 = 3$ mm, $g_3 = 16$ mm.

where $\delta = 0.1$ cm and $l = 0.1\lambda = 0.454$ cm. The frequency of adaption is taken to be 10 GHz, i.e., the center frequency. The initial geometry centers the square dielectric block (of dimension 1 cm × 1 cm) in the middle of the device.

Table II gives the final computational cost for each optimization and a speed-up factor in each case. The results show the speed-up that was attained by varying α.

Fig. 8 plots the values of the cost functions and associated costs throughout the optimization. It is apparent that great savings can be realized at any step of the optimization when the size of α is varied. (The last value of C for the $\alpha = 0.1$ optimization is actually higher than that of the previous step. This is a results of terminating the optimization within a line search. While the actual cost function may increase slightly, the gradient of the Lagrangian has satisfied the termination criterion.)

C. Two-Cavity Iris-Coupled Waveguide Filter

Fig. 9 shows a two-cavity iris-coupled filter, which is to be optimized for a given frequency band. The design and optimization of iris-coupled waveguide cavities using computer-aided design (CAD) tools is common in achieving bandpass filter characteristics [30]–[33]. Filters of this type have resonating cavities (where each guide cavity is roughly $\lambda/2$ in length), coupled by thin (or thick) irises with coupling apertures between any two

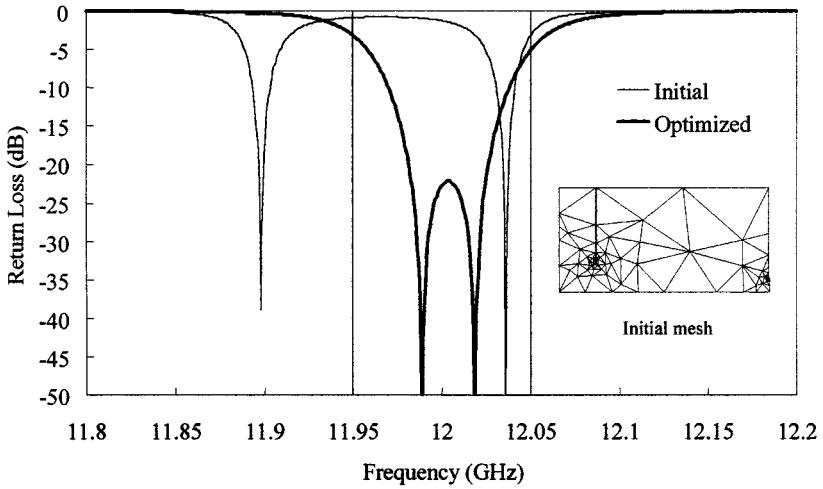

Fig. 10. Initial and optimized frequency responses of the return loss for the two-cavity filter.

cavities. While some designs make use of varying the thickness of each iris as design parameters [30], [32], the two-cavity filter described below has irises of constant small thickness.

The characteristics of a bandpass filter are to maximize transmission over a band of frequencies and stop transmission outside that band. While filter design specifications are commonly given by passband and stopband attenuation, a slightly different approach is used here: minimize reflection in the passband and minimize transmission in the stopband. The cost function for the problem is

$$C(\mathbf{g}) = \sum_{i=1}^{N_f^p} c_i^p + \sum_{i=1}^{N_f^s} c_i^s \tag{33}$$

where there are N_f^p discrete frequency points sampled in the passband and N_f^s points in the stopband. The single frequency cost functions [in the summation terms of (33)] are given by

$$\begin{aligned} c_i^p &= |S_{11}(f_i^p)|^2 \\ c_i^s &= |S_{12}(f_i^s)|^2 . \end{aligned} \tag{34}$$

The design problem has three varying geometric parameters. The aperture width of the left- and right-most irises (symmetrical) is g_1. The aperture width of the center iris is g_2. The lengths of the two symmetrical cavities is g_3. The geometric parameters and dimensions are illustrated for one-quarter of the problem in Fig. 9.

The frequency range is taken to be between 11.8–12.2 GHz with a passband (of 100-MHz bandwidth) centered at $f_a = 12$ GHz. The passband and stopband are, therefore,

$$\begin{aligned} \text{Passband:} \quad & 11.95\ \text{GHz} < f < 12.05\ \text{GHz} \\ \text{Stopbands:} \quad & 11.80\ \text{GHz} < f < 11.95\ \text{GHz} \\ & 12.05\ \text{GHz} < f < 12.20\ \text{GHz}. \end{aligned} \tag{35}$$

The total number of frequency points sampled for the CFE is $N_f = 21$, with $N_f^p = 15$ points in the passband and $N_f^s = 6$ points in the stopband. A greater number of points are taken in the passband because, in the optimal design of the filter, both cavity resonances are within this range of frequency. The FE solution is adapted at the center frequency f_a.

TABLE III
COMPUTATIONAL COSTS AND SPEED-UP FACTORS FOR DIFFERENT ACCURACY LINKS IN THE FILTER PROBLEM

Accuracy Link	Computational Cost	Speed-up Factor
Fixed accuracy	8.39E+10	1
$\alpha = 0.01$	3.01E+10	2.8
$\alpha = 0.1$	7.33E+09	11

Choosing constraints for the geometric parameters is quite simple because, as long as the parameters are positive quantities, there cannot be any overlapping of edges. Six inequality constraints can be written as

$$\begin{aligned} h_1(\mathbf{g}) &= g_1 - 10 \\ h_2(\mathbf{g}) &= 2 - g_1 \\ h_3(\mathbf{g}) &= g_2 - 10 \\ h_4(\mathbf{g}) &= 2 - g_2 \\ h_5(\mathbf{g}) &= g_3 - 20 \\ h_6(\mathbf{g}) &= 10 - g_3. \end{aligned} \tag{36}$$

The thickness of the irises was taken to be 0.1 mm.

To optimize the filter effectively, both resonant frequencies must be within the passband frequency range (also meaning they are close together because of the relatively narrow bandwidth). A poor initial choice for $\mathbf{g}$ might have resonant frequencies that are far apart and far from the passband. In addition, a resonant cavity frequency might be outside the entire range of frequencies sampled (outside the stopband as well). Using a frequency sweep for such a poor design will not work because there will be no detection of the presence of the cavity resonance in the sampled frequency range. However, microwave filter design is a good example of where an initial geometry is found by an initial design method [9] and "tuned" by the optimizer to enhance performance. The method of filter design from [9] is used to design a filter to have the frequency response within a desired initial range. The initial geometry used is $\mathbf{g} = [4.5\ \ 3\ \ 16]^{\mathrm{T}}$ mm.

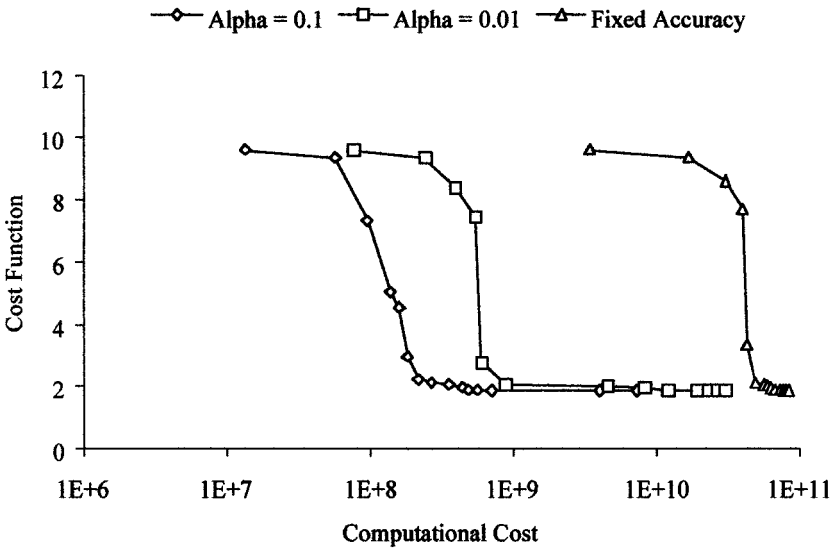

Fig. 11. "True" values of the cost function and the corresponding cumulative computational costs for different accuracy links in the optimization of the two-cavity filter problem.

The initial design and final frequency responses for the return loss of the two-cavity filter are shown in Fig. 10 (based on an evaluation of 500 discrete points in the frequency range). As expected, due to the sensitivity of the device, the geometric parameters do not change by a large amount in the optimization of the device (maximum parameter change is roughly 20%), yet produce a big difference in the performance of the device.

Table III gives the final computational cost for each optimization and a speed-up factor in each case. Once again, there is great computational cost saved when optimizing with CFEs with lower initial accuracy. Fig. 11 shows that, once again, while much of the computational effort lies in the final steps of the optimization, there is consistent savings at any point in the optimization. In addition, this figure shows that even when the accuracy demand is low, the cost function is reduced in value at roughly the same rate as the more accurate case.

VI. Conclusions

Design techniques for a wide variety of microwave devices are good enough to establish an initial design that lies reasonably close to a local optimum. Computational-field analysis can be used to move the design to that optimum, thereby maximizing device performance (at least locally). The problem is that accurate field analysis is usually expensive, perhaps prohibitively so. However, in the early stages of optimization, it is not necessary to perform highly accurate analysis; it just has to be good enough to move the design in the right direction. By providing a link from the optimization to the field analysis, the accuracy of the latter can be progressively improved during the optimization so that, in the end, the local optimum is still calculated accurately—but with much less cost along the way. The results presented above show that this approach is capable of giving an order of magnitude reduction in computational cost over the straightforward use of fixed high accuracy throughout. The results given are for two-dimensional (2-D) problems, where the analysis times are, in any case, fairly small and the cost savings perhaps not worthwhile. However, the approach applies equally to 3-D, where the costs are dramatically greater. Moreover, it is not limited to the p-adaptive method employed here, or even to the FEM. Any computationally intensive analysis scheme with a capability for accuracy-time tradeoff could benefit from the same approach.

References

[1] J. W. Bandler, R. M. Biernacki, S. H. Chen, and D. Omeragic, "Space mapping optimization of waveguide filters using finite element and mode-matching electromagnetic simulators," *Int. J. RF Microwave Computer-Aided Eng.*, vol. 9, no. 1, pp. 54–70, Jan. 1999.

[2] S. Bila, D. Baillargeat, M. Aubourg, S. Verdeyme, and P. Guillon, "A full electromagnetic CAD tool for microwave devices using a finite element method and neural networks," *Int. J. Numer. Modeling*, vol. 13, no. 2–3, pp. 167–180, Mar.–June 2000.

[3] M. Koshiba and M. Suzuki, "Finite-element analysis of H-plane waveguide junction with arbitrarily shaped ferrite post," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-34, pp. 103–109, Jan. 1986.

[4] Z. J. Cendes and J. F. Lee, "The transfinite element method for modeling MMIC devices," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 1639–1649, Dec. 1988.

[5] J. F. Lee, "Analysis of passive microwave devices by using three-dimensional tangential vector finite elements," *Int. J. Numer. Modeling*, vol. 3, no. 4, pp. 235–246, Dec. 1990.

[6] S. McFee and J. P. Webb, "Adaptive finite element analysis of microwave and optical devices using hierarchal triangles," *IEEE Trans. Magn.*, vol. 28, pp. 1708–1711, Mar. 1992.

[7] M. Salazar-Palma, T. K. Sarkar, L. E. Garcia-Castillo, T. Roy, and A. R. Djordjevic, *Iterative and Self-Adaptive Finite-Elements in Electromagnetic Modeling*. Norwood, MA: Artech House, 1998, p. 770.

[8] M. M. Gavrilovic and J. P. Webb, "A port element for p-adaptive S-parameter calculation," *IEEE Trans. Magn.*, vol. 35, pp. 1530–1533, May 1999.

[9] R. Beck and R. Hiptmair, "Multilevel solution of the time-harmonic Maxwell's equations based on edge elements," *Int. J. Numer. Methods Eng.*, vol. 45, pp. 901–920, 1999.

[10] D. K. Sun, Z. Cendes, and J. F. Lee, "Adaptive mesh refinement, h-version, for solving multiport microwave devices in three dimensions," *IEEE Trans. Magn.*, vol. 36, pp. 1596–1599, July 2000.

[11] H. Akel and J. P. Webb, "Design sensitivities for scattering-matrix calculation with tetrahedral edge elements," *IEEE Trans. Magn.*, vol. 36, pp. 1043–1046, July 2000.

[12] H. Lee and T. Itoh, "A systematic optimum design of waveguide-to-microstrip transition," *IEEE Trans. Microwave Theory Tech.*, vol. 45, pp. 803–809, May 1997.

[13] O. C. Zienkiewiesz, D. W. Kelly, J. Gago, and I. Babuska, *The Mathematics of Finite Elements and Applications*, J. R. Whiteman, Ed. New York: Academic, 1982.
[14] M. M. Gavrilovic and J. P. Webb, "An error indicator for the calculation of global quantities by the p-adaptive finite element method," *IEEE Trans. Magn.*, vol. 33, pp. 4128–4130, Sept. 1997.
[15] O. C. Zienkiewiesz and R. L. Taylor, *The Finite-Element Method*, 4th ed. New York: McGraw-Hill, 1988, ch. 14.
[16] I. Babuska, B. A. Szabo, and I. N. Katz, "The p-version of the finite-element method," *SIAM J. Numer. Anal.*, pp. 515–545, 1981.
[17] W. Daigang and J. Kexun, "P-version adaptive computation of FEM," *IEEE Trans. Magn.*, vol. 30, pp. 3515–3518, Sept. 1994.
[18] J. P. Webb and R. Abouchacra, "Hierarchal triangular elements using orthogonal polynomials," *Int. J. Numer. Methods Eng.*, vol. 38, pp. 245–257, 1995.
[19] J. F. Lee and Z. J. Cendes, "An adaptive spectral response modeling procedure for multiport microwave circuits," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-35, pp. 1240–1247, Dec. 1987.
[20] S. Kumashiro, R. A. Rohrer, and A. J. Strojwas, "Asymptotic waveform evaluation for transient analysis of 3-D interconnect structures," *IEEE Trans. Computer-Aided Design*, vol. 12, pp. 988–996, July 1993.
[21] S. V. Polstyanko, V. Sergey, R. Dyczij-Edlinger, and J. F. Lee, "Fast frequency sweep technique for the efficient analysis of dielectric waveguides," *IEEE Trans. Microwave Theory Tech.*, vol. 45, pp. 1118–1126, July 1997.
[22] R. Sanaie, E. Chiprout, M. S. Nakhla, and Q. Zhang, "A fast method for frequency and time domain simulation of high-speed VLSA interconnects," *IEEE Trans. Microwave Theory Tech.*, vol. 42, pp. 2562–2571, Dec. 1994.
[23] J. Gong and J. L. Volakis, "AWE implementation for electromagnetic FEM analysis," *Electron. Lett.*, vol. 32, pp. 2216–2217, Nov. 1996.
[24] M. M. Gavrilovic and J. P. Webb, "An error estimator for design sensitivities of microwave device parameters," *Electromagnetics*, vol. 22, no. 4, pp. 315–321, May–June 2000.
[25] A. Grace, *Optimization Toolbox for Use with MATLAB*. Natick, MA: The Math Works Inc., 1992, pp. 2-22–2-31.
[26] B. M. Irons, "A frontal solution program for finite element analysis," *Int. J. Numer. Methods Eng.*, vol. 2, pp. 5–32, 1970.
[27] H. Lee, H.-K. Jung, S. Hahn, C. Cheon, and K.-S. Lee, "Shape optimization of H-plane waveguide tee junction using edge finite element method," *IEEE Trans. Magn.*, vol. 31, pp. 1928–1931, May 1995.
[28] Ansoft Corporation, "Parametrics and optimization using Ansoft HFSS," *Microwave J.*, vol. 42, no. 11, pp. 156–166, Nov. 1999.
[29] J. Baden Fuller, *Microwaves: An Introduction to Microwave Theory and Techniques*, 2nd ed. New York: Pergamon, 1979, sec. 10.4.
[30] J. T. Alos and M. Guglielmi, "Simple and effective EM-based optimization procedure for microwave filters," *IEEE Trans. Microwave Theory Tech.*, vol. 45, pp. 856–858, June 1997.
[31] K. Shamsaifar, "Designing iris-coupled waveguide filters using the mode-matching technique," *Microwave J.*, pp. 156–154, Jan. 1992.
[32] J.-F. Liang, H.-C. Chang, and K. A. Zaki, "Design and tolerance analysis of thick iris waveguide bandpass filters," *IEEE Trans. Magn.*, vol. 29, pp. 1605–1608, Mar. 1993.
[33] F. Alessandri, M. Dionigi, and R. Sorrentino, "A full-wave CAD tool for waveguide components using a high speed direct optimizer," *IEEE Trans. Microwave Theory Tech.*, vol. 43, pp. 2046–2052, Sept. 1995.

Minya M. Gavrilovic received the B.Eng. degree in electrical engineering from Concordia University, Montreal, QC, Canada, in 1996, and the Ph.D. degree in electrical engineering from McGill University, Montreal, QC, Canada, in 2000.

Since 2000, he has been with the Antenna Research Group, EMS Technologies, Ste. Anne de Bellevue, QC, Canada. His research interests include field-based computer-aided design (CAD) and modeling of antennas and other microwave devices and reduced-time optimization of full-wave electromagnetic problems.

Jon P. Webb (M'83) received the Ph.D. degree from Cambridge University, Cambridge, U.K., in 1981.

Since 1982, he has been a Professor with the Electrical Engineering Department, McGill University, Montreal, QC, Canada. His research interest is computer methods in electromagnetics, especially the application of the FEM.