Numerical Cost of Gradient Computation Within the Method of Moments and its Reduction by Means of a Novel Boundary-Layer Concept

Smain Amari, Member, IEEE
Department of Electrical and Computer Engineering
Royal Military College of Canada, Kingston, ON
Canada, smain.amari@rmc.ca

Abstract: A rigorous investigation of the numerical cost of sensitivity analysis (gradient computation) of complex structures within Moment Method is presented. It is shown that, when the number of variables N used in the analysis is large, a common situation in complex structures, the ratio r of the number of flops required to evaluate the sensitivity of the response to structural changes to the number of flops required to determine the response at a single point is such that r=O(1/N) as long as the number of flops required to fill the matrix is not dominant. For the latter important case, a new boundary layer concept is introduced to reduce the CPU time for the gradient computation. A simple example of an iris in a rectangular waveguide is used to illustrate the concept and show its validity.

I. Introduction

The complexity of modern microwave structures makes them impossible to tackle within analytical or even semi-analytical numerical techniques. General numerical techniques such as the Method of Moment (MoM) or the Finite Element Method (FEM) are often resorted to [1].

Although design formulas for simple systems are available, actual designs of modern structures are often finalized by optimization and eventually tuned experimentally to compensate for manufacturing tolerances. It is commonly accepted that the most efficient optimization techniques are those that exploit not only the values of the cost function but its gradient as well. The computation of the gradient is often carried out using finite differences where the structure is analyzed an additional time for each independent variable. It should be obvious that this approach is viable only when the requirements of an analysis in terms of CPU times and computer memory are reasonably small.

It was recently shown that the gradient of the response of a linear systems analyzed within numerically intensive techniques can be determined directly from a single analysis [2]. A similar technique was applied to inverse scattering problem using the Finite Element Method (FEM) [3]. The adjoint network method was also used in the optimization of both radiating and guiding structures by Mongiardo and coworkers [4], [5]. In this paper, we propose to further show that the numerical cost of evaluating the gradient of an already determined response becomes negligibly small in comparison with the cost of computing the response (solution) itself when the number of variables is large provided the time it takes to fill the matrix is not dominant. For this case, a new boundary-layer concept to reduce the CPU time is introduced.

II. Statement of the Problem

We focus attention on a linear system whose response [x] (vector of length N) is related to the corresponding excitation [b] (vector of length N) by a matrix equation of the form

$$[A][x] = [b] \qquad (1)$$

The system is represented in this equation by the matrix [A] of size N x N. This formulation is quite general and is often encountered in the Method of Moments (MoM). The specific forms of the entries of this matrix depend on the physical and geometrical parameters of the system and method of analysis. The vector [x] contains the expansion coefficients in the MoM. Let us denote by $[\alpha]^t = [\alpha_1, \alpha_2, ..., \alpha_p]$ a vector of size p which contains the optimization variables α_i. We focus

specifically on the Method of Moments, or those leading to a **full** matrix [A].
The first question we are addressing is the following: **what is the additional numerical cost of evaluating the gradient of [x] at a given point with respect of [α] when N is large once the solution is known at the same point?**
We purposely consider the case of large values of N which corresponds to modern complex structures where N can be of the order of tens or hundreds of thousand and even more.

III. Cost Within Standard Finite Differences

We first assume that the gradient is determined using finite differences. Within this approach, the response is first determined at a value of the parameter vector [α]. For each of the p parameters at least an additional analysis is performed. Since each analysis requires the solution of a linear system [A][x]=[b], a $\frac{N^3}{3}$ operation, the overall additional cost is

$$C_{FD} = \frac{pN^3}{3} \qquad (2)$$

IV. Cost Using Analytical Gradient

The gradient of the response [x] can be determined from the following identity [3]

$$[A]\frac{\partial[x]}{\partial\alpha_i} + \frac{\partial[A]}{\partial\alpha_i}[x] = \frac{\partial[b]}{\partial\alpha_i} \qquad (3)$$

Since we assumed that the solution is already known, we know the LU decomposition of the matrix [A]. The total additional cost within this approach consists of the following operations:

1. Multiplication of $\frac{\partial[A]}{\partial\alpha_i}$ and [x], $\rightarrow N^2$.
2. Forward and backward substitution in equation (3), $\rightarrow N^2$.

These operations are performed once for each of the p parameters. The total additional cost is therefore,

$$C_{analyt.} = 2pN^2 \qquad (4)$$

V. Discussion

It is interesting to compare the two costs for typical values of p and N. We assume that the number of parameters is p=10 and the number of variables is N=1000. Let us denote by r the ratio

$$r = \frac{C_{analyt.}}{C_{FD}} = \frac{3}{500} = 0.006 \qquad (5)$$

In other words, the additional cost within the analytical gradient approach is negligible compared to the standard finite difference method. We stress again that these results are derived assuming that N is large and that the matrix [A] is full.

It should be mentioned that a similar analysis for electrical circuits was presented by Bandler and coworkers [6]. The present paper shows that their results can be extended to more complex structures which are analyzed by modern numerical techniques. Another important point is the applicability of the present results to methods leading to sparse matrices, such as the FEM or FDFD. For those methods, solving a linear system [A][x]=[b] is not as costly as $N^3/3$ if the sparsity of the matrix [A] is exploited. The present results may not hold in general in these sparse cases.

It may be also worth pointing out that the matrix $\frac{\partial[A]}{\partial\alpha_i}$ is usually very sparse, this feature can be used to reduce further the additional cost within the analytical approach. This, however, comes at the price of book keeping and may not be necessary unless the cost of filling the matrix [A]

is dominant. In such a case, the results presented here do not hold as they assume that the cost of filling the matrix is at most an N^2 operation. Specific techniques to compute derivatives efficiently for this case are presented below.

VI. Boundary-Layer Concept

For complex structures analyzed by the moment methods, one is often faced with the task of evaluating time consuming integrals numerically. In such cases, the dominant part of the CPU time is absorbed by the computation of the matrix [A] and not its inversion. Obviously, the computation of the gradient under these conditions can be extremely expensive in terms of CPU times. To deal with this important situation, we introduce the following idea.

Figure 1a shows a generic computational domain, which is assumed 1D for simplicity, of length $L=L_0$. In the MoM analysis of this domain, the length L is divided into N equal segments of length L/N and appropriate basis functions such as pulse or rooftop are used on each one. Let us denote by $R(L_0)$ the response of the structure when $L=L_0$.

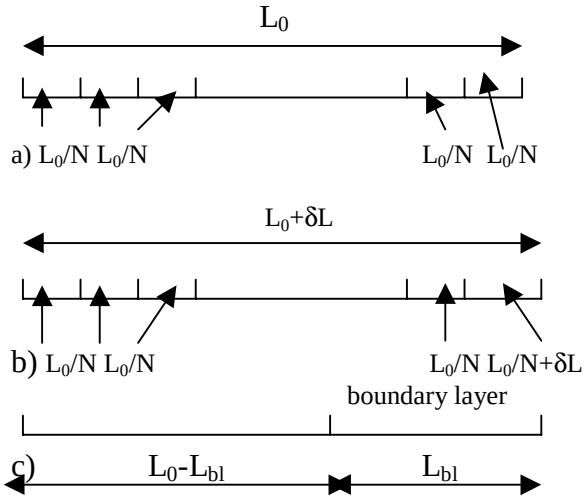

Figure 1: computational domain of length L_0.a) discretization for MoM analysis, b) inclusion of perturbation δL and c) boundary layer of thickness L_{bl}.

The question we address is the following: **is it possible to determine the response $R(L_0+\delta L)$, $\delta L \rightarrow 0$ without perturbing each of the N segments?**

Let us examine the following situation as shown in Figure 1b. Keep the support of the first N-1 segments equal to their previous value L_0/N and allow the last one to have a length $l_{new} = \frac{L_0}{N} + \delta L$. The question is then whether the response obtained from this new distribution is an expression of $R(L=L_0+\delta L)$ for $\delta L \rightarrow 0$. Obviously, the answer is in the affirmative if the discretization leads to the exact response.
However, we do not expect to have an exact description of the response (or the fields). The discretization will introduce errors in the response which may reduce the accuracy of the computed perturbed response. To reduce this error, we allow more segments to change, say those falling within a layer of thickness L_{bl} as shown in Figure 1c. We call the layer of thickness L_{bl} a **boundary layer** since most often the dominant physics takes place close to the boundaries which are changed in an optimization process. It is important to note that using the idea of a boundary layer, only a fraction of the entries of the matrix [A] will be affected by the change δL. Since we assumed that computing these entries is dominant, a tremendous reduction in CPU time can be achieved**.**

As an example, we consider an infinitely thin H-plane iris in a rectangular waveguide as shown in Figure 2.

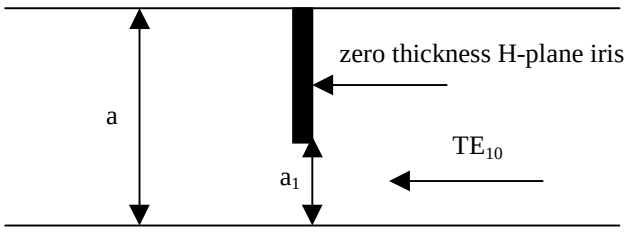

Figure 2: cross section of zero thickness H-plane iris in a rectangular waveguide.

We would like to determine the width of the aperture such that half of the power incident in the dominant TE_{10} mode is reflected at 12 GHz. The problem is formulated in terms of an integral equation for the aperture field. Pulse functions are used as basis and test functions. To determine a_1, Newton's method is used starting from the point a_1 = 0.65 a. The number of cells in the boundary layers is changed and solution recorded. The total number of basis functions is 60 and a = 19.05 mm. Table I shows the results.

Table I: start $a_1/a = 0.65$

Cells in boundary layer	Solution a_1/a
1	**0.582**
2	**0.582**
3	**0.582**
5	**0.582**
10	**0.582**
15	**0.582**

It is evident that a boundary layer which contains even one cell is sufficient to accurately determine the optimal value of a_1 using Newton's method which is very sensitive to the value of the gradient (derivative in this simple case). The approach performs not as well when the starting value is a_1 = 0.8 a instead as Table II shows.

Table II: start $a_1/a = 0.80$

Cells in boundary layer	Solution a_1/a
1	**diverges**
2	**diverges**
3	**diverges**
5	**diverges**
10	**0.582**
15	**0.582**

In this case, a boundary layer of 10 cells or more is required to achieve convergence. Of course, the potential saving in CPU time is much more substantial in complex structures where a large number of variables is used.

VII. Conclusions

The cost of computing the gradient of the response of complex linear systems analyzed by the Method of Moments can be reduced to a negligible fraction of the cost of a single analysis when the number of variables is large provided that the CPU time required to fill the matrix is not dominant. When this is not the case, the concept of a boundary layer which absorbs the differential changes was introduced.

VIII. References

1. T. Itoh Ed., Numerical Techniques for Microwave and Millimeter-Wave Passive Structures , Wiley, New York 1989.

2. S. Amari, P. Harscher, R. Vahldieck and J. Bornemann, *Novel Analytic gradient evaluation techniques for optimization of microwave structures,* IEEE MTT-S Int. Symp. Dig., Anaheim 1999, pp. 31-34.

3. I. T. Rekanos and T. D. Tsiboukis, *A combined finite element -nonlinear conjugate gradient method for the reconstruction of unknown scatterer profiles,* IEEE Tran. Magnetics, vol. 34, pp. 2829-2832, Sept. 1998.

4. M. Mongiardo and T. Ravanelli, *Automated design of corrugated feeds by the adjoint network method,* IEEE Tran. Microwave Theory Tech., vol. 45, pp. 787-792, May 1997.

5. F. Alessandri, M. Mongiardo and R. Sorrentino, *New efficient full wave optimization of microwave circuits by the adjoint network method,* IEEE Guided Wave Lett., vol, 3, pp. 414-416, Nov. 1993.

6. J. W. Bandler, S. H. Chen and S. Daijavad, *Exact sensitivity analysis for optimization of multi-coupled cavity filters,* Jour. Circuit Theory Appl., vol. 31, 63-77, 1986.