Feasible Adjoint Sensitivity Technique for EM Design Optimization

Natalia K. Georgieva, Snezana Glavic, Mohamed H. Bakr and John W. Bandler

Department of Electrical and Computer Engineering, McMaster University
Hamilton, Ontario, Canada L8S 4K1

***Abstract* — The adjoint variable method for frequency domain design sensitivity analysis is proposed for the optimization of wire and printed structures analyzed by the Method of Moments (MoM). We focus on the construction of the adjoint system using a feasible technique which requires only minor modifications of existing MoM codes. The solution to the adjoint problem is obtained with very little overhead once the original problem is solved. The gradient of the objective function is consequently computed through a single analysis regardless of the number of the design parameters. The concept is illustrated through the design of a Yagi-Uda array and a rectangular patch antenna using suitable MoM simulators.**

I. INTRODUCTION

System design sensitivity analysis (DSA) concerns the relationship between design variables, which are assigned by the engineer, and the system response (or state variables), which is determined by the laws of physics governing the system's behavior. Its purpose is to evaluate the sensitivity of the system's response to variations of the design parameters. Design sensitivity information is crucial in a number of engineering problems such as optimization, statistical and yield analysis, as well as tolerance analysis. In this paper, we focus on the implementation of the adjoint-based DSA for gradient optimization with full-wave frequency domain EM solvers.

The adjoint variable method (AVM) for DSA is an efficient design approach to complex linear and nonlinear problems. It has been proposed in areas such as structural design [1], circuit theory [2]-[6], control theory, etc. Adjoint sensitivities for circuit CAD can be found even in undergraduate courses [7]. Adjoint techniques have already been implemented in commercial structural design software based on the finite-element method (FEM) [1]. At the same time, the AVM has attracted very little attention in full-wave EM analysis.

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Grants OGP0227660-00, OGP0007239-98, STR234854-00, through the Micronet Network of Centres of Excellence and Bandler Corporation.

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

The adjoint-based DSA of microwave structures has historically been formulated in terms of circuit concepts through Tellegen's theorem rather than field concepts. It is referred to as the adjoint network method (ANM). The first applications of the ANM to microwave circuit problems were published in the early 1970s when network sensitivities were calculated on both voltage-current [3]-[5], and S-parameter bases [6], [8], [9]. Later, Alessandri *et al.* [10] applied the ANM to the analysis of microwave circuits whose subnetworks are represented by Y-parameters. Typically, the ANM considers the sensitivity of a single state variable [4], which makes its applications problem specific. It is not immediately obvious how the ANM can be utilized in a full-wave analysis.

Recently, a technique was proposed for exact sensitivity analysis with the FEM [11]. A similar approach was later applied to problems solved in terms of the MoM, and the boundary layer concept was proposed to reduce the computational load [12]. In effect, this technique is based on the *direct differentiation method* (DDM) [1], an efficient approach to the sensitivity analysis of distributed response functions. This technique stops short of defining and exploiting the concept of adjoint sensitivities.

We give the mathematical background of the AVM and discuss its implementation in exact sensitivity analysis of linear, time-harmonic EM problems. Three major issues are discussed: (i) the adjoint problem; (ii) the procedure to efficiently evaluate the gradient of the response function; and (iii) the formulation of the objective function in adjoint-based gradient optimization. The AVM approach increases substantially the efficiency of the current CAD tools based on full-wave frequency domain analysis such as the FEM and the MoM. This is due to the fact that the objective function and its gradient are computed through a single analysis.

II. ADJOINT-BASED DESIGN SENSITIVITY

Here, we present the basic concepts of the AVM for DSA in the case of a general linear problem. The importance of this discussion arises from the fact that most full-wave solvers reduce a theoretical model of the EM problem to a system of linear equations through a variety of discretization techniques. Neither the theoretical models

0-7803-7239-5/02/$10.00 © 2002 IEEE

nor the discretization techniques are discussed hereafter because the formulation of the AVM assumes that the problem has already been properly reduced to a system of linear equations. We should note that the AVM can be extended to the DSA of nonlinear systems. Nonlinear circuit sensitivities and feasible approaches to their estimation are discussed in [7],[13].

Consider the linear system of equations arising from the discretization of an EM problem

$$\boldsymbol{Z}(x)\boldsymbol{I} = \boldsymbol{V} \tag{1}$$

Here, $\boldsymbol{x}$ is the vector of designable parameters, $\boldsymbol{I}$ is the state variable vector, e.g., complex currents in the MoM. $\boldsymbol{V}$ is the global excitation vector. $\boldsymbol{Z}$ is a matrix whose coefficients depend on the structure's geometry and materials. Often, the Z-coefficients are explicit functions of the discretization grid nodes as is the case in the FEM. This can be advantageous, since it allows the computation of the exact sensitivities of the $\boldsymbol{Z}$ matrix instead of using finite-differences. Note that the solution $\boldsymbol{I}$ is an implicit function of the design parameters $\boldsymbol{x}$.

We define a general function $f(\boldsymbol{x}, \bar{\boldsymbol{I}}(\boldsymbol{x}))$, which is the *response function* of the linear system. This function has to be differentiable in all its arguments. It may have explicit dependence on the design parameters $\boldsymbol{x}$. It depends on the solution $\bar{\boldsymbol{I}}$ of (1), and therefore, has an implicit dependence on $\boldsymbol{x}$ as well. The objective is to determine the sensitivity of f with respect to the design parameters $\boldsymbol{x}$, i.e.,

$$\nabla_{\boldsymbol{x}} f \text{, subject to } \boldsymbol{ZI} = \boldsymbol{V} \tag{2}$$

where $\nabla_{\boldsymbol{x}}$ is defined as the row operator

$$\nabla_{\boldsymbol{x}} = \left[\frac{\partial}{\partial x_1} \; \frac{\partial}{\partial x_2} \; \cdots \; \frac{\partial}{\partial x_n} \right] \tag{3}$$

Assuming that the $\boldsymbol{Z}$ matrix is not singular, the following expression for $\nabla_{\boldsymbol{x}} \boldsymbol{I}$ is obtained from (1):

$$\nabla_{\boldsymbol{x}} \boldsymbol{I} = \boldsymbol{Z}^{-1} \left[\nabla_{\boldsymbol{x}} \boldsymbol{V} - \nabla_{\boldsymbol{x}} (\boldsymbol{Z}\bar{\boldsymbol{I}}) \right] \tag{4}$$

where $\boldsymbol{I}$, $\boldsymbol{V}$ and $(\boldsymbol{Z}\bar{\boldsymbol{I}})$ are column vectors, e.g.,

$$\boldsymbol{I} = [I_1 \ldots I_m]^T \tag{5}$$

In $\nabla_{\boldsymbol{x}}(\boldsymbol{Z}\bar{\boldsymbol{I}})$, $\bar{\boldsymbol{I}}$, which is the solution of (1) at the current design, is held constant during the differentiation. For clarity, (4) is rewritten as

$$\frac{\partial \boldsymbol{I}}{\partial x_i} = \boldsymbol{Z}^{-1} \left[\frac{\partial \boldsymbol{V}}{\partial x_i} - \frac{\partial \boldsymbol{Z}}{\partial x_i} \bar{\boldsymbol{I}} \right], \; i = 1, 2, \ldots, n \tag{6}$$

Equation (4) is the basis of the *direct differentiation method* (DDM) [1]. It provides the means of efficient calculation of the gradient of each of the state variables. There is no need for additional $\boldsymbol{Z}$ matrix *LU*-factorization since this has been already done at the analysis stage of the current design. The solution of (6) can be used to calculate the exact sensitivities of $f(\boldsymbol{x}, \bar{\boldsymbol{I}}(\boldsymbol{x}))$ by direct substitution in

$$\nabla_{\boldsymbol{x}} f = \nabla_{\boldsymbol{x}}^e f + \nabla_{\boldsymbol{I}} f \cdot \nabla_{\boldsymbol{x}} \boldsymbol{I} \tag{7}$$

where

$$\nabla_{\boldsymbol{I}} = \left[\frac{\partial}{\partial I_1} \; \frac{\partial}{\partial I_2} \; \cdots \; \frac{\partial}{\partial I_m} \right] \tag{8}$$

The gradient $\nabla_{\boldsymbol{x}}^e f$ reflects the explicit dependence of $f(\boldsymbol{x}, \bar{\boldsymbol{I}}(\boldsymbol{x}))$ on $\boldsymbol{x}$. The DDM, although similar to the AVM, stops short of defining the general adjoint problem; and, thus, does not make use of the associated computational benefits.

The adjoint problem can be derived from (4) and (7), which lead to

$$\nabla_{\boldsymbol{x}} f = \nabla_{\boldsymbol{x}}^e f + \nabla_{\boldsymbol{I}} f \, \boldsymbol{Z}^{-1} \left[\nabla_{\boldsymbol{x}} \boldsymbol{V} - \nabla_{\boldsymbol{x}} (\boldsymbol{Z}\bar{\boldsymbol{I}}) \right] \tag{9}$$

The vector

$$\hat{\boldsymbol{I}} = \left[\nabla_{\boldsymbol{I}} f \, \boldsymbol{Z}^{-1} \right]^T = \left[\boldsymbol{Z}^T \right]^{-1} \left[\nabla_{\boldsymbol{I}} f \right]^T \tag{10}$$

is now introduced. It is a solution to the equation

$$\boldsymbol{Z}^T \hat{\boldsymbol{I}} = \left[\nabla_{\boldsymbol{I}} f \right]^T \tag{11}$$

and is referred to as the *adjoint variable* vector. Equation (11) describes the *adjoint problem*. The factored $\boldsymbol{Z}^T$ matrix is obtained easily from the factored $\boldsymbol{Z}$ matrix of the original system. The sensitivities can now be computed in terms of the original solution $\bar{\boldsymbol{I}}$ and the adjoint solution $\hat{\boldsymbol{I}}$ as

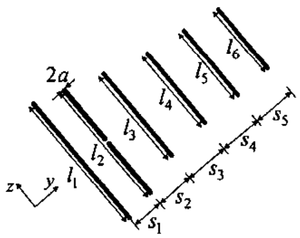
$$\nabla_{\boldsymbol{x}} f = \nabla_{\boldsymbol{x}}^e f + \hat{\boldsymbol{I}}^T \left[\nabla_{\boldsymbol{x}} \boldsymbol{V} - \nabla_{\boldsymbol{x}} (\boldsymbol{Z}\bar{\boldsymbol{I}}) \right] \tag{12}$$

Equations (11) and (12) form the basis of the AVM. The matrices $\partial \boldsymbol{Z} / \partial x_i$ $(i = 1, \ldots, n)$ in $\nabla_{\boldsymbol{x}}(\boldsymbol{Z}\bar{\boldsymbol{I}})$ may be analytically available. If this is not the case, one can always resort to the finite-difference approximation $\Delta \boldsymbol{Z} / \Delta x_i$ (i=1,…,n) [13]. This would require n additional $\boldsymbol{Z}$-matrix fills. However, the analytical evaluation of $\partial \boldsymbol{Z} / \partial x_i$ would typically be equivalent to an additional $\boldsymbol{Z}$-matrix fill. Thus, analytically available $\partial \boldsymbol{Z} / \partial x_i$ matrices are important not so much to the computational efficiency of the algorithm but rather to its accuracy.

The accuracy of the sensitivity estimation via (12) is not strongly affected if the $\partial \boldsymbol{Z} / \partial x_i$ matrices in the MoM are approximated by finite differences. This is due to the nearly linear dependence of the majority of the elements of the $\partial \boldsymbol{Z} / \partial x_i$ matrix on small perturbations Δx_i (from 1 to 5%) of a geometrical design parameter (see also [13]).

A key to the construction of the adjoint problem is the adjoint excitation vector in (11) $\hat{\boldsymbol{V}} = [\nabla_{\boldsymbol{I}} f]^T$. It is evident that the response function f has to be differentiable in I_k $(k = 1, \ldots, m)$. The accuracy of the adjoint solution $\hat{\boldsymbol{I}}$

l_1/λ	0.450
l_2/λ	0.450
l_d/λ	0.406
s_1/λ	0.200
s_d/λ	0.340
a/λ	0.003



$l_3 = l_4 = l_5 = l_6 = l_d$; $s_2 = s_3 = s_4 = s_5 = s_d$

Fig. 1. The geometry of the Yagi-Uda array.

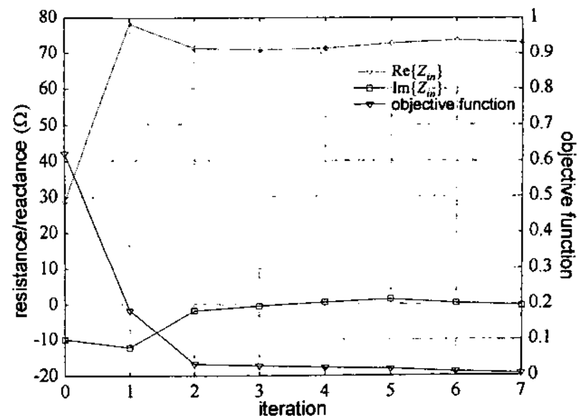


Fig. 2. The progress of the objective function and the input impedance of the Yagi-Uda design.

depends strongly on the accuracy of $\hat{V}$. Numerical tests show that inaccurate finite-difference approximations of $\hat{V}$ may result in deterioration of the sensitivity analysis via (12).

The AVM has significant computational advantages in comparison with the traditional calculation of the sensitivities through the finite-difference approach (FDA). The AVM generates the response and its sensitivities through a single analysis regardless of the number of design parameters n. Certain post-processing is required; however, its computational requirements do not exceed those of one system analysis. In contrast, the FDA performs $(n+1)$ full analyses. The AVM has better computational efficiency in comparison with the DDM as well. In the DDM, according to (6), n back substitutions of the factored $\boldsymbol{Z}$ matrix are needed to compute $\partial f/\partial \boldsymbol{x}$. In AVM, according to (11) and (12), there is only one back substitution needed regardless of n: the one used to compute $\hat{\boldsymbol{I}}$.

III. Defining an Objective Function

An objective function f may be a suitable least pth or minimax real valued function [3],[5] of the state variables I_k $(k=1,\ldots,m)$. The response in the frequency domain analysis is typically a complex valued function. The complex error $e(\omega_j)$ containing sampled frequency domain responses [3], [5] can, for example, appear in a least pth objective function as

$$f = \sum_j \frac{1}{p} | e(\omega_j) |^p \qquad (13)$$

where ω_j denotes the jth frequency of interest. Then,

$$\nabla f = \sum_j \mathrm{Re}\left\{ | e(\omega_j) |^{p-2} e^*(\omega_j) \nabla e(\omega_j) \right\} \qquad (14)$$

It is recommended that f and, therefore, $e(\omega_j)$ be analytically differentiable in I_k $(k=1,\ldots,m)$, so that the adjoint excitation $\hat{V}$ is readily computed at the current design.

IV. Results

A. Input Impedance of a Yagi-Uda Array

The first example is the optimization of the input impedance of the Yagi-Uda array, whose initial design is shown in Fig. 1. The analysis is based on Pocklington's equation using pulse subdomain basis functions. The objective function is defined as

$$f(\boldsymbol{x}) = \left| (Z_{in} - \bar{Z}) / \bar{Z} \right| \qquad (15)$$

where $\bar{Z} = 73\ \Omega$. The vector of design parameters is $\boldsymbol{x} = [l_{1n}\ s_{1n}]^T$, where $l_{1n} = l_1/\lambda$ and $s_{1n} = s_{1n}/\lambda$. The result of the optimization is shown in Fig. 2. At each iteration, only one LU-factorization of the $\boldsymbol{Z}$-matrix is performed. The adjoint excitation $\hat{\boldsymbol{V}}$ has only one nonzero element because the objective function $f(\boldsymbol{x})$ depends on a single state variable: the current at the driver's base. The optimal design is obtained as $\boldsymbol{x} = [0.5243\ \ 0.2607]^T$.

B. Input Impedance of a Rectangular Patch Antenna

The AVM technique is applied to the optimization of a microstrip-fed rectangular patch antenna with an inset, for an input impedance of 50 Ω. The design parameters are the length of the patch L, its width W and the depth of the inset S. The design problem is formulated as

$$f(\boldsymbol{x}) = \left(\mathrm{Re}\{Z_{in}\} - 50\right)^2 + \left(\mathrm{Im}\{Z_{in}\}\right)^2 \qquad (16)$$

where $\boldsymbol{x} = [L\ W\ S]^T$. The analysis is based on the electric field integral equation (EFIE). The discretization is based on triangular basis functions [14]. The progress of the design during the optimization is shown in Fig. 3. The initial design is $\boldsymbol{x} = [50\ \ 90\ \ 14]^T$ (mm). The optimal design is $\boldsymbol{x} = [51.51\ \ 96.39\ \ 15.004]^T$ (mm).

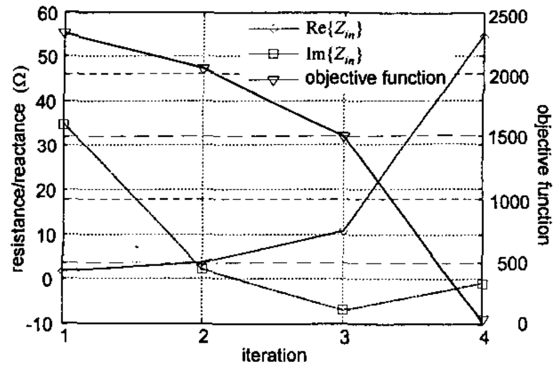


Fig. 3. The progress of the objective function and the input impedance of the patch antenna design.

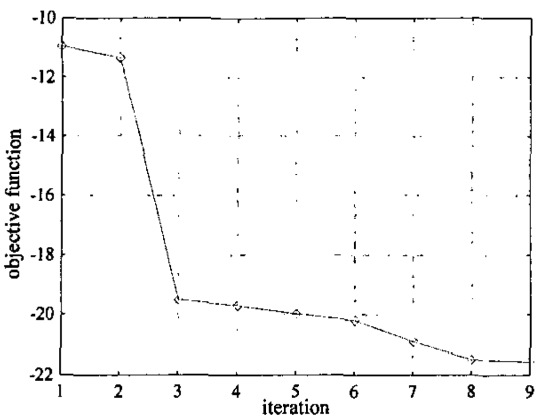


Fig. 4. The progress of the objective function for the gain optimization of the Yagi-Uda array.

C. Maximum Directivity of a Yagi-Uda Array

The directivity of the Yagi-Uda antenna of Fig. 1 is optimized by maximizing the radiation intensity in the direction of maximum radiation $(\theta = 90^\circ, \varphi = 90^\circ)$

$$f(x) = -\left|A_z(\theta = 90^\circ, \varphi = 90^\circ)\right|^2 \qquad (17)$$

where A_z is the only non-zero component of the magnetic vector potential generated by the antenna. The design space is $x = [s_{3n}\ s_{4n}\ s_{5n}]^T$, where the subscript n denotes normalization with respect to the wavelength λ. In this case, the objective function depends on all currents I_k $(k = 1,\dots,m)$ and the adjoint excitation $\hat{V}$ is a full column vector. The initial design is the one optimized for $Z_{in} = 73\ \Omega$, with $x = [0.34\ 0.34\ 0.34]^T$. The optimal design is $x = [0.3735\ 0.4471\ 0.4353]^T$. The gain of the antenna at the initial design is $G^{(0)} = 12.75$ (11.06 dB). After the optimization is completed, $G^{(9)} = 15.08$ (11.78 dB).

V. CONCLUSIONS

A feasible adjoint variable method to design sensitivity analysis with frequency domain full wave EM solvers is proposed. A theory and possible implementations of adjoint-based gradient optimization of high-frequency structures are presented. Important issues related to the formulation of the adjoint system, the accuracy of the sensitivity estimation and the objective functions are discussed and illustrated through MoM analysis.

REFERENCES

[1] E.J. Haug, K.K. Choi and V. Komkov, *Design Sensitivity Analysis of Structural Systems*, Orlando, Florida: Academic Press, Inc., 1986.

[2] S.W. Director and R.A. Rohrer, "The generalized adjoint network and network sensitivities," *IEEE Trans. Circuit Theory*, vol. CT-16, pp. 318-323, 1969.

[3] J.W. Bandler and R.E. Seviora, "Current trends in network optimization," *IEEE Trans. Microwave Theory and Tech.*, vol. MTT-18, pp. 1159-1170, 1970.

[4] V.A. Monaco and P. Tiberio, "Computer-aided analysis of microwave circuits," *IEEE Trans. Microwave Theory and Tech.*, vol. MTT-22, pp. 249-263, 1974.

[5] J.W. Bandler, "Computer-aided circuit optimization," Chapter 6 in G.C. Temes and S.K. Mitra, Eds., *Modern Filter Theory and Design*, John Wiley & Sons, 1973.

[6] K.C. Gupta, R. Garg and R. Chadha, *Computer-Aided Design of Microwave Circuits*, Dedham, MA: Artech, 1981.

[7] J.W. Bandler, *Optimization*, vol. 1, Custom Courseware for *Computer Engineering 3KB3*, 1994, McMaster University.

[8] J.W. Bandler and R.E. Seviora, "Wave sensitivities of networks," *IEEE Trans. Microwave Theory and Tech.*, vol. MTT-20, pp. 138-147, 1972.

[9] G. Iuculano, V.A. Monaco and P. Tiberio, "Network sensitivities in terms of scattering parameters," *Electronic Letters*, vol. 8, pp. 53-54, 1972.

[10] F. Alessandri, M. Mongiardo and R. Sorrentino, "New efficient full wave optimization of microwave circuits by the adjoint network method," *IEEE Microwave and Guided Wave Letters*, vol. 3, pp. 414-416, 1993.

[11] P. Harscher, S. Amari and R. Vahldieck, "Optimization of microwave circuits using analytically calculated gradients in the finite element method," in *IEEE MTT-S Int. Symp. Digest*, vol. 2, 2000, pp. 891-894.

[12] S. Amari, "Numerical cost of gradient computation within the method of moments and its reduction by means of a novel boundary-layer concept," in *IEEE MTT-S Int. Symp. Digest*, vol. 3, 2001, pp. 1945-1948.

[13] J.W. Bandler, Q.J. Zhang, J. Song and R.M. Biernacki, "FAST gradient based yield optimization of nonlinear circuits," *IEEE Trans. Microwave Theory and Tech.*, vol. MTT-38, pp. 1701-1710, 1990.

[14] S.M. Rao, D.R. Wilton and A.W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. on AP*, vol. AP-30, pp. 409-418, 1982.