RECENT TRENDS IN TIME-DOMAIN ADJOINT SENSITIVITY ESTIMATION

Payam Abolghasem [1] , Mohamed H. Bakr [2] , Natalia K. Nikolova [3]

[1] *Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, L8S 4K1,*
Email: *payam@grads.ece.mcmaster.ca*

[2] *As (1) above, but Email: mbakr@mail.ece.mcmaster.ca*

[3] *As (1) above, but Email: talia@mail.ece.mcmaster.ca*

ABSTRACT

We review recent developments in the Adjoint Variable Method (AVM) for time-domain sensitivity analysis. AVM efficiently estimates the sensitivities of a given objective function with respect to all designable parameters using only two simulations of the original and the adjoint structures. During these two simulations, the original and the adjoint field information are stored only at few nodes related to the designable parameters. Using the stored field components, the sensitivities of the objective function with respect to all designable parameters are estimated. The AVM approach is extended to calculate the sensitivities of complex valued objective functions such as the S-parameters over a wide frequency band. The AVM approach is illustrated through estimating objective function sensitivities with respect to waveguide discontinuity examples.

INTRODUCTION

The design problem of an electromagnetic structure can be formulated as

$$\boldsymbol{x}^* = \arg\{\min_{\boldsymbol{x}} F(\boldsymbol{x}, \boldsymbol{R}(\boldsymbol{x}))\} \tag{1}$$

where $\boldsymbol{x}^*$ is the set of optimal parameters, $\boldsymbol{R}(\boldsymbol{x})$ is the vector of responses, and F is the objective function. The problem (1) is usually solved using gradient-based optimizers. These optimizers require not only the structure response but also its derivatives as well. The classical approach for extracting sensitivities using finite differences can be time-intensive even for problems with a small number of designable parameters.

An adjoint variable method (AVM) has been recently developed and incorporated with time-domain TLM problems for efficient sensitivity analysis [1]. AVM estimates the gradient of a given objective function using only two simulations of the original and the adjoint problem. It is assumed that the objective function has the form

$$F = \int_0^{T_s} \psi(\boldsymbol{x}, \boldsymbol{V})\, dt \tag{2}$$

where T_s is the simulation time, and the function $\psi(\boldsymbol{x}, \boldsymbol{V})$ is the kernel of the objective function. Structures with non-dispersive boundaries or with dispersive Johns matrix boundaries are considered [2]. The approach is further developed to estimate the sensitivities of objective function with respect to dielectric discontinuities [3]. Recently, the AVM approach has been extended to calculate the S-parameter sensitivities of an N_P port network [4]. In contrast with central difference approximations that requires $2nN_P$ extra simulations, where n is the number of designable parameters, the AVM approach estimates the S-parameter sensitivities using only N_P extra simulations.

We start by a review of the AVM technique for time-domain TLM for both dispersive and non-dispersive boundaries. Different approaches for estimating sensitivities with respect to dielectric discontinuities are considered. We then discuss the estimation of S-parameter sensitivities by transforming the frequency dependant objective function into the framework of time-domain AVM. The AVM approach is illustrated through sensitivity calculations with respect to dimensions of waveguide discontinuities.

THE AVM APPROACH FOR TIME-DOMAIN TLM

The time domain TLM simulation with non-dispersive boundaries at the kth time step can be expressed as [1]

$$V_{k+1} = CSV_k + V_k^s \quad , \quad V(0) = 0 \tag{3}$$

Here, we assume that the computational domain is uniformly discretized into N nodes with node size Δl. We refer to the total number of associated TLM links as N_L. We denote the time step by Δt. In (3), $\boldsymbol{V}_k \in \Re^{N_L}$ is the vector of incident impulses for all nodes at the kth time step, $\boldsymbol{S} \in \Re^{N_L \times N_L}$ is the overall block diagonal scattering matrix whose pth block, $\boldsymbol{S}^p$ is the pth nodal scattering matrix, $\boldsymbol{C} \in \Re^{N_L \times N_L}$ is the symmetric connection matrix, and the vector $\boldsymbol{V}_k^s \in \Re^{N_L}$ is the excitation vector.

The Adjoint Variable Method (AVM) estimates the sensitivities of the real objective function (2) by carrying out the backward adjoint TLM simulation

$$\boldsymbol{\lambda}_{k-1} = \boldsymbol{S}^T(\boldsymbol{x})\boldsymbol{C}^T(\boldsymbol{x})\boldsymbol{\lambda}_k - \boldsymbol{\lambda}_k^s \quad , \quad \boldsymbol{\lambda}(T_s) = \boldsymbol{0} \tag{4}$$

where $\boldsymbol{\lambda}_k \in \Re^{N_L}$ is the vector of adjoint impulses of all nodes at the kth time step, and $\boldsymbol{\lambda}_k^s = \Delta t(\partial\psi / \partial \boldsymbol{V})_{t=k\Delta t}$ is the adjoint excitation obtained from the original TLM simulation. We notice that the scattering matrix of the adjoint simulation is the transpose of that in the original simulation. Also, the order of the scattering and the connection steps in the two simualtions is reversed. The AVM sensitivities of (2) with respect to the ith designable parameter are then obtained using [1]

$$\frac{\partial F}{\partial x_i} \approx \frac{\partial^e F}{\partial x_i} - \int_0^{T_s} \boldsymbol{\lambda} \frac{\Delta \boldsymbol{A}_i}{\Delta x_i} \boldsymbol{V} dt \approx \frac{\partial^e F}{\partial x_i} - \Delta t \sum_{k=0}^{N_t} \boldsymbol{\lambda}_k^T \boldsymbol{\eta}_k^i \tag{5}$$

where the matrix $\Delta \boldsymbol{A}_i$ is the change in the system matrix $\boldsymbol{A} = (\boldsymbol{CS} - \boldsymbol{I})/\Delta t$ due to a perturbation Δx_i of the ith designable parameter, the term $\partial^e F / \partial x_i$ denotes the explicit dependence of (2) on the ith parameter, and N_t is the number of time steps with $T_s = N_t \Delta t$. The vector $\boldsymbol{\eta}_k^i = (\Delta \boldsymbol{A}_i / \Delta x_i)\boldsymbol{V}_k$ depends on the way nodes are affected by perfectly conducting or dielectric discontinuities. We assume that the perturbation Δx_i is small enough to approximate the distribution of the incident impulses of the perturbed structure with those obtained from the unperturbed one [1]. For the cases that Δx_i results in "metallizing" or "demetallizing" certain TLM nodes, the associated changes in the system matrix, $\Delta \boldsymbol{A}_i$ is modeled by changes in the system connection matrix. In this case, the vector $\boldsymbol{\eta}_k^i$ at the kth time step is obtained as

$$\boldsymbol{\eta}_k^i = \frac{1}{\Delta x_i \Delta t} \Delta \boldsymbol{C}(\boldsymbol{x})\boldsymbol{S}(\boldsymbol{x})\boldsymbol{V}_k \tag{6}$$

When the perturbation Δx_i affects the dielectric properties of certain nodes, the changes in the system scattering matrix models the perturbation in the matrix $\boldsymbol{A}$. The vector $\boldsymbol{\eta}_k^i$ is thus given by

$$\boldsymbol{\eta}_k^i = \frac{1}{\Delta x_i \Delta t} \boldsymbol{C}(\boldsymbol{x})\Delta \boldsymbol{S}(\boldsymbol{x})\boldsymbol{V}_k \tag{7}$$

A second approach for estimating the adjoint sensitivities with dielectric discontinuities is proposed in [3]. In this technique the required approximation of the incident impulses is avoided by exploiting the analytical dependence of the scattering matrix on the material properties to obtain an exact adjoint system. It is shown that the estimated sensitivities can then be approximated by

$$\frac{\partial F}{\partial x_i} \approx \frac{\partial^e F}{\partial x_i} - \Delta t \sum_{n=l}^{l+m} \sum_{k=0}^{N_t} \boldsymbol{\lambda}_k^T \frac{\partial \boldsymbol{A}}{\partial \varepsilon_{r,n}} \boldsymbol{V}_k \frac{\Delta \varepsilon_{r,n}}{\Delta x_i} \tag{8}$$

where the first summation is carried out over the set of nodes whose dielectric properties are affected by the perturbation Δx_i.

The adjoint variable method is extended to include electromagnetic structures with dispersive Johns matrix boundaries [2]. In this case, the original TLM simulation is given by

$$\boldsymbol{V}_{k+1} = \boldsymbol{CSV}_k + \boldsymbol{V}_k^s + \sum_{k'=0}^{k} \boldsymbol{G}(k-k')\boldsymbol{V}_{k'}^R \quad , \quad \boldsymbol{V}(0) = \boldsymbol{0} \tag{9}$$

where $\boldsymbol{G}(k) \in \Re^{N_L \times N_L}$ is the kth time layer of the three-dimensional Johns matrix [2]. Similar to (4), the adjoint impulses are obtained from the backward TLM simulation

$$\boldsymbol{\lambda}_{k-1} = \boldsymbol{S}^T(\boldsymbol{x})\left[\boldsymbol{C}^T(\boldsymbol{x})\boldsymbol{\lambda}_k + \sum_{k'=k}^{N_t} \boldsymbol{G}^T(k'-k)\boldsymbol{\lambda}_{k'}\right] - \boldsymbol{\lambda}_k^s \quad , \quad \boldsymbol{\lambda}(T_s) = \boldsymbol{0} \tag{10}$$

THE AVM APPROACH FOR S-PARAMETER SENSITIVITES

An algorithm for estimating the S-parameter sensitivities within the framework of time-domain AVM has been recently developed [4]. For an N_p port electromagnetic structure, the S-parameters S_{pq}, p,q = 1, 2,…, N_P, are extracted by dividing the output spectrum of the desired mode at the pth port by the reference spectrum of the qth port. Assuming that the reference spectrum is independent of the designable parameters, the sensitivities of the S-parameters are obtained from the sensitivities of the corresponding output spectrum. For the monochromatic case, where the structure is excited with a single sinusoidal at the qth port, the real and imaginary parts of the output spectrum of the desired mode are given by

$$\begin{bmatrix} \mathrm{Re}(\tilde{E}_{pq}(f_0)) \\ \mathrm{Im}(\tilde{E}_{pq}(f_0)) \end{bmatrix} = \int_0^{T_s} \left\{ \iint_{p\text{th port}} E(t,\boldsymbol{r}) E_p(\boldsymbol{r})\, ds \right\} \begin{Bmatrix} \cos(2\pi f_0 t) \\ -\sin(2\pi f_0 t) \end{Bmatrix} dt \tag{11}$$

where $\boldsymbol{r}$ denotes the position vector, $E_p(\boldsymbol{r})$ represents the transversal mode distribution at location $\boldsymbol{r}$ of the pth port, and f_0 is the excitation frequency. Expression (11) defines two real objective functions similar to (2). Their sensitivities with respect to the designable parameters are obtained using the adjoint excitation vectors [4]

$$\begin{pmatrix} \boldsymbol{\lambda}_{\mathrm{Re},k}^{S} \\ \boldsymbol{\lambda}_{\mathrm{Im},k}^{S} \end{pmatrix} = \Delta t \Delta s \boldsymbol{a} E_p(\boldsymbol{r}_j) \begin{Bmatrix} \cos(2\pi f_0 k \Delta t) \\ -\sin(2\pi f_0 k \Delta t) \end{Bmatrix} \tag{12}$$

where Δs is the port surface element, and the vector $\boldsymbol{a}$ relates the incident impulses to the transversal electric field components [4]. We notice that the only difference between the adjoint excitations in (12) is a constant phase shift equal to $\pi/2$. For a sufficient number of time steps, it is shown that one of the adjoint simulations in (12) can be avoided by taking into account only the steady state part of the adjoint impulses [4]. It follows that the sensitivities of the real and imaginary parts of the output spectrum of the pth port due to the excitation of the qth port are

$$\frac{\partial\, \mathrm{Re}(\tilde{E}_{pq}(f_0))}{\partial x_i} \approx -\Delta t \sum_{k=0}^{N_t} \tilde{\boldsymbol{\lambda}}_{i,k}^{p,\mathrm{Re}} \boldsymbol{\eta}_{i,k}^{q} \quad , \quad \frac{\partial\, \mathrm{Im}(\tilde{E}_{pq}(f_0))}{\partial x_i} \approx -\Delta t \sum_{k=T}^{N_t} \tilde{\boldsymbol{\lambda}}_{i,k}^{p,\mathrm{Re}} \boldsymbol{\eta}_{i,k}^{q} \tag{13}$$

where T corresponds to a shift in the sampling index of the adjoint impulses associated with the $\pi/2$ phase shift in adjoint excitations. Expression (13) indicates that only one additional adjoint simulation is required to estimate the sensitivities of a given S-parameter at a specific port for a single frequency. For an N_p port network, the algorithm requires N_P adjoint simulations to estimate all the S-parameter sensitivities with respect to designable parameters.

For the multi-frequency cases where the S-parameters sensitivities over a range of frequencies are desired, the discussed algorithm has the disadvantage of requiring a large number of adjoint simulations. This problem can be avoided by using a wideband adjoint excitation that covers all the frequencies within the desired frequency range. Discrete Fourier Transform (DFT) is then utilized to decompose the adjoint impulses into their spectral components [4].

EXAMPLE

We apply the AVM approach to estimate the S-parameter sensitivities of a single-resonator filter (See Fig. 1). A square TLM cell of dimension 1.0 mm is utilized. The vector of designable parameters is $\boldsymbol{x} = [D \quad W]^T$. The waveguide is excited at the input port with a Gaussian-modulated sinusoidal signal centered at frequency $f_0 = 4$ GHz. We allow for 4000 time steps. Figures 2 and 3 illustrate the sensitivities of both real and imaginary parts of S_{21} over a sweep of frequencies for $\boldsymbol{x} = [D \quad W]^T = [0.028 \quad 0.013]^T$ using the adjoint-variable method and central finite differences. The results in both cases show good match.

CONCLUSIONS

The adjoint variable method for TLM is reviewed in this paper. In comparison with central-finite difference approximations, AVM enjoys the advantage of estimating the sensitivities of objective function in a mush shorter simulation time, especially when the number of designable parameters are large. This approach is well developed not only to estimate the gradient of real objective function in time-domain, but also complex objective functions such as S-parameters. The AVM approach is illustrated through the estimation of objective function sensitivities with respect to physical dimensions of waveguide discontinuities. Excellent match is obtained with the accurate but time-intensive central-finite-differences.

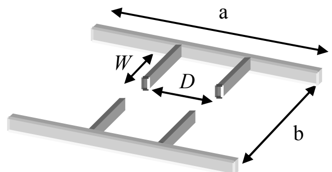

Fig. 1. Single-Resonator filter.

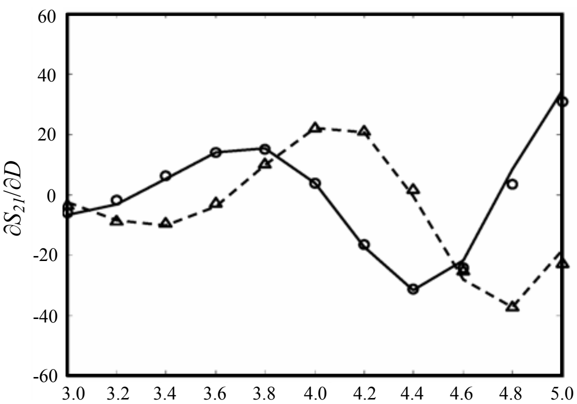

Fig. 2. The sensitivities of S_{21} with respect to D over a sweep of frequencies for [D W] = [28Δl 13Δl]; the imaginary part sensitivities obtained through AVM (—); the imaginary part sensitivities obtained through central differences (o); the real part sensitivities obtained through AVM (--); and the real part sensitivities obtained using central differences(Δ).

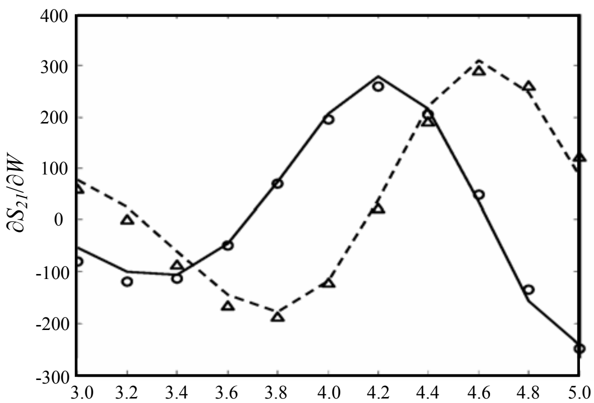

Fig. 3. The sensitivities of S_{21} with respect to W over a sweep of frequencies for [D W] = [28Δl 13Δl]; the imaginary part sensitivities obtained through AVM (—); the imaginary part sensitivities obtained through central differences (o); the real part sensitivities obtained through AVM (--); and the real part sensitivities obtained using central differences(Δ).

REFERENCES

[1] M.H. Bakr and N.K. Nikolova, "An adjoint variable method for time domain TLM with fixed structured grids," *IEEE Trans.Microwave Theory Tech*., vol. 52, Feb. 2004, pp. 554-559.

[2] M.H. Bakr and N.K. Nikolova, "An adjoint variable method for time domain TLM with wideband Johns matrix boundaries," *IEEE Trans. Microwave Theory Tech*., vol. 52, February 2004, pp. 678-685.

[3] P.A.W. Basl, M.H. Bakr and N.K. Nikolova, "Efficient estimation of sensitivities in TLM with dielectric discontinuities," *Microwave and Wireless Components Letters*, vol 15, No. 2, February 2005, pp. 89-91.

[4] M.H. Bakr and N.K. Nikolova, "Efficient estimation of adjoint-variable *S*-parameter sensitivities with time domain TLM," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 18, No. 2, March 2005, pp. 171-187.

[5] A.D. Belegundu and T.R. Chandrupatla, *Optimization Concepts and Applications in Engineering*. Upper Saddle River, NJ: Prentice-Hall Inc., 1999.