Feasible Adjoint Sensitivity Technique for EM Design Optimization

Natalia K. Georgieva, *Member, IEEE*, Snezana Glavic, Mohamed H. Bakr, *Member, IEEE*, and John W. Bandler, *Fellow, IEEE*

***Abstract*—An adjoint-variable approach to frequency-domain design sensitivity analysis is proposed for the optimization of high-frequency structures with full-wave electromagnetic solvers. We investigate sensitivity estimations based on a feasible perturbation technique which is versatile and requires only minor modifications of existing analysis algorithms. It extends the feasible adjoint-sensitivity technique previously applied in nonlinear microwave circuits to full-wave electromagnetic analysis. 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 sensitivity analysis and the design of a Yagi–Uda array and a rectangular patch antenna using suitable method of moments simulators.**

***Index Terms*—Adjoint techniques, antenna design, computer-aided design, design automation, frequency-domain analysis, optimization, sensitivity.**

I. Introduction

The purpose of system design sensitivity analysis is to evaluate the sensitivity of the response of a system to variations of its design parameters. The design sensitivity is represented by the gradient of a given response function in the design parameter space. In high-frequency structure analysis, the design parameters typically describe the structure's geometry and the electromagnetic (EM) properties of the media involved. The system response may be defined as 1) a distributed response represented by the state variables such as current or field distributions; 2) a set of engineering parameters describing the structure's performance such as S or Z parameters; and 3) a single scalar function, which represents some kind of a global performance measure, such as the objective function in an optimization problem. 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 design sensitivity analysis for gradient optimization with full-wave frequency domain EM solvers.

The adjoint-variable method (AVM) for design sensitivity analysis 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]–[7], control theory, etc. Adjoint sensitivities for circuit computer-aided design (CAD) can be found even in undergraduate courses [8]. 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 with applications almost exclusively limited to finite-element analysis [9]–[11].

The adjoint-based design sensitivity analysis 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*. The first applications of the adjoint network method 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], [12], [13]. Later, Alessandri *et al.* [14] applied the adjoint network method to the analysis of microwave circuits whose subnetworks were represented by Y parameters. Typically, the adjoint network method considers the sensitivity of a response with respect to a single state variable [4], which makes its applications problem specific. It is not immediately obvious how it can be utilized in a full-wave analysis.

Recently, an exact sensitivity technique was proposed for applications with the method of moments (MoM) [15] and the boundary layer concept was proposed to reduce the computational load associated with overhead computations related to derivative estimations. In effect, this technique is based on the *direct differentiation method* [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 review the mathematical background of the AVM and discuss feasible implementations in the sensitivity analysis of linear, time-harmonic EM problems. Three major issues are discussed: 1) the adjoint problem; 2) the procedure to efficiently evaluate the gradient of the response function; and 3) the formulation of the objective function in adjoint-based gradient optimization. The proposed approach can substantially increase the efficiency of current CAD tools based on full-wave frequency domain analysis, as it allows the computation of the objective function and its gradient through a single analysis. Unlike previous work on EM-based exact sensitivities, the algorithm is not limited to a specific numerical technique, or to a specific class of high-frequency structures, or to a specific type of response

Manuscript received April 5, 2002; revised August 13, 2002. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Grant OGP0227660-00, Grant OGP0007239-98, and Grant STR234854-00 through the Micronet Network of Centres of Excellence and Bandler Corporation.

N. K. Georgieva, S. Glavic, and M. H. Bakr are with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada L8S 4K1.

J. W. Bandler is with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada L8S 4K1 and also with the Bandler Corporation, Dundas, ON, Canada L9H 5E7.

Digital Object Identifier 10.1109/TMTT.2002.805131

functions. It constitutes a versatile CAD approach which is compatible with existing EM solvers and which requires only minor additions to their respective software.

II. Overview of the Adjoint Techniques for Linear Systems

Here, we review the basic concepts of the AVM for design sensitivity analysis in the case of a general linear problem [1], [7]. 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 nor the discretization techniques are discussed hereafter, because the feasible adjoint sensitivity technique (FAST) does not require analytical derivatives of the coefficients of the matrices associated with a specific computational EM method. We should note that the AVM can be extended to the design sensitivity analysis of nonlinear systems. Nonlinear circuit sensitivities and feasible approaches to their estimation are discussed in [8] and [16].

A. Sensitivity of the Solution via Direct Differentiation

Using notations typical for the MoM analysis, a linear EM problem is represented by

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

Here, $\boldsymbol{x}$ is the vector of design parameters, $\boldsymbol{I}$ is the state-variable vector, e.g., complex currents or current densities in the MoM, or field distribution in the FEM, and $\boldsymbol{V}$ is the global excitation vector which, in general, depends on the sources and the boundary conditions of the problem. Since the $\boldsymbol{Z}$ matrix depends on the structure's geometry and materials, the solution $\boldsymbol{I}$ is an implicit function of the design parameters $\boldsymbol{x}$. In the following discussion, the sensitivity of the $\boldsymbol{Z}$ matrix with respect to the design parameters $\boldsymbol{x}$ plays an important role. The $\boldsymbol{Z}$ coefficients may be 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 with respect to the node's coordinates. In MoM solvers, however, the $\boldsymbol{Z}$ coefficients are integrals of a specific Green's function, which may depend on the user-defined design parameters in a complicated way. In such a case, it is preferable to approximate the $\boldsymbol{Z}$ matrix sensitivities using finite differences, as we discuss later. This ensures the versatility and feasibility of the proposed design-sensitivity-analysis technique.

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, always has an implicit dependence on $\boldsymbol{x}$. The objective is to determine the sensitivity of the response function with respect to the design parameters $\boldsymbol{x}$, i.e.,

$$\nabla_x f, \text{ subject to } \boldsymbol{Z}\boldsymbol{I} = \boldsymbol{V} \tag{2}$$

where ∇_x is the row operator

$$\nabla_x = \left[\frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2}\cdots\frac{\partial}{\partial x_n}\right]. \tag{3}$$

Assuming that the $\boldsymbol{Z}$ matrix is not singular, $\nabla_x \boldsymbol{I}$ is obtained from (1) as

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

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

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

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

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

Equation (4) is the basis of the *direct differentiation method* [1]. It provides the means of efficient calculation of the gradient of the state-variable vector. There is no need for additional $\boldsymbol{Z}$ matrix *LU* factorization, since this has already been done at the analysis stage of the current design. The overhead due to the computation of $\nabla_x \boldsymbol{I}$, in addition to the solution of (1), is associated with: 1) the computation of the $\boldsymbol{Z}$ matrix sensitivities $\partial\boldsymbol{Z}/\partial x_i \quad (i = 1, \dots, n)$ and 2) n back substitutions of the *LU*-factored $\boldsymbol{Z}$ matrix in (6).

B. Sensitivity of the Response Function Using the Adjoint Solution

The solution of (6) can be subsequently used to calculate the exact sensitivities of $f(\boldsymbol{x}, \bar{\boldsymbol{I}}(\boldsymbol{x}))$ by the substitution of the computed state-variable sensitivities $\nabla_x \boldsymbol{I}$ into

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

where $\nabla_{\boldsymbol{I}}$ is a row operator analogous to ∇_x [see (3)]. The gradient $\nabla_x^e f$ reflects the explicit dependence of $f(\boldsymbol{x}, \bar{\boldsymbol{I}}(\boldsymbol{x}))$ on $\boldsymbol{x}$.

In an optimization problem, however, we are interested in the sensitivity of the response function $\nabla_x f$, rather than the sensitivity of the state-variables $\nabla_x \boldsymbol{I}$. In such a case, the AVM provides the most computationally efficient sensitivity estimation. Substituting (4) into (7) leads to

$$\nabla_x f = \nabla_x^e f + \nabla_{\boldsymbol{I}} f \boldsymbol{Z}^{-1}\left[\nabla_x \boldsymbol{V} - \nabla_x(\boldsymbol{Z}\bar{\boldsymbol{I}})\right]. \tag{8}$$

Here, the vector

$$\hat{\boldsymbol{I}} = \left[\nabla_{\boldsymbol{I}} f \boldsymbol{Z}^{-1}\right]^T = \left[\boldsymbol{Z}^T\right]^{-1}\left[\nabla_{\boldsymbol{I}} f\right]^T \tag{9}$$

is introduced. It is a solution to

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

and is referred to as the *adjoint variable* vector. The right-hand side of (10)

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

is the *adjoint excitation*, which is the gradient of the response function in the state-variable space. Equation (10) describes the so-called *adjoint problem*. The factored $\boldsymbol{Z}^T$ matrix is obtained easily from the factored $\boldsymbol{Z}$ matrix of the original system. The response sensitivities can now be computed in terms of the original solution $\bar{\boldsymbol{I}}$ and the adjoint solution $\hat{\boldsymbol{I}}$ as

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

Equations (10) and (12) form the basis of the AVM.

C. Computational Advantages and Feasibility of the AVM

In the AVM, the overhead due to the computation of $\nabla_x f$ in addition to the solution of (1) is associated with: 1) the computation of the $\boldsymbol{Z}$ matrix sensitivities $\partial \boldsymbol{Z}/\partial x_i \quad (i = 1, \ldots, n)$ needed in (12) and 2) one back substitution of the *LU*-factored $\boldsymbol{Z}^T$ matrix in (9).

The AVM has significant computational advantage in comparison with the traditional calculation of the sensitivities through a finite-difference approach (FDA). The FDA applies finite differences directly to the gradient of the response function. For example, the forward finite-difference approximation of a response derivative is

$$\frac{\partial f\left(x_i^{(k)}\right)}{\partial x_i} \simeq \frac{f\left(x_i^{(k)} + \Delta x_i^{(k)}\right) - f\left(x_i^{(k)}\right)}{\Delta x_i^{(k)}}, \quad i = 1, \ldots, n \tag{13}$$

where $x_i^{(k)}$ is the value of the x_i design parameter at the current (kth) design iteration and $\Delta x_i^{(k)}$ is a specified perturbation. In this case, the FDA performs $(n+1)$ full analyses in order to generate both the response function and its sensitivity. This involves $(n+1)$ $\boldsymbol{Z}$ matrix fills, factorizations, and back substitutions in order to derive all $(n+1)$ solutions to (1). Obviously, the FDA is computationally inefficient in problems involving multiple-design parameters. It becomes even more computationally demanding if higher order approximations of the sensitivities are used. In contrast, the AVM generates the response and its sensitivities through a single analysis, regardless of the number of design parameters n. The feasible approach presented here requires $(n+1)$ $\boldsymbol{Z}$ matrix fills, one matrix factorization, and one additioinal back substitution.

The AVM has better computational efficiency in comparison with the direct differentiation method as well. In the direct differentiation method, according to (6), n back substitutions of the factored $\boldsymbol{Z}$ matrix are needed to compute $\nabla_x \boldsymbol{I}$ and, subsequently, to obtain $\nabla_x f$. In the AVM, according to (10) and (12), there is only one back substitution needed regardless of n: the one used to compute $\hat{\boldsymbol{I}}$.

The advantage of the FDA is its simplicity and the ease of implementation with any EM solver. This makes it the most popular (if not the only) approach currently used in commercial high-frequency CAD tools together with other more sophisticated (quadratic and cubic) locally valid approximations of the response used to obtain sensitivity estimates. Here, we propose a feasible technique based on adjoint sensitivities, which is equally easy to implement and in the same time retains high computational efficiency. From (10) and (12), it is clear that there are two types of derivatives involved in the computation of $\nabla_x f$ through the AVM; namely, the matrices $\partial \boldsymbol{Z}/\partial x_i$ $(i = 1, \ldots, n)$ in $\nabla_x(\boldsymbol{Z}\bar{\boldsymbol{I}})$ and the adjoint excitation $\hat{\boldsymbol{V}} = [\nabla_{\boldsymbol{I}} f]^T$. The matrices $\partial \boldsymbol{Z}/\partial x_i \quad (i = 1, \ldots, n)$ may be analytically available. However, as discussed before, their evaluation is not only solver specific, but also far from trivial. Besides, the analytical evaluation of the $\boldsymbol{Z}$ matrix sensitivities does not improve the computational efficiency in comparison with the finite-difference approximations $\Delta \boldsymbol{Z}/\Delta x_i \quad (i = 1, \ldots, n)$ [16]. These approximations require only minor simple modifications in an existing EM code, which do not depend on the nature of the numerical algorithm. The important issue here is whether these approximations can affect the accuracy of the sensitivity estimation via the AVM. As we demonstrate later, the accuracy of the sensitivity estimation via (12) is preserved. This is due to the nearly linear dependence of the majority of the elements of the $\boldsymbol{Z}$ matrix on a geometrical design parameter x_i for small perturbations (Δx_i from 1% to 5%) (see also [16]).

The construction of the adjoint problem requires the computation of the adjoint excitation vector $\hat{\boldsymbol{V}} = [\nabla_{\boldsymbol{I}} f]^T$, which depends entirely on the way the user defines the response function $f(\boldsymbol{x}, \bar{\boldsymbol{I}}(\boldsymbol{x}))$. In general, it is desirable that $f(\boldsymbol{x}, \bar{\boldsymbol{I}}(\boldsymbol{x}))$ is analytically differentiable in I_k $(k = 1, \ldots, m)$ because the accuracy of the adjoint solution $\hat{\boldsymbol{I}}$ through (10) depends strongly on the accuracy of $\hat{\boldsymbol{V}}$. Our numerical tests show that inaccurate finite-difference approximations of $\hat{\boldsymbol{V}}$ may result in deterioration of the sensitivity analysis via (10) and (12).

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 can, for example, appear in a least pth objective function as

$$f = \sum_j \frac{1}{p} |e(\omega_j)|^p \tag{14}$$

where ω_j denotes the jth frequency of interest. Then [3]

$$\nabla_{\boldsymbol{I}} f = \sum_j \mathbf{Re}\left\{|e(\omega_j)|^{p-2} e^*(\omega_j) \nabla_{\boldsymbol{I}} e(\omega_j)\right\}. \tag{15}$$

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 $\hat{\boldsymbol{V}}$ is computed accurately.

IV. Results and Discussion

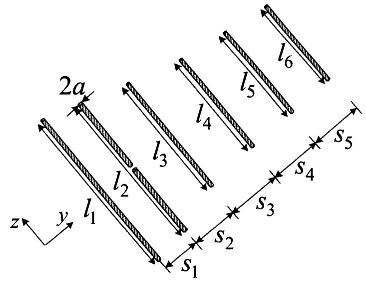
A. Finite-Difference Approximations and the Accuracy of the Adjoint Sensitivity Analysis

We illustrate the adjoint sensitivity analysis through the computation of the input impedance and gain sensitivities of a Yagi–Uda array whose nominal geometry is shown in Fig. 1. The analysis is based on Pocklington's equation. The discretization is carried out via the MoM using pulse expansion functions and a point matching technique. The complex system of (1) is reduced to a system of real-valued equations

$$\begin{bmatrix} \boldsymbol{Z}_{RR} & \boldsymbol{Z}_{IR} \\ \boldsymbol{Z}_{RI} & \boldsymbol{Z}_{II} \end{bmatrix} \begin{bmatrix} \mathbf{Re}\{\boldsymbol{I}\} \\ \mathbf{Im}\{\boldsymbol{I}\} \end{bmatrix} = \begin{bmatrix} \mathbf{Re}\{\boldsymbol{V}\} \\ \mathbf{Im}\{\boldsymbol{V}\} \end{bmatrix} \tag{16}$$

where $\boldsymbol{Z}_{RR} = \boldsymbol{Z}_{II} = \mathbf{Re}\{\boldsymbol{Z}\}$ and $\boldsymbol{Z}_{RI} = -\boldsymbol{Z}_{IR} = \mathbf{Im}\{\boldsymbol{Z}\}$. The symmetry of the structure is used and, thus, only half

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. Geometry of the Yagi–Uda array.

of the array is actually discretized. Every half-array element is discretized into 45 segments. All design parameters in the Yagi–Uda array example are derived by normalizing the geometrical dimensions with respect to the wavelength, e.g., $x_{i_n} = x_i/\lambda$, where λ is the wavelength.

We show the results for the derivative of the input impedance with respect to the separation distance s_{1_n}, $\partial \boldsymbol{Z}_{\text{in}}/\partial s_{1_n}$. It is first computed with the FDA using central finite differences with 1% perturbation of the current value of the design parameter

$$\begin{aligned}\frac{\partial Z_{\text{in}}}{\partial s_{1_n}} &\simeq \frac{Z_{\text{in}}\left(s_{1_n}^{(k)} + \Delta s_{1_n}^{(k)}\right) - Z_{\text{in}}\left(s_{1_n}^{(k)} - s_{1_n}^{(k)}\right)}{2\Delta s_{1_n}^{(k)}}\\ \Delta s_{1_n}^{(k)} &= 0.01 s_{1_n}^{(k)}.\end{aligned} \tag{17}$$

Here, $s_{1_n}^{(k)}$ is the current value of the design parameter s_{1_n}. Second, we compute $\partial Z_{\text{in}}/\partial s_{1_n}$ using the AVM with analytically calculated adjoint excitation. In this case, the response function Z_{in} is a complex function which depends on a single state variable, namely, the phasor of the current at the base of the driver element I_b, where b is the index of the respective subsection. The sensitivities of $R_{\text{in}} = \mathbf{Re}\{Z_{\text{in}}\}$ and $X_{\text{in}} = \mathbf{Im}\{Z_{\text{in}}\}$ are evaluated. Since the driver is excited with an input voltage whose phasor is set as $V_b = 1$ V, $Z_{\text{in}} = 1/I_b$. The adjoint currents are computed from

$$\begin{bmatrix} \boldsymbol{Z}_{RR} & \boldsymbol{Z}_{IR} \\ \boldsymbol{Z}_{RI} & \boldsymbol{Z}_{II} \end{bmatrix}^T \begin{bmatrix} \hat{\boldsymbol{I}}_R \\ \hat{\boldsymbol{I}}_I \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{V}}_R \\ \hat{\boldsymbol{V}}_I \end{bmatrix}. \tag{18}$$

When the sensitivity of R_{in} is estimated, the adjoint excitation is

$$\begin{aligned}\hat{\boldsymbol{V}}_{Rb} &= \frac{\partial R_{\text{in}}}{\partial \mathbf{Re}\{I_b\}} = \left.\frac{\mathbf{Im}\{I_b\}^2 - \mathbf{Re}\{I_b\}^2}{|I_b|^4}\right|_{I_b=\bar{I}_b}\\ \hat{\boldsymbol{V}}_{Rj} &= 0,\;\; j \neq b\\ \hat{\boldsymbol{V}}_{Ib} &= \frac{\partial R_{\text{in}}}{\partial \mathbf{Im}\{I_b\}} = \left.\frac{-2\mathbf{Re}\{I_b\}\mathbf{Im}\{I_b\}}{|I_b|^4}\right|_{I_b=\bar{I}_b}\\ \hat{\boldsymbol{V}}_{Ij} &= 0,\;\; j \neq b.\end{aligned} \tag{19}$$

The input resistance sensitivity is finally computed as

$$\frac{\partial R_{\text{in}}}{\partial s_{1_n}} = -\begin{bmatrix} \hat{\boldsymbol{I}}_R^T & \hat{\boldsymbol{I}}_I^T \end{bmatrix} \left(\frac{\partial}{\partial s_{1_n}} \begin{bmatrix} \boldsymbol{Z}_{RR} & \boldsymbol{Z}_{IR} \\ \boldsymbol{Z}_{RI} & \boldsymbol{Z}_{II} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{Re}\{\bar{\boldsymbol{I}}\} \\ \mathbf{Im}\{\bar{\boldsymbol{I}}\} \end{bmatrix} \right). \tag{20}$$

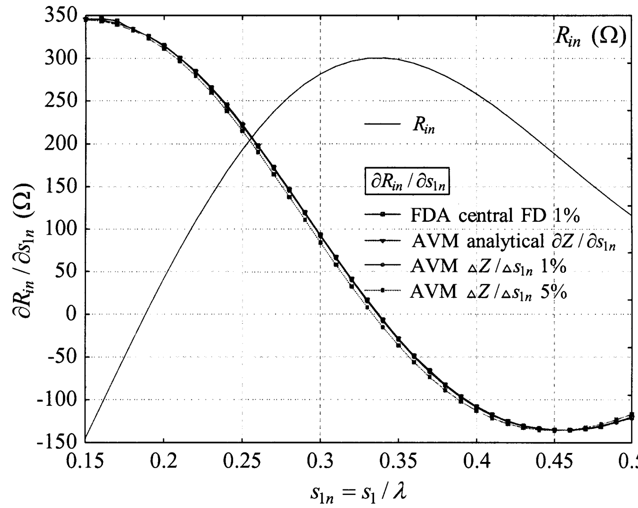


Fig. 2. Input resistance of the Yagi–Uda array and its sensitivity with respect to the normalized separation s_{1n}.

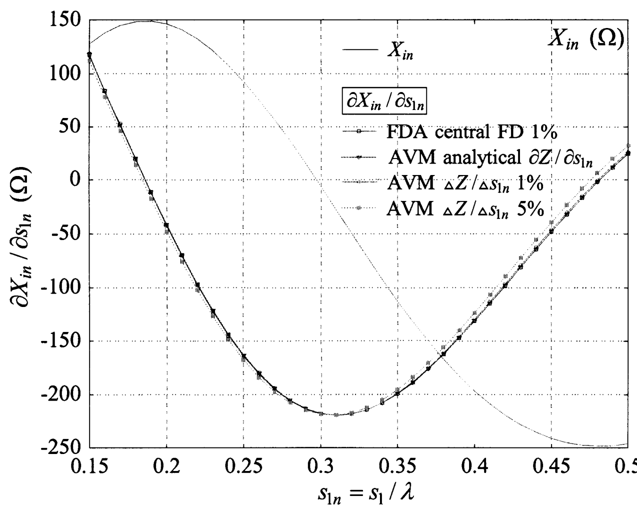


Fig. 3. Input reactance of the Yagi–Uda array and its sensitivity with respect to the normalized separation s_{1n}.

The X_{in} sensitivity is calculated in an analogous manner. Notice that $\partial^e R_{\text{in}}/\partial s_{1_n}$ has been set equal to zero as the input impedance has no explicit dependence on the separation s_{1_n}. The excitation vector does not depend on s_{1_n} either, which sets $\partial \boldsymbol{V}/\partial s_{1_n} = \mathbf{0}$.

The matrix $\partial \boldsymbol{Z}/\partial s_{1_n}$ is calculated in three different ways: 1) analytically; 2) with forward finite differences using 1% perturbation; and 3) with forward finite differences using 5% perturbation. The input resistance and reactance sensitivities computed with the FDA and with the three implementations of the AVM are plotted in Figs. 2 and 3, respectively. The excellent match between the results obtained with the FDA (central finite differences) and the results of the AVM with the analytical $\partial \boldsymbol{Z}/\partial s_{1_n}$ matrix confirms the validity and the accuracy of the proposed technique. It is also evident that the FAST, which relies on the finite-difference approximation of the matrices $\partial \boldsymbol{Z}/\partial x_i$, yields very accurate results. At the same time, its implementation is straightforward as one does not need the derivatives of the specific Green's function with respect to a given design parameter.

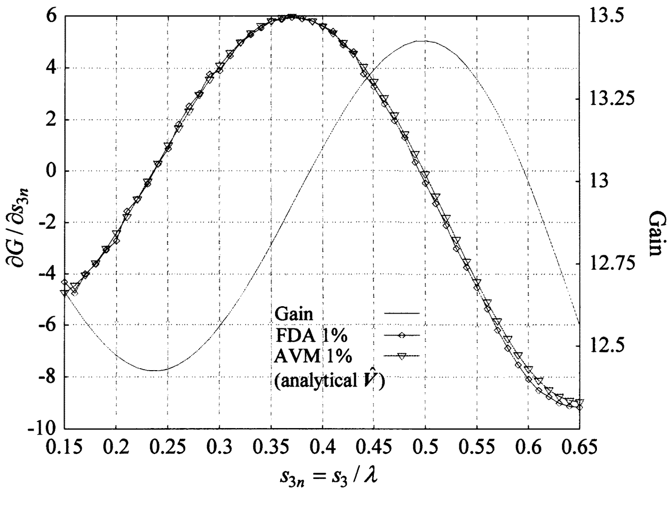


Fig. 4. Gain and gain sensitivity of the Yagi–Uda array with respect to the separation s_3.

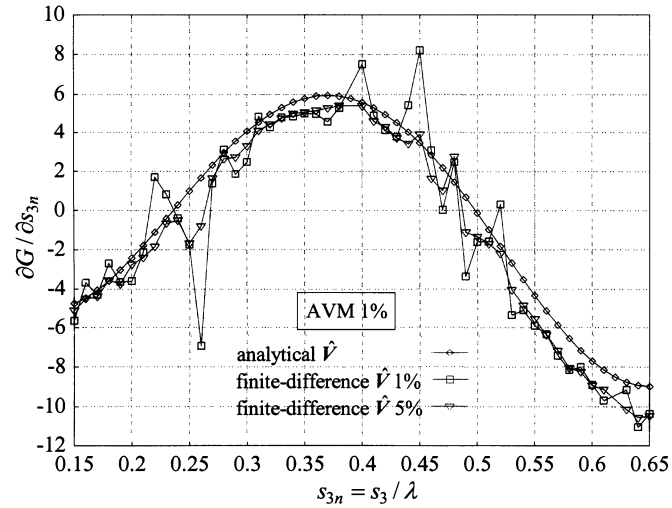


Fig. 5. Gain sensitivity of the Yagi–Uda array with respect to the separation s_3 calculated with the AVM using $\Delta s_{3n} = 0.01 s_{3n}$ perturbations in the calculation of the $\Delta Z / \Delta s_{3n}$ matrix. Adjoint excitation $\hat{V}$ is calculated analytically and with finite differences.

Next, we compute the sensitivities of the antenna gain G with respect to the normalized separation distances s_{i_n} $(i = 2, \ldots, 5)$ and investigate the influence of the finite-difference approximations of the adjoint excitation on the overall accuracy of the sensitivity estimation. The sensitivities are computed for a range of values of s_{i_n} while keeping the rest of the design parameters at the following fixed values: $l_{1_n} = 0.5243$, $l_{2_n} = 0.45$, $l_{3_n} = l_{4_n} = l_{5_n} = l_{6_n} = 0.406$, $s_{l_n} = 0.2607$, and $s_{2_n} = s_{3_n} = s_{4_n} = s_{5_n} = 0.34$. We first obtain a reference gain sensitivity solution applying the FDA with 1% perturbations for each normalized separation distance. The antenna gain G and its derivative with respect to the normalized separation s_{3_n} as a function of s_{3_n} are plotted in Fig. 4. The gain sensitivities are next estimated with the AVM.

The response function in this case is the gain G. Thus, the adjoint excitation vector is

$$\hat{\boldsymbol{V}} = \begin{bmatrix} \frac{\partial G}{\partial \mathbf{Re}\{\boldsymbol{I}\}} \\ \frac{\partial G}{\partial \mathbf{Im}\{\boldsymbol{I}\}} \end{bmatrix}_{\boldsymbol{I}=\bar{\boldsymbol{I}}}. \tag{21}$$

This time, the adjoint excitation (21) has no zero elements because the antenna gain depends on all state variables. The adjoint variable vector $\hat{\boldsymbol{I}}$ is a solution to

$$\begin{bmatrix} \boldsymbol{Z}_{RR} & \boldsymbol{Z}_{IR} \\ \boldsymbol{Z}_{RI} & \boldsymbol{Z}_{II} \end{bmatrix}^T \begin{bmatrix} \hat{\boldsymbol{I}}_R \\ \hat{\boldsymbol{I}}_I \end{bmatrix} = \begin{bmatrix} \frac{\partial G}{\partial \mathbf{Re}\{\boldsymbol{I}\}} \\ \frac{\partial G}{\partial \mathbf{Im}\{\boldsymbol{I}\}} \end{bmatrix}_{\boldsymbol{I}=\hat{\boldsymbol{I}}}. \tag{22}$$

According to (12), the sensitivity $\partial G/\partial s_{i_n}$ with respect to a given separation si_n is computed by

$$\frac{\partial G}{\partial s_{i_n}} = \frac{\partial^e G}{\partial s_{i_n}} - \begin{bmatrix} \hat{\boldsymbol{I}}_R^T & \hat{\boldsymbol{I}}_I^T \end{bmatrix} \times \left(\frac{\partial}{\partial s_{i_n}} \begin{bmatrix} \boldsymbol{Z}_{RR} & \boldsymbol{Z}_{IR} \\ \boldsymbol{Z}_{RI} & \boldsymbol{Z}_{II} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{Re}\{\bar{\boldsymbol{I}}\} \\ \mathbf{Im}\{\bar{\boldsymbol{I}}\} \end{bmatrix} \right), \quad i = 2, \ldots, 5. \tag{23}$$

Notice that the excitation vector $\boldsymbol{V}$ is independent of the separation distances; thus, $\partial \boldsymbol{V}/\partial s_{i_n} = \mathbf{0}$. The gain has an explicit dependence on the separations because they affect the relative positions of the currents which generate the radiated far field. Thus, the derivative $\partial^e G/\partial s_{i_n}$ is nonzero and is evaluated using finite differences (while keeping the current solution $\bar{\boldsymbol{I}}$ fixed).

The adjoint excitation (21) is first calculated using analytical derivatives. The result is substituted in (22) to produce the adjoint variable-solution $\hat{\boldsymbol{I}}$. Finally, the sensitivities $\partial G/\partial s_{i_n}$ $(i = 2, \ldots, 5)$ are calculated via (23). The result for $\partial G/\partial s_{3_n}$ is plotted in Fig. 4, together with the reference solution. Both sensitivity curves are in excellent agreement. This shows that the finite-difference approximations of the $\boldsymbol{Z}$ matrix sensitivities in the AVM do not lead to deterioration of the accuracy of the response sensitivity estimation. Similar results are observed for the gain sensitivity, as well as for the antenna input impedance sensitivities, with respect to any other design parameter.

We now investigate the possibility of applying finite differences to the calculation of the adjoint excitation (21). Instead of using the analytical gradient $\nabla_{\boldsymbol{I}} G$, we compute the adjoint excitation as

$$\hat{\boldsymbol{V}} = \begin{bmatrix} \frac{\Delta G}{\Delta \mathbf{Re}\{\boldsymbol{I}\}} \\ \frac{\Delta G}{\Delta \mathbf{Im}\{\boldsymbol{I}\}} \end{bmatrix}_{\boldsymbol{I}=\bar{\boldsymbol{I}}} \tag{24}$$

applying perturbations to the state variables as: 1) $\Delta\mathbf{Re}\{I_k\} = 0.01\mathbf{Re}\{\bar{I}_k\}$, $\Delta\mathbf{Im}\{I_k\} = 0.01\mathbf{Im}\{\bar{I}_k\}$ and 2) $\Delta\mathbf{Re}\{I_k\} = 0.05\mathbf{Re}\{\bar{I}_k\}$, $\Delta\mathbf{Im}\{I_l\} = 0.05\mathbf{Im}\{\bar{I}_k\}$ $(k = 1, \ldots, m)$. The gain sensitivity $\partial G/\partial s_{3_n}$ results are plotted in Fig. 5 together with the sensitivity generated by the adjoint technique, which uses the analytical gradient $\Delta_{\boldsymbol{I}} G$. It is obvious that the accuracy of the sensitivity estimation is strongly affected by the finite-difference approximation of the adjoint excitation. Moreover, it shows poor convergence since the deviation from the reference solution is stronger for smaller (1%) perturbations of the state variables. This is associated with truncation errors in the floating point representation of numbers. The largest differences between the adjoint excitation values generated by (21) and by (24) appear where the state variables have their smallest absolute values (below 10^{-5} relative to maximum). In all numerical examples, single precision is used.

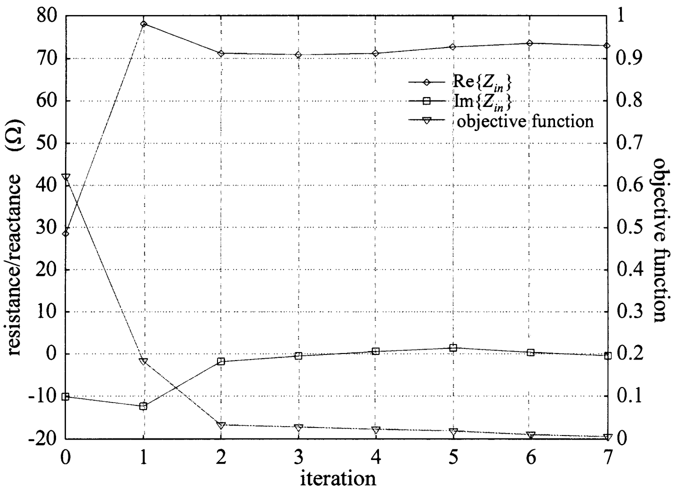


Fig. 6. Progress of the objective function and the input impedance of the Yagi–Uda design.

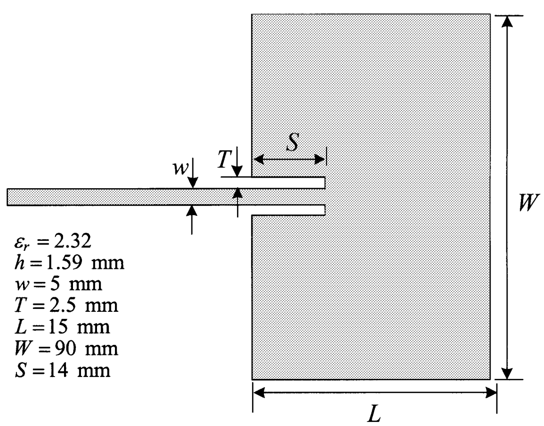


Fig. 7. Geometry of the rectangular patch antenna.

We should note that the finite-difference approximation of the adjoint excitation does not necessarily lead to deterioration of the sensitivity analysis. For example, applying finite differences to the adjoint excitation in (18) for the estimation of the input impedance sensitivities yields very accurate results. This is typical for the cases where the response function depends on a single state variable and is thus not prone to roundoff errors. Nonetheless, it is recommended that the user formulates the response function in such a way that it is analytically or semianalytically differentiable in all state variables.

B. Optimization of the Input Impedance of a Yagi–Uda Array

We use the FAST described above to optimize the input impedance of the Yagi–Uda array shown in Fig. 1. The objective function is defined as

$$f(\boldsymbol{x}) = \left| \frac{Z_{\rm in} - \bar{Z}}{\bar{Z}} \right| \tag{25}$$

where $\bar{Z} = 73\ \Omega$ and $Z_{\rm in}$ is the input impedance of the antenna. The vector of design parameters is $\boldsymbol{x} = [l_{1_n}\ s_{1_n}]^T$. The objective function (25) depends on a single complex-valued current, the current at the driver's base I_b. This dependence follows from $Z_{\rm in} = V_b/I_b$, where $V_b = 1$ (V). The complex error

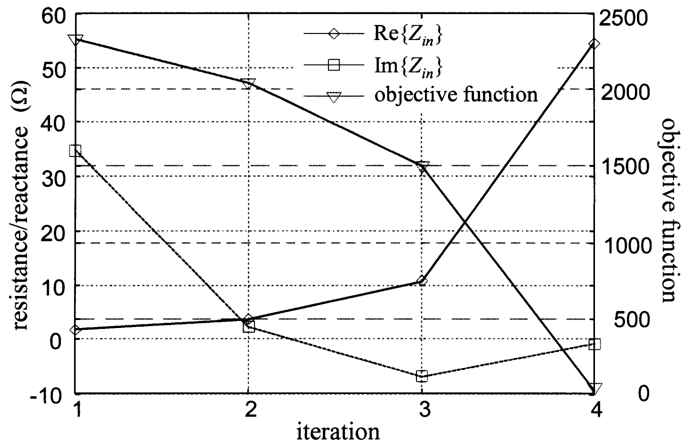


Fig. 8. Progress of the objective function and the input impedance of the patch antenna design.

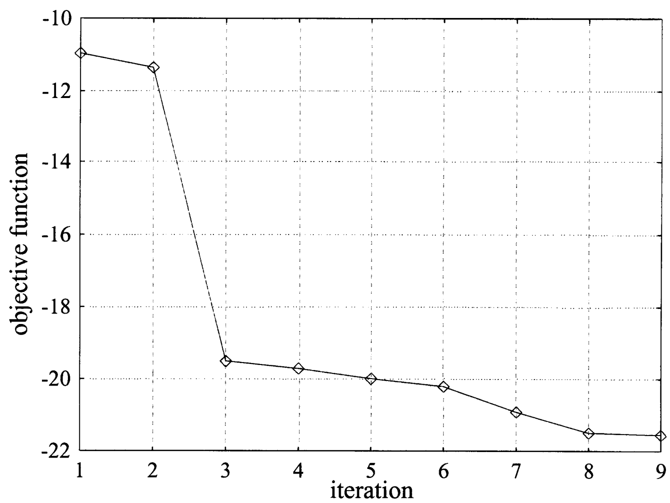


Fig. 9. Progress of the objective function during the gain optimization of the Yagi–Uda array.

$e = (Z_{\rm in} - 73)/73$ is computed at a single frequency, which determines the wavelength used to normalize the geometrical design parameters. Using (15), the gradient of the objective function is obtained as

$$\begin{aligned} \frac{\partial f}{\partial \mathbf{Re}\{I_b\}} &= \mathbf{Re}\left\{ \frac{e^*}{|e|} \frac{\partial e}{\partial \mathbf{Re}\{I_b\}} \right\} \\ \frac{\partial f}{\partial \mathbf{Im}\{I_b\}} &= \mathbf{Re}\left\{ \frac{e^*}{|e|} \frac{\partial e}{\partial \mathbf{Im}\{I_b\}} \right\}. \end{aligned} \tag{26}$$

The derivatives of the complex error function e are easily identified as

$$\frac{\partial e}{\partial \xi} = \frac{1}{73} \left(\frac{\partial R_{\rm in}}{\partial \xi} + j \frac{\partial X_{\rm in}}{\partial \xi} \right), \quad \xi = \mathbf{Re}\{I_b\}, \mathbf{Im}\{I_b\} \tag{27}$$

where

$$\frac{\partial X_{\rm in}}{\partial \mathbf{Re}\{I_b\}} = -\frac{\partial R_{\rm in}}{\partial \mathbf{Im}\{I_b\}} \text{ and } \frac{\partial X_{\rm in}}{\partial \mathbf{Im}\{I_b\}} = \frac{\partial R_{\rm in}}{\partial \mathbf{Re}\{I_b\}}. \tag{28}$$

The respective derivatives of $R_{\rm in}$ have been already given in (19). Thus, the adjoint excitation $\hat{\boldsymbol{V}}$ has only two nonzero elements corresponding to $\partial f/\partial \mathbf{Re}\{I_b\}$ and $\partial f/\partial \mathbf{Im}\{I_b\}$ given by (26).

TABLE I
DESIGN PARAMETERS, INPUT IMPEDANCE, AND GAIN OF THE YAGI–UDA ARRAY

	s_{1n}	s_{2n}	s_{3n}	s_{4n}	s_{5n}	R_{in}	X_{in}	G
1	0.2607	0.3400	0.3735	0.4471	0.4353	47.10	-4.15	15.08
2	0.3455	0.4050	0.3301	0.3853	0.3765	77.77	-23.52	11.08
3	0.3544	0.4294	0.3639	0.4122	0.3544	81.02	-16.26	11.19
4	0.3158	0.3720	0.4229	0.4591	0.4158	73.33	13.25	13.08
5	0.3086	0.3613	0.4232	0.4519	0.4023	65.85	11.18	13.92
6	0.3450	0.3744	0.3953	0.4204	0.3909	70.23	-5.99	12.87
7	0.3214	0.3986	0.3535	0.4653	0.3432	72.36	-5.77	12.91
8	0.3062	0.3923	0.3844	0.4822	0.3362	72.85	5.46	13.41
9	0.2531	0.4357	0.3794	0.3607	0.3645	75.63	0.97	13.25
10	0.2999	0.4061	0.3777	0.4205	0.3627	72.99	-1.00	13.45
11	0.2874	0.4193	0.3685	0.4057	0.3825	72.26	-1.27	13.59
12	0.2884	0.4175	0.3749	0.4064	0.3937	71.51	0.92	13.77
13	0.2906	0.4168	0.3771	0.4046	0.3966	71.80	0.38	13.75

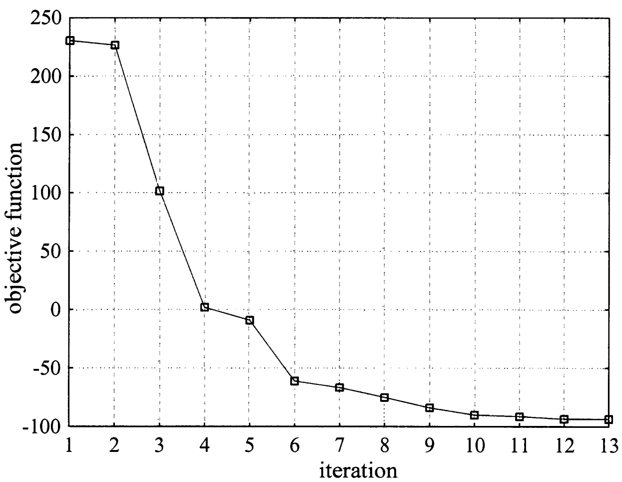


Fig. 10. Progress of the objective function during the optimization of the input impedance and the gain of the Yagi–Uda array.

The result of the optimization of the input impedance of the Yagi–Uda array is shown in Fig. 6. The optimal design is obtained as $\mathbf{x} = [0.5243\ 0.2607]^T$. At each iteration, only one *LU* factorization of the $\boldsymbol{Z}$ matrix is performed. In addition, one back substitution of the *LU*-factored $\boldsymbol{Z}^T$ matrix is needed in order to compute the adjoint vector $\hat{\boldsymbol{I}}$.

C. Input Impedance of a Rectangular Patch Antenna

The adjoint sensitivity technique is applied to the optimization of a microstrip-fed rectangular patch antenna with an inset, for an input impedance of 50 Ω at the operating frequency $f_0 = 2$ GHz. The geometry of the structure is given in Fig. 7. The antenna is printed on a dielectric substrate of height $h = 1.59$ mm and relative permittivity $\varepsilon_r = 2.32$. The design parameters are the length of the patch L, its width W, and the depth of the inset S. The objective function is formulated as

$$f(\boldsymbol{x}) = (\mathbf{Re}\{Z_{\text{in}}\} - 50)^2 + (\mathbf{Im}\{Z_{\text{in}}\})^2 \tag{29}$$

where $\mathbf{x} = [L\ W\ S]^T$. The analysis is based on the electric field integral equation. The discretization uses triangular basis functions [17]. The progress of the design during the optimization is shown in Fig. 8. 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).

D. Maximum Gain of the Yagi–Uda Array

The gain of the Yagi–Uda antenna of Fig. 1 is optimized by maximizing the radiation intensity in the direction of maximum radiation $(\theta = 90°, \varphi = 90°)$

$$f(\boldsymbol{x}) = -|A_z(\theta = 90°, \phi = 90°)|^2 \tag{30}$$

where A_z is the only nonzero component of the magnetic vector potential generated by the antenna. The design space is $\boldsymbol{x} = [s_{3_n}\ s_{4_n}\ s_{5_n}]^T$. The initial design is the one optimized for $Z_{\text{in}} = 73\ \Omega$, with $l_{1_n} = 0.5243$, $s_{1_n} = 0.2607$ and $\boldsymbol{x} = [0.34\ 0.34\ 0.34]^T$. The optimal design is $\boldsymbol{x} = [0.3735\ 0.4471\ 0.4353]^T$. 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). The progress of the objective function is given in Fig. 9.

E. Yagi–Uda Array Design for Optimum Input Impedance and Gain

We now consider a more practical design problem where both the input impedance and the gain of the Yagi–Uda array are to be optimized. The design parameters are all five separation distances, i.e., $\boldsymbol{x} = [s_{1_n}\ s_{2_n}\ s_{3_n}\ s_{4_n}\ s_{5_n}]^T$. As an initial design, we use the optimized array from the previous example where $l_{1_n} = 0.532\,43$ and $\boldsymbol{x} = [0.2607\ 0.3400\ 0.3735\ 0.4471\ 0.4353]^T$. The objective function is now formulated as

$$f(\boldsymbol{x}) = 0.5\left[(\mathbf{Re}\{Z_{\text{in}}\} - 73)^2 + (\mathbf{Im}\{Z_{\text{in}}\})^2\right] - 0.5G^2. \tag{31}$$

The values of the design parameters, as well as the values of the input impedance and the gain at each design iteration, are listed in Table I. The progress of the objective function is given in Fig. 10.

V. CONCLUSIONS

A feasible AVM 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, FL: Academic, 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, Aug. 1969.

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

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

[5] J. W. Bandler, "Computer-aided circuit optimization," in *Modern Filter Theory and Design*, G. C. Temes and S. K. Mitra, Eds. New York: Wiley, 1973, ch. 6.

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

[7] J. Vlach and K. Singhal, *Computer Methods for Circuit Analysis and Design*. New York: Van Nostrand, 1983.

[8] J. W. Bandler, *Optimization*. Hamilton, ON, Canada: McMaster Univ., 1994, vol. 1, Custom Courseware for Computer Engineering 3 KB3.

[9] 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.

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

[11] J. P. Webb, "Design sensitivity of frequency response in 3-D finite-element analysis of microwave devices," *IEEE Trans. Mag.*, vol. 38, pp. 1109–1112, Mar. 2002.

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

[13] G. Iuculano, V. A. Monaco, and P. Tiberio, "Network sensitivities in terms of scattering parameters," *Electron. Lett.*, vol. 7, pp. 53–55, Jan. 1971.

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

[15] 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. Dig.*, vol. 3, 2001, pp. 1945–1948.

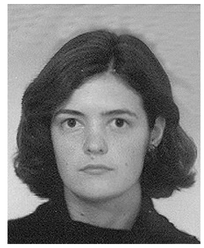
[16] J. W. Bandler, Q. J. Zhang, J. Song, and R. M. Biernacki, "FAST gradient based yield optimization of nonlinear circuits," *IEEE Trans. Microwave Theory Tech.*, vol. 38, pp. 1701–1710, Nov. 1990.

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

Natalia K. Georgieva (S'96–M'97) received the Dipl. Eng. degree from the Technical University of Varna, Varna, Bulgaria, in 1989, and the Ph.D. degree from the University of Electro-Communications, Tokyo, Japan, in 1997.

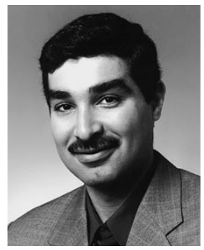
From 1998 to 1999, she held a Post-Doctoral Fellowship with the Natural Sciences and Engineering Research Council of Canada (NSERC), during which she was with the Microwave and Electromagnetics Laboratory, DalTech, Dalhousie University, Halifax, NS, Canada, and then the Simulation Optimization Systems Research Laboratory, McMaster University, Hamilton, ON, Canada. In July 1999, she joined the Department of Electrical and Computer Engineering, McMaster University, where she is an Assistant Professor. Her research interests include theoretical and computational electromagnetism, high-frequency analysis techniques, as well as computer-aided design methods for high-frequency structures and antennas.

Dr. Georgieva received the NSERC University Faculty Award in 2000.



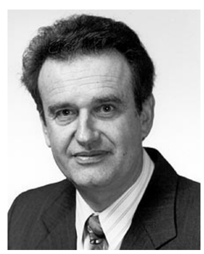
Snezana Glavic received the B.Sc. degree from the University of Belgrade, Belgrade, Yugoslavia, in 1996 and the M.A.Sc. degree from McMaster University, Hamilton, ON, Canada, in 2002, both in electrical engineering.

During 1996–1999, she was a Telecommunication Engineer for the national telecom operator of Yugoslavia, Telekom Srbija. From 2000 to 2002, she was a Teaching and Research Assistant at McMaster University. Her research interests include computational electromagnetics and CAD methods for high-frequency structures and antennas.



Mohamed H. Bakr (S'98–M'01) received the B.Sc. degree (with honors) in electronics and communications engineering in 1992 and the Master's degree in engineering mathematics in 1996, both from Cairo University, Cairo, Egypt, and the Ph.D. degree in 2000 from the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada.

During 1997, he was an intern with Optimization Systems Associates Inc., Dundas, ON, Canada. From 1998 to 2000, he was a Research Assistant with the Simulation Optimization Systems Research Laboratory, McMaster University. In November 2000, he joined the Computational Electromagnetics Research Laboratory (CERL), University of Victoria, Victoria, BC, Canada, as an Natural Sciences and Engineering Research Council of Canada Post-Doctoral Fellow. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, McMaster University. His areas of research include optimization methods, CAD and modeling of microwave circuits, neural network applications, smart analysis of microwave circuits and efficient optimization using time-/frequency-domain methods.



John W. Bandler (S'66–M'66–SM'74–F'78) was born in Jerusalem on November 9, 1941. He studied at Imperial College of Science and Technology, London, U.K., from 1960 to 1966. He received the B.Sc. (Eng.), Ph.D., and D.Sc. (Eng.) degrees from the University of London, London, U.K., in 1963, 1967, and 1976, respectively.

In 1966, he joined Mullard Research Laboratories, Redhill, Surrey, U.K. From 1967 to 1969, he was a Post-Doctorate Fellow and Sessional Lecturer at the University of Manitoba, Winnipeg, MB, Canada. In 1969, he joined McMaster University, Hamilton, ON, Canada, where he has served as Chairman of the Department of Electrical Engineering and Dean of the Faculty of Engineering. He is currently Professor Emeritus in Electrical and Computer Engineering, and directs research in the Simulation Optimization Systems Research Laboratory. He was President of Optimization Systems Associates Inc. (OSA), which he founded in 1983, until November 20, 1997, the date of acquisition of OSA by the Hewlett-Packard Company (HP). OSA implemented a first-generation yield-driven microwave CAD capability for Raytheon in 1985, followed by further innovations in linear and nonlinear microwave CAD technology for the Raytheon/Texas Instruments Joint Venture MIMIC Program. OSA introduced the CAE systems RoMPE in 1988, HarPE in 1989, OSA90 and OSA90/hope in 1991, Empipe in 1992, and Empipe3D and EmpipeExpress in 1996. OSA created ***empath*** in 1996, marketed by Sonnet Software Inc. He is currently the President of the Bandler Corporation, Dundas, ON, Canada, which he founded in 1997. He has authored or co-authored over 340 papers. He joined the Editorial Boards of the *International Journal of Numerical Modeling* in 1987, the *International Journal of Microwave and Millimeterwave Computer-Aided Engineering* in 1989, and *Optimization and Engineering* in 1998.

Dr. Bandler was an Associate Editor of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES (1969–1974), and has continued serving as a member of the IEEE Microwave Theory and Techniques Society (IEEE MTT-S) Editorial Board. He is currently the co-chair of the IEEE MTT-1 Technical Committee on Computer-Aided Design. He is a Fellow of the Royal Society of Canada, the Institution of Electrical Engineers (IEE), U.K., and the Engineering Institute of Canada, a member of the Association of Professional Engineers of the Province of Ontario, Canada, and a Member of the Massachusetts Institute of Technology (MIT) Electromagnetics Academy. He received the Automatic Radio Frequency Techniques Group (ARFTG) Automated Measurements Career Award in 1994.