Adjoint Techniques for Sensitivity Analysis in High-Frequency Structure CAD

Natalia K. Nikolova, *Member, IEEE*, John W. Bandler, *Fellow, IEEE*, and Mohamed H. Bakr, *Member, IEEE*

Abstract—There is a revival of the interest in adjoint sensitivity analysis techniques. This is partly because current computer-aided-design software based on full-wave electromagnetic (EM) solvers remains too slow for the purposes of practical high-frequency structure design despite the increasing capacity of computers. The adjoint-variable methods for design sensitivity analysis offer computational speed and accuracy. They can be used for efficient gradient-based optimization, in tolerance and yield analysis. Adjoint-based sensitivity analysis for circuits has been well studied and extensively covered in the microwave literature. In comparison, sensitivities with full-wave analysis techniques have attracted little attention, and there have been few applications into feasible and versatile algorithms. We review adjoint-variable methods used in high-frequency structure design with both circuit analysis techniques and full-wave EM analysis techniques. A brief discussion on adjoint-based sensitivity analysis for nonlinear dynamic systems is also included.

Index Terms—Adjoint networks, computer-aided design (CAD), electromagnetic (EM) modeling, microwave circuits, optimization, sensitivity analysis.

I. INTRODUCTION

COMPUTER-AIDED engineering for high-frequency structures (microwave and millimeter-wave circuits and antennas) originated in the early 1950s with the advent of first-generation computers. Historically, RF and microwave modeling started with the representation of the complex electromagnetic (EM) environment in terms of circuit equivalents—lumped elements and transmission lines. This is still a major branch in the family of computational approaches used by microwave and antenna engineers because of the physical insight, manageable storage requirements, and high computational speed. It is, therefore, not surprising that the first significant advances in the direction of automated design of high-frequency structures in the late 1960s and early 1970s were made with microwave circuit modeling in mind.

As computing resources became more powerful and widely available, computational electromagnetics emerged and spurred

N. K. Nikolova and M. H. Bakr are with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada L8S 4K1 (e-mail address: talia@mcmaster.ca).

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

a variety of numerical algorithms for full-wave EM analysis. They offer superior accuracy and complete field representation—as long as the theoretical model includes all EM field interactions. On the other hand, unlike the equivalent-circuit models, these algorithms are extremely demanding in terms of computer memory and time. Even today, full-wave analysis appears prohibitively slow for the purposes of modeling and design of a complete microwave circuit. EM field analysts—preoccupied with the complexity of full-wave simulation—are the last to appreciate the requirements of automated EM design. The problem of efficient sensitivity estimation and optimization with full-wave EM analysis remains a challenge, especially from the viewpoint of implementation into robust and feasible algorithms.

Here, we focus on adjoint-based approaches to design sensitivity analysis whose purpose is to evaluate the gradient of the system response in the design parameter space. We consider both network- and field-based analyses. The system response may be the solution at the current design given by the state variables, e.g., currents or voltages of a circuit and field or current density distribution in a high-frequency structure. This is the system distributed response. The response may also be defined in terms of multiport network parameters, e.g., S-parameters. For optimization purposes, the response of interest is a real-valued objective function (or cost function). Design sensitivity information is crucial in engineering problems such as optimization, statistical, yield, and tolerance analysis.

The adjoint-variable approach yields the response and its sensitivity through two analyses: of the original structure or circuit and of an auxiliary (or adjoint) structure or circuit. In general, two system analyses are sufficient regardless of the number of design parameters. As we describe below, the adjoint system analysis is often obtained with very little or no overhead once the solution to the original problem is carried out.

Adjoint-based sensitivity analysis for networks is attributed to Director and Rohrer [1], [2] who derived a sensitivity expression based on Tellegen's theorem [3], [4]. An adjoint network is constructed through relatively straightforward rules and solved to produce the adjoint voltages and currents, which are subsequently substituted in the sensitivity expression. Generalized network sensitivities based on a somewhat simpler derivation—a direct manipulation of the system matrix equations—were suggested by Branin soon after [5]. Theoretically, this formalism is equivalent to the adjoint-network analysis of Director and Rohrer. It is shown that a simple mathematical relation exists between the original circuit and the adjoint circuit through the transpose of the system matrix [5], [6]. Consequently, the sensitivity analysis with multiple design

parameters can involve as little as a matrix transposition and a few matrix multiplications in addition to the original network analysis.

The microwave literature in the 1970s and 1980s abounds with techniques and applications of adjoint-network sensitivity analysis using S-parameters, voltage–current variables, and branched cascaded topologies [7]–[12]. Some comprehensive reviews can be found in [12]–[15]. Exact sensitivities were developed for harmonic-balance analyses [15] and feasible implementations were proposed for nonlinear circuits [16]. Later, Alessandri *et al.* applied the adjoint network method to the sensitivity analysis of waveguide structures composed of subnetworks whose admittance matrices are analytically available [17]. Mongiardo and Ravanelli applied the adjoint network sensitivities to the design of corrugated feeds [18]. Exact sensitivities for circuit-based analyses are applied to problems involving coupled resonator filters [19] and high-speed very large scale integration (VLSI) interconnects [20]–[25]. Exact adjoint sensitivities have been recently considered in neural-based modeling and design [26].

Until the 1990s, computational EM analysis has been almost exclusively focused on the problem of finding the response of a given structure and has hardly addressed sensitivity estimations for changes of the shape or the materials involved. When integrated into an automated design environment, full-wave simulators are closed "black boxes," which communicate with the optimization algorithms through a set of responses. The lack of sensitivity information is typically compensated for by gradient approximations [27] and multidimensional response surface approximations [28]. Thus, the integration of conventional full-wave simulators into a design environment is numerically inefficient for the purposes of gradient-based optimization. Consider the simplest implementation—forward or backward finite-difference approximation of the response derivatives in addition to the evaluation of the response itself. The current design is slightly perturbed for each of the design parameters to produce N additional structures where N is the number of design parameters. This procedure requires $N + 1$ full-wave analyses per design iteration. Since optimization typically requires hundreds of iterations, the overall numerical effort for EM-based design may easily become unacceptable.

The need for adjoint-variable sensitivity analysis with full-wave EM solvers seems obvious, and yet, to our knowledge, commercial computer-aided design (CAD) software based on full-wave analysis has not adopted adjoint-based techniques for response derivative estimations. A possible reason is that the implementation of adjoint sensitivities with complex numerical EM algorithms is far from trivial. In the microwave literature, we have seen only a few developments, which, until recently, almost exclusively addressed the finite-element method (FEM). Webb *et al.* considered exact sensitivities with the FEM and developed element representations, which allow analytical derivatives of the FEM system matrix with respect to the Cartesian coordinates of the vertices of the finite-element mesh [29]–[34]. Another more specific application with the FEM was reported in [35]. More recent developments involve problems analyzed with the finite-element time-domain (FETD) method, as well as with the finite-difference

time-domain (FDTD) method on unstructured grids [36]–[38]. Currently, efforts are directed toward the development of feasible adjoint-based sensitivity analysis approaches, which are easy to implement with a given class of EM solvers. Feasible implementations are proposed with the method of moments (MoM) [39], where the derivatives of the system matrix are approximated using finite differences. A reduction of the computational effort associated with the computation of the derivatives of the MoM matrix through the perturbation of the mesh boundary layers was suggested in [40]. Adjoint techniques are under development for full-wave algorithms on fixed structured grids such as the frequency-domain transmission-line method (FD–TLM) [41], the time-domain TLM [42], and the FDTD method [43].

We start with a brief outline of the basics of the adjoint network method in microwave circuit sensitivity analysis. Subsequently, we proceed with the adjoint variable method in full-wave EM analysis and its applications with different classes of EM solvers. We discuss difficulties arising in implementations into practical algorithms and review approaches dealing with these difficulties. The scope of this discussion includes both frequency- and time-domain full-wave techniques.

It would be impossible to present here all work related to adjoint sensitivity analysis in its entirety or in detail. We would rather address only a few basic concepts, which serve as a starting point toward future developments.

II. ADJOINT NETWORK METHODS: TELLEGEN APPROACH

The adjoint network method and its applications in microwave circuit design have been discussed extensively in the literature. Comprehensive reviews can be found, for example, in [9], [12], and [44]. Reference [12] also provides a useful comparison between the adjoint network method based on Tellegen's theorem on the one hand and adjoint-based methods such as the direct method and the transpose-matrix methods on the other.

We do not repeat these discussions here. We highlight some basic concepts, which may be useful if applied with full-wave analyses, especially with techniques representing the EM environment in terms of distributed circuits, e.g., TLM methods. Thus, we hope to address a broader audience including not only experts in microwave CAD, but also experts in computational EM. The latter are, ultimately, the people capable of integrating adjoint sensitivity techniques with full-wave EM simulators.

A. Tellegen's Theorem

Tellegen's theorem—well known to electrical engineers—is often stated in matrix form as[1] $\boldsymbol{v}^T \boldsymbol{i} = 0$, where $\boldsymbol{v}$ and $\boldsymbol{i}$ are vectors containing the time-dependent branch voltages and currents of a circuit. This is known as the "actual-power" Tellegen theorem. It follows from Kirchhoff's circuit laws. Since the phasors of the circuit voltages and currents $\boldsymbol{V}$ and $\boldsymbol{I}$ obey Kirchhoff's laws, Tellegen's theorem applies to them as well.

---

[1]Throughout this paper, matrices and vectors are in bold italic, while vectors in three-dimensional (3-D) space are in bold.
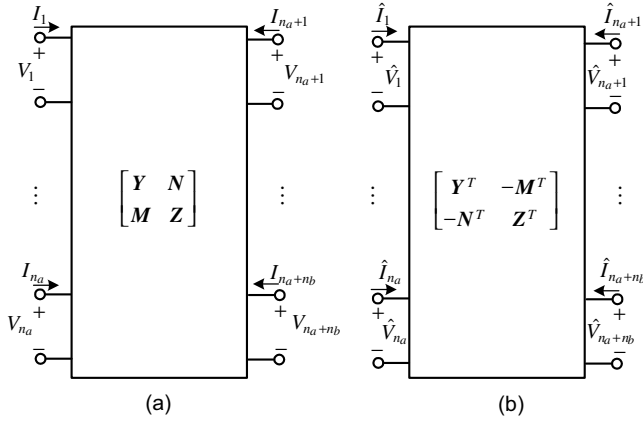
Fig. 1. (a) Circuit element represented by its hybrid matrix. (b) Respective element in the adjoint circuit.

More generally, Tellegen's theorem proves for any two topologically identical circuits [45]–[47] whose voltages and currents are given by the vectors $\boldsymbol{V}, \boldsymbol{I}$ and $\hat{\boldsymbol{V}}, \hat{\boldsymbol{I}}$, respectively, that

$$\boldsymbol{V}^T \hat{\boldsymbol{I}} = 0 \quad \boldsymbol{I}^T \hat{\boldsymbol{V}} = 0 \qquad (1)$$

(the "quasi-power" Tellegen theorem). Therefore, Tellegen's theorem can be written in the difference form

$$\boldsymbol{V}^T \hat{\boldsymbol{I}} - \boldsymbol{I}^T \hat{\boldsymbol{V}} = 0. \qquad (2)$$

Perturbing elements in the original circuit, and applying (2) to the perturbed original circuit and an auxiliary circuit represented by $\hat{\boldsymbol{V}}, \hat{\boldsymbol{I}}$, leads to

$$\Delta \boldsymbol{V}^T \hat{\boldsymbol{I}} - \Delta \boldsymbol{I}^T \hat{\boldsymbol{V}} = 0 \qquad (3)$$

where $\Delta \boldsymbol{V}$ and $\Delta \boldsymbol{I}$ represent the incremental changes in the voltages and currents of the original circuit. Equation (3) is commonly referred to as the perturbed difference form of Tellegen's theorem.

According to Penfield *et al.* [47], all forms of Tellegen's theorem are valid for circuits obeying Kirchhoff's laws: linear or nonlinear, reciprocal or nonreciprocal, passive or active, hysteretic or nonhysteretic.

B. Circuit Sensitivities

Consider a multiport component of a circuit described by its hybrid matrix [see Fig. 1(a)]

$$\begin{bmatrix} \boldsymbol{I}_a \\ \boldsymbol{V}_b \end{bmatrix} = \begin{bmatrix} \boldsymbol{Y} & \boldsymbol{N} \\ \boldsymbol{M} & \boldsymbol{Z} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_a \\ \boldsymbol{I}_b \end{bmatrix} \qquad (4)$$

where

$$\boldsymbol{I}_a = [I_1 \dots I_{n_a}]^T \quad \boldsymbol{V}_b = [V_{n_a+1} \dots V_{n_a+n_b}]^T \qquad (5)$$

and

$$\boldsymbol{V}_a = [V_1 \dots V_{n_a}]^T \quad \boldsymbol{I}_b = [I_{n_a+1} \dots I_{n_a+n_b}]^T. \qquad (6)$$

The hybrid matrix offers a generic representation, which includes the impedance matrix $\bar{\boldsymbol{Z}}$ and admittance matrix $\boldsymbol{Y}$ as special cases.

A perturbation of some of the elements of this component leads to changes in all its branch currents and voltages. Neglecting higher order terms, the changes of the currents and voltages at its ports are expressed as

$$\begin{bmatrix} \Delta \boldsymbol{I}_a \\ \Delta \boldsymbol{V}_b \end{bmatrix} = \begin{bmatrix} \boldsymbol{Y} & \boldsymbol{N} \\ \boldsymbol{M} & \boldsymbol{Z} \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{V}_a \\ \Delta \boldsymbol{I}_b \end{bmatrix} + \begin{bmatrix} \Delta \boldsymbol{Y} & \Delta \boldsymbol{N} \\ \Delta \boldsymbol{M} & \Delta \boldsymbol{Z} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_a \\ \boldsymbol{I}_b \end{bmatrix}. \qquad (7)$$

We write the product terms in (3) involving only the voltages and currents of our circuit component and its auxiliary counterpart as

$$\Delta G = \begin{bmatrix} \Delta \boldsymbol{I}_a \\ \Delta \boldsymbol{V}_b \end{bmatrix}^T \begin{bmatrix} -\hat{\boldsymbol{V}}_a \\ \hat{\boldsymbol{I}}_b \end{bmatrix} + \begin{bmatrix} \Delta \boldsymbol{V}_a \\ \Delta \boldsymbol{I}_b \end{bmatrix}^T \begin{bmatrix} \hat{\boldsymbol{I}}_a \\ -\hat{\boldsymbol{V}}_b \end{bmatrix}. \qquad (8)$$

Notice that (3) applies to the whole circuit, while (8) represents just a portion of its left-hand side. Thus, in general, $\Delta G \neq 0$. We refer to ΔG as the *component sensitivity*. The substitution of (7) in the right-hand side of (8) leads to [9]

$$\Delta G = \begin{bmatrix} \boldsymbol{V}_a \\ \boldsymbol{I}_b \end{bmatrix}^T \begin{bmatrix} -\Delta \boldsymbol{Y}^T & \Delta \boldsymbol{M}^T \\ -\Delta \boldsymbol{N}^T & \Delta \boldsymbol{Z}^T \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{V}}_a \\ \hat{\boldsymbol{I}}_b \end{bmatrix}$$
$$+ \begin{bmatrix} \Delta \boldsymbol{V}_a \\ \Delta \boldsymbol{I}_b \end{bmatrix}^T \left(\begin{bmatrix} \boldsymbol{Y} & \boldsymbol{N} \\ \boldsymbol{M} & \boldsymbol{Z} \end{bmatrix}^T \begin{bmatrix} -\hat{\boldsymbol{V}}_a \\ \hat{\boldsymbol{I}}_b \end{bmatrix} - \begin{bmatrix} -\hat{\boldsymbol{I}}_a \\ \hat{\boldsymbol{V}}_b \end{bmatrix} \right). \qquad (9)$$

The expression in (9) is significantly simplified if we eliminate the second term, which involves the variations of the voltages and currents $[\Delta \boldsymbol{V}_a^T \ \Delta \boldsymbol{I}_b^T]^T$ at the ports of the circuit element. After imposing the relation

$$\begin{bmatrix} \hat{\boldsymbol{I}}_a \\ \hat{\boldsymbol{V}}_b \end{bmatrix} = \begin{bmatrix} \boldsymbol{Y}^T & -\boldsymbol{M}^T \\ -\boldsymbol{N}^T & \boldsymbol{Z}^T \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{V}}_a \\ \hat{\boldsymbol{I}}_b \end{bmatrix} \qquad (10)$$

between the currents and the voltages of the respective element in the auxiliary circuit, we obtain

$$\Delta G = \begin{bmatrix} \boldsymbol{V}_a \\ \boldsymbol{I}_b \end{bmatrix}^T \begin{bmatrix} -\Delta \boldsymbol{Y}^T & \Delta \boldsymbol{M}^T \\ -\Delta \boldsymbol{N}^T & \Delta \boldsymbol{Z}^T \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{V}}_a \\ \hat{\boldsymbol{I}}_b \end{bmatrix}. \qquad (11)$$

The voltage–current relation in (10) defines the state equations for the component in the auxiliary circuit [see Fig. 1(b)]. The auxiliary circuit is not only topologically identical to the original circuit, but it consists of elements whose hybrid matrices relate to the hybrid matrices of the respective elements of the original circuit in a straightforward manner. This auxiliary circuit is referred to as the *adjoint circuit*.

The component sensitivity in (11) may, in general, be a result of perturbations in all design parameters $p_n, \ n = 1, \dots, N$

$$\Delta G = \sum_{n=1}^{N} \frac{\partial G}{\partial p_n} \Delta p_n \qquad (12)$$

where

$$\frac{\partial G}{\partial p_n} = \begin{bmatrix} \boldsymbol{V}_a \\ \boldsymbol{I}_b \end{bmatrix}^T \cdot \left(\frac{\partial}{\partial p_n} \begin{bmatrix} -\boldsymbol{Y}^T & \boldsymbol{M}^T \\ -\boldsymbol{N}^T & \boldsymbol{Z}^T \end{bmatrix} \right) \cdot \begin{bmatrix} \hat{\boldsymbol{V}}_a \\ \hat{\boldsymbol{I}}_b \end{bmatrix}. \qquad (13)$$

In anticipation of our further discussion on alternative matrix formulations of adjoint-based sensitivity analysis, we rewrite (11) as

$$\Delta G = \hat{\boldsymbol{X}}^T \Delta \boldsymbol{H} \boldsymbol{X} \qquad (14)$$

where

$$\hat{X} = \begin{bmatrix} -\hat{V}_a \\ \hat{I}_b \end{bmatrix} \quad \Delta H = \begin{bmatrix} \Delta Y & \Delta N \\ \Delta M & \Delta Z \end{bmatrix} \quad X = \begin{bmatrix} V_a \\ I_b \end{bmatrix}. \tag{15}$$

The adjoint component equations (10) are modified accordingly:

$$\begin{bmatrix} -\hat{I}_a \\ \hat{V}_b \end{bmatrix} = \begin{bmatrix} Y & N \\ M & Z \end{bmatrix}^T \begin{bmatrix} -\hat{V}_a \\ \hat{I}_b \end{bmatrix}. \tag{16}$$

The latter equation represents a circuit component, which, despite the sign differences in the hybrid matrix, is equivalent to the one in Fig. 1(b). In fact, it shows that the sign reversal in the adjoint hybrid matrix in (10) is superfluous [5]—notice that the signs of both the voltages and currents at the first n_a ports of the adjoint circuit element described by (16) are reversed.

Let us now assume that our circuit consists of M elements each defined by its hybrid matrix. Each of these elements has a sensitivity term ΔG_m, $m = 1, \ldots, M$. The circuit also contains external sources, which are independent of any element changes. These are given by a vector of current sources I_I and a vector of voltage sources V_V. We now apply (3) to all branches of the complete network. Keeping in mind that the sources are independent of parameter perturbations, i.e., $\Delta I_I = 0$ and $\Delta V_V = 0$, we obtain

$$\Delta I_V^T \hat{V}_V - \Delta V_I^T \hat{I}_I = \sum_{m=1}^{M} \Delta G_m = \sum_{m=1}^{M} \hat{X}_m^T \Delta H_m X_m. \tag{17}$$

Here, $\hat{I}_I$ and $\hat{V}_V$ are the current and voltage sources in the adjoint circuit, respectively, ΔV_I is the change in the voltages at the current sources in the original circuit due to the perturbations, and ΔI_V is the change of the currents through the voltage sources in the original circuit. The vector $\hat{X}_m$ is defined according to (15) for each adjoint network element. It is a subset of the solution $\hat{X}$ containing the voltages and currents of the whole adjoint network. Similarly, X_m is a subset of X, which contains the voltages and currents of the whole original network. Equation (17) suggests a way to compute the circuit sensitivity to parameter perturbations through the sensitivity of its response at the source locations in terms of ΔV_I and ΔI_V.

We note that often a perturbation of a designable circuit parameter Δp_n belonging to the mth network component affects only the mth component matrix. This leaves only one nonzero sensitivity term with respect to Δp_n to be calculated on the right-hand side of (17). Component sensitivities for all fundamental circuit elements can be found, for example, in [2] and [44].

C. Original and Adjoint System Analyses

The adjoint network is assembled from the respective adjoint circuit components in a similar fashion as the original network. In fact, a reciprocal original circuit has an adjoint circuit, which is identical to it. Take as an example a single-port passive circuit component [see Fig. 2(a)]. The Z-matrix representation is a special case of the hybrid matrix formulation used above and,
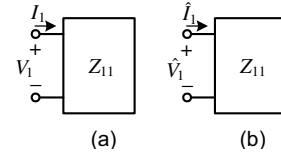

Fig. 2. (a) Passive one-port circuit element. (b) Its identical counterpart in the adjoint circuit.

in the case of a one-port circuit element, it has one element $Z_{11} = R_{11} + jX_{11}$ so that $V_1 = Z_{11}I_1$. The respective adjoint circuit component—represented by the transpose Z-matrix according to (10)—is identical to the original one, i.e., $\hat{V}_1 = Z_{11}\hat{I}_1$ [see Fig. 2(b)]. Thus, a resistor in the original circuit corresponds to the same resistor in the adjoint circuit. The same holds for inductors and capacitors.

In fact, for any reciprocal multiport circuit component, $Z_m^T = Z_m$ or $Y_m^T = Y_m$ holds, meaning that its counterpart in the adjoint circuit is identical. For example, a piece of uniform transmission line of characteristic impedance Z_0 and electrical length θ is a reciprocal two-port element because its Z-matrix is symmetric as follows:

$$Z^T = Z = Z_0 \begin{bmatrix} \coth\theta & \operatorname{csch}\theta \\ \operatorname{csch}\theta & \coth\theta \end{bmatrix}. \tag{18}$$

Thus, its adjoint counterpart is an identical transmission line. Tables with correspondence between common circuit elements and their adjoint counterparts can be found, for example, in [2] and [44].

Note that the right-hand side of (17) requires the solution of the original circuit and that of the adjoint circuit X and $\hat{X}$. These analyses can be carried out with any preferable technique. In fact, the adjoint circuit analysis requires very few computations. It consists of elements whose matrices are the transpose of those of the original circuit elements. As a result, it can be shown that the system matrix of the overall adjoint network—obtained, for example, through a nodal formulation—is the transpose of the original system matrix A resulting from a nodal analysis. The matrix A is factorized $(A = LU)$ when the original network is analyzed. Its LU factors are then rearranged to produce the LU factors of A^T as follows:

$$A^T = U^T L^T \tag{19}$$

which are used to compute the adjoint solution through forward and backward substitutions. Details on efficient numerical algorithms for obtaining, storing, and rearranging the LU factors used in sensitivity analysis can be found in [12] and [48]. The computational overhead associated with the sensitivity estimation is negligible in comparison with the computational requirements of the original circuit analysis. Moreover, it does not depend on the number of design parameters.

D. Response Sensitivity and Adjoint Excitation

There is one issue related to the adjoint network, which still remains to be resolved, namely, the adjoint sources $\hat{I}_I$ and $\hat{V}_V$. We are free to choose them so that the left-hand side of (17) yields the derivative of the response function of interest with

respect to a given design parameter. An excitation vector and a response vector can be introduced as

$$\hat{E} = \begin{bmatrix} \hat{I}_I \\ \hat{V}_V \end{bmatrix} \quad R = \begin{bmatrix} V_I \\ -I_V \end{bmatrix} \tag{20}$$

respectively, so that (17) becomes

$$\Delta R^T \hat{E} = -\sum_{m=1}^{M} \Delta G_m. \tag{21}$$

Assume that a scalar response function is given as follows:

$$f(R(p), p) \tag{22}$$

where $p = [p_1 \cdots p_N]^T$ is the vector of design parameters. The response function depends implicitly on p through the circuit response R. It may also have explicit dependence on p. We assume that f is analytically differentiable in all elements of R and p.

Consider first the simplest case—the response function f is identical with the circuit reaction (voltage across current source or current through a voltage source) at the kth port[2] $f = R_k$. To find its sensitivity, we set all elements of the adjoint excitation vector $\hat{E}$ equal to zero, except for $\hat{E}_k = 1$. For example, if the kth port corresponds to a current source in the original circuit, then its sensitivity is $\Delta R_k = \Delta V_{Ik}$. Consequently, a unit current source must be connected to the kth port of the adjoint circuit, i.e., $\hat{I}_{Ik} = 1$. All remaining adjoint circuit ports, which correspond to current sources in the original network, are open circuited, i.e., $\hat{I}_{Il} = 0$ and $l \neq k$. All adjoint circuit ports, which correspond to voltage sources, are short circuited, i.e., $\hat{V}_V = 0$.

Once the adjoint sources are set, the adjoint network is analyzed. According to (21), the desired response sensitivity is calculated from

$$\Delta R_k = \Delta f = -\sum_{m=1}^{M} \Delta G_m. \tag{23}$$

The sensitivity with respect to any design parameter p_n is

$$\frac{\partial f}{\partial p_n} = -\sum_{m=1}^{M} \hat{X}_m^T \frac{\partial H_m}{\partial p_n} X_m, \qquad n = 1, \dots, N. \tag{24}$$

Notice that, in (24), the only terms that need to be recalculated for different n are the matrix derivatives $\partial H_m / \partial p_n, n = 1, \dots, N$. The solutions of the original circuit X and the adjoint circuit $\hat{X}$ do not depend on any perturbation, i.e., they are independent of n and are computed once.

In a slightly different scenario, we may be interested in a response, which is not simply equal to the voltage–current response of the circuit at its kth port, but is some function of it. In microwave circuit design, we may be interested in minimizing the reflection at this port [9], assuming that the loads of all other ports are known. The reflection coefficient

$$\Gamma_k = \frac{Z_{\text{in}_k} - R_{g_k}}{Z_{\text{in}_k} + R_{g_k}} = 1 - \frac{2R_{g_k}}{Z_{\text{in}_k} + R_{g_k}} = 1 + \frac{2R_{g_k} I_{Vk}}{V_{Vk}} \tag{25}$$

is, in general, a complex number. In (25), $Z_{\text{in}_k} = V_k / I_k$ is the input impedance of the circuit at the kth port, R_{g_k} is the source impedance, V_{Vk} is the voltage of the source, and I_{Vk} is the current through the voltage source. A circuit diagram showing the

[2]Hereafter, we refer to the location of the external sources in the network as ports.
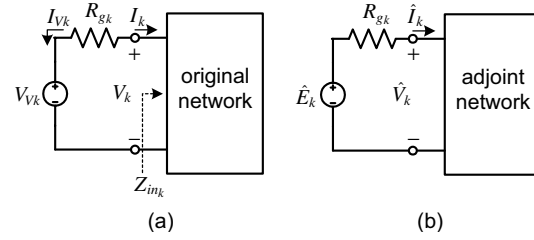

Fig. 3. Original and adjoint networks in the sensitivity analysis of the kth port reflection.

arrangement at the kth port is shown in Fig. 3(a). If we are interested in the sensitivity at a single frequency ω_0, the objective function f to be minimized may be defined as

$$f(p, \omega = \omega_0) = |\Gamma_k(p, \omega_0)|^2. \tag{26}$$

The derivative of (26) is

$$\frac{\partial f}{\partial p_n} = \Re \left\{ 2\Gamma_k^* \frac{\partial \Gamma_k}{\partial p_n} \right\} = \Re \left\{ \frac{4R_{g_k}}{V_{Vk}} \Gamma_k^* \frac{\partial I_{Vk}}{\partial p_n} \right\}. \tag{27}$$

where $\Re$ denotes the real part of a complex-valued quantity.

At the same time, using (21) and setting a nonzero excitation only at the kth port of the adjoint circuit [see Fig. 3(b)], we can express $\partial I_{Vk} / \partial p_n$ as

$$\frac{\partial I_{Vk}}{\partial p_n} = \frac{1}{\hat{E}_k} \sum_{m=1}^{M} \hat{X}_m^T \frac{\partial H_m}{\partial p_n} X_m. \tag{28}$$

The substitution of (28) in (27) yields the sensitivity with respect to any of the design parameters as

$$\frac{\partial f}{\partial p_n} = \Re \left\{ \frac{4R_{g_k}}{V_{Vk}\hat{E}_k} \Gamma_k^* \sum_{m=1}^{M} \hat{X}_m^T \frac{\partial H_m}{\partial p_n} X_m \right\}, \qquad n = 1, \dots, N. \tag{29}$$

We are at liberty to set $\hat{E}_k = V_{Vk}$. In the case of a reciprocal network, this makes the adjoint circuit identical to the original circuit. Thus, $\hat{X} = X$, and the analysis of the adjoint circuit is avoided.

The adjoint network method is often thought to be convenient just in the case of a response function depending on the circuit output at a single port. This, however, is not the case. The sensitivity estimation of response functions, which depend on the circuit output at all ports, is possible and it requires just as many computations as in the case of an objective function depending on a single port. Consider the following case, which arises in gradient-based optimization. The response (or objective) function f is a real-valued quantity, which may depend on the circuit reaction at all ports of the circuit. Neglecting higher order terms, the sensitivity of f can be expressed as

$$\Delta f = \sum_{k=1}^{N_I} \left(\frac{\partial f}{\partial \Re V_{Ik}} \Delta \Re V_{Ik} + \frac{\partial f}{\partial \Im V_{Ik}} \Delta \Im V_{Ik} \right)$$
$$+ \sum_{k=N_I+1}^{N_I+N_V} \left(\frac{\partial f}{\partial \Re I_{Vk}} \Delta \Re I_{Vk} + \frac{\partial f}{\partial \Im I_{Vk}} \Delta \Im I_{Vk} \right) \tag{30}$$

where $\Re$ and $\Im$ denote the real and the imaginary parts of a complex-valued quantity, respectively; N_I denotes the number of the external current sources, and N_V is the number of the external voltage sources. A comparison between (30) and the left-hand side of (21), which we can expand as

$$\Delta \boldsymbol{R}^T \hat{\boldsymbol{E}} = \sum_{k=1}^{N_I} (\Delta \Re V_{Ik} + j\Delta \Im V_{Ik}) \cdot \left(\Re \hat{I}_{Ik} + j\Im \hat{I}_{Ik}\right)$$
$$- \sum_{k=N_I+1}^{N_I+N_V} (\Delta \Re I_{Vk} + j\Delta \Im I_{Vk}) \cdot \left(\Re \hat{V}_{Vk} + j\Im \hat{V}_{Vk}\right) \tag{31}$$

reveals that setting the elements of the adjoint excitation $\hat{\boldsymbol{E}}$ as

$$\hat{I}_{I_k} = \frac{\partial f}{\partial \Re V_{I_k}} - j\frac{\partial f}{\partial \Im V_{I_k}}, \qquad k = 1, \dots, N_I$$
$$\hat{V}_{V_k} = \frac{\partial f}{\partial \Re I_{V_k}} - j\frac{\partial f}{\partial \Im I_{V_k}}, \qquad k = N_I+1, \dots, N_I+N_V \tag{32}$$

leads to

$$\Delta f = -\sum_{m=1}^{M} \Re\left(\hat{\boldsymbol{X}}_m^T \Delta \boldsymbol{H}_m \boldsymbol{X}_m\right). \tag{33}$$

In (32), $j = \sqrt{-1}$. Thus, the sensitivity expression becomes

$$\frac{\partial f}{\partial p_n} = -\Re \sum_{m=1}^{M} \hat{\boldsymbol{X}}_m^T \frac{\partial \boldsymbol{H}_m}{\partial p_n} \boldsymbol{X}_m, \qquad n = 1, \dots, N. \tag{34}$$

Now, every port of the adjoint network has an external source, as per (32).

III. ADJOINT NETWORK S-PARAMETER SENSITIVITIES BASED ON WAVE VARIABLES

The sensitivities of the S-parameters of microwave circuits have been discussed extensively in the microwave literature; see, e.g., [7]–[11]. Here, we only show that the same principles apply as in the case of current–voltage responses. For that, it would suffice to state the generalized Tellegen theorem in wave variables [8], [11], [45] for any two topologically identical networks

$$\boldsymbol{b}^T \hat{\boldsymbol{a}} - \boldsymbol{a}^T \hat{\boldsymbol{b}} = 0. \tag{35}$$

Here, $\boldsymbol{a}$ and $\boldsymbol{b}$ are vectors containing the incident and reflected waves, respectively, at all junctions in the first network, and $\hat{\boldsymbol{a}}$ and $\hat{\boldsymbol{b}}$ contain the corresponding wave variables of the second network. The voltage–current Tellegen theorem (2) can be easily transformed into (35) by a substitution of the relation between the voltage V_k and current I_k of a branch and its wave variables a_k and b_k [45] as follows:

$$V_k = \frac{(a_k Z_n^* + b_k Z_n)}{\sqrt{\Re Z_n}}$$
$$I_k = \frac{(a_k - b_k)}{\sqrt{\Re Z_n}} \tag{36}$$

where Z_n is a normalization impedance.

Formally, (35) is identical to the voltage–current formulation of Tellegen's theorem (2). Thus, all voltage–current expressions can be adapted to operate on the wave variables and the S-matrices of the original and the adjoint circuits.

Assume that a microwave network is composed of M subnetworks, each described by

$$\boldsymbol{b}_m = \boldsymbol{S}_m \boldsymbol{a}_m, \qquad m = 1, \dots, M. \tag{37}$$

The corresponding adjoint subnetwork S-parameters are given by $\boldsymbol{S}_m^T$ so that

$$\hat{\boldsymbol{b}}_m = \boldsymbol{S}_m^T \hat{\boldsymbol{a}}_m, \qquad m = 1, \dots, M. \tag{38}$$

The external ports of the original and adjoint networks are excited by sources, which are considered independent of any parameter changes in the networks. The wave variables at the external ports are given by $\boldsymbol{a}_E$ and $\boldsymbol{b}_E$ for the original circuit and by $\hat{\boldsymbol{a}}_E$ and $\hat{\boldsymbol{b}}_E$ for the adjoint circuit. For an adjoint circuit whose components are defined by (38), the sensitivity expression is in the form [11]

$$\left(\frac{\Delta \boldsymbol{b}_E}{\Delta p_n}\right)^T \hat{\boldsymbol{a}}_E - \left(\frac{\Delta \boldsymbol{a}_E}{\Delta p_n}\right)^T \hat{\boldsymbol{b}}_E = \sum_{m=1}^{M} \hat{\boldsymbol{a}}_m^T \frac{\Delta \boldsymbol{S}_m}{\Delta p_n} \boldsymbol{a}_m,$$
$$n = 1, \dots, N. \tag{39}$$

Notice that the incident waves on the external ports $\boldsymbol{a}_E$ may still depend on a varying circuit parameter p_n because they are determined not only by the external source, but also by the impedance matching at that port. If all ports are matched, $\Delta \boldsymbol{a}_E / \Delta p_n = \boldsymbol{0}$. The choice of the adjoint external waves $\hat{\boldsymbol{a}}_E$ is made in a manner similar to the case considered before so that the left-hand side of (39) yields the sensitivity of the response function $\Delta f / \Delta p_n$.

The sensitivity of an S-parameter of the overall circuit is often required. Since

$$\frac{\Delta S_{kj}}{\Delta p_n} = \left.\frac{\Delta b_k}{\Delta p_n}\right|_{\substack{a_j=1 \\ a_l=0, l\neq j, l=1,\dots,L}} \tag{40}$$

we single out the $\Delta b_k / \Delta p_n$ element from the $\Delta \boldsymbol{b}_E^T / \Delta p_n$ row in (39) by setting the incident adjoint wave at the kth port as $\hat{a}_k = 1$, while all other ports of the adjoint circuit have no excitation and are matched, i.e., $\hat{a}_l = 0$ and $l \neq k$. The arrangement of the original and the adjoint circuit for the calculation of the S_{kj}-parameter and its sensitivity are given in Fig. 4. Notice that when the network is reciprocal, $\boldsymbol{S}^T = \boldsymbol{S}$, the adjoint network is the same as the original one, and there is no need to solve an adjoint problem, in the sense that the arrangement in Fig. 4(b) can be set with the original circuit.

IV. MATRIX APPROACHES TO ADJOINT SENSITIVITY ANALYSIS

Thus far, we have used first-order approximations to the derivatives of the state variables of our problem and those of the response function. Such analysis follows from the perturbation difference form of Tellegen's theorem in a straightforward manner and it leads to a definition of the adjoint network and its components. It features clarity and physical insight.
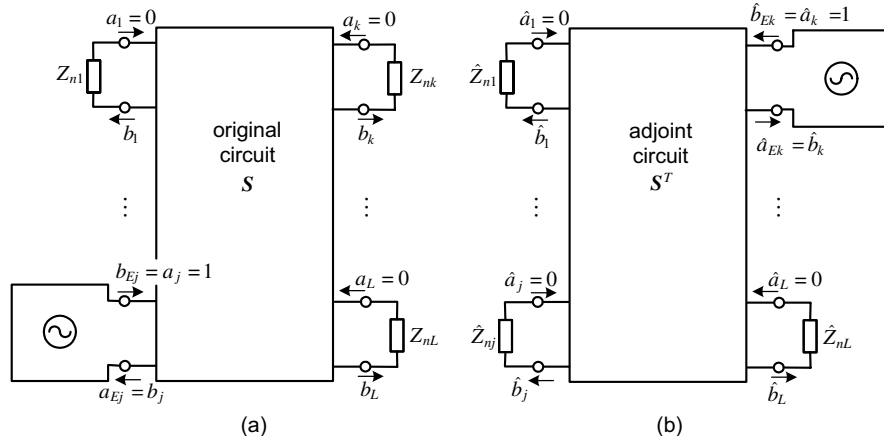
Fig. 4.   Original and adjoint circuits with single-port excitations when the sensitivity of the S_{kj}-parameter is calculated.

However, one could apply direct differentiation to the circuit equations. Such an approach is often referred to as the *direct differentiation method* [12], [49]. This method is general in the sense that it applies to any problem, which is cast in the form of a system of linear or nonlinear equations. Thus, it is applicable to the system of equations arising in computational EM algorithms as well.

With further matrix manipulations, the direct differentiation method leads to an adjoint-based sensitivity analysis approach, which is mathematically equivalent to the adjoint-network method in the case of circuit sensitivities. In control theory, this approach is usually referred to as the *adjoint-variable method* [49], [50], but it is also sometimes called the *transpose-matrix method* [12]. Below, we outline the application of the direct differentiation and the adjoint-variable methods to linear and nonlinear problems.

A. Direct Differentiation Method for Linear Systems

Assume that a linear circuit or EM field problem is cast in the form of a system of equations

$$A(p)x = b(p) \tag{41}$$

where A is the system matrix, $x = [x_1 \ldots x_M]^T$ is the vector of state variables, and $b = [b_1 \ldots b_M]^T$ is the excitation vector. The state variables may be voltages, currents, or normalized waves in circuit analysis. In EM field analysis, x may represent the field solution (e.g., the **E**-field distribution resulting from FEM analysis) or a solution in terms of current densities (as in the case of MoM analysis). In high-frequency structure design, the design parameters p would typically relate to the structure's geometry and the constitutive parameters of the materials. The system matrix A is clearly dependent on p and, thus, the solution x is an implicit function of p. The excitation vector b, in general, depends on the sources and boundary conditions of the problem. It may depend on p as well.

We define the response function $f(p, x(p))$ as in (22). The objective is to determine its sensitivity with respect to all p_n, $n = 1, \ldots, N$ as follows:

$$\nabla_p f, \text{ subject to } Ax = b \tag{42}$$

where the gradient ∇_p is defined as a row operator

$$\nabla_p = \left[\frac{\partial}{\partial p_1} \cdots \frac{\partial}{\partial p_N} \right]. \tag{43}$$

A direct differentiation of (41) with respect to each design variable p_n, $n = 1, \ldots, N$ yields

$$A\frac{\partial x}{\partial p_n} = \frac{\partial b}{\partial p_n} - \frac{\partial A}{\partial p_n}x, \qquad n = 1, \ldots, N. \tag{44}$$

Using the definition of the gradient in (43) and assuming that A is not singular, (44) is written as

$$\nabla_p x = A^{-1} \left[\nabla_p b - \nabla_p (A\bar{x}) \right] \tag{45}$$

where

$$\nabla_p x = \begin{bmatrix} \dfrac{\partial x_1}{\partial p_1} & \cdots & \dfrac{\partial x_1}{\partial p_N} \\ \vdots & & \vdots \\ \dfrac{\partial x_M}{\partial p_1} & \cdots & \dfrac{\partial x_M}{\partial p_N} \end{bmatrix}. \tag{46}$$

In (45), $(A\bar{x})$ is a vector obtained from the solution at the current design $\bar{x}$, which is held constant during the differentiation, see (44). We note parenthetically that, in high-frequency design, the EM sources are usually independent of the design parameters, i.e., $\nabla_p b = 0$.

Expression (45) gives the sensitivities of all state variables. Once they are known, the sensitivity of the response function (42) can be obtained by substituting $\nabla_p x$ into

$$\nabla_p f = \nabla_p^e f + \nabla_x f \cdot \nabla_p x. \tag{47}$$

Here, the gradient operator ∇_x is defined as in (43), and $\nabla_p^e f$ reflects the explicit dependence of $f(p, x(p))$ on p. Often, $\nabla_p^e f = 0$.

The procedure summarized by (45) and (47) is what is usually referred to as sensitivity analysis via direct differentiation. The computational overhead associated with it depends on the method of solution of (41). If *LU* factoring of the A matrix is used, then the overhead is small because the A matrix *LU* factors

are reused in the forward–backward substitutions in (45). Obviously, the number of forward–backward substitutions is equal to the number of design parameters N. On the other hand, if $\boldsymbol{A}$ is very large and sparse (as, for example, is the case with the FEM), the preferred method of solution may be an iterative solver. Neither $\boldsymbol{A}^{-1}$, nor its *LU* factors are then available after the system analysis (41) is completed. The solution for the state variable sensitivities via (45) now requires N linear system solutions—each equivalent to that in (41). This is a very substantial overhead, which renders the technique as inefficient as the perturbation technique where finite differences are applied at the level of the response.

Another issue relates to the computation of the derivatives of the system matrix with respect to all design parameters, $\partial \boldsymbol{A}/\partial p_n$, $n = 1, \ldots, N$ in (44) or (45). In linear circuits, this is usually not a problem, as these derivatives are analytically available. In full-wave analysis, however, these derivatives pose a major difficulty. We will return later to this problem when we discuss adjoint-based sensitivities with EM solvers.

The direct differentiation method has been applied to a number of practical microwave design examples [19], [40], [51].

B. Adjoint-Variable Method for Linear Systems

In optimization, we are usually interested in the sensitivity of the objective function $\nabla_{\boldsymbol{p}} f$ rather than the sensitivities of all state variables. Thus, the calculation of $\nabla_{\boldsymbol{p}} \boldsymbol{x}$ via (45) seems somewhat superfluous. If (45) is substituted in (47), the response sensitivity is obtained as

$$\nabla_{\boldsymbol{p}} f = \nabla_{\boldsymbol{p}}^e f + \nabla_{\boldsymbol{x}} f \boldsymbol{A}^{-1} \left[\nabla_{\boldsymbol{p}} \boldsymbol{b} - \nabla_{\boldsymbol{p}}(A\bar{\boldsymbol{x}}) \right]. \tag{48}$$

We can now introduce the adjoint-variable vector as [11], [12], [49]

$$\hat{\boldsymbol{x}}^T = \nabla_{\boldsymbol{x}} f \cdot \boldsymbol{A}^{-1}. \tag{49}$$

This defines the adjoint problem as

$$\boldsymbol{A}^T \hat{\boldsymbol{x}} = [\nabla_{\boldsymbol{x}} f]^T \tag{50}$$

analogously to (16) and (38). With the introduction of the adjoint solution, the sensitivity expression (48) becomes

$$\nabla_{\boldsymbol{p}} f = \nabla_{\boldsymbol{p}}^e f + \hat{\boldsymbol{x}}^T \left[\nabla_{\boldsymbol{p}} \boldsymbol{b} - \nabla_{\boldsymbol{p}}(A\bar{\boldsymbol{x}}) \right]. \tag{51}$$

For clarity, we rewrite (51) as

$$\frac{\partial f}{\partial p_n} = \frac{\partial^e f}{\partial p_n} + \hat{\boldsymbol{x}}^T \left[\frac{\partial \boldsymbol{b}}{\partial p_n} - \frac{\partial \boldsymbol{A}}{\partial p_n} \cdot \bar{\boldsymbol{x}} \right], \qquad n = 1, \ldots, N. \tag{52}$$

The computational overhead of the adjoint-variable method is approximately N times smaller than that of the sensitivity analysis via direct differentiation. If the *LU* factors of the $\boldsymbol{A}$ matrix are available, only one forward–backward substitution is performed to solve the adjoint problem (50). Similarly, in the case of an iterative linear system solver, only one additional system analysis is performed when solving (50). Thus, the efficiency

of the adjoint variable method is independent of the number of design parameters. In the worst case, its overhead is roughly equivalent to one additional system analysis.

C. Sensitivities of Complex Linear Systems

A note of caution should be made. Strictly speaking, the derivation above refers to real-number linear systems. In the case of complex linear systems, the transposition of the system matrix $\boldsymbol{A}$ also involves conjugation. Such a matrix is referred to as Hermitian conjugate, and is sometimes denoted as $\boldsymbol{A}^H$ (or $\boldsymbol{A}^{T^*}$) instead of $\boldsymbol{A}^T$, which may be misleading in the case of complex analysis. From the theory of complex matrices, it is known that the Hermitian conjugate of a complex-number matrix has properties analogous to the transpose of a real-number matrix, e.g., $(\boldsymbol{AB})^H = \boldsymbol{B}^H \boldsymbol{A}^H$; see, e.g., [52]. This applies to every analysis, which requires the transposition of a complex-valued system matrix. In fact, the adjoint sensitivity expressions in the case of complex systems are readily obtained from those for real-valued systems.

Consider the following representation of the complex-valued linear system (41) in a real-number form

$$\begin{bmatrix} \Re\boldsymbol{A} & -\Im\boldsymbol{A} \\ \Im\boldsymbol{A} & \Re\boldsymbol{A} \end{bmatrix} \begin{bmatrix} \Re\boldsymbol{x} \\ \Im\boldsymbol{x} \end{bmatrix} = \begin{bmatrix} \Re\boldsymbol{b} \\ \Im\boldsymbol{b} \end{bmatrix}. \tag{53}$$

The application of the standard adjoint-based approach to (53) leads to the sensitivity expression

$$\nabla_{\boldsymbol{p}} f = \nabla_{\boldsymbol{p}}^e f + \begin{bmatrix} \Re\hat{\boldsymbol{x}} \\ \Im\hat{\boldsymbol{x}} \end{bmatrix}^T \cdot \nabla_{\boldsymbol{p}} \left(\begin{bmatrix} \Re\boldsymbol{b} \\ \Im\boldsymbol{b} \end{bmatrix} - \begin{bmatrix} \Re\boldsymbol{A} & -\Im\boldsymbol{A} \\ \Im\boldsymbol{A} & \Re\boldsymbol{A} \end{bmatrix} \begin{bmatrix} \Re\bar{\boldsymbol{x}} \\ \Im\bar{\boldsymbol{x}} \end{bmatrix} \right) \tag{54}$$

where the real-valued adjoint vector is a solution to

$$\begin{bmatrix} \Re\boldsymbol{A}^T & \Im\boldsymbol{A}^T \\ -\Im\boldsymbol{A}^T & \Re\boldsymbol{A}^T \end{bmatrix} \begin{bmatrix} \Re\hat{\boldsymbol{x}} \\ \Im\hat{\boldsymbol{x}} \end{bmatrix} = [\nabla_{\Re\boldsymbol{x}} f \quad \nabla_{\Im\boldsymbol{x}} f]^T. \tag{55}$$

Notice that the response f is assumed to be a real quantity. It is now obvious that (54) and (55) can be cast in equivalent complex-number representations

$$\nabla_{\boldsymbol{p}} f = \nabla_{\boldsymbol{p}}^e f + \Re \left[\hat{\boldsymbol{x}}^H \cdot \nabla_{\boldsymbol{p}}(\boldsymbol{b} - \boldsymbol{A}\bar{\boldsymbol{x}}) \right] \tag{56}$$

$$\boldsymbol{A}^H \hat{\boldsymbol{x}} = [\nabla_{\boldsymbol{x}} f]^H \tag{57}$$

where $\nabla_{\boldsymbol{x}} f = \nabla_{\Re\boldsymbol{x}} f - j\nabla_{\Im\boldsymbol{x}} f$. Note that the right-hand side of (57) has to be conjugated when the row of complex derivative elements is transposed to produce a column, i.e.,

$$[\nabla_{\boldsymbol{x}} f]^H = \begin{bmatrix} \dfrac{\partial f}{\partial \Re x_1} + j\dfrac{\partial f}{\partial \Im x_1} \\ \vdots \\ \dfrac{\partial f}{\partial \Re x_M} + j\dfrac{\partial f}{\partial \Im x_M} \end{bmatrix}. \tag{58}$$

The same rule of conjugation applies to $\hat{\boldsymbol{x}}^H$ in (56) when the adjoint-variable vector $\hat{\boldsymbol{x}}$ is converted into a row.

We should point out that the sensitivity analysis represented by (56)–(58) can be extended to complex response functions. To find the sensitivity of the real and imaginary parts of the

response function $f = f_R + j f_I$, we still need only one complex linear system analysis. Using Cauchy–Riemann equations,

$$\nabla_{\Re x} f_R = \nabla_{\Im x} f_I = \Re \nabla_x f$$
$$-\nabla_{\Im x} f_R = \nabla_{\Re x} f_I = \Im \nabla_x f \qquad (59)$$

we can show that if the complex adjoint problem is defined as in (57) with $f = f_R$, then the complex sensitivity expression becomes

$$\nabla_p f = \nabla_p^e f + \hat{x}^H \cdot \nabla_p (b - A\bar{x}). \qquad (60)$$

Another equivalent complex-number representation, which avoids the conjugation of the system matrix, is

$$\nabla_p f = \nabla_p^e f + (\hat{x}^*)^T \cdot \nabla_p (b - A\bar{x}) \qquad (61)$$
$$A^T \hat{x}^* = [\nabla_x f]^T \qquad (62)$$

where $[\nabla_x f]^T$ is the conjugate of (58) with $f = f_R$. This is the case with the adjoint network method when it is applied to complex circuit analysis, where $\hat{x}^*$ is considered to be the adjoint system solution instead of $\hat{x}$.

The adjoint-variable method is the most computationally efficient approach to design sensitivity analysis of linear systems. The CPU-time reduction in comparison with direct forward or backward finite-difference sensitivity approximations at the level of the response, e.g.,

$$\frac{\partial f\left(p_n^{(k)}\right)}{\partial p_n} \approx \frac{f\left(p_n^{(k)} + \Delta p_n^{(k)}\right) - f\left(p_n^{(k)}\right)}{\Delta p_n^{(k)}},$$
$$n = 1, \ldots, N \qquad (63)$$

is very significant. It is of the order of $(N + 1)$.

The adjoint-network method for circuit sensitivities presented in Sections II and III and the adjoint-variable approach to linear system sensitivities presented here are conceptually identical. They rely on an adjoint system analysis whose system matrix is the transpose of the original system matrix, and their sensitivity expressions utilize the derivatives of the matrices describing the circuit or system. The difference is in the fact that the adjoint-network method computes the overall circuit sensitivity as a summation over the individual sensitivities of the circuit components (34), (39), each of these sensitivities being dependent on the derivative of the component matrix (a component-level approach). The particular solution method used to carry out the circuit analysis and the resulting system matrix are of no importance to the sensitivity expression as long as they provide the accurate voltage/current distribution in the original and adjoint circuits.

On the other hand, the adjoint-variable method uses the derivative of the system matrix to compute the overall sensitivity directly (a system-level approach) (56), (60). Strictly speaking, this system matrix must be the one arising from the chosen analysis method. When the solution, however, is known to be unique, it is possible to use the derivative of a system matrix arising from a given analysis method and a vector of state variables (i.e., the solution) obtained by another method.

For example, in [37] and [38], the sensitivity expression uses the derivatives of a matrix arising from a finite-element discretization and a solution obtained via the FDTD method.

D. Adjoint-Variable Method for Nonlinear Systems [15], [16]

The previously discussed matrix approach can be extended to the sensitivity analysis of nonlinear systems. Our previous results then appear as a special case of the nonlinear system analysis. This analysis has been applied mainly with microwave circuits, e.g., [15] and [16], since most of the full-wave EM solvers assume a linear medium. However, the trend of integrating lumped nonlinear components with full-wave EM solvers justifies an overview of the subject.

Assume a problem is represented by the system of nonlinear equations

$$F(p, x) = 0. \qquad (64)$$

As before, $p = [p_1 \ldots p_N]^T$ represents the vector of design parameters, and $x = [x_1 \ldots x_M]^T$ is the vector of state variables. Differentiating (64) directly with respect to a given design parameter p_n gives

$$\frac{\partial F}{\partial p_n} = -J \cdot \frac{\partial x}{\partial p_n} \qquad (65)$$

where

$$J = \nabla_x F = \begin{bmatrix} \dfrac{\partial F_1}{\partial x_1} & \cdots & \dfrac{\partial F_1}{\partial x_M} \\ \vdots & & \vdots \\ \dfrac{\partial F_M}{\partial x_1} & \cdots & \dfrac{\partial F_M}{\partial x_M} \end{bmatrix} \qquad (66)$$

is the Jacobian of (64). If J is a nonsingular matrix, the sensitivities of the state variables are found as

$$\nabla_p x = -J^{-1} \cdot \nabla_p F. \qquad (67)$$

Equation (67) is a generalization of the linear case in (45) with $F = Ax - b$ and $J = A$. Its substitution in (47) results in

$$\nabla_p f = \nabla_p^e f + \nabla_x f \cdot \left(-J^{-1} \nabla_p F\right) \qquad (68)$$

which is used to introduce the adjoint problem as the solution to the linear system

$$J^T \hat{x} = [\nabla_x f]^T. \qquad (69)$$

Finally, the sensitivity expression is obtained as

$$\nabla_p f = \nabla_p^e f - \hat{x}^T \cdot \nabla_p F(\bar{x}). \qquad (70)$$

Here, $\bar{x}$ denotes the solution at the current design. It is kept constant when calculating the gradient of F on the right-hand side of (70).

The nonlinear system (64) corresponding to the current design $F(p, x) = 0$ of the optimization can be solved using Newton's update [15]

$$x^{k+1} = x^k - J^{-1} F(x^k) \qquad (71)$$

where k denotes the iteration of the update. At each iteration, the Jacobian is updated and factored in order to compute $\boldsymbol{x}^{k+1}$. Therefore, its *LU* factors are available after the last update and can be used for the sensitivity estimation via (69) and (70) at the current design.

Nonlinear circuit sensitivities with harmonic-balance simulations of microwave circuits such as mixers and frequency doublers have been applied in yield analysis and optimization [15], [16]. Therein, a detailed description of related numerical algorithms and nonlinear circuit analysis approaches can be found.

V. ADJOINT SENSITIVITIES IN EM ANALYSIS

The theory of adjoint sensitivities for linear systems discussed in the previous section, in principle, applies to EM analysis in linear media. However, to our knowledge, commercial EM-based CAD software does not employ adjoint sensitivities for optimization purposes. A major difficulty in EM-based sensitivity analysis is the availability of the derivatives of the system matrix with respect to the design parameters, $\partial \boldsymbol{A} / \partial p_n$, $n = 1, \ldots, N$ [see (52)]. When the design parameters describe variations in the shape and size of objects, they may relate to the solver's discrete mesh, its expansion, and weighting functions in a very complicated manner. This prevents applications versatile enough to allow for an arbitrary user-defined set of design parameters.

Recently, there have been efforts to integrate adjoint sensitivities with certain EM computational approaches. We outline these achievements and discuss issues related to their applicability, efficiency, and accuracy.

A. Adjoint Network Applications With Mode-Matching (MM) Analysis [17]

The MM method—particularly well suited for waveguide problems—utilizes modal expansions of the EM field in properly defined sub-domains of the computational volume. The boundary conditions at a junction plane between two neighboring sub-domains lead to a linear system of equations in the modal expansion coefficients. Traditional MM techniques rely on segmentation such that the sub-domains are waveguide sections where the field expansion is in terms of the intrinsic waveguide modes. The equivalence between a waveguide mode—characterized by its wave impedance and propagation constant—and a transmission line is used to develop microwave network formalism where every mode is represented by a transmission line connecting multiport networks corresponding to respective discontinuities at junction planes. Each port of these networks corresponds to a given mode. Thus, the microwave structure is represented by an equivalent network whose complexity depends on the number of modes taken into account in the field expansion at the discontinuities. This is why the adjoint network approach to design sensitivity analysis is particularly well suited for application with MM techniques.

In [17], the authors exploit the fact that the admittance matrix for a subnetwork corresponding to a microwave component bounded by metal walls is evaluated analytically and is, thus, differentiable in the geometrical parameters of this component. Using an efficient cellular segmentation technique [53] and a generalized admittance matrix formulation [54], they cast the problem in the form

$$\boldsymbol{I}_E = \boldsymbol{Y} \boldsymbol{V}_E \tag{72}$$

where $\boldsymbol{I}_E$ and $\boldsymbol{V}_E$ are vectors consisting of the currents and voltages at the external ports of the microwave network. Due to the reciprocity of the waveguide structures of interest, the adjoint network is identical to the original one. Further, the derivatives of the system matrix $\boldsymbol{Y}$ with respect to a design parameter p_n are expressed in terms of the analytically available derivatives of the subnetworks' $\boldsymbol{Y}$-matrices as

$$\frac{\partial Y_{ij}}{\partial p_n} = \sum_{m=1}^{M} \boldsymbol{V}_{mi}^T \frac{\partial \boldsymbol{Y}_m^T}{\partial p_n} \boldsymbol{V}_{mj} \tag{73}$$

where $\boldsymbol{V}_{mi}$ are the voltages at the ports of the subnetwork m when the circuit is excited from the external port i and $\boldsymbol{Y}_m$ is the admittance matrix of the mth subnetwork.

The voltages at the external ports are known as they are set equal to zero or one depending on which port is excited. From them, the voltages at the ports of the internal (or connected) ports are determined. For that, a matrix equation is constructed making use of the connection matrix, which expresses the topological relations of the connected ports. Finding the internal-port voltages requires just one additional system analysis. Once the sensitivity of the system $\boldsymbol{Y}$-matrix is found from (73), the sensitivity of the $\boldsymbol{S}$ matrix is readily computed [13], [17].

This technique has been used in the design of waveguide filters.

B. Adjoint Sensitivities With the FEM

The FEM offers very good flexibility and accuracy for adjoint-based sensitivity analysis. This is due to the fact that the elements of the FEM system matrix are analytically differentiable with respect to the coordinates of the vertices of the finite-element mesh. This is why sensitivities have been first computed with FEM analysis in structural and electrical engineering (see, e.g., [29], [49], [55], and [56]).

It appears that Lee and Itoh were first to apply adjoint variables with the FEM to high-frequency structure design [35]. Later, the approach was generalized to handle an arbitrary variation in the geometry [30]–[33]. Akel and Webb [30] and Webb [31] developed a systematic way of computing the derivatives of the FEM system matrix with respect to the rectangular coordinates of the vertices of the mesh. The sensitivity estimates showed excellent accuracy when compared with the results of a central finite-difference calculation using the response directly. The same authors also showed that there is no need for an adjoint-system analysis when the sensitivities of the structure's S-parameters are desired [30] provided that the port geometries do not change under variations of the design variables. These sensitivities with respect to a design parameter p_n are obtained as

$$\frac{\partial S_{jk}}{\partial p_n} = -\frac{1}{2} \bar{\boldsymbol{E}}^{(j)T} \frac{\partial \boldsymbol{A}}{\partial p_n} \bar{\boldsymbol{E}}^{(k)} \tag{74}$$

where $\bar{\bar{E}}^{(j)}$ and $\bar{\bar{E}}^{(k)}$ are vectors containing the **E**-field solution at port j and port k, respectively.

Further, Webb developed an efficient technique for the calculation of the S-parameter sensitivities in a broad frequency range using a single-frequency analysis and employing Padé expansion with complex frequencies [32]. The accuracy control of the FEM solution for both the S-parameters and their sensitivities is considered in [33] and [34].

C. Adjoint Sensitivities With the MoM

Exact sensitivities of charge and current densities computed with the MoM have been considered in [57] and [58]. In [57], electrostatic problems for planar conducting surfaces embedded in planar structures are solved by applying the MoM to the integral equation relating the potential on the conductor's surface to the surface charge distribution. A new integral equation is developed for the derivative of the potential with respect to the design (shape) parameter. The two integral equations are solved with the MoM using the same set of basis functions. A similar approach is applied with the mixed-potential integral equation for the analysis of planar microstrip structures, where two coupled integral equations are solved [58]. The sensitivities of the S-parameters are calculated from the sensitivity of the current densities. These techniques, although formally different from the adjoint variable methods, provide accurate derivative estimation whose computational requirements are independent of the number of design parameters. The drawback is that their implementation into a design environment would require programming of the analysis software practically from scratch.

Recently, a feasible adjoint-based technique was proposed for use with full-wave frequency-domain solvers [39]. More specifically, it considers applications with the MoM. There are a number of MoM techniques currently in use in microwave and antenna engineering. They differ significantly in the choice of expansion and test functions. To reduce the computational effort, most MoM techniques utilize specialized Green's functions; e.g., planar MoM solvers compute a Green's function, which satisfies the boundary conditions of a multilayer dielectric medium with specific heights and dielectric constants of the layers. Unlike in the FEM, these functions lead to a system matrix whose elements depend on the coordinates of the point of observation and the point of integration in a very complicated manner. With very few exceptions, straightforward analytical differentiation is not feasible. Moreover, it would require complete reprogramming of the existing codes, which is an insurmountable complication for most researchers and commercial software developers.

The feasible technique proposed in [39] applies finite-difference approximation to the derivatives of the MoM system matrix, thus avoiding analytical preprocessing. It requires only minor additions to the existing analysis algorithms. Feasible techniques utilizing finite differences for the approximation of the Jacobian of the system equations have been previously applied to the sensitivity analysis of nonlinear circuits [16]. The accuracy of the technique with the MoM is carefully investigated, and it is shown that perturbations of the design parameters in the range of 1%–5% yields excellent accuracy when compared with both exact sensitivities and central finite-difference approximations utilizing the response directly. The adjoint system analysis reuses the *LU* factors of the system matrix obtained in the solution of the original problem. The overhead associated with the sensitivity estimation is thus due mainly to the computation of the derivatives of the system matrix, which requires N additional matrix fills at each design iteration. Keeping in mind that in a large problem a matrix fill requires a small portion of the overall CPU time needed for one system analysis, this overhead is significantly less than N additional full analyses. Thus, the technique offers a reduction of the computational cost of the design process by a factor of approximately N.

The feasible technique has been applied to the sensitivity analysis and the design of wire antennas (dipoles and Yagi–Uda arrays) and printed patch antennas [39]. The response functions include the antenna input impedance and antenna gain. The technique is simple to implement, robust, and accurate. It is versatile as its implementation is practically independent of the type of the MoM algorithm or its Green's functions.

Possible venues for this line of research concern the application of other computationally more efficient gradient approximation techniques to the estimation of the derivatives of the system matrix [59]. A number of such techniques have been developed for the purposes of gradient-free optimization, and they can be applied at the level of the system matrix rather than the level of the response itself. Such an approach would reduce the computational requirements of the design process in comparison with the feasible adjoint technique. At the same time, it would improve the accuracy of the sensitivity estimation in comparison with approximations applied to the response directly. This is due to the weakly nonlinear dependence of the elements of the system matrix on geometrical perturbations, unlike the response function, which may exhibit highly nonlinear behavior at certain points of the design parameter space.

D. Adjoint Sensitivities With Fixed Structured Grids

Finite-difference or transmission-line modeling methods for EM analysis typically utilize fixed structured grids in Cartesian coordinates or orthogonal curvilinear coordinates. Their equations can be cast in a matrix form with a system matrix, which is very sparse. The elements of the system matrix, however, are not analytical functions of the coordinates of the grid nodes and, therefore, they are not differentiable with respect to design parameters related to the structure's geometry. We note parenthetically that the matrix elements are differentiable with respect to the constitutive parameters and, thus, allow for a straightforward computation of material-related adjoint sensitivities.

We illustrate the adjoint-based sensitivity analysis on discrete structured grids with a recent application with the FD–TLM [41]. Let us consider a rectangular mesh and metallic boundaries perturbed as shown in Fig. 5. The designable parameters (e.g., L and W) can assume only a discrete set of values that are multiples of the cell size in the respective direction. The perturbation shown in Fig. 5(b) (referred to as "metallization") requires that the respective tangential field components (total nodal voltages in the case of TLM) be set equal to zero in the "metallized" cells. This effectively reduces the number of unknowns, which
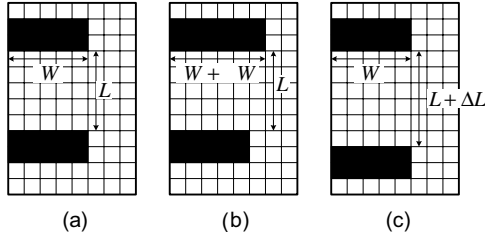
Fig. 5.   Discrete perturbations of metallic boundaries that result in cells being "metallized" and "de-metallized." (a) Nominal structure. (b) Perturbation ΔW results in two cells being "metallized." (c) Perturbation ΔL results in five cells being "metallized" and other five being "de-metallized."

the system is solved for. In general, a perturbation may require previously perfectly conducting cells to be "de-metallized" and others to be "metallized" [see Fig. 5(c)]. When cells are being de-metallized, new unknowns are added to the vector of state variables. In any case, the perturbation results in a new system matrix with a different size because variables (incident voltages or field components) are added or removed at the location of perturbation. Thus, a meaningful definition of a derivative of the system matrix is not possible unless an enlarged system matrix is introduced in anticipation of possible geometrical modifications.

When a cell is metallized, the coefficients in the system matrix multiplying its vanishing field components must change from a fixed known value to zero. Alternatively, when a cell is de-metallized, the respective coefficients change from zero to known fixed values. This simple "binary" behavior in the case of perfectly conducting walls is used to calculate the finite-difference approximation $\Delta \boldsymbol{A}/\Delta p_n$, $n = 1, \ldots, N$. The design parameter perturbation is limited to the smallest on-grid perturbation along the respective axis $\Delta p_n = \Delta h$, where Δh is the size of the grid step. It is obvious that the difference matrices $\Delta \boldsymbol{A}/\Delta p_n$ have very few nonzero elements corresponding to the locations at which the boundary conditions change.

We should emphasize that the perturbed structure is never actually simulated. We consider it only because we need an estimate of the derivatives of the system matrix with respect to all design parameters.

Such a "discrete" adjoint-based approach has to take into account second-order terms in the perturbed system of equations describing the linear EM problem (41)

$$(\boldsymbol{A} + \Delta \boldsymbol{A})(\bar{\boldsymbol{x}} + \Delta \boldsymbol{x}) = \boldsymbol{b}. \tag{75}$$

In (75), $\bar{\boldsymbol{x}}$ represents the solution of the problem at the current design. Simplifying (75),

$$\boldsymbol{A}\Delta \boldsymbol{x} = \Delta \boldsymbol{b} - (\Delta \boldsymbol{A}\bar{\boldsymbol{x}} + \Delta \boldsymbol{A}\Delta \boldsymbol{x}). \tag{76}$$

Notice that (76) is not just the finite-difference approximation of (44) obtained from the direct differentiation of (41). In (76), the second-order term $\Delta \boldsymbol{A}\Delta \boldsymbol{x}$ is retained. This is necessary in order to achieve reasonable accuracy when the perturbation $\Delta \boldsymbol{A}$ is actually of the same magnitude as $\boldsymbol{A}$ itself. The technique further approximates the elements of the vector $\Delta \boldsymbol{x} \approx \Delta \tilde{\boldsymbol{x}}$ on the right-hand side of (76), which correspond to the few nonzero columns of $\Delta \boldsymbol{A}$. The components of $\Delta \tilde{\boldsymbol{x}}$ corresponding to the

metallized cells are set to $\Delta \tilde{\boldsymbol{x}}_k = -\bar{\boldsymbol{x}}_k$ as the incident voltages vanish for these cells. For the de-metallized nodes, the corresponding components of $\Delta \tilde{\boldsymbol{x}}$ are approximated by the values of the incident voltages of neighboring nodes along the perturbation direction [41]. The adjoint-based sensitivity expression is now given as

$$\frac{\partial f}{\partial p_n} \approx \frac{\Delta^e f}{\Delta p_n} + \hat{\boldsymbol{x}}^T \left(\frac{\Delta \boldsymbol{b}}{\Delta p_n} - \frac{\Delta \boldsymbol{A}}{\Delta p_n} \tilde{\boldsymbol{x}} \right) \tag{77}$$

where $\tilde{\boldsymbol{x}} = \bar{\boldsymbol{x}} + \Delta \tilde{\boldsymbol{x}}$. The adjoint solution $\hat{\boldsymbol{x}}$ is obtained in the usual way from the adjoint system (50).

The approach described above is, in principle, applicable with any frequency- or time-domain EM technique, which uses a fixed structured mesh.

E. Adjoint Sensitivities With Time-Domain Techniques

The subject of dynamic response sensitivity has been given very little consideration in the microwave literature. Some applications with the FETD and FDTD methods have been reported in [36]–[38]. These applications are based on a variational theory developed for second-order time-dependent problems [49]. We present this theory in the context of time-domain EM analysis.

Consider the vector wave equation for the electric field vector

$$\nabla \times \frac{1}{\mu_r} \nabla \times \mathbf{E} + \mu_0 \sigma \frac{\partial \mathbf{E}}{\partial t} + \frac{\varepsilon_r}{c_0^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = -\mu_0 \frac{\partial \mathbf{J}}{\partial t} \tag{78}$$

where $c_0 = (\mu_0 \varepsilon_0)^{-1/2}$, μ_0 and ε_0 are the magnetic permeability and dielectric permittivity of vacuum, respectively, μ_r is the relative permeability, ε_r is the relative permittivity, and σ is the conductivity of the medium. It can be discretized (e.g., using finite elements) to construct a system of equations

$$\boldsymbol{M}\ddot{\boldsymbol{x}} + \boldsymbol{N}\dot{\boldsymbol{x}} + \boldsymbol{K}\boldsymbol{x} = \boldsymbol{Q} \tag{79}$$

with initial conditions

$$\boldsymbol{x}(0) = \boldsymbol{0}$$
$$\dot{\boldsymbol{x}}(0) = \boldsymbol{0}. \tag{80}$$

In the matrix (79), the first- and second-order time derivatives are denoted as $\dot{\boldsymbol{x}}$ and $\ddot{\boldsymbol{x}}$, respectively. The vector $\boldsymbol{x}$ contains the tangential components of the $\mathbf{E}$-field along the edges of the finite-element mesh.

The response function can be defined as a general integral functional of the form

$$F = \int_0^{T_{\max}} \iiint_\Omega f(\mathbf{E}, \boldsymbol{p}) d\Omega dt \tag{81}$$

where Ω is the volume of observation and $T_{\max}$ is the period of observation. The volume of observation must be contained within the computational volume, and $T_{\max}$ is usually the same as the computational time required by the time-domain simulation. F is an implicit function of $\boldsymbol{p}$ through the electric field solution $\mathbf{E}$ (or equivalently through $\boldsymbol{x}$), but it may also have an explicit dependence on the design variables $\boldsymbol{p}$.

Using the chain rule of differentiation, the derivative of F with respect to the design parameter p_n is obtained as

$$\frac{\partial F}{\partial p_n} = \int_0^{T_{\max}} \left(\nabla_{\boldsymbol{x}} f \cdot \frac{\partial \boldsymbol{x}}{\partial p_n} \right) dt + \frac{\partial^e F}{\partial p_n} \quad (82)$$

where $\partial^e / \partial p_n$ relates to the explicit dependence only of F on p_n. The evaluation of (82) is subject to the solution of the system (79) complemented by the initial conditions (80) and by specified boundary conditions.

We multiply a residual vector

$$\boldsymbol{R} = \boldsymbol{M}\ddot{\boldsymbol{x}} + \boldsymbol{N}\dot{\boldsymbol{x}} + \boldsymbol{K}\boldsymbol{x} - \boldsymbol{Q} = \boldsymbol{0} \quad (83)$$

with the transpose of an auxiliary vector $\hat{\boldsymbol{x}}$ and integrate in time

$$\Psi = \int_0^{T_{\max}} \hat{\boldsymbol{x}}^T \cdot (\boldsymbol{M}\ddot{\boldsymbol{x}} + \boldsymbol{N}\dot{\boldsymbol{x}} + \boldsymbol{K}\boldsymbol{x} - \boldsymbol{Q}) dt = 0. \quad (84)$$

The above expression must hold for an arbitrary vector $\hat{\boldsymbol{x}}$, which is *independent* of any changes occurring in the original problem. Thus, differentiating Ψ with respect to p_n yields

$$\frac{\partial \Psi}{\partial p_n} = \int_0^{T_{\max}} \hat{\boldsymbol{x}}^T \cdot \left(\frac{\partial \boldsymbol{M}}{\partial p_n}\ddot{\boldsymbol{x}} + \frac{\partial \boldsymbol{N}}{\partial p_n}\dot{\boldsymbol{x}} + \frac{\partial \boldsymbol{K}}{\partial p_n}\boldsymbol{x} - \frac{\partial \boldsymbol{Q}}{\partial p_n} \right) dt$$
$$+ \int_0^{T_{\max}} \hat{\boldsymbol{x}}^T \cdot \left(\boldsymbol{M}\frac{\partial \ddot{\boldsymbol{x}}}{\partial p_n} + \boldsymbol{N}\frac{\partial \dot{\boldsymbol{x}}}{\partial p_n} + \boldsymbol{K}\frac{\partial \boldsymbol{x}}{\partial p_n} \right) dt$$
$$= 0. \quad (85)$$

The first variation of Ψ in (85) is zero when the vectors $\boldsymbol{x}$, $\dot{\boldsymbol{x}}$, and $\ddot{\boldsymbol{x}}$ represent the solution to the original problem (79). In the second integral of (85), we integrate by parts twice the term $(\hat{\boldsymbol{x}}^T \boldsymbol{M} \partial \ddot{\boldsymbol{x}}/\partial p_n)$, as well as we integrate by parts once the term $(\hat{\boldsymbol{x}}^T \boldsymbol{N} \partial \dot{\boldsymbol{x}}/\partial p_n)$. This shifts the time derivatives onto the auxiliary vector $\hat{\boldsymbol{x}}$. Note that the matrices $\boldsymbol{M}, \boldsymbol{N}, \boldsymbol{K}$, and $\boldsymbol{Q}$ are independent of time. We also introduce the notation

$$\frac{\partial \boldsymbol{R}(\bar{\boldsymbol{x}}, \bar{\dot{\boldsymbol{x}}}, \bar{\ddot{\boldsymbol{x}}})}{\partial p_n} = \frac{\partial \boldsymbol{M}}{\partial p_n}\bar{\ddot{\boldsymbol{x}}} + \frac{\partial \boldsymbol{N}}{\partial p_n}\bar{\dot{\boldsymbol{x}}} + \frac{\partial \boldsymbol{K}}{\partial p_n}\bar{\boldsymbol{x}} - \frac{\partial \boldsymbol{Q}}{\partial p_n}. \quad (86)$$

The bar in $\bar{\boldsymbol{x}}$ and its derivatives emphasizes that the field solution is held constant during the differentiation of $\boldsymbol{R}$ and contains the values obtained from the analysis at the current design. Thus, (85) is transformed into

$$\frac{\partial \Psi}{\partial p_n} = \int_0^{T_{\max}} \hat{\boldsymbol{x}}^T \cdot \frac{\partial \boldsymbol{R}(\bar{\boldsymbol{x}}, \bar{\dot{\boldsymbol{x}}}, \bar{\ddot{\boldsymbol{x}}})}{\partial p_n} dt$$
$$+ \int_0^{T_{\max}} \left(\ddot{\hat{\boldsymbol{x}}}^T \boldsymbol{M} - \dot{\hat{\boldsymbol{x}}}^T \boldsymbol{N} + \hat{\boldsymbol{x}}^T \boldsymbol{K} \right) \cdot \frac{\partial \boldsymbol{x}}{\partial p_n} dt$$
$$+ \hat{\boldsymbol{x}}^T \cdot \left(\boldsymbol{M}\frac{\partial \dot{\boldsymbol{x}}}{\partial p_n} + \boldsymbol{N}\frac{\partial \boldsymbol{x}}{\partial p_n} \right) \Big|_0^{T_{\max}} - \dot{\hat{\boldsymbol{x}}}^T \cdot \boldsymbol{M}\frac{\partial \boldsymbol{x}}{\partial p_n} \Big|_0^{T_{\max}}$$
$$= 0. \quad (87)$$

We now define the auxiliary field solution $\hat{\boldsymbol{x}}$ so that it relates (87) to (82). The zero initial conditions for the original problem (80) together with zero terminal conditions for the auxiliary field problem

$$\hat{\boldsymbol{x}}(T_{\max}) = \boldsymbol{0}$$
$$\dot{\hat{\boldsymbol{x}}}(T_{\max}) = \boldsymbol{0} \quad (88)$$

eliminate the last two terms in (87). Next, we choose an auxiliary vector $\hat{\boldsymbol{x}}$ such that

$$\ddot{\hat{\boldsymbol{x}}}^T \boldsymbol{M} - \dot{\hat{\boldsymbol{x}}}^T \boldsymbol{N} + \hat{\boldsymbol{x}}^T \boldsymbol{K} = \nabla_{\boldsymbol{x}} f. \quad (89)$$

This makes the second integral of (87) identical to the first term in the right-hand side of (82). Thus,

$$\frac{\partial \Psi}{\partial p_n} = \int_0^{T_{\max}} \hat{\boldsymbol{x}}^T \cdot \frac{\partial \boldsymbol{R}(\bar{\boldsymbol{x}}, \bar{\dot{\boldsymbol{x}}}, \bar{\ddot{\boldsymbol{x}}})}{\partial p_n} dt + \frac{\partial F}{\partial p_n} - \frac{\partial^e F}{\partial p_n} = 0. \quad (90)$$

We can now express the response derivative as

$$\frac{\partial F}{\partial p_n} = \frac{\partial^e F}{\partial p_n} - \int_0^{T_{\max}} \hat{\boldsymbol{x}}^T \cdot \frac{\partial \boldsymbol{R}(\bar{\boldsymbol{x}}, \bar{\dot{\boldsymbol{x}}}, \bar{\ddot{\boldsymbol{x}}})}{\partial p_n} dt. \quad (91)$$

Finally, we can generalize (91) to include all derivatives of the response gradient in the design parameter space

$$\nabla_{\boldsymbol{p}} F = \nabla_{\boldsymbol{p}}^e F - \int_0^{T_{\max}} \hat{\boldsymbol{x}}^T \cdot \nabla_{\boldsymbol{p}} \boldsymbol{R}(\bar{\boldsymbol{x}}, \bar{\dot{\boldsymbol{x}}}, \bar{\ddot{\boldsymbol{x}}}) dt. \quad (92)$$

The sensitivity expression above does not contain derivatives of the field solution with respect to the design parameters. It requires only the derivatives of the system matrices $\partial \boldsymbol{M}/\partial p_n$, $\partial \boldsymbol{N}/\partial p_n$, $\partial \boldsymbol{K}/\partial p_n$, $\partial \boldsymbol{Q}/\partial p_n$, $n = 1, \ldots, N$. It also requires the solution of (89), which we write as

$$\boldsymbol{M}^T \ddot{\hat{\boldsymbol{x}}} - \boldsymbol{N}^T \dot{\hat{\boldsymbol{x}}} + \boldsymbol{K}^T \hat{\boldsymbol{x}} = [\nabla_{\boldsymbol{x}} f]^T. \quad (93)$$

Equation (93) defines the adjoint problem. As before, the system matrices of the adjoint problem are equal to the transposed matrices of the original problem, and the adjoint excitation is the gradient of the local response function f with respect to the state variables $\boldsymbol{x}$. Note, however, the sign reversal associated with the first-order time derivative. The quasi-EM problem defined by (93) has the same boundary conditions as the original problem. Instead of zero initial conditions, it has zero terminal conditions (88). Thus, in an explicit scheme, the simulation runs backward in time.

The adjoint sources are located at the points of observation belonging to Ω where the local response is recorded. Typically, these points are located at a port of the microwave circuit. This is where the excitation of the adjoint circuit exists. For example, in [36], the response function is defined to be proportional to the average power leaving port 2 of a Ka-band unilateral fin line to waveguide transition

$$F = \int_0^{T_{\max}} \iint_{S_{\text{port}}} E_x^2(x, y, z_{P2}, t) dx dy \, dt \quad (94)$$

where $E_x(x, y, z_{P2}, t)$ is the total instantaneous transmitted electric field at the observation plane, which is port 2 at $z = z_{P2}$. The z-axis is along the direction of propagation. The local response is then

$$f(x, y, t) = E_x^2(x, y, t), \qquad z = z_{P2} \qquad (95)$$

which defines the adjoint excitation at each point of port 2 as

$$g(x, y, t) = \frac{\partial f}{\partial E_x} = 2E_x(x, y, t), \qquad z = z_{P2}. \qquad (96)$$

The original problem corresponding to (79) and the adjoint problem corresponding to (93) can be solved either by the FETD method based on the **E**-field wave equation or by Yee's FDTD algorithm based on the Maxwell curl equations. The details about the conversion of the adjoint excitation (96) into adjoint current densities for use with the adjoint Maxwell equations are given in [36] and [38]. In this particular application, the authors consider FDTD solutions to both the original and adjoint problems. To accommodate fine tuning of the grid nodes, the FDTD algorithm utilizes unstructured grids, which are regenerated with each design iteration. This makes the implementation rather complicated.

The adjoint problem, which is cast in the form of either a wave equation or a system of quasi-Maxwell equations (loss terms have their signs reversed), needs absorbing boundary conditions (ABCs) if the original problem itself contains open boundaries (matched ports or radiation into open space). A procedure for the derivation of various ABCs for adjoint problems is provided in [60].

The similarity between static/time–harmonic and dynamic system sensitivity analysis is obvious when (51) and (50) are compared with (92) and (93), respectively. In any case, two system analyses are sufficient to obtain the gradient of the response, regardless of the number of the design parameters. While the time–harmonic adjoint system analysis is practically avoided by the reuse of the *LU* factors of the original system matrix, the dynamic adjoint system analysis seems imperative.

The derivatives of the $\boldsymbol{M}, \boldsymbol{N}, \boldsymbol{K}$, and $\boldsymbol{Q}$ matrices in (86) are analytically available when finite-element discretization is applied. These FEM derivative matrices are used for sensitivity estimations with the FDTD method as well [36], [38]. The FDTD method is used only as an analysis tool to obtain the solution of the original and adjoint problems. As these solutions are unique, the FDTD solution should be identical with that of the FETD method.

F. Adjoint Sensitivities for Nonlinear Dynamic Systems

The full-wave EM analysis in the time domain usually assumes linear properties of the media involved. For the majority of practical microwave and antenna designs, this assumption holds. However, there are classes of EM problems, which are inherently nonlinear, e.g., involving ferromagnetic materials, light generation in semiconductors, nonlinear lumped devices, etc. For this reason, as well as for completeness, we briefly outline the sensitivity analysis of a class of nonlinear dynamic systems.

We assume that the time-dependent nonlinear problem is cast in the form of a system of M nonlinear equations of M state variables

$$\mathfrak{F}(\boldsymbol{x}, \dot{\boldsymbol{x}}, \ddot{\boldsymbol{x}}, \boldsymbol{p}) = \boldsymbol{0} \qquad (97)$$

where, as in the linear case, the vectors $\boldsymbol{x}$, $\dot{\boldsymbol{x}}$, and $\ddot{\boldsymbol{x}}$ represent the EM field solution and its time derivatives. We assume zero initial conditions as in (80). We emphasize that $\mathfrak{F}$ depends on time through the field solution only. We also mention that the following discussion is not limited to second-order systems. We assume the form in (97) since it is most often encountered in EM analysis. Its extension to higher order systems is straightforward.

An application of the direct differentiation method to the sensitivity analysis of (97) is described in [61]. The sensitivity of the state variables is found by solving N linear time-dependent systems of equations

$$\ddot{\boldsymbol{J}}\frac{\partial \ddot{\boldsymbol{x}}}{\partial p_n} + \dot{\boldsymbol{J}}\frac{\partial \dot{\boldsymbol{x}}}{\partial p_n} + \boldsymbol{J}\frac{\partial \boldsymbol{x}}{\partial p_n} = -\frac{\partial \mathfrak{F}\left(\bar{\boldsymbol{x}}, \dot{\bar{\boldsymbol{x}}}, \ddot{\bar{\boldsymbol{x}}}\right)}{\partial p_n}, \qquad n = 1, \dots, N. \qquad (98)$$

obtained by the direct chain differentiation of (97). Here, the Jacobian matrices are $\boldsymbol{J} = \nabla_{\boldsymbol{x}}\mathfrak{F}$, $\dot{\boldsymbol{J}} = \nabla_{\dot{\boldsymbol{x}}}\mathfrak{F}$, and $\ddot{\boldsymbol{J}} = \nabla_{\ddot{\boldsymbol{x}}}\mathfrak{F}$. The right-hand derivative term in (98) relates to the explicit dependence of $\mathfrak{F}$ on p_n, i.e., the state variables and their time derivatives are held constant during the differentiation. In the special case when $\mathfrak{F}$ represents the linear problem defined in (79) and (80), $\ddot{\boldsymbol{J}} = \boldsymbol{M}$, $\dot{\boldsymbol{J}} = \boldsymbol{N}$, $\boldsymbol{J} = \boldsymbol{K}$, and $\partial \mathfrak{F}(\bar{\boldsymbol{x}}, \dot{\bar{\boldsymbol{x}}}, \ddot{\bar{\boldsymbol{x}}})/\partial p_n = \partial \boldsymbol{R}(\bar{\boldsymbol{x}}, \dot{\bar{\boldsymbol{x}}}, \ddot{\bar{\boldsymbol{x}}})/\partial p_n$ defined in (86).

Here, we show that the adjoint-variable approach is applicable to nonlinear dynamic systems and that it needs only one additional linear dynamic system analysis. The methodologies for the solution of nonlinear time-dependent problems and their relation to the system sensitivity analysis are not discussed here. Excellent introduction into this subject can be found in [61] and [62].

The response function is defined as in (81). The sensitivity expression is obtained from the first variation of the functional

$$\Psi = \int_0^{T_{\max}} \hat{\boldsymbol{x}}^T \cdot \mathfrak{F}(\boldsymbol{x}, \dot{\boldsymbol{x}}, \ddot{\boldsymbol{x}}) dt = 0 \qquad (99)$$

with respect to a given design parameter p_n, which we write as

$$\frac{\partial \psi}{\partial p_n} = \int_0^{T_{\max}} \hat{\boldsymbol{x}}^T \cdot \left[\ddot{\boldsymbol{J}}\frac{\partial \ddot{\boldsymbol{x}}}{\partial p_n} + \dot{\boldsymbol{J}}\frac{\partial \dot{\boldsymbol{x}}}{\partial p_n} + \boldsymbol{J}\frac{\partial \boldsymbol{x}}{\partial p_n} + \frac{\partial \mathfrak{F}\left(\bar{\boldsymbol{x}}, \dot{\bar{\boldsymbol{x}}}, \ddot{\bar{\boldsymbol{x}}}\right)}{\partial p_n}\right] dt = 0. \qquad (100)$$

Following the same approach (as in the linear case), we integrate by parts and arrive at the intermediate result

$$\hat{\boldsymbol{x}}^T \cdot \left(\ddot{\boldsymbol{J}}\frac{\partial \dot{\boldsymbol{x}}}{\partial p_n} + \dot{\boldsymbol{J}}\frac{\partial \boldsymbol{x}}{\partial p_n}\right)\Bigg|_0^{T_{\max}} - \left(\dot{\hat{\boldsymbol{x}}}^T \cdot \ddot{\boldsymbol{J}}\frac{\partial \boldsymbol{x}}{\partial p_n}\right)\Bigg|_0^{T_{\max}}$$

$$+ \int_0^{T_{\max}} \left(\ddot{\hat{\boldsymbol{x}}}^T \ddot{\boldsymbol{J}} - \dot{\hat{\boldsymbol{x}}}^T \dot{\boldsymbol{J}} + \hat{\boldsymbol{x}}^T \boldsymbol{J}\right) \cdot \frac{\partial \boldsymbol{x}}{\partial p_n} dt$$

$$+ \int_0^{T_{\max}} \hat{\boldsymbol{x}}^T \cdot \frac{\partial \mathfrak{F}\left(\bar{\boldsymbol{x}}, \dot{\bar{\boldsymbol{x}}}, \ddot{\bar{\boldsymbol{x}}}\right)}{\partial p_n} dt = 0. \qquad (101)$$

We eliminate the first two terms by imposing zero terminal conditions on $\hat{\boldsymbol{x}}$ and $\dot{\hat{\boldsymbol{x}}}$. The third term is used to define the adjoint problem

$$\ddot{\boldsymbol{J}}^T \hat{\boldsymbol{x}} - \dot{\boldsymbol{J}}^T \dot{\hat{\boldsymbol{x}}} + \boldsymbol{J}^T \ddot{\hat{\boldsymbol{x}}} = [\nabla_{\boldsymbol{x}} f]^T. \tag{102}$$

Notice that the adjoint system (102) is linear since the matrices $\ddot{\boldsymbol{J}}^T$, $\dot{\boldsymbol{J}}^T$, and $\boldsymbol{J}^T$ are independent of the adjoint vector and its derivatives. We have already encountered a similar situation in the analysis of steady-state nonlinear problems [see (69)]. Thus, (101) becomes

$$\int_0^{T_{\max}} \left(\nabla_{\boldsymbol{x}} f \cdot \frac{\partial \boldsymbol{x}}{\partial p_n} \right) dt = -\int_0^{T_{\max}} \hat{\boldsymbol{x}}^T \cdot \frac{\partial \mathfrak{F} \left(\bar{\boldsymbol{x}}, \dot{\bar{\boldsymbol{x}}}, \ddot{\bar{\boldsymbol{x}}} \right)}{\partial p_n} dt. \tag{103}$$

Taking into account (82), we finally obtain the sensitivity expression

$$\frac{\partial F}{\partial p_n} = \frac{\partial^e F}{\partial p_n} - \int_0^{T_{\max}} \hat{\boldsymbol{x}}^T \cdot \frac{\partial \mathfrak{F} \left(\bar{\boldsymbol{x}}, \dot{\bar{\boldsymbol{x}}}, \ddot{\bar{\boldsymbol{x}}} \right)}{\partial p_n} dt \tag{104}$$

which is analogous to (70) and (91).

VI. TELLEGEN APPROACH TO EM SENSITIVITY ANALYSIS

It has been long ago noted that Tellegen's theorem can be extended to EM fields [47]. In fact, the reaction concept introduced by Rumsey [63] is very similar and can be considered as the analog of the Tellegen's theorem in electromagnetics. The derivation of the reaction equations follows the well-known procedure used to derive the expressions for the conservation of EM power (Poynting's theorem) [64]. Two sets of sources, $\mathbf{J}_e$, $\mathbf{J}_m$ and $\hat{\mathbf{J}}_e$, $\hat{\mathbf{J}}_m$, generate two fields

$$\nabla \times \mathbf{H} = j\omega \mathbf{D} + \mathbf{J}_e$$
$$-\nabla \times \mathbf{E} = j\omega \mathbf{B} + \mathbf{J}_m \tag{105}$$

and

$$\nabla \times \hat{\mathbf{H}} = j\omega \hat{\mathbf{D}} + \hat{\mathbf{J}}_e$$
$$-\nabla \times \hat{\mathbf{E}} = j\omega \hat{\mathbf{B}} + \hat{\mathbf{J}}_m. \tag{106}$$

Here, $\mathbf{J}_e$ and $\hat{\mathbf{J}}_e$ denote electric current densities of external sources, $\mathbf{J}_m$ and $\hat{\mathbf{J}}_m$ are the magnetic current densities of external sources, $\mathbf{E}$ and $\hat{\mathbf{E}}$ are electric field vectors, $\mathbf{H}$ and $\hat{\mathbf{H}}$ are magnetic field vectors, $\mathbf{D}$ and $\hat{\mathbf{D}}$ are electric flux density vectors, and $\mathbf{B}$ and $\hat{\mathbf{B}}$ are magnetic flux density vectors.

Take the dot product of $\hat{\mathbf{E}}$ with the first equation of (105) and the dot product of $\mathbf{H}$ with the second equation of (106). Summing up the resulting equations and integrating over the volume of interest yields

$$-\oiint_{S_\Omega} \hat{\mathbf{E}} \times \mathbf{H} \cdot d\mathbf{s} = \int\int_\Omega \int \left[(j\omega \mathbf{D} + \mathbf{J}_e) \cdot \hat{\mathbf{E}} \right.$$
$$\left. + (j\omega \hat{\mathbf{B}} + \hat{\mathbf{J}}_m) \cdot \mathbf{H} \right] d\Omega. \tag{107}$$

A similar manipulation of the other two equations in (105) and (106) gives

$$-\oiint_{S_\Omega} \mathbf{E} \times \hat{\mathbf{H}} \cdot d\mathbf{s} = \int\int_\Omega \int \left[(j\omega \hat{\mathbf{D}} + \hat{\mathbf{J}}_e) \cdot \mathbf{E} \right.$$
$$\left. + (j\omega \mathbf{B} + \mathbf{J}_m) \cdot \hat{\mathbf{H}} \right] d\Omega. \tag{108}$$

We can subtract (107) from (108) to obtain

$$-\oiint_{S_\Omega} \left(\mathbf{E} \times \hat{\mathbf{H}} - \hat{\mathbf{E}} \times \mathbf{H} \right) \cdot d\mathbf{s}$$
$$= \int\int_\Omega \int \left(\hat{\mathbf{J}}_e^t \cdot \mathbf{E} - \hat{\mathbf{J}}_m^t \cdot \mathbf{H} - \mathbf{J}_e^t \cdot \hat{\mathbf{E}} + \mathbf{J}_m^t \cdot \hat{\mathbf{H}} \right) d\Omega \tag{109}$$

where $\hat{\mathbf{J}}_e^t = j\omega \hat{\mathbf{D}} + \hat{\mathbf{J}}_e$, $\mathbf{J}_e^t = j\omega \mathbf{D} + \mathbf{J}_e$, $\hat{\mathbf{J}}_m^t = j\omega \hat{\mathbf{B}} + \hat{\mathbf{J}}_m$, and $\mathbf{J}_m^t = j\omega \mathbf{B} + \mathbf{J}_m$ are respective total current densities. The analogy between (109) and (2) should be noted. If the boundary conditions or material parameters are perturbed in the original problem described by (105), an equation for the incremental changes of the perturbed field $\Delta \mathbf{E}$ and $\Delta \mathbf{H}$ can be obtained as follows:

$$-\oiint_{S_\Omega} (\Delta \mathbf{E} \times \hat{\mathbf{H}} - \hat{\mathbf{E}} \times \Delta \mathbf{H}) \cdot d\mathbf{s}$$
$$= \int\int_\Omega \int \left\langle \hat{\mathbf{J}}_e^t \cdot \Delta \mathbf{E} - \hat{\mathbf{J}}_m^t \cdot \Delta \mathbf{H} - j\omega \Delta \mathbf{D} \cdot \hat{\mathbf{E}} + j\omega \Delta \mathbf{B} \cdot \hat{\mathbf{H}} \right\rangle d\Omega \tag{110}$$

which is analogous to (3). In (110), it is assumed that the external sources of the original problem are independent of the changes occurring in the boundary conditions or the constitutive parameters.

Using a formulation of Tellegen's theorem in electromagnetics similar to (110), Dyck and Lowther [65] have developed a sensitivity expression with regard to changes in the constitutive parameters and have applied it to an inverse problem—determining the shape of a defect in an aluminum block [65].

There is little doubt that the Tellegen approach to EM sensitivity analysis can be applied to more general problems also involving variations in the structure's geometry.

VII. CONCLUSION

We have reviewed a variety of techniques for exact and approximate sensitivity estimation with circuit- and EM-based analysis. We have shown that adjoint-network and adjoint-variable sensitivity analyses offer the best computational efficiency and can drastically reduce the computer time required by gradient-based optimization, tolerance, and yield analyses. We have seen very limited applications of adjoint-based techniques with full-wave EM analysis in high-frequency structure CAD. There is a need for feasible and robust algorithms that can be readily implemented in existing numerical analysis algorithms in the time and frequency domains. EM simulators

should no longer be viewed as "black boxes" linked to the optimizers through a set of response functions. With clever use of adjoint-sensitivity approaches, new features can be added to them, which would provide accurate and low-cost gradient information. Future developments in this area will greatly benefit both research and commercial applications for fully automated high-frequency structure design.

REFERENCES

[1] 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.

[2] ——, "Automated network design—The frequency-domain case," *IEEE Trans. Circuit Theory*, vol. CT-16, pp. 330–337, Aug. 1969.

[3] B. D. H. Tellegen, "A general network theorem, with applications," *Philips Res. Rep.*, vol. 7, pp. 259–269, 1959.

[4] P. Penfield, Jr., R. Spence, and S. Duinker, "A generalized form of Tellegen's theorem," *IEEE Trans. Circuit Theory*, vol. CT-17, pp. 302–305, Aug. 1970.

[5] F. H. Branin, Jr., "Network sensitivity and noise analysis simplified," *IEEE Trans. Circuit Theory*, vol. CT-20, pp. 285–288, May 1973.

[6] S. W. Director, "*LU* factorization in network sensitivity computations," *IEEE Trans. Circuit Theory*, vol. CT-18, pp. 184–185, Jan. 1971.

[7] V. A. Monaco and P. Tiberio, "A method for automatic scattering matrix computation of a lumped, linear network," *Alta Freq.*, vol. XXXVIII, pp. 906–912, Nov. 1969.

[8] ——, "On linear network scattering matrix sensitivity," *Alta Freq.*, vol. XXXIX, pp. 193–195, Feb. 1970.

[9] 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.

[10] G. Iuculano, V. A. Monaco, and P. Tiberio, "Network sensitivities in terms of scattering parameters," *Electron. Lett.*, vol. 7, pp. 54–55, Jan. 1971.

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

[12] 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.

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

[14] J. W. Bandler and S. H. Chen, "Circuit optimization: The state of the art," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 424–442, Feb. 1988.

[15] J. W. Bandler, Q.-J. Zhang, and R. M. Biernacki, "A unified theory for frequency-domain simulation and sensitivity analysis of linear and nonlinear circuits," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 1661–1669, Dec. 1988.

[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] 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.

[18] M. Mongiardo and R. Ravanelli, "Automated design of corrugated feeds by the adjoint network method," *IEEE Trans. Microwave Theory Tech.*, vol. 45, pp. 787–793, May 1997.

[19] S. Amari, "Sensitivity analysis of coupled resonator filters," *IEEE Trans. Circuits Syst. II*, vol. 47, pp. 1017–1022, Oct. 2000.

[20] Q.-J. Zhang, S. Lum, and M. S. Nakhla, "Minimization of delay and crosstalk in high-speed VLSI interconnects," *IEEE Trans. Microwave Theory Tech.*, vol. 40, pp. 1555–1563, July 1992.

[21] R. Liu, Q.-J. Zhang, and M. S. Nakhla, "A frequency-domain approach to performance optimization of high-speed VLSI interconnects," *IEEE Trans. Microwave Theory Tech.*, vol. 40, pp. 2403–2411, Dec. 1992.

[22] S. Lum, M. S. Nakhla, and Q.-J. Zhang, "Sensitivity analysis of lossy coupled transmission lines with nonlinear terminations," *IEEE Trans. Microwave Theory Tech.*, vol. 42, pp. 607–615, Apr. 1994.

[23] N. Liu, M. Nakhla, and Q.-J. Zhang, "Time domain sensitivity of high-speed VLSI interconnects," *Int. J. Circuit Theory Applicat.*, vol. 22, pp. 479–511, Nov. 1994.

[24] R. Khazaka, P. Gunupudi, and M. Nakhla, "Efficient sensitivity analysis of transmission-line networks using model reduction techniques," *IEEE Trans. Microwave Theory Tech.*, vol. 38, pp. 2345–2351, Dec. 2000.

[25] A. Dounavis, R. Achar, and M. Nakhla, "Efficient sensitivity analysis of lossy multiconductor transmission lines with nonlinear terminations," *IEEE Trans. Microwave Theory Tech.*, vol. 49, pp. 2292–2299, Oct. 2001.

[26] J. Xu, M. C. E. Yagoub, R. Ding, and Q.-J. Zhang, "Exact adjoint sensitivity analysis for neural-based microwave modeling and design," *IEEE Trans. Microwave Theory Tech.*, vol. 51, pp. 226–237, Jan. 2003.

[27] J. W. Bandler, S. H. Chen, S. Daijavad, and K. Madsen, "Efficient optimization with integrated gradient approximations," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 444–455, Feb. 1988.

[28] J. W. Bandler, R. M. Biernacki, S. H. Chen, L. W. Hendrick, and D. Omeragić, "Electromagnetic optimization of 3-D structures," *IEEE Trans. Microwave Theory Tech.*, vol. 45, pp. 770–779, May 1997.

[29] P. Garcia and J. P. Webb, "Optimization of planar devices by the finite element method," *IEEE Trans. Microwave Theory Tech.*, vol. 38, pp. 48–53, Jan. 1990.

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

[31] J. P. Webb, "Design sensitivities using high-order tetrahedral vector elements," *IEEE Trans. Magn.*, vol. 37, pp. 3600–3603, Sept. 2001.

[32] ——, "Design sensitivity of frequency response in 3-D finite-element analysis of microwave devices," *IEEE Trans. Magn.*, vol. 38, pp. 1109–1112, Mar. 2002.

[33] M. M. Gavrilovic and J. P. Webb, "Accuracy control in the optimization of microwave devices by finite-element methods," *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 1901–1911, Aug. 2002.

[34] ——, "An error estimator for design sensitivities of microwave device parameters," *Electromagnetics*, vol. 22, pp. 315–321, May–June 2002.

[35] 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.

[36] Y.-S. Chung, C. Cheon, I. Park, and S. Hahn, "Optimal shape design of microwave device using FDTD and design sensitivity analysis," *IEEE Trans. Microwave Theory Tech.*, vol. 48, pp. 2289–2296, Dec. 2000.

[37] Y. S. Chung, C. Cheon, I. H. Park, and S. Y. Hahn, "Optimal design method for microwave device using time domain method and design sensitivity analysis—Part I: FETD case," *IEEE Trans. Magn.*, vol. 37, pp. 3289–3293, Sept. 2001.

[38] ——, "Optimal design method for microwave device using time domain method and design sensitivity analysis—Part II: FDTD case," *IEEE Trans. Magn.*, vol. 37, pp. 3255–3259, Sept. 2001.

[39] N. K. Georgieva, S. Glavic, M. H. Bakr, and J. W. Bandler, "Feasible adjoint sensitivity technique for EM design optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 2751–2758, Dec. 2002.

[40] 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. Microwave Symp. Dig.*, vol. 3, 2001, pp. 1945–1948.

[41] M. H. Bakr and N. K. Nikolova, "An adjoint variable method for frequency domain TLM problems with conducting boundaries," *IEEE Microwave Wireless Comp. Lett.*, vol. 13, pp. 408–410, Sept. 2003.

[42] ——, "An adjoint variable method for time domain TLM with fixed structured grids," in *IEEE MTT-S Int. Microwave Symp. Dig.*, vol. 2, Philadelphia, PA, June 2003, pp. 1121–1124.

[43] G. Shen, H. W. W. Tam, N. K. Nikolova, and M. H. Bakr, "Adjoint sensitivity technique for FDTD methods on structured grids," in *Proc. IEEE/URSI Int. Antennas Propagation Symp.*, Columbus, OH, June 2003, pp. 746–749.

[44] 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.

[45] B. D. H. Tellegen, "A general network theorem, with applications," *Philips Res. Rep.*, vol. 7, pp. 259–269, Aug. 1952.

[46] ——, "A general network theorem, with applications," *Proc. Inst. Radio Eng. Australia*, vol. 14, pp. 265–270, Nov. 1953.

[47] P. Penfield Jr., R. Spence, and S. Duinker, *Tellegen's Theorem and Electrical Networks*. Cambridge, MA: MIT Press, 1970, p. 18.

[48] F. H. Branin, "Computer methods of network analysis," *Proc. IEEE*, vol. 55, pp. 1787–1801, Nov. 1967.

[49] E. J. Haug, K. K. Choi, and V. Komkov, *Design Sensitivity Analysis of Structural Systems*. Orlando, FL: Academic, 1986.

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

[51] P. Harscher, S. Amari, and R. Vahldieck, "A fast finite-element-based field optimizer using analytically calculated gradients," *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 433–439, Feb. 2002.

[52] M. D. Greenberg, *Advanced Engineering Mathematics*. Upper Saddle River, NJ: Prentice-Hall, 1998, p. 605.

[53] F. Alessandri, M. Mongiardo, and R. Sorrentino, "Transverse segmentation: A novel technique for the efficient CAD of $2N$-port branch-guide couplers," *IEEE Microwave Guided Wave Lett.*, vol. 1, pp. 204–207, Aug. 1991.

[54] ——, "A technique for the fullwave automatic synthesis of waveguide components: Application to fixed phase shifters," *IEEE Trans. Microwave Theory Tech.*, vol. 40, pp. 1484–1495, July 1992.

[55] A. Marrocco and O. Pironneau, "Optimum design with Lagrangian finite elements: Design of an electromagnet," *Comput. Methods Appl. Mech. Eng.*, vol. 15, pp. 277–308, 1978.

[56] J. L. Coulomb, G. Meunier, and J. C. Sabonnadiere, "Energy methods for the evaluation of global quantities and integral parameters in a finite elements analysis of electromagnetic devices," *IEEE Trans. Magn.*, vol. MAG-21, pp. 1817–1822, Sept. 1985.

[57] J. Ureel and D. De Zutter, "Shape sensitivities of capacitances of planar conducting surfaces using the method of moments," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 198–207, Feb. 1996.

[58] ——, "A new method for obtaining the shape sensitivities of planar microstrip structures by a full-wave analysis," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 249–260, Feb. 1996.

[59] R. Safian, N. K. Nikolova, M. H. Bakr, and J. W. Bandler, "Feasible adjoint sensitivity technique for EM design exploiting Broyden's update," in *IEEE MTT-S Int. Microwave Symp. Dig.*, Philadelphia, PA, June 2003, pp. 299–302.

[60] Y. S. Rickard, N. K. Georgieva, and H. W. Tam, "Absorbing boundary conditions for adjoint problems in the design sensitivity analysis with the FDTD method," *IEEE Trans. Microwave Theory Tech.*, vol. 51, pp. 526–529, Feb. 2003.

[61] R. Tomović, *Sensitivity Analysis of Dynamic Systems*. New York: McGraw-Hill, 1963.

[62] ——, *Introduction to Nonlinear Automatic Control Systems*. New York: Wiley, 1966.

[63] V. H. Rumsey, "Reaction concept in electromagnetic theory," *Phys. Rev.*, vol. 94, pp. 1483–1491, June 1954.

[64] R. F. Harrington, *Time–Harmonic Electromagnetic Fields*. New York: McGraw-Hill, 1961, p. 117.

[65] D. N. Dyck and D. A. Lowther, "A method of computing the sensitivity of electromagnetic quantities to changes in materials and sources," *IEEE Trans. Magn.*, vol. 30, pp. 3415–3418, Sept. 1994.

Natalia K. Nikolova (S'93–M'97) received the Ph.D. degree from the University of Electro-Communications, Tokyo, Japan, in 1997.

From 1998 to 1999, she was with the Natural Sciences and Engineering Research Council of Canada (NSERC), during which time she was initially with the Microwave and Electromagnetics Laboratory, DalTech, Dalhousie University, Halifax, NS, Canada. For a year, she was then with 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 currently an Assistant Professor. Her research interests include theoretical and computational electromagnetism, high-frequency analysis techniques, as well as CAD methods for high-frequency structures and antennas.

Dr. Nikolova was the recipient of an NSERC Post-Doctoral Fellowship from 1998 to 1999. She currently holds the 2000 NSERC University Faculty Award.

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 with 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 computer-aided engineering (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 President of Bandler Corporation, Dundas, ON, Canada, which he founded in 1997. He has authored or coauthored over 350 papers from 1965 to 2003. He contributed to *Modern Filter Theory and Design* (New York: Wiley-Interscience, 1973) and *Analog Methods for Computer-aided Analysis and Diagnosis* (New York: Marcel Dekker Inc., 1988). Four of his papers have been reprinted in *Computer-Aided Filter Design* (New York: IEEE Press, 1973), one in each of *Microwave Integrated Circuits* (Norwood, MA: Artech House, 1975), *Low-Noise Microwave Transistors and Amplifiers* (New York: IEEE Press, 1981), *Microwave Integrated Circuits, 2nd ed.*(Norwood, MA: Artech House, 1985), *Statistical Design of Integrated Circuits* (New York: IEEE Press, 1987), and *Analog Fault Diagnosis* (New York: IEEE Press, 1987). He joined the Editorial Boards of the *International Journal of Numerical Modeling* (1987), the *International Journal of Microwave and Millimeterwave Computer-Aided Engineering* (1989), and *Optimization Eng.* in 1998. He was Guest Editor of the *International Journal of Microwave and Millimeter-Wave Computer-Aided Engineering* Special Issue on Optimization-Oriented Microwave CAD (1997). He was Guest Co-Editor of the *Optimization Eng.* Special Issue on Surrogate Modelling and Space Mapping for Engineering Optimization (2001).

Dr. Bandler is a Fellow of the Canadian Academy of Engineering, the Royal Society of Canada, the Institution of Electrical Engineers (U.K.), and the Engineering Institute of Canada. He is 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 was an associate editor of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES (1969–1974), and has continued serving as a member of the Editorial Board. He was guest editor of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES Special Issue on Computer-Oriented Microwave Practices (1974) and guest co-editor of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES Special Issue on Process-Oriented Microwave CAD and Modeling (1992). He was guest editor of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES Special Issue on Automated Circuit Design Using Electromagnetic Simulators (1997). He is guest co-editor of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES Special Issue on Electromagnetics-Based Optimization of Microwave Components and Circuits (2004). He has served as chair of the MTT-1 Technical Committee on Computer-Aided Design. He was the recipient of the 1994 Automatic Radio Frequency Techniques Group (ARFTG) Automated Measurements Career Award.

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

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