INCREMENTAL CIRCUIT SENSITIVITY COMPUTATION

Chandu Visweswariah

April 26, 2002

ONTWERPTECHNOLOGIE 5L050 HANDOUT 4
Class 4, lecture 1: Sensitivity analysis by the direct method

1 Motivation

The circuit analysis techniques we have learned so far can be used to create a circuit simulation software package, which is an invaluable design aid. For a given circuit topology, we can change component values such as resistance or capacitance values or transistor sizes, and re-simulate the circuit to judge the effect of the change on the behavior of the circuit. But if the circuit simulator could predict the change of circuit behavior with respect to changes in design parameters, this information could be tremendously valuable during design. Such an analysis is called a sensitivity analysis. Of course, sensitivities or gradients can be obtained by a *finite difference* procedure whereby the parameter of interest is changed by a small amount and the circuit re-simulated. But finite difference methods suffer from some serious problems, and we will see more about that later. This lecture will describe *incremental* sensitivity computation methods that overcome the problems of finite difference methods. It is assumed that the perturbations of interest are infinitesimally small and do not change the topology of the circuit.

Large-change sensitivity analyses and changes that involve topology modification have also been studied (see [1] for a discussion), but are not topics of this set of lectures.

2 Introduction

The sensitivity or gradient of a function ("behavior," "response" or "measurement" of a circuit) f with respect to a parameter p is the partial derivative of f with respect to p, measured at some nominal value of the parameter p (see Fig. 1). It is written as $\partial f / \partial p$ or $\nabla_p f$. For a general nonlinear function f, the gradient is valid in a small region around the nominal value of the parameter, and hence is referred to as *small-change sensitivity.*

3 The importance of sensitivities

Sensitivity analysis is a crucial part of tolerance analysis, design centering, circuit optimization, reliability analysis, periodic steady-state analysis, and so on. Variations are an inevitable part of manufacturing.

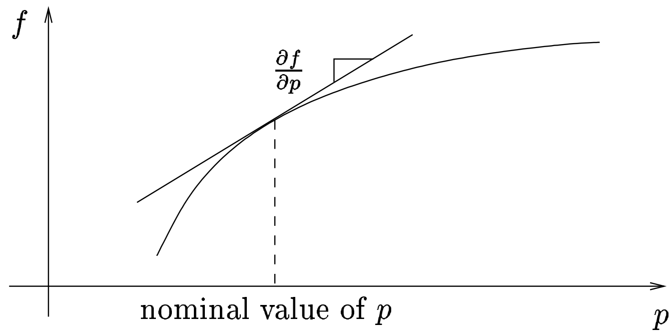

Figure 1: The gradient is valid in a small region around the nominal value of the parameter p.

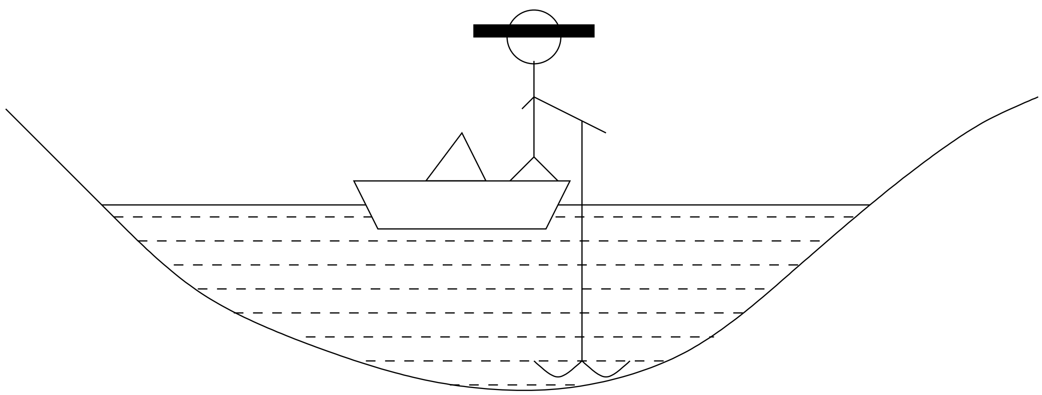

Figure 2: The importance of sensitivity information in optimization.

Operating conditions such as temperature and power supply voltage can vary, too. Sensitivities help in understanding the impact of these uncertainties on circuit performance. Sensitivities are also useful in optimizing circuit parameters such as component values or transistor sizes. It is well known that solving general nonlinear optimization problems with more than a few 10s of variables is almost impossible without sensitivity information. If we desire to robustly and repeatably solve problems in large dimensional spaces, efficient and accurate sensitivity analysis is a pre-requisite.

It is instructive to consider the analogy of a blindfolded person searching for the deepest point of a lake by taking depth measurements with an anchor (see Fig. 2). Nonlinear optimization algorithms typically only have local function information (hence the blindfolds) and history information from previous measurements. The person would have to repeatedly take measurements in different directions to find the direction in which the bed of the lake is sloping and sail in that direction in an attempt to find the deepest point. On the other hand, if the anchor could automatically sense the direction in which the bed was sloping, finding the deepest point would be much easier. The difference made by having sensitivities is clear in this simple two-dimensional problem; the difference is much more dramatic in problems with higher dimensionality. In practice, we wish to solve problems in thousands and even tens of thousands of variables. Therefore sensitivity information is invaluable in the context of circuit optimization.

4 Incremental sensitivity analysis

Let us say we wish to determine the sensitivity of a delay d with respect to a transistor width w. One way to do this would be rerun our simulator with a perturbation of Δw on the transistor width, and use a finite-difference formula to compute the sensitivity.

$$\frac{\partial d}{\partial w} \approx \frac{d(w+\Delta w) - d(w)}{\Delta w} \tag{1}$$

One advantage of this method is that the sensitivity of several functions with respect to a single parameter can be obtained at once. A second advantage is that the simulator can be treated as a black box. However, it suffers from two main disadvantages.

1. The procedure incurs a 100% computational overhead for each parameter of interest.
2. The choice of Δw is fraught with danger. If Δw is too small, we may be trapped in the inherent numerical noisiness of the simulator and obtain an inaccurate gradient, especially because of finite-precision arithmetic in computers. If Δw is too large, we obtain a secant to the curve of d versus w. In particular, if $d(w)$ is approaching a stationary point, the choice of a positive or negative Δw may lead to quite different results. To address this problem, the transistor width can be perturbed in both directions and the average of the two finite-difference sensitivities used. However, the computational overhead would then be 200%!

If the motivation is to solve optimization problems of high dimensionality, it becomes rapidly apparent that we need more efficient methods than finite-difference sensitivity computation. This lecture focuses on incremental sensitivity methods wherein we take advantage of our knowledge of the solution of the nominal (unperturbed) circuit and the fact that we are in a circuit simulation environment. The goal, in other words, is to compute sensitivities with an overhead that is significantly lower than the 100% overhead incurred by finite difference, and to devise methods that scale well to large numbers of functions and parameters.

The two well-known methods for incremental sensitivity analysis are the direct [2] and adjoint [3] methods.

5 The direct method: DC analysis

The direct method of sensitivity analysis is based on direct differentiation of the circuit equations, including the branch constitutive relations (BCRs) of individual circuit elements. First, we will consider the DC solution of circuits consisting of only resistors and voltage sources.

5.1 KCL and KVL

We begin by differentiating KCL and KVL with respect to some parameter p, which could be temperature, or the width of a transistor or wire, or resistance or capacitance value. If KCL is expressed as

$$A i_b = 0 \tag{2}$$

where A is the reduced incidence matrix and i_b the vector of branch currents, then we obtain

$$A \frac{\partial i_b}{\partial p} = 0. \tag{3}$$

Let us represent the unknown voltage and current sensitivities with the notation

$$\frac{\partial i_b}{\partial p} = \hat{i}_b, \tag{4}$$

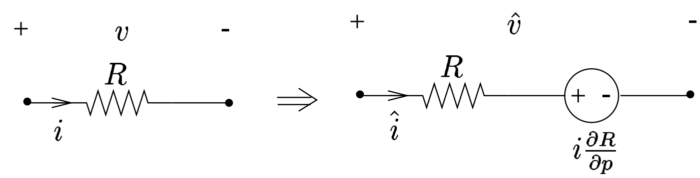

Figure 3: A resistor in the direct method.

$$\frac{\partial v_b}{\partial p} = \hat{v}_b, \text{ and} \tag{5}$$

$$\frac{\partial v_n}{\partial p} = \hat{v}_n \tag{6}$$

where v_b and v_n are the branch voltage and node voltage vectors, respectively.

If KVL is expressed as

$$A^T v_n = v_b \tag{7}$$

then

$$A^T \frac{\partial v_n}{\partial p} = \frac{\partial v_b}{\partial p}, \tag{8}$$

or

$$A^T \hat{v}_n = \hat{v}_b. \tag{9}$$

As we shall see later, it is most convenient to compute the required sensitivities as the solution of a circuit whose branch voltages, branch currents and node voltages are $\hat{v}_b$, $\hat{i}_b$ and $\hat{v}_n$, respectively. It is clear then that no matter how KCL and KVL are expressed for the nominal (original) circuit, they are also satisfied for the new circuit.

5.2 Resistor

Consider a resistor R (see Fig. 3). The BCR for the resistor (Ohm's Law) is

$$v = iR. \tag{10}$$

The sensitivity parameter of interest p could be the value of the resistance itself. Differentiating,

$$\frac{\partial v}{\partial p} = \hat{v} = i\frac{\partial R}{\partial p} + R\frac{\partial i}{\partial p} = i\frac{\partial R}{\partial p} + R\hat{i}. \tag{11}$$

In the above equation, $\frac{\partial R}{\partial p}$ is known and is equal to 1.0 if the parameter of interest is the resistance value R and zero if R does not depend on p. In general, it is obtained by differentiating the model $R(p)$ with respect to p. The nominal current i is known after simulating the original circuit. Thus, he first term of the right-hand side of (11) is known and can be evaluated as a constant. Since the sensitivity analysis is performed in the context of a circuit simulator, it makes eminent sense to cast the differentiated relation in the form of a circuit equation. Interpreting (11) as the Branch Constitutive Relation of a new circuit element, it can be seen to represent a Thevenin equivalent (Fig. 3). Note the orientation of the voltage source: the voltage drop across the composite element is the sum of the drop across the resistor and the independent voltage source. Of course, the Thevenin equivalent can be converted to an electrically identical Norton equivalent.

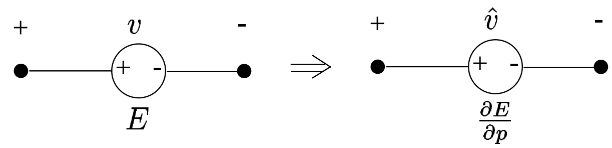

Figure 4: An independent voltage source in the direct method.

5.3 Independent voltage source

The analogous equations for an independent voltage source are simple.

$$v = E \tag{12}$$
$$\frac{\partial v}{\partial p} = \hat{v} = \frac{\partial E}{\partial p}, \tag{13}$$

as depicted in Fig. 4. Similarly, if we can successfully replace every element of the nominal circuit with a new element that reflects the differentiation of its Branch Constitutive Relation, we will obtain a new circuit called the *sensitivity circuit.* The unknowns of the sensitivity circuit are the $\hat{v}$ and $\hat{i}$ variables, which are the required sensitivities. The sensitivity circuit satisfies KCL and KVL. Hence the goal is to create a new circuit, the solution of which yields all the required sensitivities.

5.4 The direct method: a simple example

Consider the circuit on the left of Fig. 5. The solution of the circuit is

$$i = \frac{E}{R_1 + R_2} \tag{14}$$
$$v_{R1} = \frac{ER_1}{R_1 + R_2} \tag{15}$$
$$v_{R2} = \frac{ER_2}{R_1 + R_2}. \tag{16}$$

Let us assume the parameter of interest is R_1. We can write down the sensitivities by inspection to be

$$\frac{\partial i}{\partial R_1} = \hat{i} = \frac{-E}{(R_1+R_2)^2} \tag{17}$$
$$\frac{\partial v_{R1}}{\partial R_1} = \hat{v}_{R1} = \frac{(R_1+R_2)E - ER_1}{(R_1+R_2)^2} = \frac{ER_2}{(R_1+R_2)^2} \tag{18}$$
$$\frac{\partial v_{R2}}{\partial R_1} = \hat{v}_{R2} = \frac{-ER_2}{(R_1+R_2)^2}. \tag{19}$$

The sensitivity of the current i to R_1 is, as expected, negative. That is, the more the resistance, the less the current. The larger the resistance R_1, the more of E falls across it, so v_{R1} has positive sensitivity to R_1. Likewise, v_{R2} is the remaining voltage, and so it has a negative sensitivity to R_1. In this simple but illustrative case, the sensitivities could be derived by analytically differentiating the circuit solution. Of course, such a simple method cannot be applied to larger and more complex circuits.

Now let us attempt to derive the same result by the direct method. Replacing each element by its sensitivity equivalent, we obtain the sensitivity circuit shown on the right side of Fig. 5. Solving this circuit, we obtain

$$\hat{i} = \frac{-E}{(R_1+R_2)^2} \tag{20}$$
$$\hat{v}_{R1} = \hat{i}_{R1} + \frac{E}{R_1+R_2} = \frac{-ER_1}{(R_1+R_2)^2} + \frac{E}{R_1+R_2} = \frac{ER_2}{(R_1+R_2)^2} \tag{21}$$
$$\hat{v}_{R2} = \frac{-ER_2}{(R_1+R_2)^2}, \tag{22}$$

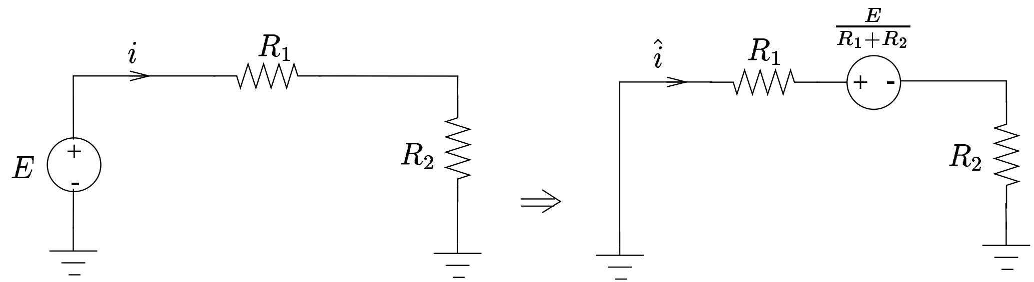

Figure 5: A simple example of the direct method.

the same results as before. Thus, we obtained the desired sensitivities as the solution of an associated sensitivity circuit.

Of course, not all circuits consist of just resistors and independent voltage sources. Before we extend the method to deal with other element types, let us consider how we could compute sensitivities efficiently in the direct method.

5.5 Direct method: the key idea

So far, the nominal and sensitivity circuits differ only in the position and value of the independent sources. Circuit equations are typically formulated as

$$Yv = b \tag{23}$$

where Y is the nodal admittance matrix, v the vector of unknowns (typically the vector of node voltages augmented by the vector of voltage source currents) and b the vector of external stimuli. The sensitivity circuit equations, on the other hand, are of the form

$$Y\hat{v} = \hat{b} \tag{24}$$

with the *same* admittance matrix. The differences in the independent sources are reflected on the right hand side, in $\hat{b}$. Thus, if the nominal circuit were solved by obtaining the LU factors of Y, the factors can be re-used to solve the sensitivity circuit. Therein lies the incremental nature of the direct method of sensitivity computation!

A practical procedure to obtain sensitivities by the direct method consists of the following steps.

1. Since resistors that can vary (i.e., resistors whose value depends on any of the parameters of interest) will have voltage sources in series in the sensitivity circuit, introduce zero-valued voltage sources in these positions *before* solving the nominal circuit. Alternatively, use the Norton equivalent model of the sensitivity BCR for the resistor which adds current sources but no additional nodes to the circuit. Again, adding zero-valued sources as place-holders in the nominal circuit will simplify implementation of the sensitivity computation.
2. Solve the nominal circuit and save the LU factors.
3. Replace all elements by new values as dictated by differentiating the nominal BCRs with respect to the parameter of interest; recalculate the right hand vector $\hat{b}$ based on the values of these sources.
4. Solve the sensitivity circuit by re-using the LU factors. Only forward- and back-substitution are necessary.

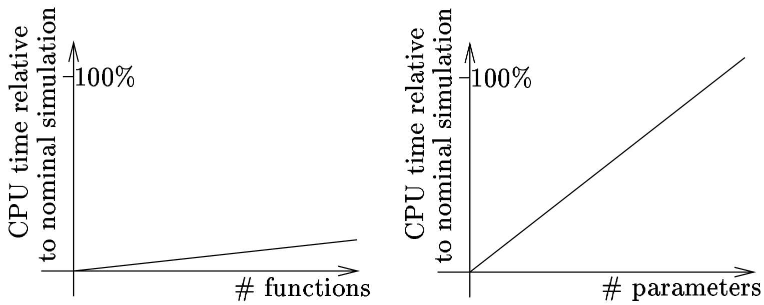

Figure 6: Computational cost of the direct method.

5. Pick up the gradients of *all* measurements (currents and voltages) with respect to the parameter of interest.
6. Repeat steps 3, 4 and 5 for each parameter.

The gradients of any number of functions with respect to one parameter can be obtained by a simple forward- and back-substitution step. Therefore, gradients can be obtained at a small overhead over and above the nominal simulation. Additional functions are accommodated inexpensively, since the gradients of all functions are obtained simultaneously. Additional parameters involve repetition of steps 3, 4 and 5 above. Each additional parameter costs a small overhead over the nominal simulation, but a large number of parameters could cause significant overhead. Nonetheless, the required sensitivities can be computed orders of magnitude faster than using finite-difference methods!

Fig. 6 depicts the dependence of the computational cost of the direct method of sensitivity analysis on the number of functions (measurements) of interest and the number of parameters of interest. Obviously, the gradients of any number of functions with respect to one parameter are obtained with a single forward- and back-substitution. Additional parameters required a new forward- and back-solve, and so the computational cost of accommodating an additional parameter is much more than an additional function, but still a small overhead over the nominal circuit analysis.

5.6 Other elements

Now we extend the analysis of the previous sections to other element types.

5.6.1 Conductances

Fig. 7 shows a conductance G and its sensitivity circuit counterpart. Differentiating the BCR

$$i = Gv, \tag{25}$$

we obtain

$$\frac{\partial i}{\partial p} = G\frac{\partial v}{\partial p} + v\frac{\partial G}{\partial p} \tag{26}$$

$$\hat{i} = G\hat{v} + v\frac{\partial G}{\partial p}. \tag{27}$$

Thus, a conductance is represented by an identical conductance in the sensitivity circuit, with an additional current source in parallel if the conductance value depends on the parameter p of interest.

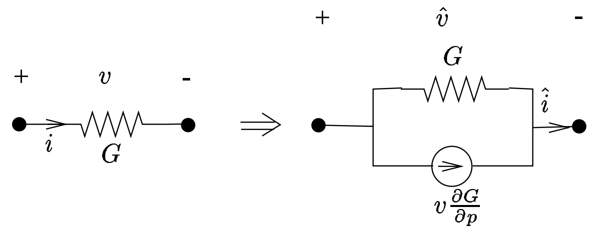

Figure 7: A conductance in the direct method.

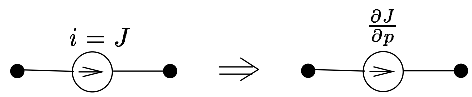

Figure 8: An independent current source in the direct method.

5.6.2 Independent current sources

Fig. 8 shows an independent current source J and it sensitivity circuit counterpart. Differentiating the BCR

$$i = J \tag{28}$$

we obtain

$$\frac{\partial i}{\partial p} = \frac{\partial J}{\partial p} \tag{29}$$

$$\hat{i} = \frac{\partial J}{\partial p}. \tag{30}$$

Thus, an independent current source is represented by an independent current source in the sensitivity circuit. Of course, if J does not depend on the parameter of interest p, then the current source is zero-valued (i.e., the element is an open circuit) in the sensitivity circuit.

5.6.3 Linear current-controlled current source (CCCS)

Fig 9 shows a linear current-controlled current source (CCCS). The controlling branch of each current-controlled source is modeled as a zero-valued voltage source ("ideal ammeter") in series with the controlling element, so as to simplify the derivation. Likewise, the controlling branch of each voltage-controlled source is modeled as a zero-valued current source ("ideal voltmeter") in parallel with the controlling element. The equations for the two branches of a CCCS in the sensitivity circuit are

$$\frac{\partial v_1}{\partial p} = 0 \tag{31}$$

$$\hat{v}_1 = 0 \tag{32}$$

and

$$\frac{\partial i_2}{\partial p} = h\frac{\partial i_1}{\partial p} + i_1\frac{\partial h}{\partial p} \tag{33}$$

$$\hat{i}_2 = h\hat{i}_1 + i_1\frac{\partial h}{\partial p}. \tag{34}$$

Thus, a CCCS is represented by an identical CCCS in the sensitivity circuit, but with an additional constant current of $i_1\frac{\partial h}{\partial p}$ in the controlled branch. The contribution of the controlled source to the left-hand-side of the (modified nodal, for example) equations is thus unchanged in the sensitivity circuit.

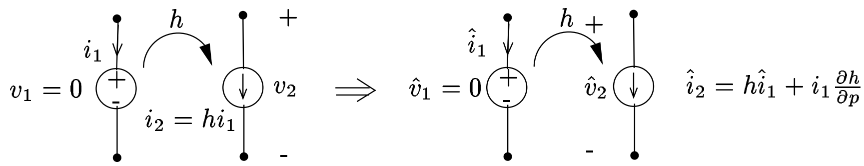

Figure 9: A linear current-controlled current source in the direct method.

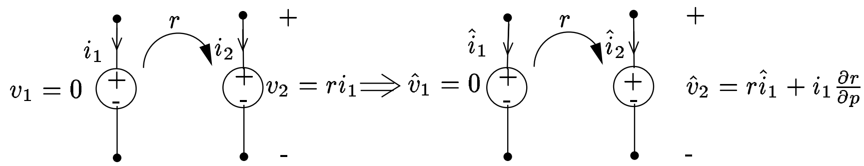

Figure 10: A linear current-controlled voltage source in the direct method.

5.6.4 Linear current-controlled voltage source (CCVS)

Fig 10 shows a linear current-controlled voltage source (CCVS) with a canonical zero-valued voltage source ("ideal ammeter") as the controlling element. Differentiating the BCRs with respect to the parameter of interest, we obtain the equations for the two branches of a CCVS in the sensitivity circuit.

$$\frac{\partial v_1}{\partial p} = 0 \tag{35}$$

$$\hat{v}_1 = 0 \tag{36}$$

and

$$\frac{\partial v_2}{\partial p} = r\frac{\partial i_1}{\partial p} + i_1\frac{\partial r}{\partial p} \tag{37}$$

$$\hat{v}_2 = r\hat{i}_1 + i_1\frac{\partial r}{\partial p}. \tag{38}$$

Thus, a CCVS is represented by an identical CCVS in the sensitivity circuit, but with an additional constant voltage of $i_1 \frac{\partial r}{\partial p}$ across the controlled branch. The contribution of the controlled source to the left-hand-side of the (modified nodal, for example) equations is thus unchanged in the sensitivity circuit.

5.6.5 Linear voltage-controlled voltage source (VCVS)

Fig 11 shows a linear voltage-controlled voltage source (VCVS) with a canonical zero-valued current source ("ideal voltmeter") as the controlling element. Differentiating the BCRs with respect to the parameter of interest, we obtain the equations for the two branches of a VCVS in the sensitivity circuit.

$$\frac{\partial i_1}{\partial p} = 0 \tag{39}$$

$$\hat{i}_1 = 0 \tag{40}$$

and

$$\frac{\partial v_2}{\partial p} = \mu\frac{\partial v_1}{\partial p} + v_1\frac{\partial \mu}{\partial p} \tag{41}$$

$$\hat{v}_2 = \mu\hat{v}_1 + v_1\frac{\partial \mu}{\partial p}. \tag{42}$$

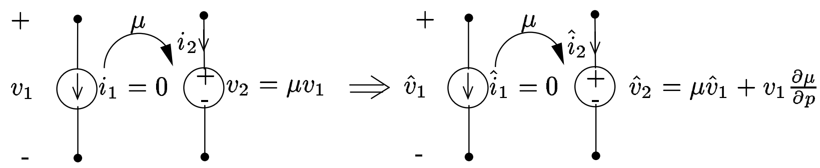

Figure 11: A linear voltage-controlled voltage source in the direct method.

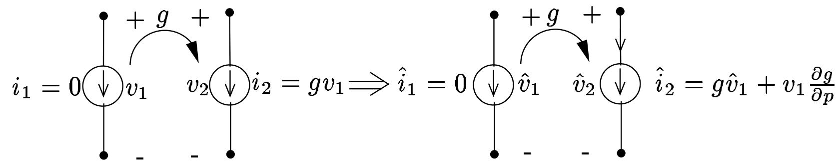

Figure 12: A linear voltage-controlled current source in the direct method.

Thus, a VCVS is represented by an identical VCVS in the sensitivity circuit, but with an additional constant voltage of $v_1 \frac{\partial \mu}{\partial p}$ across the controlled branch. The contribution of the controlled source to the left-hand-side of the (modified nodal, for example) equations is thus unchanged in the sensitivity circuit.

5.6.6 Linear voltage-controlled current source (VCCS)

Fig 12 shows a linear voltage-controlled current source (VCCS) with a canonical zero-valued current source ("ideal voltmeter") as the controlling element. Differentiating the BCRs with respect to the parameter of interest, we obtain the equations for the two branches of a VCCS in the sensitivity circuit.

$$\frac{\partial i_1}{\partial p} = 0 \quad (43)$$

$$\hat{i}_1 = 0 \quad (44)$$

and

$$\frac{\partial i_2}{\partial p} = g \frac{\partial v_1}{\partial p} + v_1 \frac{\partial g}{\partial p} \quad (45)$$

$$\hat{i}_2 = g \hat{v}_1 + v_1 \frac{\partial g}{\partial p}. \quad (46)$$

Thus, a VCCS is represented by an identical VCCS in the sensitivity circuit, but with an additional constant current of $v_1 \frac{\partial g}{\partial p}$ in the controlled branch. The contribution of the controlled source to the left-hand-side of the (modified nodal, for example) equations is thus unchanged in the sensitivity circuit.

In conclusion, resistances, conductances, and all types of independent and linear dependent sources can be accommodated by the direct method. In all cases, the left-hand-side of the circuit equations of the nominal and sensitivity circuits are identical, and hence only one LU factorization is required per unique parameter.

5.7 Nonlinear circuits

So far, we have applied the direct method of sensitivity computation to linear circuits that contain no energy-storage elements (such as capacitances and inductances). In the nonlinear case, the solution of the nominal circuit yields a linearization of each nonlinear element that is valid at the solution of the nominal circuit.

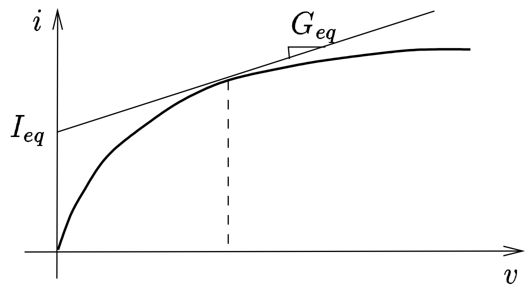

Figure 13: Linearization of a nonlinear element at the solution.

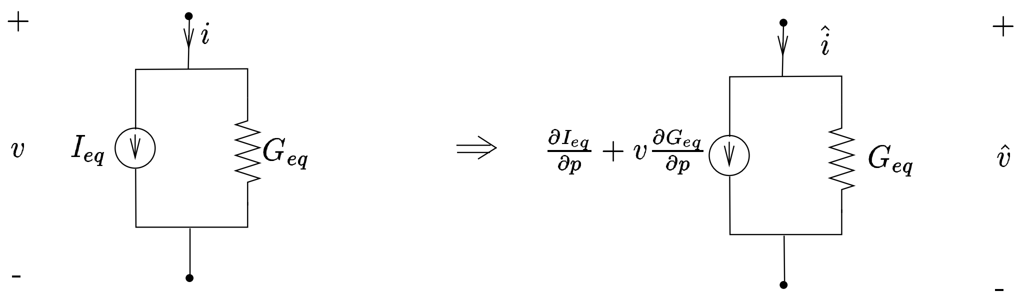

Figure 14: The linear equivalent of a nonlinear element in the nominal circuit and the sensitivity circuit in the direct method.

Since we are interested in first-order sensitivity, we can assume that the linearization of the nominal circuit at the solution does not change due to a small perturbation in each parameter of interest. More precisely, the effect of the change in the measurements due to changes in the linearization of nonlinear elements, in turn due to small perturbations of the parameters of interest are second order and can therefore be neglected for our purposes.

Thus, all we are required to compute is the sensitivity of the linear circuit obtained by linearizing the nominal circuit about its nominal solution. If Newton's method (also called "Newton-Raphson") was used to solve the nominal circuit, this simply means taking the linearized circuit that was solved in the last Newton iteration and finding the sensitivity of that linear circuit.

The current-voltage characteristics of a two-terminal element and its linearization are shown in Fig. 13. In the nominal circuit, the nonlinear element is represented by a Norton equivalent shown on the left of Fig. 14. The sensitivity circuit equivalent is derived by differentiating the BCR of this generic linearized companion model.

$$\frac{\partial i}{\partial p} = \frac{\partial I_{eq}}{\partial p} + v\frac{\partial G_{eq}}{\partial p} + G_{eq}\frac{\partial v}{\partial p} \tag{47}$$

$$\hat{i} = \frac{\partial I_{eq}}{\partial p} + v\frac{\partial G_{eq}}{\partial p} + G_{eq}\hat{v}. \tag{48}$$

Again, we observe that the only change in circuit equations compared to the nominal circuit would occur in the right-hand-side, allowing the LU factors of the nominal circuit to be re-used.

In general, for a nonlinear element,

$$i = f(v,p) \tag{49}$$

which leads to

$$\frac{\partial i}{\partial p} = \frac{\partial f}{\partial v}\frac{\partial v}{\partial p} + \frac{df}{dp} \tag{50}$$

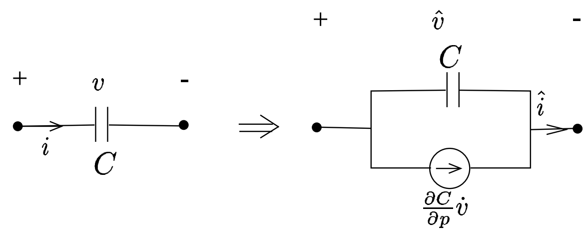

Figure 15: Linear capacitance in the direct method.

$$\hat{i} = G_{eq}\hat{v} + \hat{I}_{eq}, \tag{51}$$

and thus the nonlinear element can be represented by a linearized Norton equivalent in the sensitivity circuit, with the same conductance and therefore the same contribution to the circuit equations as the nominal circuit linearized about its solution.

As a practical matter, nonlinear circuits work exactly the same as linear circuits for the purpose of sensitivity analysis since we find the sensitivity of the linearized companion network obtained at the solution of the nominal circuit.

5.8 Transient case

Next, we will apply the direct method to the case of time-domain or transient simulation of a circuit that contains energy-storage elements such as capacitances and inductances.

5.8.1 Linear capacitances

Consider a linear capacitance C as shown in Fig. 15. Differentiating the BCR for the capacitance, we obtain

$$i = C\dot{v} \tag{52}$$

$$\frac{\partial i}{\partial p} = \frac{\partial C}{\partial p}\dot{v} + C\frac{\partial \dot{v}}{\partial p} \tag{53}$$

$$\hat{i} = \frac{\partial C}{\partial p}\dot{v} + C\dot{\hat{v}}. \tag{54}$$

Thus, the corresponding element in the sensitivity circuit is an identical capacitance in parallel with an independent current source as shown in Fig. 15. So the capacitance can be handled by the same integration methods that are used for the nominal circuit, and, provided the same time steps are chosen for the sensitivity circuit, will have the same matrix “stamps” or “stencils” as the nominal circuit in the left-hand-side of the circuit equations. Thus, the LU factorization used to solve the nominal circuit at any particular time step can be re-used to solve the sensitivity circuit, with just a different right-hand side. Further, nonlinear capacitances are not treated any differently, since their linearization at the solution of the nominal circuit is what matters for the purposes of sensitivity computation.

5.8.2 Linear inductances

Consider a linear inductance C as shown in Fig. 16. Differentiating the BCR for the inductance, we obtain

$$v = L\dot{i} \tag{55}$$

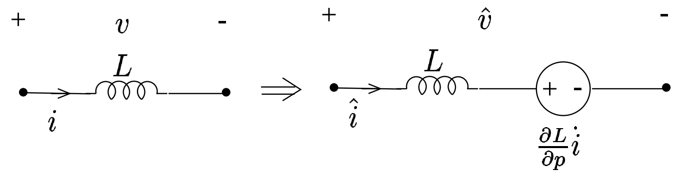

Figure 16: Linear inductance in the direct method.

$$\frac{\partial v}{\partial p} = \frac{\partial L}{\partial p}\dot{i} + L\frac{\partial \dot{i}}{\partial p} \tag{56}$$

$$\hat{v} = \frac{\partial L}{\partial p}\dot{i} + L\hat{\dot{i}}. \tag{57}$$

Thus, the corresponding element in the sensitivity circuit is an identical inductance in series with an independent voltage source as shown in Fig. 16. So the inductance can be handled by the same integration methods that are used for the nominal circuit, and, provided the same time steps are chosen for the sensitivity circuit, will have the same matrix "stamps" or "stencils" as the nominal circuit in the left-hand-side of the circuit equations. Thus, the LU factorization used to solve the nominal circuit at any particular time step can be re-used to solve the sensitivity circuit, with just a different right-hand side. Further, nonlinear inductances are not treated any differently, since their linearization at the solution of the nominal circuit is what matters for the purposes of sensitivity computation.

5.9 Dependent parameters

It is often the case that two or more parameters are related to each other. For example, two transistors may be constrained to have the same width or may be constrained to have widths that conform to a certain ratio. For example, suppose transistor t_1 has a width W_1 and transistor t_2 has a width W_2, but $W_2 = kW_1$. In this case, we can define an independent parameter p such that $W_1 = p$ and $W_2 = kp$ and then find the sensitivity of the circuit with respect to p in one sensitivity analysis. In other words, we will be driving the sensitivity circuit with two simultaneous stimuli, one corresponding to the stimulus we would use if W_1 were the only parameter of interest, and the other corresponding to k times the stimulus to be used if W_2 were the only parameter of interest. We are therefore exploiting superposition in the linear time-varying sensitivity circuit to compute the sensitivity with respect to p in a single sensitivity analysis.

5.10 Direct method conclusions

The direct method is an efficient method to compute the gradients of a circuit, provided there are not too many parameters of interest. In the linear DC case, the left-hand-side of the sensitivity circuit equations are identical to the nominal circuit, so the LU factors can simply be re-used with a different right-hand-side. Nonlinear circuits do not add any complications, since the linearization at the solution of the nominal circuit (and the LU factors thereof) are re-used for the sensitivity circuit. In the transient analysis case, provided the same time steps are chosen for the nominal and sensitivity circuit, the LU factors of the last Newton iteration at each time step can be re-used to solve the sensitivity circuit. The sensitivity circuit is a linear time-varying circuit. The one practical difficulty involved is to make sure that the choice of time step yields an acceptable local truncation error (LTE) for both the nominal and sensitivity circuits. As soon as the LU factors at a given time step are re-used for the sensitivity circuit, they can be deleted (or that section of memory re-used). In other words, the solution of the nominal circuit and as many sensitivity circuits as there are independent parameters can be conducted in lock step before proceeding to the next time point, to avoid expensive storage of LU factors.

The main disadvantage of the direct method is that it can be inefficient when the number of independent parameters is large. This problem is addressed by the adjoint method (but the adjoint method has its own disadvantages, as we will see in the subsequent section).

ONTWERPTECHNOLOGIE 5L050 HANDOUT 4
Class 4, lecture 2: Sensitivity analysis by the adjoint method

6 The adjoint method

6.1 Tellegen's theorem as a backdrop to adjoint sensitivity

We begin by reviewing Tellegen's theorem [4]. Given two networks N and $\hat{N}$ with the same topology (i.e., the same reduced incidence matrix A), Tellegen's theorem states that

$$\sum_{\text{all branches}} \hat{i}_{b_k} v_{b_k} = \hat{i}_b^T v_b = \hat{i}_b^T A^T v_n = [A\hat{i}_b]^T v_n = 0 \tag{58}$$

and

$$\sum_{\text{all branches}} \hat{v}_{b_k} i_{b_k} = \hat{v}_b^T i_b = [A^T \hat{v}_n]^T i_b = \hat{v}_n^T A i_b = 0, \tag{59}$$

where i_b, v_b and v_n are the vectors of branch currents, branch voltages and node voltages of the first network, and $\hat{i}_b$, $\hat{v}_b$ and $\hat{v}_n$ are the corresponding vectors of the second network. Stated in words, the pairwise products of branch voltages of the first network and corresponding branch currents of the second network (and vice versa) sum up to zero. We will use this remarkable theorem to derive sensitivities by the adjoint method.

6.2 Basic derivation

Consider a nominal circuit with branch voltages v_b and branch currents i_b. Now simultaneously perturb all the parameters $p_1, p_2, \ldots, p_n$ of interest by small amounts $\delta p_1, \delta p_2, \ldots, \delta p_n$ to obtain a new circuit as shown in Fig. 17. Assume that the branch voltage and branch current vectors for the perturbed circuit are $(v_b + \delta v_b)$ and $(i_b + \delta i_b)$, respectively. Also consider a new circuit which we will call the *adjoint* circuit, with the same topology as the nominal circuit but possibly different BCRs. Assume that the branch voltage and current vectors of the adjoint circuit are $\hat{v}_b$ and $\hat{i}_b$, respectively.

Writing Tellegen's theorem for the nominal and adjoint networks, we obtain:

$$\sum_{\text{all branches}} \hat{i}_b v_b = \sum_{\text{all branches}} i_b \hat{v}_b = 0. \tag{60}$$

Writing Tellegen's theorem for the perturbed and adjoint networks, we obtain:

$$\sum_{\text{all branches}} \hat{i}_b (v_b + \delta v_b) = \sum_{\text{all branches}} (i_b + \delta i_b) \hat{v}_b = 0. \tag{61}$$

Subtracting (61) from (60) we obtain

$$\sum_{\text{all branches}} (\delta v_b \hat{i}_b - \delta i_b \hat{v}_b) = 0. \tag{62}$$

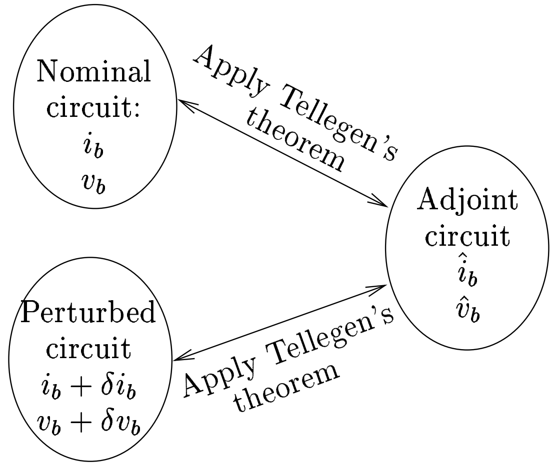

Figure 17: Basic derivation of the adjoint method.

The term $(\delta v_b \hat{i}_b - \delta i_b \hat{v}_b)$ is called the *typical term* and the summation of the typical term over all branches amounts to zero.

Before going further, let us motivate the above manipulation. The formulas were derived with no assumption other than that the topology of the nominal and adjoint circuits are identical. The actual electrical characteristics (or BCRs) of the individual elements can be chosen arbitrarily. Now suppose we have a single function or measurement f of interest

$$f = f(p_1, p_2, \ldots, p_n) \tag{63}$$

then to first order the perturbation in the function can be written as

$$\delta f \approx \frac{\partial f}{\partial p_1}\delta p_1 + \frac{\partial f}{\partial p_2}\delta p_2 + \ldots + \frac{\partial f}{\partial p_n}\delta p_n. \tag{64}$$

Suppose we can manipulate (62) into (64), then we can pick off *all* the gradients of a single function f with respect to all parameters at once.

6.2.1 Resistance

The BCR for a resistor is

$$v_R = R i_R \tag{65}$$

and therefore, to first order

$$\delta v_R = R \delta i_R + i_R \delta R. \tag{66}$$

For a resistance,

$$\begin{aligned} \text{typical term} &= \delta v_R \hat{i}_R - \delta i_R \hat{v}_R & (67)\\ &= R \delta i_R \hat{i}_R + i_R \delta R \hat{i}_R - \delta i_R \hat{v}_R. & (68) \end{aligned}$$

We can choose *any* BCR for the resistance in the adjoint circuit. An inspection of (68) reveals that we are interested in the δR term for sensitivity purposes, but not the others. By choosing the BCR of the corresponding adjoint element to be

$$\hat{v}_R = R \hat{i}_R, \tag{69}$$

we obtain

$$\text{typical term} \quad = \quad i_R \hat{i}_R \delta R. \tag{70}$$

In other words, we choose to replace a resistance in the nominal circuit by an identical resistance of equal value in the adjoint circuit to obtain the typical term in the desired form. (Thus, a resistance is said to be *self-adjoint.*)

6.2.2 Independent current sources

The BCR for an independent current source is

$$i_J = J \tag{71}$$

and the perturbation of the BCR is simply

$$\delta i_J = \delta J \tag{72}$$

resulting in

$$\text{typical term} = \delta v_J \hat{i}_J - \delta J \hat{v}_J. \tag{73}$$

6.2.3 Circuits with resistances and independent current sources only

Consider the DC analysis of a circuit that consists of only resistances and current sources. The sum of the typical terms is zero, so we obtain

$$\sum_{all\ R} i_R \hat{i}_R \delta R + \sum_{all\ J} (\delta v_J \hat{i}_J - \delta J \hat{v}_J) = 0, \tag{74}$$

or

$$\sum_{all\ J} \delta v_J \hat{i}_J = \sum_{all\ J} \delta J \hat{v}_J - \sum_{all\ R} i_R \hat{i}_R \delta R. \tag{75}$$

Assume that f is the voltage across the k^{th} current source. Choose

$$\hat{i}_{J_k} = 1 \tag{76}$$

and all other $\hat{i}_J = 0$. Then

$$\delta f = \delta v_{Jk} = \sum_{all\ J} \delta J \hat{v}_J - \sum_{all\ R} i_R \hat{i}_R \delta R. \tag{77}$$

Compare to

$$\delta f = \sum_{all\ J} \delta J \frac{\partial f}{\partial J} + \sum_{all\ R} \frac{\partial f}{\partial R} \delta R. \tag{78}$$

Now we can pick off *all* the required sensitivities by inspection as

$$\frac{\partial f}{\partial J} = \hat{v}_J \tag{79}$$

and

$$\frac{\partial f}{\partial R} = -i_R \hat{i}_R. \tag{80}$$

Thus, if we solved the adjoint circuit, we could compute all the desired sensitivities at once!

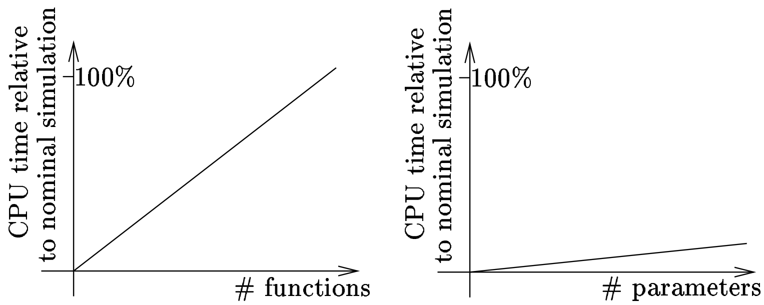

Figure 18: Growth of adjoint method run time.

6.2.4 Observations

- We assumed that $f = v_k$, the voltage across the k^{th} current source. If the function of interest was not across a current source, we could easily introduce a zero-valued current source ("ideal voltmeter") in parallel with the element whose voltage we want to measure, and then use the voltage of the current source as the function of interest.
- We picked the BCRs of the current sources in such a way that the left-hand side of of the "typical terms" equation (75) represented the perturbation of the function of interest (δf).
- The only change between the nominal and adjoint circuits is the value of the current sources, which only appear in the right-hand side of the circuit equations. Hence the LU factors obtained from the solution of the nominal circuit can be re-used and only forward-substitution and back-substitution are necessary to solve the adjoint circuit. We can therefore compute the sensitivity of one function with respect to any number of parameters in a single adjoint analysis at a small fraction of the computational cost of the nominal analysis.
- If we are interested in a second function, we will have to "activate" a second current source in the adjoint circuit and solve this new circuit. Again, we can re-use the LU factors of the nominal circuit, but a new forward and back-solve will be required. The computational overhead of additional parameters is almost negligible, whereas the computational overhead of additional functions is non-trivial, but still small compared to the original nominal circuit analysis. Fig. 18 shows the growth of CPU time with the number of parameters and functions. It is instructive to compare this situation with the corresponding CPU time growth in the direct method.

6.3 An example

Consider the simple circuit of Fig. 19. The currents through the resistances R_1 and R_2 are

$$i_{R1} = \frac{JR_2}{R_1 + R_2} \tag{81}$$

$$i_{R2} = \frac{JR_1}{R_1 + R_2}. \tag{82}$$

Suppose the function of interest were

$$f = v_{R2} \tag{83}$$

and the parameters of interest were R_1, R_2 and J.

Let us first compute the sensitivities manually.

$$f = \frac{JR_1R_2}{R_1 + R_2}, \tag{84}$$

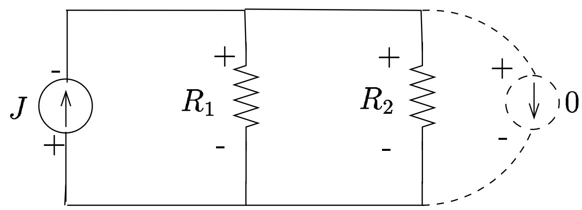

Figure 19: Sample circuit to demonstrate the adjoint method.

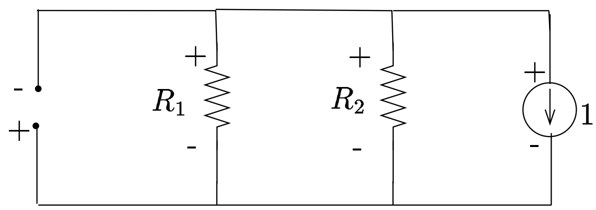

Figure 20: Adjoint circuit for the example of Fig. 19.

$$\frac{\partial f}{\partial R_1} = \frac{(R_1+R_2)JR_2 - JR_1R_2}{(R_1+R_2)^2} = J\left(\frac{R_2}{R_1+R_2}\right)^2, \tag{85}$$

$$\frac{\partial f}{\partial R_2} = \frac{(R_1+R_2)JR_1 - JR_1R_2}{(R_1+R_2)^2} = J\left(\frac{R_1}{R_1+R_2}\right)^2, \tag{86}$$

and

$$\frac{\partial f}{\partial J} = \frac{R_1R_2}{R_1+R_2}. \tag{87}$$

As can be expected, the sensitivity of f with respect to each of R_1, R_2 and J is positive assuming J is positive.

Now let us try to solve the same problem using adjoint sensitivity analysis. First, we introduce a zero-valued current source to measure the voltage of interest, as shown in the dotted lines in Fig. 19. Then, we replace each circuit by its adjoint circuit equivalent, as shown in Fig. 20. The resistances are unchanged. The current source J is replaced by a zero-valued source or open circuit (i.e., we choose $\hat{i}_J = 0$ since the voltage across J is not the function of interest). The zero-valued source gets activated with a unit current source in the adjoint circuit. The solution of the adjoint circuit gives us

$$\hat{i}_{R1} = \frac{-R_2}{R_1+R_2} \tag{88}$$

$$\hat{i}_{R2} = \frac{-R_1}{R_1+R_2} \tag{89}$$

$$\hat{v}_J = \frac{R_1R_2}{R_1+R_2}. \tag{90}$$

Then we can obtain the required sensitivities as follows.

$$\frac{\partial f}{\partial R_1} = -i_{R1}\hat{i}_{R1} = J\left(\frac{R_2}{R_1+R_2}\right)^2, \tag{91}$$

$$\frac{\partial f}{\partial R_2} = -i_{R2}\hat{i}_{R2} = J\left(\frac{R_1}{R_1+R_2}\right)^2, \tag{92}$$

and

$$\frac{\partial f}{\partial J} = \hat{v}_J = \frac{R_1R_2}{R_1+R_2}. \tag{93}$$

We obtain the same answer as before! Of course, in large circuits, it is not possible to write down an analytic solution and differentiate. The adjoint method is a very efficient numerical procedure to determine all the gradients of a function or measurement with respect to multiple parameters by means of a single circuit analysis.

6.3.1 Adjoint method procedure

For the simple case of circuits that contain only resistances and independent current sources, we can lay out a procedure to compute sensitivities by the adjoint method.

1. Introduce zero-valued current sources ("ideal voltmeters") in parallel with the elements whose voltages are functions of interest.
2. Solve the nominal circuit and save the LU factors as well as the currents through all resistors whose resistance values are parameters of interest.
3. Zero out the original current sources. Change the zero-valued source corresponding to the first function of interest to unity.
4. Solve the resulting adjoint circuit efficiently by re-using the LU factors of the nominal circuit.
5. Pick off the required sensitivities of the function with respect to all parameters.
6. Move on to the next function and repeat steps 3, 4 and 5 until all functions have been exhausted.

Again, it is instructive to compare this procedure to the direct method procedure.

6.4 Other elements

So far, we have only considered circuits with resistances and independent current sources. In this section, we will extend the discussion to voltage sources, conductances, and all four types of linear controlled sources.

6.4.1 Voltage source

The BCR for a voltage source is

$$v_E = E \tag{94}$$

and its perturbation is simply

$$\delta v_E = \delta E. \tag{95}$$

Hence,

$$\begin{aligned} \text{typical term} &= \delta v_E \hat{i}_E - \delta i_E \hat{v}_E & (96) \\ &= \delta E \hat{i}_E - \delta i_E \hat{v}_E. & (97) \end{aligned}$$

The first term tells us that the sensitivity with respect to the voltage of the independent voltage source is simply the current through the voltage source in the adjoint circuit. The second term allows us to "activate" this source in the adjoint circuit by choosing $\hat{v}_E = -1$ so that the function of interest is the current through the independent voltage source. So any time we have a measurement which is a branch current, we can introduce a zero-valued voltage source in series with it in the nominal circuit, and activate the voltage source to have a value of -1 in the adjoint circuit.

6.4.2 Conductance

The BCR for a conductance is

$$i_G = Gv_G \tag{98}$$

and its first-order perturbation is

$$\delta i_G = G\delta v_G + v_G \delta G. \tag{99}$$

Hence,

$$\begin{aligned}\text{typical term} &= \delta v_G \hat{i}_G - \delta i_G \hat{v}_G &(100)\\ &= \delta v_G \hat{i}_G - G\delta v_G \hat{v}_G - v_G \delta G \hat{v}_G. &(101)\\ & &(102)\end{aligned}$$

We obviously want to retain the δG term to obtain sensitivity with respect to the conductance value, so we choose the BCR for the adjoint circuit as

$$\hat{i}_G = G\hat{v}_G \tag{103}$$

to obtain

$$\text{typical term} = -v_G \hat{v}_G \delta G. \tag{104}$$

Thus, a conductance is replaced by an identical conductance in the adjoint circuit (conductances are *self-adjoint*) and sensitivity with respect to a conductance value is simply the product of the nominal and adjoint voltages across the conductance (we lose the negative sign since the typical term gets shifted to the other side in the final sensitivity expression).

6.4.3 Linear current-controlled current source (CCCS)

The nominal and adjoint elements of a linear current-controlled current source are shown in Fig. 21. The BCRs of the nominal circuit elements are

$$\begin{aligned} v_1 &= 0 &(105)\\ i_2 &= hi_1 &(106)\end{aligned}$$

and hence the typical term for the controlled source is

$$\begin{aligned}\text{typical term} &= \delta v_1 \hat{i}_1 - \delta i_1 \hat{v}_1 + \delta v_2 \hat{i}_2 - \delta i_2 \hat{v}_2 &(107)\\ &= 0 - \delta i_1 \hat{v}_1 + \delta v_2 \hat{i}_2 - h\delta i_1 \hat{v}_2 - \delta h i_1 \hat{v}_2. &(108)\end{aligned}$$

We are only interested in the δh term to obtain the sensitivity with respect to the gain of the controlled source. Hence we choose

$$\begin{aligned} \hat{i}_2 &= 0 &(109)\\ \hat{v}_1 &= -h\hat{v}_2. &(110)\end{aligned}$$

Thus, we find that the control of the controlled source is reversed in the adjoint circuit! With this choice of BCRs for the adjoint circuit,

$$\text{typical term} = -i_1 \hat{v}_2 \delta h. \tag{111}$$

Therefore, the sensitivity with respect to h is the product of the nominal controlling current and adjoint controlling voltage (remember that the typical term gets shifted to the other side in the final sensitivity expression). The CCCS is not self-adjoint, since the adjoint element is not identical to the nominal element.

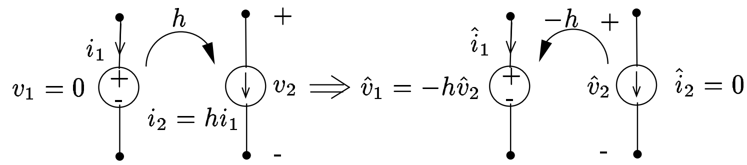

Figure 21: Adjoint method for a linear current-controlled current source.

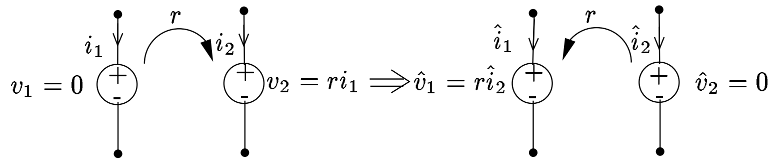

Figure 22: Adjoint method for a linear current-controlled voltage source.

6.4.4 Linear current-controlled voltage source (CCVS)

The nominal and adjoint elements of a linear current-controlled voltage source are shown in Fig. 22. The BCRs of the nominal circuit elements are

$$v_1 = 0 \tag{112}$$
$$v_2 = ri_1 \tag{113}$$

and hence the typical term for the controlled source is

$$\begin{aligned}\text{typical term} &= \delta v_1 \hat{i}_1 - \delta i_1 \hat{v}_1 + \delta v_2 \hat{i}_2 - \delta i_2 \hat{v}_2 && (114)\\ &= 0 - \delta i_1 \hat{v}_1 + r\delta i_1 \hat{i}_2 + \delta r i_1 \hat{i}_2 - \delta i_2 \hat{v}_2. && (115)\end{aligned}$$

We are only interested in the δr term to obtain the sensitivity with respect to the gain of the controlled source. Hence we choose

$$\hat{v}_2 = 0 \tag{116}$$
$$\hat{v}_1 = r\hat{i}_2. \tag{117}$$

Thus, we find that the control of the controlled source is reversed in the adjoint circuit! With this choice of BCRs for the adjoint circuit,

$$\text{typical term} = i_1 \hat{i}_2 \delta r. \tag{118}$$

Therefore, the sensitivity with respect to r is the negative product of the nominal controlling current and adjoint controlling current (remember that the typical term gets shifted to the other side in the final sensitivity expression). The CCVS is not self-adjoint, since the adjoint element is not identical to the nominal element.

6.4.5 Linear voltage-controlled voltage source (VCVS)

The nominal and adjoint elements of a linear voltage-controlled voltage source are shown in Fig. 23. The BCRs of the nominal circuit elements are

$$i_1 = 0 \tag{119}$$
$$v_2 = \mu v_1 \tag{120}$$

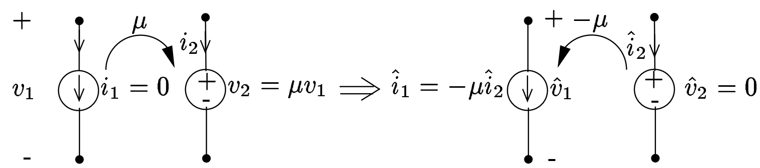

Figure 23: Adjoint method for a linear voltage-controlled voltage source.

and hence the typical term for the controlled source is

$$\begin{aligned} \text{typical term} &= \delta v_1 \hat{i}_1 - \delta i_1 \hat{v}_1 + \delta v_2 \hat{i}_2 - \delta i_2 \hat{v}_2 & (121) \\ &= \delta v_1 \hat{i}_1 - 0 + \mu \delta v_1 \hat{i}_2 + \delta \mu v_1 \hat{i}_2 - \delta i_2 \hat{v}_2. & (122) \end{aligned}$$

We are only interested in the $\delta\mu$ term to obtain the sensitivity with respect to the gain of the controlled source. Hence we choose

$$\begin{aligned} \hat{v}_2 &= 0 & (123) \\ \hat{i}_1 &= -\mu \hat{i}_2. & (124) \end{aligned}$$

Thus, we find that the control of the controlled source is reversed in the adjoint circuit! With this choice of BCRs for the adjoint circuit,

$$\text{typical term} = v_1 \hat{i}_2 \delta\mu. \qquad (125)$$

Therefore, the sensitivity with respect to μ is the negative product of the nominal controlling voltage and adjoint controlling current (remember that the typical term gets shifted to the other side in the final sensitivity expression). The VCVS is not self-adjoint, since the adjoint element is not identical to the nominal element.

6.4.6 Linear voltage-controlled current source (VCCS)

The nominal and adjoint elements of a linear voltage-controlled current source are shown in Fig. 24. The BCRs of the nominal circuit elements are

$$\begin{aligned} i_1 &= 0 & (126) \\ i_2 &= g v_1 & (127) \end{aligned}$$

and hence the typical term for the controlled source is

$$\begin{aligned} \text{typical term} &= \delta v_1 \hat{i}_1 - \delta i_1 \hat{v}_1 + \delta v_2 \hat{i}_2 - \delta i_2 \hat{v}_2 & (128) \\ &= \delta v_1 \hat{i}_1 - 0 + \delta v_2 \hat{i}_2 - \delta g v_1 \hat{v}_2 - g \delta v_1 \hat{v}_2. & (129) \end{aligned}$$

We are only interested in the δg term to obtain the sensitivity with respect to the gain of the controlled source. Hence we choose

$$\begin{aligned} \hat{i}_2 &= 0 & (130) \\ \hat{i}_1 &= g \hat{v}_2. & (131) \end{aligned}$$

Thus, we find that the control of the controlled source is reversed in the adjoint circuit! With this choice of BCRs for the adjoint circuit,

$$\text{typical term} = -v_1 \hat{v}_2 \delta h. \qquad (132)$$

Therefore, the sensitivity with respect to h is the product of the nominal controlling voltage and adjoint controlling voltage (remember that the typical term gets shifted to the other side in the final sensitivity expression). The VCCS is not self-adjoint, since the adjoint element is not identical to the nominal element.

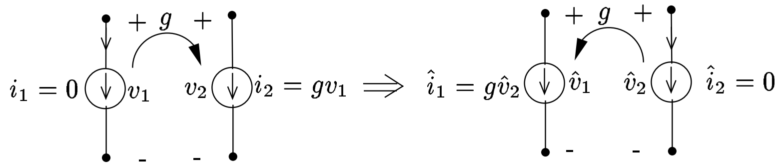

Figure 24: Adjoint method for a linear voltage-controlled current source.

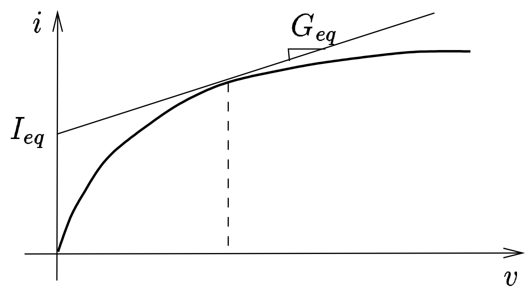

Figure 25: Linearization of a nonlinear element.

6.5 Fast solution of the adjoint circuit

The contribution of controlled sources to the left-hand side of the circuit equations is different in the nominal and adjoint circuits. Does this mean that LU factors cannot be re-used in circuits with controlled circuits? It turns out that in every case, the contribution or stamps of the controlled sources in the adjoint circuit are the transpose of the corresponding stamps in the nominal circuit equations. In other words, *the left-hand side matrix of the circuit equations of the adjoint circuit is the transpose of the left-hand side matrix of the nominal equations!* Of course, we now see that self-adjoint elements have a symmetric contribution to the circuit equations, and transposing does not change them. So indeed the LU factors can be re-used since if

$$A = LU \tag{133}$$

then

$$A^T = U^T L^T \tag{134}$$

with U^T being lower triangular and L^T being upper triangular. Hence in all cases, the LU factors of the nominal circuit can be re-used for the adjoint circuit.

6.6 Nonlinear circuits

As in the case of the direct method, the linearization at the solution of the nominal circuit is valid in the face of infinitesimally small perturbations of the sensitivity parameters. Thus, we take the linearized equivalent circuit obtained at the solution of the nominal circuit and compute its sensitivities. As a practical matter, nonlinear circuits pose no additional theoretical complexity.

Fig. 25 shows the linearization of a nonlinear circuit at its solution, represented by

$$i = i_1 + i_2 = G_{eq}v + I_{eq}. \tag{135}$$

This model is represented by a Norton equivalent as shown in Fig. 26. For this composite branch,

$$\begin{aligned} \text{typical term} &= \delta v\hat{i}_1 - \delta i_1\hat{v} + \delta v\hat{i}_2 - \delta i_2\hat{v} && (136)\\ &= \delta v\hat{i}_1 - G_{eq}\delta v\hat{v} - v\delta G_{eq}\hat{v} + \delta v\hat{i}_2 - \delta I_{eq}\hat{v}. && (137) \end{aligned}$$

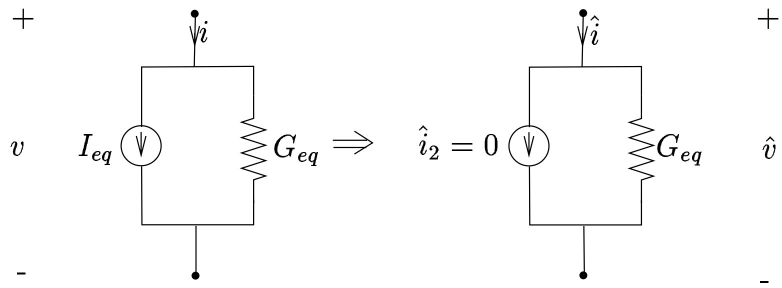

Figure 26: Nominal and adjoint circuit elements for a nonlinear element.

Since we are only interested in the terms that contain δI_{eq} and δG_{eq}, we choose as BCRs for the adjoint circuit

$$\hat{i}_1 = G_{eq}\hat{v} \tag{138}$$

$$\hat{i}_2 = 0. \tag{139}$$

Then, for the composite branch

$$\text{typical term} = -v\hat{v}\delta G_{eq} - \hat{v}\delta I_{eq}. \tag{140}$$

In other words, the adjoint element is chosen to be a simple conductance of value equal to the conductance of the nominal linearized element. The sensitivity with respect to the equivalent conductance and current intercept are given by (140). These can be chain-ruled to compute sensitivity with respect to other model parameters (such as model coefficients or device sizes) since I_{eq} and G_{eq} are expressed in terms of these parameters in the device model.

6.7 Summary of DC case

For circuits without energy-storage elements, the adjoint equations that we have derived so far can be summarized by a single equation

$$\sum_J \delta v_J \hat{i}_J - \sum_E \delta i_E \hat{v}_E = \sum_J \hat{v}_J \delta J - \sum_E \hat{i}_E \delta E - \sum_R i_R \hat{i}_R \delta R + \sum_G v_G \hat{v}_G \delta G + \sum_{CCCS} i_1 \hat{v}_2 \delta h \tag{141}$$

$$- \sum_{CCVS} i_1 \hat{i}_2 \delta r - \sum_{VCVS} v_1 \hat{i}_2 \delta \mu + \sum_{VCCS} v_1 \hat{v}_2 \delta g + \sum_{NL} (v\hat{v}\delta G_{eq} + \hat{v}\delta I_{eq}). \tag{142}$$

By appropriate choice of $\hat{i}_J$ and $\hat{v}_E$ (the independent sources of the adjoint circuit whose value we can pick in any way we wish), the left hand side can be manipulated into the perturbation of the function of interest. This excitation is applied, and the adjoint circuit solved by re-using the LU factors of the nominal solution. Then the solution of the nominal and adjoint circuits are combined as dictated by the right-hand side of (142) and (142) to obtain the sensitivity of the function with respect to all parameters at once.

6.8 Scalar functions

Suppose we are interested in a function or measurement that is more complex than just a single voltage or current. For example,

$$f = f(i_E, v_J) \tag{143}$$

where f is an arbitrary (possibly nonlinear) continuously differentiable function of some current and voltage measurements of the circuit. Then

$$\delta f = \sum_E \frac{\partial f}{\partial i_E} \delta i_E + \sum_J \frac{\partial f}{\partial v_J} \delta v_J. \tag{144}$$

For the adjoint circuit, for each voltage source we choose

$$\hat{v}_E = -\frac{\partial f}{\partial i_E} \tag{145}$$

and for each current source, we choose

$$\hat{i}_J = \frac{\partial f}{\partial v_J}. \tag{146}$$

The right-hand side of the above two equations is known once the nominal circuit is solved. So we simultaneously drive the adjoint circuit by multiple excitations given by the above equation, and then pick off the sensitivities of f with respect to all parameters of interest in a single adjoint analysis. We therefore have the ability to compute the sensitivity of any arbitrary *scalar* function of currents and voltages with respect to any number of parameters by means of a single adjoint analysis [5, 6].

If the sensitivities are being computed for the purposes of optimization, it is often the case that the optimization package employed minimizes a single *cost function* or *merit function* such as a Lagrangian or augmented Lagrangian. In this case, the optimization package does not need to know the sensitivities of the individual measurements, it only needs to know the gradients of the merit function. In this situation, all the gradients necessary for optimization can be computed in a single adjoint analysis, irrespective of the number of individual measurements and irrespective of the number of parameters of interest!

6.9 Matrix interpretation of adjoint sensitivity

Adjoint sensitivity analysis can be understood from a "matrix" or "system" theory perspective, as described in this section. Consider a linear (or linearized) system of equations

$$Ax = b. \tag{147}$$

Perturb parameters $p_1, p_2, \ldots, p_n$ simultaneously. Assume that x, A and b have perturbations of δx, δA and δb, respectively. Then, to first order,

$$A\delta x + \delta A x = \delta b \tag{148}$$

or

$$\delta x = A^{-1}[\delta b - \delta A x]. \tag{149}$$

Consider a scalar function $f(x)$. Then

$$\delta f = \left[\frac{\partial f}{\partial x}\right]^T \delta x = \left[\frac{\partial f}{\partial x}\right]^T A^{-1}[\delta b - \delta A x]. \tag{150}$$

Postulate an adjoint system and its solution vector ζ such that

$$A^T \zeta = \left[\frac{\partial f}{\partial x}\right]. \tag{151}$$

Since the system matrix of the adjoint system is the transpose of the nominal system, LU factors can be re-used to solve (151). Then

$$\zeta^T A = \left[\frac{\partial f}{\partial x}\right]^T, \tag{152}$$

and finally

$$\delta f = \zeta^T \left[\delta b - \delta A x\right]. \tag{153}$$

The system described by (151) has the transposed system matrix compared to the nominal system, and a right-hand side that is eerily identical to the excitations proposed by (145 and 146). Equation (153) says to combine the results of the nominal and adjoint systems to obtain all the required sensitivities at once. Of course, depending on the type of circuit element, the pattern of the perturbation of A and b is determined, which in turn determines how the final sensitivity is computed.

ONTWERPTECHNOLOGIE 5L050 HANDOUT 4

Class 4, lecture 3: Transient sensitivity analysis by the adjoint method

7 Transient case

The most interesting part of the derivation of adjoint sensitivity analysis is its application to transient simulation. Consider a linear(ized) capacitance C. Its BCR is

$$i_C = C\dot{v}_C, \tag{154}$$

the perturbation of which yields

$$\delta i_C = C\delta\dot{v}_C + \delta C\dot{v}_C \tag{155}$$

leading to

$$\begin{aligned} \text{typical term} &= \delta v_C \hat{i}_C - \delta i_C \hat{v}_C & (156)\\ &= \delta v_C \hat{i}_C - C\delta\dot{v}_C\hat{v}_C - \delta C\dot{v}_C\hat{v}_C. & (157)\end{aligned}$$

It is clear that we have to generalize our application of Tellegen's theorem to deal with energy-storage elements such as capacitances. Let us modify our application of Tellegen's theorem from

$$\sum_{\text{all branches}} (\delta v_b \hat{i}_b - \delta i_b \hat{v}_b) = 0 \tag{158}$$

to

$$\sum_{\text{all branches}} \int_{t_0}^{t_f} \left[\delta v_b(t)\hat{i}_b(\tau) - \delta i_b(t)\hat{v}_b(\tau)\right] dt = 0 \tag{159}$$

where t is the time in the nominal circuit, t_0 is the start time of the simulation, t_f is the end time of the simulation, and τ is an as-yet-unchosen choice of time for the adjoint circuit. Then,

$$\begin{aligned} \text{new typical term} &= \int_{t_0}^{t_f} \left[\delta v_C(t)\hat{i}_C(\tau) - C\delta\dot{v}_C(t)\hat{v}_C(\tau) - \delta C\dot{v}_C(t)\hat{v}_C(\tau)\right] dt & (160)\\ &= \int_{t_0}^{t_f} \delta v_C(t)\hat{i}_C(\tau)dt - C\hat{v}_C(\tau)\delta v_C(t)\big|_{t_0}^{t_f} & (161)\\ &\quad + \int_{t_0}^{t_f} C\delta v_C(t)\dot{\hat{v}}_C(\tau)dt - \int_{t_0}^{t_f} \delta C\dot{v}_C(t)\hat{v}_C(\tau)dt & (162)\\ &= \int_{t_0}^{t_f} \delta v_C(t)\left[\hat{i}_C(\tau) + C\dot{\hat{v}}_C(\tau)\right] dt - C\hat{v}_C(\tau)\delta v_C(t)\big|_{t=t_f} & (163)\\ &\quad + C\hat{v}_C(\tau)\delta v_C(t)\big|_{t=t_0} - \left[\int_{t_0}^{t_f} \dot{v}_C(t)\hat{v}_C(\tau)dt\right]\delta C. & (164)\end{aligned}$$

In the above equations, the second term of the new typical term was integrated by parts using the familiar formula

$$\int_a^b u dv = uv|_a^b - \int_a^b v du. \qquad (165)$$

The formula for the typical term left us with four terms. Of these, the fourth term is welcome since it has a δC that will enable us to compute sensitivity with respect to the capacitance value. We need to get rid of the first, second and third terms by a suitable choice of BCR for the capacitance in the adjoint circuit, and perhaps even obtain sensitivity with respect to the initial condition of the capacitance. To get rid of the first term, we require

$$\hat{i}_C(\tau) + C\hat{\dot{v}}_C(\tau) = 0 \qquad (166)$$

which would seem to require a negative capacitance in the adjoint circuit, a decidedly ugly and unnatural choice. Instead, *we choose to run time backwards in the adjoint circuit.* In other words, we choose a relationship between t and τ as follows:

$$\tau = t_0 + t_f - t, \qquad (167)$$

and we also choose the BCR for the adjoint element to be identical to the nominal BCR, i.e., a capacitance remains a capacitance of the same value in the adjoint circuit.

With these choices,

$$\hat{i}_C(\tau) = C\frac{d\hat{v}_C}{d\tau} = -C\frac{d\hat{v}_C}{dt} = -C\hat{\dot{v}}_C \qquad (168)$$

and the first term of (164) disappears! Then we choose

$$\hat{v}_C|_{t=t_f \text{ or } \tau=t_0} = 0 \qquad (169)$$

as the initial condition of the adjoint circuit and the second term of (164) vanishes. Finally, we are left with a typical term of

$$-\delta C \int_{t_0}^{t_f} \dot{v}_C(t)\hat{v}_C(\tau)dt + C\hat{v}_C(\tau = t_f)\delta v_C(t_0). \qquad (170)$$

This remarkable equation tells us that the sensitivity with respect to the capacitance value is the convolution integral of the slope of the nominal capacitance voltage waveform and the adjoint capacitance voltage waveform, and the sensitivity with respect to the *initial condition* of the capacitance is the negative of the capacitance times the *final condition* of the capacitance in the adjoint circuit!

7.1 Transient sensitivity for a resistance

Since we have now redefined the typical term, we have to revisit the typical term for every element type. We will derive the equations for a resistance and then re-write the summary equation for the adjoint method accordingly.

For a resistance, the BCR is

$$v_R = i_R R \qquad (171)$$

and the perturbation of the BCR is

$$\delta v_R = i_R \delta R + \delta i_R R \qquad (172)$$

leading to

$$\text{typical term} = \int_{t_0}^{t_f} \left[i_R(t)\delta R\hat{i}_R(\tau) + \delta i_R(t)R\hat{i}_R(\tau) - \delta i_R(t)\hat{v}_R(\tau) \right] dt. \qquad (173)$$

For the adjoint BCR, we choose

$$\hat{v}_R(\tau) = R\hat{i}_R(\tau) \tag{174}$$

for all τ. Thus,

$$\text{typical term} = \delta R \int_{t_0}^{t_f} i_R(t)\hat{i}_R(\tau)dt \tag{175}$$

$$= \delta R \int_{t_0}^{t_f} i_R(t)\hat{i}_R(t_0 + t_f - t)dt. \tag{176}$$

We therefore find that for all element types, the sensitivities are expressed as convolution integrals.

7.2 Basic relation for transient sensitivity

Extending the derivation of the previous two sections to all types of linear and nonlinear elements, we obtain the following general formula that summarizes the adjoint method in the time-domain.

$$\sum_J \int_{t_0}^{t_f} \delta v_J(t)\hat{i}_J(\tau)dt - \sum_E \int_{t_0}^{t_f} \delta i_E(t)\hat{v}_E(\tau)dt = \sum_J \delta J \int_{t_0}^{t_f} \hat{v}_J(\tau)dt \tag{177}$$

$$-\sum_E \delta E \int_{t_0}^{t_f} \hat{i}_E(\tau)dt \tag{178}$$

$$-\sum_R \delta R \int_{t_0}^{t_f} i_R(t)\hat{i}_R(\tau)dt \tag{179}$$

$$+\sum_G \delta G \int_{t_0}^{t_f} v_G(t)\hat{v}_G(\tau)dt \tag{180}$$

$$+\sum_{CCCS} \delta h \int_{t_0}^{t_f} i_1(t)\hat{v}_2(\tau)dt \tag{181}$$

$$-\sum_{CCVS} \delta r \int_{t_0}^{t_f} i_1(t)\hat{i}_2(\tau)dt \tag{182}$$

$$-\sum_{VCVS} \delta \mu \int_{t_0}^{t_f} v_1(t)\hat{i}_2(\tau)dt \tag{183}$$

$$+\sum_{VCCS} \delta g \int_{t_0}^{t_f} v_1(t)\hat{v}_2(\tau)dt \tag{184}$$

$$+\sum_{NL} \left(\delta G_{eq} \int_{t_0}^{t_f} v(t)\hat{v}(\tau)dt + \delta I_{eq} \int_{t_0}^{t_f} \hat{v}(\tau)dt\right) \tag{185}$$

$$+\sum_C \left(\delta C \int_{t_0}^{t_f} \dot{v}_C(t)\hat{v}_C(\tau)dt - C\hat{v}_C(t_f)\delta v_C(t_0)\right) \tag{186}$$

$$-\sum_L \left(\delta L \int_{t_0}^{t_f} \dot{i}_L(t)\hat{i}_L(\tau)dt - L\hat{i}_L(t_f)\delta i_L(t_0)\right), \tag{187}$$

where L indicates linear inductances and NL are the nonlinear elements.

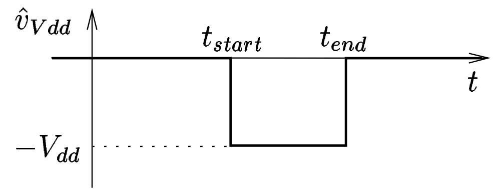

Figure 27: Adjoint excitation to express a power function.

7.3 Choice of input excitation

In the transient case, it is our goal to pick the BCRs for independent voltage and current sources so that the perturbation of the function of interest

$$\delta f = \sum_J \int_{t_0}^{t_f} \delta v_J(t)\hat{i}_J(\tau)dt - \sum_E \int_{t_0}^{t_f} \delta i_E(t)\hat{v}_E(\tau)dt. \tag{188}$$

While this may seem like a daunting task at first, it is really quite simple and the paradigm allows for the expression of a vast variety of measurements.

7.3.1 Example 1: power

Suppose we are interested in the sensitivity of power through an independent voltage source V_{dd} from $t = t_{start}$ to t_{end}. Then

$$f = \int_{t_{start}}^{t_{end}} i_{Vdd}V_{dd}dt \tag{189}$$

and

$$\delta f = V_{dd}\int_{t_{start}}^{t_{end}} \delta i_{Vdd}dt. \tag{190}$$

To express this function, we set all independent current and voltage sources in the adjoint circuit to zero, except for V_{dd} whose waveform is chosen to be as shown in Fig. 27. It can be seen that by plugging this waveform into (188), we obtain the necessary function (190).

7.4 Example 2: voltage at a particular time t'

In this case, we choose the current source in the adjoint circuit at the measurement port to have a unit Dirac impulse at $t = t'$ (see Fig. 28) and all other independent sources to be zero for all time. Thus,

$$\delta f = \int_{t_0}^{t_f} \delta v_J(t)\delta(t - t')dt = \delta v_J(t'). \tag{191}$$

The Dirac impulse is not difficult to handle during the analysis of the adjoint circuit. For example, if the node to which the impulse is applied has a grounded capacitance, the voltage of that node will instantaneously increase at t' by a step amount equal to the height of the impulse divided by the capacitance value.

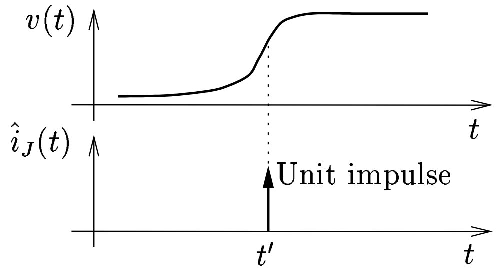

Figure 28: Adjoint excitation for sensitivity of a voltage measurement at a particular time.

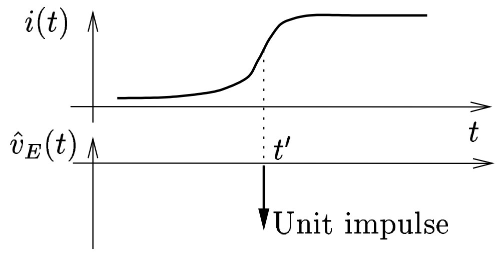

Figure 29: Adjoint excitation for sensitivity of a current measurement at a particular time.

7.5 Example 3: current at a particular time t'

In this case, we choose the voltage source in the adjoint circuit at the measurement port to have a negative unit Dirac impulse at $t = t'$ (see Fig. 29) and all other independent sources to be zero for all time. Thus,

$$\delta f = -\int_{t_0}^{t_f} \delta i_E(t)\{-\delta(t-t')\}dt = \delta i_E(t'). \tag{192}$$

7.6 Example 4: area under a noise bump

Suppose the nominal waveform at some measurement port is as shown in Fig. 30. Suppose we do not want the voltage to exceed a certain threshold voltage V_{thresh} for noise reasons. Then we would like to compute the sensitivity of the area under the noise bump. This is easily done by selecting the waveform of the current source at that measurement port in the adjoint circuit as shown in the lower waveform of Fig. 30. It is clear that

$$\begin{aligned} \delta f &= \int_{t_0}^{t_f} \delta v_J(t)\hat{i}_J(\tau)dt && (193)\\ &= \int_{t_{start}}^{t_{end}} \delta v_J(t)dt && (194)\\ &= \delta\left(\int_{t_{start}}^{t_{end}} v_J(t)dt\right) && (195) \end{aligned}$$

and therefore represents the perturbation of our function of interest.

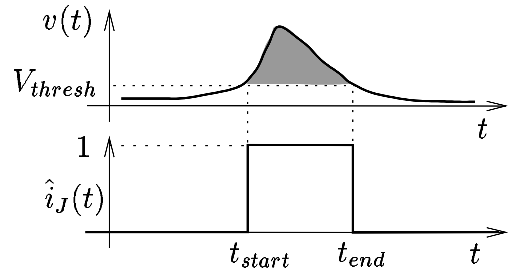

Figure 30: Adjoint sensitivity of a noise bump.

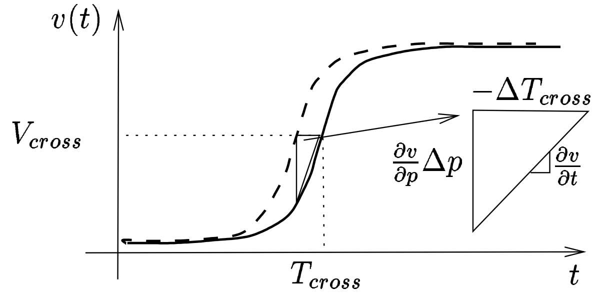

Figure 31: Adjoint sensitivity of a crossing time.

7.7 Example 5: crossing time or delay measurement

Suppose we are interested in the time at which a voltage crosses a certain crossing voltage V_{cross}. In digital circuits, we often measure the time at which an output waveform crosses 10%, 50% and 90% of V_{dd}. Fig. 31 shows a nominal waveform which crosses V_{cross} at time T_{cross}. We are interested in computing the sensitivity of T_{cross}. Starting with

$$v(t,p)|_{t=T_{cross}} = V_{cross}, \tag{196}$$

we differentiate with respect to a parameter p to obtain

$$\frac{\partial v}{\partial t}\frac{\partial t}{\partial p}\bigg|_{t=T_{cross}} + \frac{\partial v}{\partial p}\bigg|_{t=T_{cross}} = 0 \tag{197}$$

which leads to the final result

$$\frac{\partial T_{cross}}{\partial p} = \frac{-\left.\frac{\partial v}{\partial p}\right|_{t=T_{cross}}}{\left.\frac{\partial v}{\partial t}\right|_{t=T_{cross}}}. \tag{198}$$

The required sensitivity is simply the negative ratio of the sensitivity of the voltage at the crossing time to the voltage slope of the crossing in the nominal circuit. It makes sense that the nominal voltage slope is in the denominator since the "steeper" the nominal crossing, the smaller the magnitude of sensitivity that we expect. Fig. 31 makes a self-evident geometrical argument to obtain the same result, with the dotted line representing a perturbed version of the voltage waveform whose crossing time is being measured.

7.8 Example 6: scalar function

As in the DC case, we can express scalar functions of multiple measurements by simultaneously applying scaled versions of the excitations that would be applied to compute the sensitivity of the individual measurements. Even in the time-domain, we can compute the sensitivity of a scalar function of any number of measurements to any number of parameters in a single adjoint analysis.

7.9 Adjoint sensitivity transient procedure

In the transient case, the following steps constitute the procedure to compute sensitivities by the adjoint method.

1. Introduce zero-valued voltage source ("ideal ammeters") and zero-valued current sources ("ideal voltmeters") at current and voltage measurement ports, respectively.
2. Start the nominal circuit with the initial conditions provided by the user. Solve the nominal circuit and save the LU factors at each time step. Also save the currents/voltage waveforms of all circuit elements that contribute sensitivity parameters, as prescribed by (177) to (187).
3. Start the adjoint circuit with zero initial conditions.
4. Determine the excitation waveform(s) of the adjoint circuit that express the first sensitivity function of interest. Time can be "fast-forwarded" to the time of the first excitation in τ (last excitation in t) since the circuit has no inputs or initial conditions until then.
5. Solve the adjoint circuit, running time backwards. At each time point, since we are solving a transposed system, re-use the LU factors of the last Newton iteration of the corresponding time step in the nominal solution of the circuit.
6. Keep the convolution integrals of (177) to (187) up-to-date as time progresses.
7. When the adjoint simulation is complete, the convolution integrals yield the sensitivity of the function of interest with respect to *all* parameters.
8. Repeat steps 3, 4, 5, 6 and 7 for each function of interest.

7.10 Conclusions and discussion

We finally understand the cryptic statement, "In the adjoint method, time and control are reversed!"

The adjoint method has the following advantages:

1. Clearly, in the situation where we have a large number of parameters, gradients with respect to all of them are computed simultaneously.
2. Scalar functions of multiple measurements also require just a single adjoint analysis.
3. Nominal LU factors at each time point can be re-used.

The adjoint method has the following disadvantages:

1. Since we are running time backwards, nominal LU factors and nominal waveforms required for convolution must be saved. This can be a large memory burden.

2. The convolution integrals add an overhead to the adjoint method that is not required in the direct method.

3. The adjoint method cannot compute the "full transient" sensitivity of a function, i.e., $\frac{\partial f(t)}{\partial p}$ for a range of t, whereas this type of computation is possible in the direct method.

4. If sensitivity parameters are related (e.g., if two parameters bear a fixed ratio to each other), the effective number of parameters is reduced in the direct method, but there is no easy way to take advantage of this "grouping" in the adjoint method.

In conclusion, we note that when the number of sensitivity functions outnumbers the parameters, the direct method is generally advantageous. Whereas, when the number of parameters outnumbers the functions, the adjoint method is likely to be more efficient.

8 Frequency-domain sensitivity analysis

The methods developed above for the DC case are directly applicable in the frequency-domain. Frequency-domain application is briefly described in this section.

8.1 Frequency-domain direct method sensitivity analysis

As an example, consider a linear capacitance. The BCR in the frequency domain is

$$I_C = j\omega C V_C \tag{199}$$

where $j = \sqrt{-1}$ and ω is the frequency of interest. Directly differentiating with respect to a parameter p of interest,

$$\frac{\partial I_C}{\partial p} = j\omega C \frac{\partial V_C}{\partial p} + j\omega \frac{\partial C}{\partial p} V_C + j \frac{\partial \omega}{\partial p} C V_C \tag{200}$$

$$\hat{I}_C = j\omega C \hat{V}_C + j\omega \frac{\partial C}{\partial p} V_C + j \frac{\partial \omega}{\partial p} C V_C. \tag{201}$$

The BCR in the sensitivity circuit is that of an ordinary capacitance, but with two additional current sources in parallel. The value of these current sources is known after the nominal circuit analysis. Thus we can compute sensitivity both with respect to circuit component values and with respect to the frequency ω, in a manner quite analogous to the DC case.

8.2 Frequency-domain adjoint method sensitivity analysis

As an example, consider a linear capacitance. The BCR in the frequency domain is

$$I_C = j\omega C V_C \tag{202}$$

and its perturbation is

$$\delta I_C = j\omega C \delta V_C + j\omega \delta C V_C + j\delta\omega C V_C \tag{203}$$

leading to

$$\text{typical term} = \delta V_C \hat{I}_C - j\omega C \delta V_C \hat{V}_C - j\omega \delta C V_C \hat{V}_C - j\delta\omega C V_C \hat{V}_C. \tag{204}$$

We are interested in preserving the δC and $\delta\omega$ terms. For the adjoint circuit, we choose the BCR

$$\hat{I}_C = j\omega C\hat{V}_C \tag{205}$$

or in other words the capacitance is replaced by an identical capacitance in the adjoint circuit. Then

$$\text{typical term} = -j\omega V_C\hat{V}_C\delta C - jCV_C\hat{V}_C\delta\omega. \tag{206}$$

Thus, we can compute sensitivity both with respect to circuit component values and with respect to the frequency ω, in a manner quite analogous to the DC case.

References

[1] L. T. Pillage, R. A. Rohrer, and C. Visweswariah, *Electronic circuit and system simulation methods.* McGraw-Hill, 1995.

[2] D. A. Hocevar, P. Yang, T. N. Trick, and B. D. Epler, "Transient sensitivity computation for MOSFET circuits," *IEEE Transactions on Computer-Aided Design of ICs and Systems*, vol. CAD-4, pp. 609–620, October 1985.

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

[4] B. D. H. Tellegen, "A general network theorem, with applications," *Philips Research Reports*, vol. 7, pp. 259–269, 1952.

[5] A. R. Conn, R. A. Haring, C. Visweswariah, and C. W. Wu, "Circuit optimization via adjoint Lagrangians," *IEEE International Conference on Computer-Aided Design*, pp. 281–288, November 1997.

[6] A. R. Conn, P. K. Coulman, R. A. Haring, G. L. Morrill, C. Visweswariah, and C. W. Wu, "JiffyTune: circuit optimization using time-domain sensitivities," *IEEE Transactions on Computer-Aided Design of ICs and Systems*, vol. 17, pp. 1292–1309, December 1998.