Improvement in Norm-Reducing Newton Methods for Circuit Simulation

by
H. R. Yeager and R. W. Dutton
Stanford Electronics Laboratory
Stanford University
Stanford, California 94305

Address:
Hal R. Yeager
AEL 231D
Stanford University
Stanford, California 94305
(415) 857-7384
(415) 723-1482

Abstract

Two basic convergence problems exist when implementing high gain device models such as that used for HEMT transistors into SPICE[1,2] . The first is the classic concave function problem encounted with all FET and bipolar drain current characteristics. The case has long been know to cause failure in global convergence for the undamped Newton method. The solution to this problem has been to develop various damping strategies for Newton's method, of which the norm-reducing methods [3,4] are the most general and mathematically sound. The second is the transconductance case and is, unfortunately, peculiar to norm-reducing methods that employ the L_2-norm. The solution presented here is to replace the L_2-norm with a different norm, the Nu-norm. From a circuit point of view, the Nu-norm determines which unknowns should be converged first, prioritizes those unknowns, and guides the damping of the Newton updates accordingly. From a mathematical point of view, the steepest descent direction in the Nu-norm is parallel to each Newton update at the iterate point and, therefore, allows more effective damping of the updates. The most tangible results of employing the Nu-norm have been an order of magnitude reduction in Newton iterations for high voltage and current gain circuits.

In this paper, we consider solving the homeomorphic system of non-linear equations $G(u) = 0$ with Newton's Method. The system $G(u)$ is the set of node and branch equations that arise from a given circuit topology and model equations and the components of u are node voltages and branch currents. We employ the following form of Newton's method

$$u_{k+1} = u_k + T_k \cdot \delta u_k \qquad [1.a]$$

$$\mathbf{H}_k \cdot \delta u_k = -G(u_k) \qquad [1.b]$$

where u_{k+1} and u_k are the $k+1^{th}$ and k^{th} iterates in the unknowns, δu_k is the update direction between iterates, $T_k \in (0,1]$ is a scalar damping parameter, $G(u_k)$ is the nonlinear system evaluated at the u_k iterate, and $\mathbf{H}_k$ is the Jacobian of $G(u_k)$ with respect to u_k. The parameter T_k is chosen such that the norm of $G(u_{k+1})$ decreases from that of $G(u_k)$. The specific choice of T_k is guided by the particular norm-reducing technique employed and its values are well defined for homeomorphic mappings [4] .

The L_2-norm is by far the most common choice in many applications for measuring $G(\cdot)$ owing to its ease of calculation. In the one-dimensional case, the choice of norm is not critical because the update-direction and the direction of steepest-descent in any norm will always be the same. In two or more dimensions, however, the update-direction and steepest-descent direction of any norm may not always coincide and there may be as much as a 90° angular difference between them. For non-linear transconductive elements, this situation occurs often with the application of the L_2-norm. We introduce a new norm, the 'Nu'-norm, which has its steepest-descent direction coincident with the Newton-update.

The Nu-norm is given by equation 2 and the gradient with respect to u is given by equation 3:

$$N_{Nu}(G(u)) = \left[G^T(u)\mathbf{H}^{-T}(u_k)\mathbf{H}^{-1}(u_k)G(u)\right]^{\frac{1}{2}} \qquad [2]$$

$$\nabla_u N_{Nu}(G(u)) = \frac{1}{\left[N_{Nu}(G(u))\right]^{\frac{1}{2}}}\mathbf{H}^T(u)\mathbf{H}^{-T}(u_k)\mathbf{H}^{-1}(u_k)G(u) \qquad [3]$$

where $\mathbf{H}(u)$ and $\mathbf{H}(u_k)$ are the Jacobians of $G(\cdot)$ evaluated at the points u and u_k, respectively. When evaluated at the iterate ($u = u_k$) the steepest descent direction of $N_{Nu}(G(u_k))$ coincidents with δu_k. However, the steepest descent in L_2-norm is coincident to the update only when $G(u_k)$ is proportional to one of the eigenvectors of $\mathbf{H}^{-T}(u_k)\mathbf{H}^{-1}(u_k)$.

Figure 1 shows a transconductance topology commonly used to model field-effect and bipolar transistor devices. This circuit contains two nodes with an input voltage V_{app} applied to node 1 via a conductance g_{in}. An output current which is functionally related to the input voltage V_1 is generated at node 2 by the transconductance generator $I(V_1)$. A load conductance g_l is connected between node 2 and ground, producing a voltage gain at node 2. This circuit is unidirectional in behavior and an applied voltage at node 2 does not produce a current or a voltage gain at node 1.

Applying Kirchoff's current law at each node, the non-linear system of equations is extracted as:

$$G(V_1,V_2) = \begin{bmatrix} g_1(V_1,V_2) \\ g_2(V_1,V_2) \end{bmatrix} = \begin{bmatrix} gin^*(V_1 - V_{app}) \\ I(V_1) + gl^*V_2 \end{bmatrix} \qquad [5]$$

where gin is the input conductance, gl is the output conductance, and $I(V_1)$ is a non-linear monotonic function given by $I(V_1) = 0$ for $V_1 < 0$ and $I(V_1) = GmV_1^3$ for $V_1 > 0$. Figure 2 shows the contour graph of $G(\cdot)$ in the L_2-norm. (This and the following contour graphs employ a gray-scale to indicate the height of the contour lines. A darker line corresponds to a higher contour line.) A steep canyon structure is present in this graph and curves in a smooth trajectory towards the solution at $(V_1,V_2) = (1,-4)$. The large arrows on the graph in Figure 2 are the Newton-update directions for the initial guess and for the iterates #10, #20, #30, #40 and #45. The iterates started with the initial guess of (V_1,V_2)=(0,0) and each subsequent iterate was chosen such that the L_2-norm was minimized along the Newton-update direction specified by the previous iterate. The small arrow accompanying each Newton-update direction is the steepest-descent direction in the L_2-norm. On the initial iterate, the two directions coincide but become nearly orthogonal at subsequent iterates, differing by an average of 89 degrees. In general, the first update moves the first iterate into the floor of the canyon and the remaining updates move the iterates along the floor until the solution point is found. For this and other non-linear problems, the curved trajectory of the canyon floor severly limits progression along the update direction. Only a small value of T_k is required to send the progression quickly up the steep canyon walls.

Applying the same initial guess and iterate generation scheme in the Nu-norm results in five iterations. Figures 3a-b show the contour graphs of the Nu-norm for the initial guess and the second iterate. The canyon structure still appears in the these contour graphs but, because steepest-descent and Newton update are always coincident at the iterate, each update-direction faces into the canyon and never lies on the canyon floor.

Figure 4 shows three test circuits used in our comparison of the L_2- and Nu-norms with SPICE. Each transistor is modeled with a three terminal, seven node topology detailed in reference [5] . Circuit #1 is a simple driver connected to a low conductance load, Circuit #2 is a series of five chained E/D inverter stages, and Circuit #3 is composed of ten enhancement transistors stacked on top of one another (drain to source) with a common gate potential (Only five transistors are depicted for a more compact presentation). Circuits #2 and #3 represent some of the most common circuit configurations as well as some of the most difficult in terms of convergence.

Table 1 compares the performance of the two norms and the standard source-stepping technique on each of the test circuits. Each circuit is converged to sixteen different bias conditions and the minimum, the average and the maximum number of iterations for each circuit is given by Table 1.

Comparison of Methods			
Method	Iterations <minimum, average, maximum>		
	Circuit #1	Circuit #2	Circuit #3
Source Stepping	< 3, 165, 648>	< 5, 21, 35>	< 3, 3860,31230>
Norm Reducing (N_{L_2})	< 2, 7, 12>	< 2, 47, 200>	< 2, 14, 53>
Norm Reducing (N_{NU})	< 2, 9, 12>	< 2, 15, 31>	< 2, 11, 28>

Table 1: Comparison of Methods

The source-stepping approach gives the worst overall performance for all three circuits and the Nu-norm based method give the best. The greatest improvement occurs in the high-gain configuration of circuit #2 where the Nu-norm takes a third to a seventh of the Newton steps required by the L_2-norm.

References

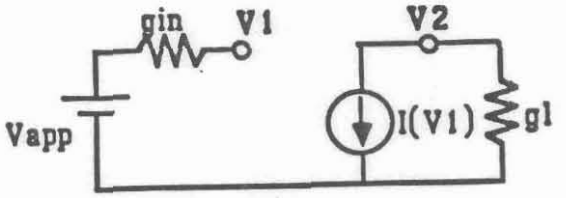
1. Y. C. Yuan, D. Divekar, and R. I. Dowell, "Modular-SPICE -- A Modular Circuit Simulation Program," *Proceedings of the ICCAD*, p. 250, Nov. 1983. (Program developed for Hewlett-Packard internal use.)

2. T. Quarles, A. R. Newton, D. O. Pederson, and A. Sangiovanni-Vincentelli, *SPICE 3A7 User's Guide*, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, April 1986.

3. F. H. Branin, G. R. Hogsett, R. L. Lunde, and L. E. Kugel, "ECAP-II -- A New Electronic Circuit Analysis Program," *IEEE Journal of Solid-State Circuits*, vol. SC-6, pp. 146-166, IEEE, 1971.

4. R. E. Banks and D. J. Rose, "Global Approximate Newton Methods," *Numersiche Mathematik*, vol. 37, pp. 279-295, 1981.

5. Hal R. Yeager and R. W. Dutton, "Circuit Simulation Models for the High Electron Mobility Transistor," *IEEE Transactions on Electron Devices*, vol. ED-33, no. 5, pp. 682-691, IEEE, May 1986.



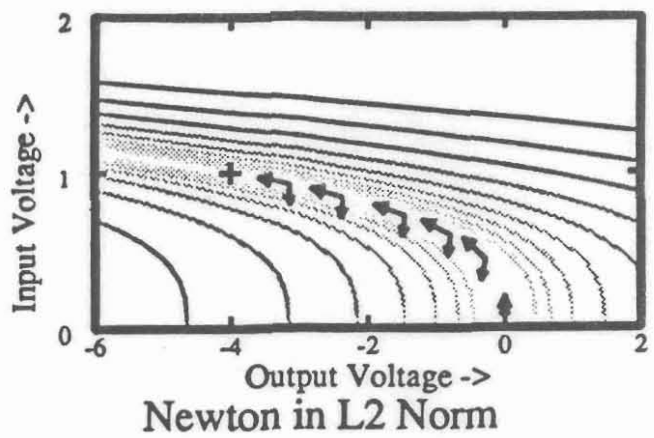
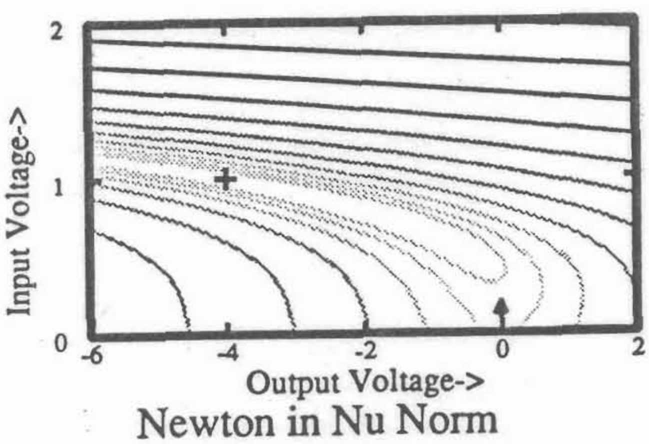


Figure 1: Common transconductance topology used to model FET and bipolar transistor current gain.

Figure 2: Contour graph of the L_2-norm with the nonlinear transconductance model as a function of the input (V_1) and output (V_2) voltages. The contour lines are draw at logarithmic intervals. Large arrows are the Newton update directions and the small arrows are the steepest descent directions.



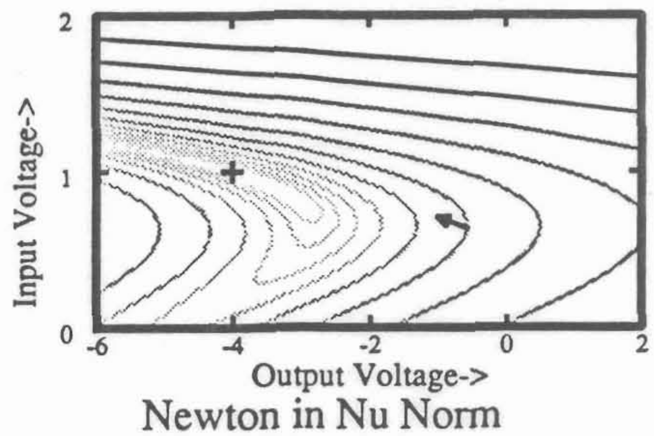


Figure 3a: Contour graph of the Nu-norm for the nonlinear transconductance model as seen at the initial guess. The norm is presented as a function of the input (V_1) and output (V_2) voltages. The contour lines are draw at logarithmic intervals.

Figure 3b: Contour graph of the Nu-norm for the nonlinear transconductance model as seen at the second iterate. The norm is presented as a function of the input (V_1) and output (V_2) voltages. The contour lines are draw at logarithmic intervals.

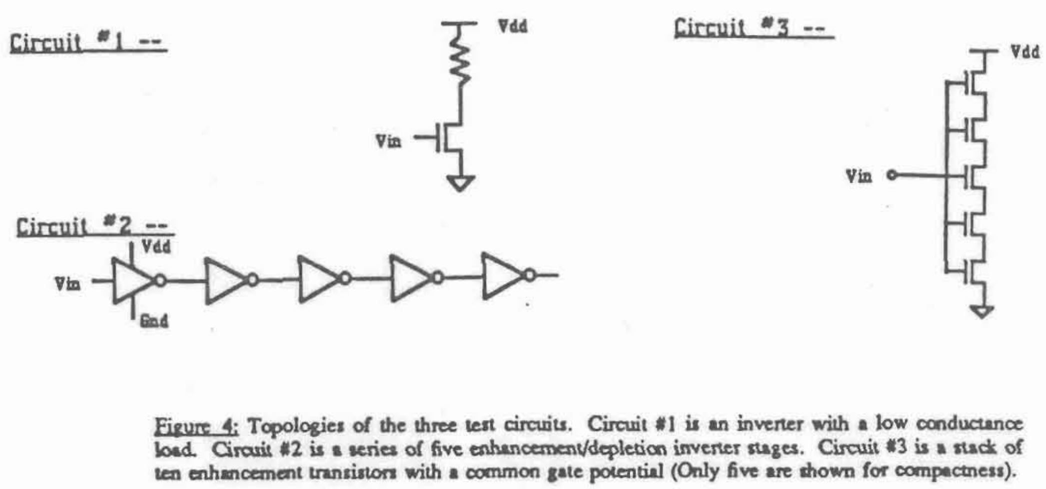


Figure 4: Topologies of the three test circuits. Circuit #1 is an inverter with a low conductance load. Circuit #2 is a series of five enhancement/depletion inverter stages. Circuit #3 is a stack of ten enhancement transistors with a common gate potential (Only five are shown for compactness).