

BRASIL: The Braunschweig Mixed-Mode-Simulator for Integrated Circuits

Ulrich Bretthauer and Ernst-Helmut Horneber
Institute for Network Theory and Circuit Design
Technical University of Braunschweig
Langer Kamp 19c, 38106 Braunschweig, Germany
U.Bretthauer@tu-bs.de, EH.Horneber@tu-bs.de

Abstract

BRASIL consists of a timing simulator for digital MOS circuits coupled with an algorithm for circuit simulation. The timing simulation is based upon a fast macromodelling approach and the calculation of time-variant RC networks. The circuit simulator takes advantage out of structuring the system of nodal equations. With BRASIL a fast simulation of large circuits, with special regards to systems with the need of higher accuracy, is possible.

1. Introduction

In the field of simulation of integrated circuits the timing simulation became essential for the verification of digital systems on the transistor level. A major drawback of most algorithms is that they are restricted to circuits with driver load structure. For circuits or subsystems which can not suitably be modelled on the timing level like analog sub-circuits, a second simulator is needed, which is in the most cases a SPICE-like circuit simulator. To calculate the whole system in one environment, a coupling of these simulators is necessary. The external coupling of a timing simulator with a circuit simulator seems not to be very effective because timing algorithms control the simulation flow with an event scheduler, whereas circuit simulators use time step mechanisms. The BRAunschweig-SImuLator BRASIL is based upon the internal coupling of a switch level timing algorithm with a circuit simulator. For the calculation of delay-times of logic gates, a fast macromodelling approach can be used. For more detailed information or more flexible circuit configurations, a second timing-mode builds a network of capacitances and time-variant conductances. The resulting system of differential equations is solved with numerical integration. To obtain a higher flexibility a circuit simulator is implemented in BRASIL, which directly interacts with the timing simulation modes. In this paper we present a survey of the simulation modes and the concept behind the coupling of the algorithms in BRASIL. Typical

results are shown to demonstrate the capability of calculating large circuits.

2. Timing simulation with BRASIL

The timing algorithms in BRASIL are suited for the calculation of digital MOS circuits. In a presimulation phase a switch-level-simulator similar to MOSSIM II by R. E. Bryant [1] determines the steady state of the circuit. With the knowledge of the logical states of each node in the network all transistors in the circuit are set to their corresponding states. With this information, a dynamic partitioning breaks down the circuit into a number of small networks or stages, bounded by gate terminals of further transistors, voltage sources and drain-source-channels of transistors in the cut-off region.

For the application of the macromodelling approach a stage has to consist of a pull-up and a pull-down network. A number of modelling parameters have to be extracted and the delay-time of the stage can be approximated. With the logical trigger voltage V_{tr} , the rise time t_{in} of the input-signal, the load capacitance C_L , the pull-up resistor R_L and the pull-down resistor R_D of the stage, the delay-time t_d can be normalized by [4]:

$$t'_d = \frac{t_d(t'_{in})}{f_d \left(C_L^{ecl_d} R_L^{erl_d} R_D^{erd_d} V_{tr}^{evtr_d} \right)}, \quad (1)$$

using the empirical exponents exx_d . The rise time of the input-signal t_{in} has to be normalized as well:

$$t'_{in} = \frac{t_{in}}{f_{in} \left(C_L^{ecl_{in}} R_L^{erl_{in}} R_D^{erd_{in}} V_{tr}^{evtr_{in}} \right)}. \quad (2)$$

The result of this normalization is the transformation of the wide range of possible delay-times into a narrow band. Figure 1 shows as an example the delay-time plotted against the rise time of a single inverter, with the load capacitance as a parameter. In figure 2 the effect of the normalization is shown, that is that the curves are shrunk in a bounded range, for which a mean value can be stored as reference curve.

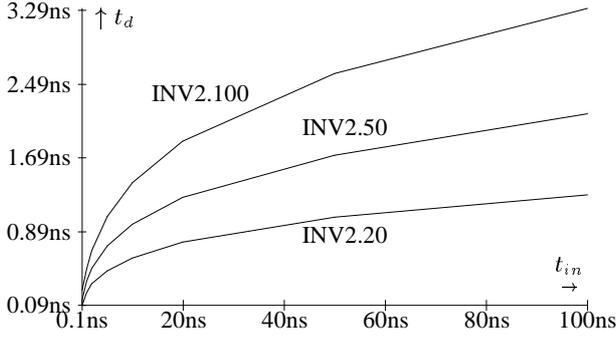


Figure 1. Delay-time for different capacitive loads

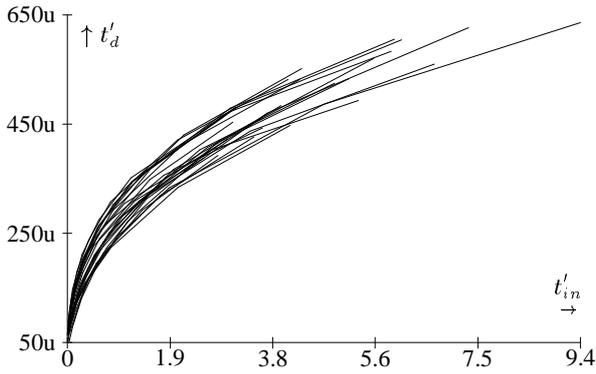


Figure 2. Normalized Delay-time

As soon as the empirical exponents $exx_{[d,in]}$ for a specific technology are evaluated, the macromodelling approach can be applied to digital MOS circuits. BRASIL examines the class of the circuit, NMOS or CMOS, checks the assumptions for the macromodelling and determines the modelling parameters. With a valid reference curve and a denormalization of t'_d an approximation of the delay-time is obtained.

For circuits which don't match the assumptions for this macromodelling procedure, BRASIL applies a more flexible approach to the timing analysis [5]. For this algorithm BRASIL replaces the drain-source-channel of every single transistor in the subcircuit by a time-variant conductance. It uses the equation for the drain current for a transistor in the linear region of the SPICE-level 1-model [3], to write the conductance of the channel as:

$$G(v_{gs}, v_{ds}) = K \left[v_{gs} - V_t - \frac{1}{2} v_{ds} \right] \quad (3)$$

in terms of the gate-source-voltage v_{gs} , the drain-source-voltage v_{ds} , the threshold-voltage V_t and a transconductance parameter K . With the assumption, that all input signals of the stage are known the gradients of $G(v_{gs}, v_{ds})$ are evaluated. All input voltages are approximated by ramps,

for which the starting-points, the end-points and the slopes are calculated. With these values the conductances of the drain-source-channels in the stage are modeled by time-variant conductances $G(t)$:

$$G(t) = K \left[v_{gs}(t) - V_t - \frac{1}{2} v_d(t) - \frac{1}{2} v_s(t) \right]. \quad (4)$$

If the subcircuit contains any transistors with unknown gate-voltages, for example depletion-type transistors whose gates are shorted to their sources, the conductance of the drain-source-channel has to be adjusted during the simulation.

In figure 3 this approach is compared with the real waveform calculated by SPICE. The upper diagram shows the shape of the gate-voltage of an NMOS-transistor. The ramp fits the curve where $v_g = 0.4$ VDD and $v_g = 0.6$ VDD and determines the starting-point $t_{r0,[f,r]}$ and the ending-point $t_{r1,[f,r]}$. The ramp for the conductance $G(t)$ is assumed to start simultaneously if the voltage is decreasing. The end of the ramp $t_{g1,f}$ is calculated with the aid of the known voltages at the end of the transition and the duration of the voltage ramp. For a rising voltage, the edge of the ramp for $G(t)$ ends at the same time $t_{r1,r} = t_{g1,r}$ and the starting point $t_{g0,r}$ is calculated backwards.

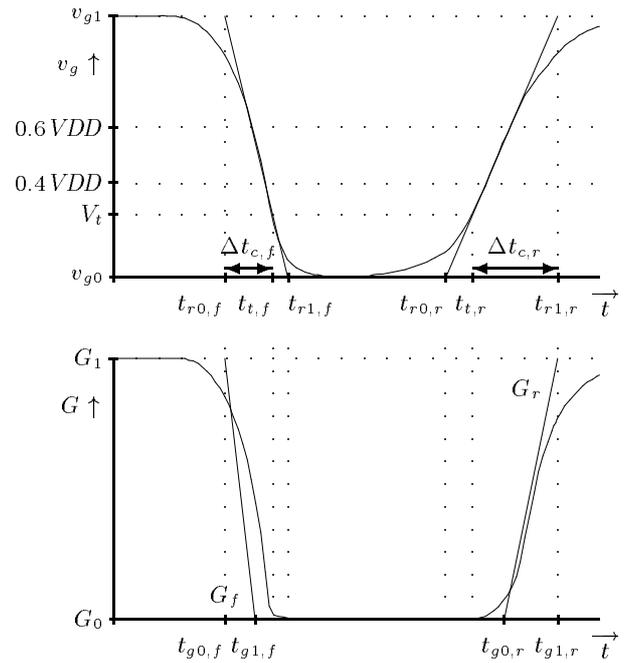


Figure 3. Scope of gate-voltage and drain-source-conductance of a NMOS-FET

This approach of modelling the drain-source-channel of switching transistors with time-variable conductances leads to a linear time-variant GC-network. The capacitances are

the sums of the parasitic capacitances of the transistors and all capacitances tied to a node. Figure 4 shows a typical cut-out for one single node i .

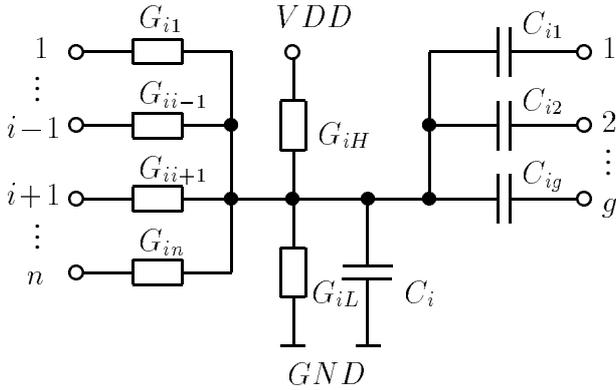


Figure 4. Vicinity of node i in the GC-network

Since the dynamic partitioning of BRASIL builds small subcircuits of conducting drain-source-channels, the capacitances C_{ig} provide coupling to external nodes. These nodes are assumed to be already simulated, therefore they are substituted by voltage sources. Performing a nodal analysis on the whole network results in a linear time-variant system of differential equations:

$$C_i \dot{v}_i + G_{i,P} v_i = G_{i,H} V_{DD} + \sum_{\substack{j=1 \\ j \neq i}}^n G_{i,j} v_j + \sum_{j=1}^g C_{i,j} \dot{v}_j, \quad (5)$$

$$G_{i,P} = G_{i,H} + G_{i,L} + \sum_{\substack{j=1 \\ j \neq i}}^n G_{i,j}$$

for $i \in 1, 2, \dots, g$. BRASIL solves this system of differential equations by numerical integration.

With the combination of the macromodelling and the timing-analysis, BRASIL is about 2 orders of magnitude faster than SPICE, with the ability of simulating circuits with more than 100 000 MOS-transistors. The timing-error is typically less than 15 %, compared to SPICE.

3. Circuit simulation in BRASIL

For the simulation of mixed digital-analog systems or circuits with bipolar transistors, BRASIL disposes of an internal circuit simulator. While the timing algorithms are intended for very large circuits, the circuit simulator is designed for the calculation of circuits and subsystems of medium size. A particular scheme is recognized regarding the pattern of a transistor in the matrix of a nodal analysis. The stamp for the Ebers-Moll-model with parasitic series resistances is shown in figure 5 a). The inner nodes B' , C' and E' have to be added for every transistor with

finite conductance of its base-, collector- or emitter-region respectively. The tight coupling between these nodes via the inner transistor leads to a completely filled 3×3 block and the connection to their corresponding outer nodes results in three 3×3 diagonal-matrices.

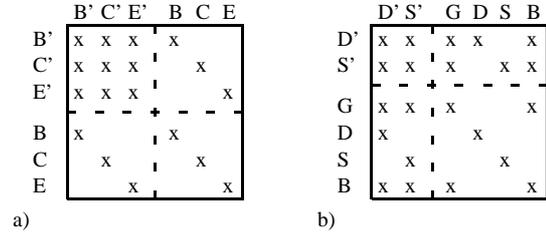


Figure 5. Matrix stamps for a) bipolar transistors and b) MOS transistors

For MOS transistors the structure of the matrix entries is similar. Because of the infinite input resistance of a MOS transistor, there is no inner node for the gate terminal, but for drain and source two additional nodes appear. Similarly the bulk region has no parasitic resistance, since it is connected by reverse-biased diodes. In figure 5 b) there is once again a completely filled 2×2 block, derived from the entries of the inner nodes. Their connection to their corresponding outer nodes, to the gate and the bulk leads to the two 2×4 blocks. The remaining 4×4 block doesn't show a diagonal form, because of capacitive coupling of gate and bulk, resulting in two entries.

Ordering the system of nodal equations following these patterns leads to a bordered block-diagonal matrix. The LU-factorization of such a partitioned system of equations can be done by solving four partial matrices [2]:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \quad (6)$$

with \mathbf{A}_{11} as block-diagonal-matrix and \mathbf{A}_{12} , \mathbf{A}_{21} and \mathbf{A}_{22} as corresponding parts of the border. The LU-factorization now solves the particular blocks:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{11} & \\ & \mathbf{L}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} \\ & \mathbf{U}_{22} \end{bmatrix}, \quad (7)$$

with \mathbf{L}_{11} and \mathbf{L}_{22} in lower, \mathbf{U}_{11} and \mathbf{U}_{22} in upper triangular-form. The decomposition can be done by a LU-factorization of \mathbf{A}_{11} , two forward substitutions to obtain \mathbf{U}_{12} and \mathbf{L}_{21} and another LU-factorization of $\mathbf{A}_{22} - \mathbf{L}_{21}\mathbf{U}_{12}$. In addition, the matrix \mathbf{A}_{11} is of blockdiagonal form, with block-sizes of 2×2 and 3×3 , so this transformation can also be done in several independent LU-factorizations, which were fixed implemented due to the known size of the blocks.

This formulation and solving procedure for a system of equations in blockdiagonal form has advantages over sparse-matrix-algorithms calculating circuits up to 50 MOS-transistors or 100 transistors in BiCMOS-technology respectively. So it is well suited for the fast simulation of subnetworks in a mixed-mode-simulator like BRASIL.

Nonlinear circuit elements are iteratively linearized by the well known Newton-Raphson method, which can show convergence problems. The simulation of small typical examples has shown, that in practical circuits the strong nonlinear equations of a bipolar transistor speed up the process of convergence, whereas CMOS inverters with their quadratic characteristic need more iterations. For that reason, BRASIL uses the NR method together with a damping scheme. The common definition of the (n+1)-th iteration of the NR-algorithm for a systems of nonlinear equations $\mathbf{f}(\mathbf{x}) = 0$ is:

$$\mathbf{J}^{(n)}\mathbf{x}^{(n+1)} = -\mathbf{f}(\mathbf{x}^{(n)}) + \mathbf{J}^{(n)}\mathbf{x}^{(n)}, \quad (8)$$

with $\mathbf{x}^{(n+1)} = \mathbf{x}^{(n)} + \Delta\mathbf{x}^{(n+1)}$ and \mathbf{J} as the Jacobian matrix of $\mathbf{f}(\mathbf{x})$. An approach for controlling the size of $\Delta\mathbf{x}$ is the multiplication with a diagonal matrix \mathbf{D} with damping or accelerating factors for the components of \mathbf{x} : $\mathbf{x}^{(n+1)} = \mathbf{x}^{(n)} + \mathbf{D}\Delta\mathbf{x}^{(n+1)}$. Thus the iterative equation has to be written as

$$\mathbf{J}^{(n)}\mathbf{D}^{-1}\mathbf{x}^{(n+1)} = -\mathbf{f}(\mathbf{x}^{(n)}) + \mathbf{J}^{(n)}\mathbf{D}^{-1}\mathbf{x}^{(n)}. \quad (9)$$

The coefficients $d_k^{(n)}$ are evaluated empirically for every node connected to a MOS or bipolar transistor. They gradually approach 1 in the iteration progress. Experiments showed that for the exponential characteristics of bipolar transistors damping is not necessary. For circuits in CMOS technology however damping factors greater than 1 speed up convergence. It should be mentioned, that with this modified NR-algorithm the stamps for linear network elements must also be rewritten.

4. Coupling of the simulation algorithms

For the coupling of the different algorithms in BRASIL an effective interaction of the event-driven timing simulator and the circuit simulator with its time step controlling scheme is essential. In the pre-simulation phase it must be determined which parts of the circuit require calculation to a higher accuracy. Since BRASIL dispose of an interface for netlists in SPICE-language [3], the user specifies the nodes or the subcircuit he wishes to simulate with the circuit simulator within the circuit description. Furthermore a rule-based investigation of the transistor functions separates logic gates, transfer gates and analog blocks. The analog blocks and subcircuits with bipolar transistors are automatically labeled for the circuit simulator. A partitioning

is also applied, which uses well defined borders to prevent these parts from spreading over the entire network. The timing simulator performs a dynamic partitioning during the simulation. Boundaries of a subcircuit are defined by input nodes, gate terminals of MOS transistors and drain- and source-terminals of transistors in the "OFF"-region with $v_{gs} < V_t$. It is reasonable to adopt this procedure for the static partitioning at the beginning of the simulation. Because this distribution is valid throughout the whole simulation, drain-source-channels are rejected as borders between the simulation-levels.

The circuit elements at the borders of the partitions have to be modeled with network elements adapted to both algorithms. For the timing algorithms, the gates of MOS transistors in the area of the circuit simulation can be treated as constant capacitors. For the more accurate circuit simulation, this approximation may cause undesirable errors, but in the decision either to enlarge the domain for the circuit simulator or to tolerate this loss of exactness, the latter is preferred.

Since the timing simulator calculates node voltages as well as logic states, there is no need to transform different representations of signals. Only a denormalization is necessary, because timing algorithms work with voltages normalized to the supply voltage VDD. In the circuit simulation, these voltages are acting as piecewise-linear voltage sources. The relationships getting logic states out of node voltages are straight forward, a simple conversion from circuit- to logic-timing-level is all that is required.

The controlling of the simulation sequence in BRASIL is done by an event-driven scheduler, whereas other circuit simulators work with time-step-controlling mechanisms. While the timing algorithms remain the main simulators, the event-driven scheme is preserved. If an event for the input of a subcircuit has been marked for the circuit simulator, an operation-point analysis has to be performed. A second operation-point analysis calculates the steady-state of the subcircuit. This information is needed, to terminate the simulation of subcircuits. After all node voltages have reached their final value, the control of the simulation flow is given back to the event scheduler. This procedure has the advantage that phases, in which the network is idle, don't have to be calculated.

5. Examples

As an example to show the ability to calculate circuits with MOS and bipolar transistors, a small logic circuit in BiCMOS technology is simulated. Figure 6 shows the schematic containing two NAND gates and two inverters. One of those (inv1) is realized in BiCMOS technology. In the prephase of the simulation a subcircuit is build up, containing the bipolar transistors of *inv1* with their surrounding MOS-transistors. Due to the rules for the partitioning, the subcircuit matches *inv1*. In figure 7 the voltage at the out-

put node *out1* is printed. $/OUT1$ is calculated by BRASIL and $V(OUT1)$ by HSPICE. BRASIL needs 2.4 seconds, whereas HSPICE takes 5.7 seconds of user-time.

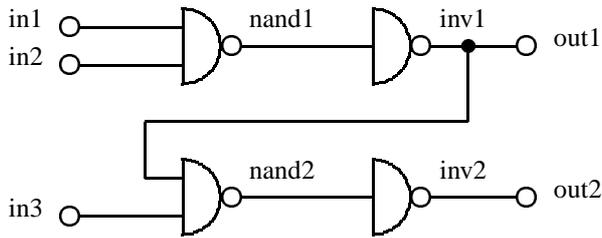


Figure 6. Logic-circuit with BiCMOS-inverter

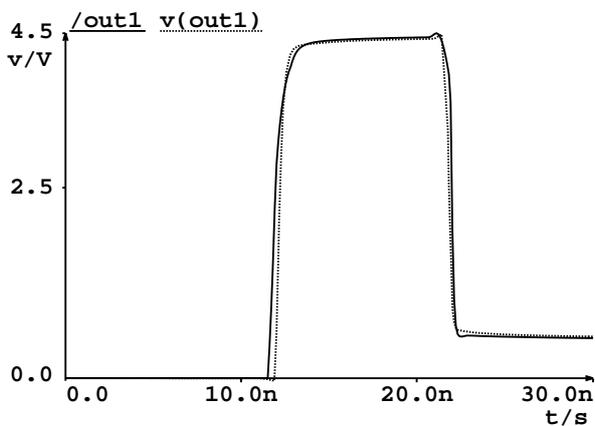


Figure 7. Voltage at node *out1*

Consider a larger circuit, a 16-bit multiplier in CMOS technology. It is built of 8 half- and 48 full-adders, with 2464 transistors and 1953 nodes. For BRASIL 20 subcircuits together with 238 transistors are chosen to be simulated with the circuit simulator. In figure 8 a section of the voltage curve at an output node of the multiplier is shown. $/P10$ is calculated by the circuit simulator in BRASIL and $V(P10)$ results from a simulation by HSPICE. Although the signals have to pass up to ten gates, before $/P10$ is generated, no substantial error is apparent. The hazard, when the voltage at $P10$ drops to $0V$, before it returns to $5V$, it's final value, is detected and a potential source of logical malfunctions can be removed. On a DECstation 5000/200 under ULTRIX 4.2 HSPICE took 18508.4 seconds user-time and 7018 kB memory, whereas a timing simulation needs 10.0 seconds and 1557 kB. The mixed-mode-simulation taking 1937 seconds and 2049 kB occupies the middle position.

6. Conclusion

Traditional timing simulators are limited to digital circuits of MOS transistors. This restriction is overcome by the

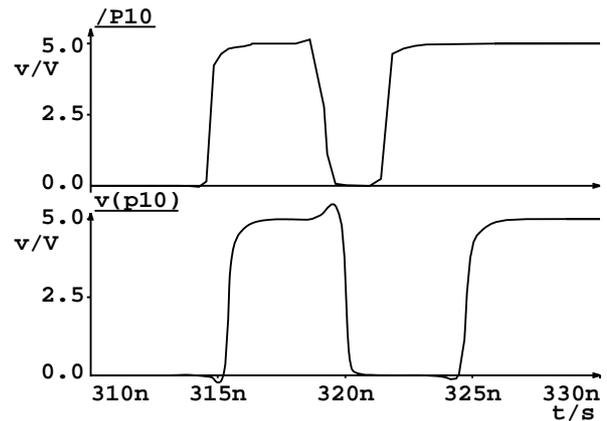


Figure 8. Results of a mixed-mode-simulation by BRASIL

combination of a switch level timing algorithm with a circuit simulator in BRASIL. The delay times of logic gates can be estimated with a fast macromodelling approach, which do not need a library because the model parameters are empirical for every special technology. The timing behavior of digital MOS circuits can be evaluated with an algorithm, using time-variant GC-networks. This algorithm accomplishes flexible circuit-designs e.g. circuits with transfer gates or pass-transistors. The timing error is typically less than 15 % when compared to SPICE, with a gain in CPU-time of 2 orders of magnitude. With an algorithm for circuit simulation BRASIL has the ability for calculating subcircuits with higher accuracy.

7. Acknowledgment

The authors would like to thank Dr. D. Sass, who contributed significantly to the current version of the program. Thanks also to M. Leberecht and U. Klaperski for their work on the circuit simulator.

References

- [1] R. E. Bryant. A switch-level model and simulator for mos digital systems. *IEEE Transactions on Computers*, c-33(2):160–177, February 1984.
- [2] I. S. Duff, A. Erisman, and M. K. Reid. *Direct methods for sparse matrices*. Oxford University Press, Oxford, 1992.
- [3] L. W. Nagel. *SPICE2: A computer program to simulate semiconductor circuits*. Rep. UCB/ERL M520, Univ. of Calif, Berkeley, 1975.
- [4] D. Sass. *Logik- und Timing-Simulation digitaler MOS-Schaltungen auf Transistorebene*. PhD thesis, Technical University of Braunschweig, 1994. Available as Fortschritt-Berichte VDI, VDI Verlag, Düsseldorf, FRG.
- [5] D. Sass and E.-H. Horneber. Modeling digital mos circuits for timing simulation using time-variant rc networks. In *European Conference on Circuit Theory and Design*, Davos, 1993.