Extracting Circuit Models for Large RC Interconnections that are Accurate up to a Predefined Signal Frequency

P.J.H. Elias and N.P. van der Meijs
Faculty of Electrical Engineering
Delft University of Technology
P.O. Box 5031, 2600 GA Delft,The Netherlands
elias@cas.et.tudelft.nl

Abstract

This paper presents a technique that transforms large and complex RC networks into much smaller, physically realizable RC networks. These models reflect the transmission behavior of the initial network accurately for frequencies up to a user-defined maximal signal frequency. This technique has been incorporated in a layout-to-circuit extractor, using a scan-line approach. The method guarantees numerical stability and performs excellently in modeling RC interconnects.

1 Introduction

This paper presents a solution to the major problem how to obtain a simple, passive and physically realizable RC network that matches the transmission behavior of a given very large and complex RC network up to a predefined maximal signal frequency. This problem is particularly acute in (3-D) layout-to-circuit extraction, where appropriate models for the IC interconnections in VLSI designs have to be generated.

In most VLSI circuits, the IC interconnections and their parasitics play an essential role in the overall behavior of the circuits. Therefore accurate modeling is essential. Commonly, in existing extractors, the interconnections are subdivided into elements and each element is replaced by a lumped RC section. Initially, the number of sections must be sufficiently large to guarantee that the distributed properties of the interconnections are reflected accurately by the resulting network. This can be achieved by using a Finite Element Mesh, on which the Laplace equation is solved in order to obtain values for the resistors and capacitors at each section (e.g. [1]). However the total number of sections extracted in VLSI requires an enormous amount of memory, making an efficient timing analysis and verification of the circuit afterwards virtually impossible. To deal with this problem, a model of the interconnections has to be obtained that has a much lower complexity but that displays approximately the same transmission behavior. A method like AWE [2] in theory offers a solution to the problem: by reducing the large set of system poles to a limited set of approximated poles, the circuit is simplified, while its behavior essentially is retained. However AWE has two important draw-backs. First, poles are obtained, not an approximating circuit. It is possible to obtain macro-models from these poles [3], but these models cannot be used in standard circuit simulators. An actual circuit model would be much more attractive. Not only could such a model be used directly in standard circuit simulators, it would also give information that is much easier to interpret by a designer. A second important draw-back is that AWE cannot guarantee numerical stability.

A method that does not have the disadvantages of AWE but yields only an adequate low-frequency model, has been presented in [1] and [4]. This method yields a simple full-graph network between the terminals. To guarantee a close resemblance between the transmission behavior of the initial and the final network, this method preserves the values of the Elmore delay time, the total resistance between each pair of terminal nodes and the total ground capacitance. Therefore, this method will be referred to as the Elmore delay Preserving Reduction (EPR) method.

The major disadvantage of EPR is that it yields only a first-order approximation (low frequency) of the initial network. At high signal frequencies this approximation may not be sufficient. This is a severe limitation as high-frequency effects become increasingly relevant. A second disadvantage, especially from a designers point of view, is that EPR always yields a full graph between the terminals, loosing the original physical structure completely. All this is due to the fact that EPR eliminates *all* non-terminal nodes, and in many cases that simply is too many.

In this paper we present a method that *selectively* removes non-terminal nodes: only nodes that are considered not essential in order to describe the network up to the desired accuracy at a given frequency will be eliminated. This way also the original interconnection topology can be retained, as will be shown further on. In order to decide which nodes are essential in a given configuration, the method calculates for each non-terminal node an estimation of the relative error made if that node should be eliminated. Only nodes of which that error does not exceed a predefined tolerance actually are eliminated. By continuously recalculating the errors of the remaining nodes, the method takes into account

33rd Design Automation Conference ®

DAC 96 - 06/96 Las Vegas, NV, USA

that the relative importance of the nodes increases with decreasing number of remaining nodes. Finally, the error of all remaining nodes will exceed the tolerance, and these nodes will be considered essential in the approximating model.

As will be shown in the next section, the relative error of each node is a function of the frequency at which the circuit will be operated. Now the maximal signal frequency f_s is introduced as a parameter, and the method will guarantee that the model extracted is an accurate representation of the original large network, as long as the frequency at which the circuit will be operated does not exceed f_s. This means in general that at high f_s a more detailed circuit will be extracted than at low f_s.

Because non-terminal nodes are eliminated selectively, a large interconnection network always will yield another, much smaller passive network, of which the system poles are good approximations of the original system poles. Although in principle the method is also capable of dealing with inductive networks, we will limit ourselves in this paper to RC circuits. Like EPR, this method also preserves the Elmore time-delay, the total resistance between terminals, the total ground capacitance of each line, and the total coupling capacitance between different lines.

2 Error Function

The algorithm proposed essentially is as follows:

1. Calculate for each node of a given initial network an estimation of the relative error made if that node should be eliminated. This error is a function of the maximal signal frequency f_s.
2. The node with the lowest error is eliminated.
3. Due to the elimination of a single node of the network, the relative importance of the remaining nodes increases. In other words, the relative error made if one of the remaining nodes should be eliminated increases. For this reason it is essential that, after a node has been eliminated, the increased relative errors of all nodes connected to the eliminated node are recalculated. By continuously updating the relative errors, we guarantee the relative error of the remaining nodes to be relevant with respect to the *original* network.
4. Steps (2) and (3) are repeated until the error of all remaining non-terminal nodes exceeds a predefined value, the tolerance δ_{max}. These remaining nodes are considered to be essential in order to describe the network up to the desired accuracy at given f_s.

In practice, for very large networks, this basic algorithm needs several refinements in order to be efficient in computer memory and cpu-time. That will be dealt with in the next section. In this section we will concentrate on the calculation of the estimated relative error assigned to each node, and the elimination procedure itself.

The error function uses higher-order moments theory and can be thought of as follows: suppose an admittance Y_{ij} is present between two coupled nodes i and j. This admittance is represented as a series expansion in s about zero. If this admittance is simply a conductor, only the zeroth-order moment is non-zero (and positive); if the admittance is a capacitor, only the first-order moment is non-zero; and if a conductor and a capacitor are in parallel, both the zeroth and the first-order moments are non-zero (and positive). Now consider a third node k that is connected to both i and j. If node k is eliminated using Gaussian elimination, an extra shunt $\tilde{Y}_{ij}$ parallel to Y_{ij} is generated in order to conserve the node information. In general this shunt is a rational function in s, of which a MacLaurin series expansion can be given. As the interconnections are considered to be essentially low pass, the lower-order moments dominate the higher-order moments. We want the interconnection to be represented by resistors and capacitors only, i.e. the zeroth and first-order moments. However, the higher-order moments in general have non-zero values, so an error is being made if the higher-order moments of the series expansion are neglected completely. An estimation of the relative importance of this error will be called the shunt error δ_{ij}, and the following definition will be used:

$$\delta_{ij} = \lim_{s \to j\omega_s} \mid \frac{s^2 \tilde{M}_{ij}^{(2)}}{\overline{M}_{ij}^{(0)} + s\overline{M}_{ij}^{(1)}} \mid = \frac{\omega_s^2 \mid \tilde{M}_{ij}^{(2)} \mid}{\left(\mid \overline{M}_{ij}^{(0)} \mid^2 + \omega_s^2 \mid \overline{M}_{ij}^{(1)} \mid^2\right)^{\frac{1}{2}}} \quad (1)$$

with $\omega_s = 2\pi f_s$, $\tilde{M}_{ij}^{(l)}$ the l-th moment of $\tilde{Y}_{ij}$, and $\overline{M}_{ij}^{(l)}$ the l-th moment of $Y_{ij} + \tilde{Y}_{ij}$. These moments can be extracted from the network by using a specific node elimination procedure that preserves the moments exactly, up to a given order $\mathcal{N}$. This method has been described in detail elsewhere [5], and is summarized in the Appendix. In the problem at hand, only moments up to the second-order are needed, i.e. $\mathcal{N} = 2$.

Different and more detailed definitions than (1) are possible. The reasons for choosing this one are as follows:

- In definition (1) δ_{ij} is defined relative to the zeroth and first-order moments only, because, in principle, those can be represented by resistances and capacitances.
- Only the second-order moment is used in the error function. This is a first-order approximation of the error, minimizing the extra CPU time and memory needed.
- Experiments, as are given in the next sections, demonstrate this definition to be efficient and adequate.

If N_k is the degree of node k, i.e. node k is connected to N_k other nodes, $\frac{1}{2} N_k (N_k - 1)$ shunts will be generated by the elimination of k. For each shunt the error can be calculated using definition (1). These errors are considered to be the elements of the node error vector $\overline{\delta}_k$. Now the node error δ_k is defined as the norm of the vector $\overline{\delta}_k$. In this paper we will use the L_∞ norm, i.e. the maximal absolute value of the elements of $\overline{\delta}_k$. Again, other norm definitions are possible, and there seems no theoretical reason to prefer any of them. Our preference for the L_∞ norm is mainly based on its simplicity. Experiments show that $\delta_{max} = 0.05$ is a very good choice in combination with this norm.

If $\delta_k < \delta_{max}$, node k is eliminated, using the already mentioned moments preserving elimination technique (Appendix and [5]). If $\delta_k \geq \delta_{max}$, node k is considered to be essential in order to represent the original network.

After elimination of all non-terminal nodes having $\delta_k < \delta_{max}$, the following conclusions remain valid, due to the elimination procedure used:

- The relation between the steady-state potentials and the steady-state currents at the terminals of the final network is the same as the currents and potential at the terminals of the initial network.
- The total capacitance to ground for each conductor in the network is unchanged and the total coupling capacitance between two conductors is unchanged. The term *conductor* here refers to a subset of nodes within the initial network. In that subset there is at least one path via resistances between any two nodes, and there is no resistive path between two nodes situated in different subsets.
- For each directed pair of nodes that remain in the final network and that are on the same conductor, the value of the Elmore delay time is preserved.
- The moments obtained also preserve the multiple time-constants that are typical for charge-sharing models [6].

One more note has to be made. The purpose of this work is to find a small passive circuit that optimally represents a very large passive circuit, possibly of hundred of thousands of nodes and components. In that case, the calculation of the node error vectors demands a large amount of computational effort, especially while the dimension of the error vectors is quadratic in the degree of the node. Since, during the elimination process the degree of the remaining nodes may increase strongly, this can cause severe problems. By taking only the shunt errors connected to the ground node into account, the dimension of the error vector becomes linear in the total number of non-terminal nodes. Experiments show that an order of magnitude in time reduction can be obtained, without any loss in applicability of the method.

3 Implementation

The reduction method described in this paper has been implemented in the layout-to-circuit extractor Space. In the implementation, resistors and capacitors represent zeroth and first-order moments, respectively. To calculate the relative node errors, second-order moments are needed. By representing these second-order moments by second-order admittances parallel to the capacitors, no extra data-structures and little extra overhead is needed.

Space uses a scan-line technique to efficiently extract RC models. As the scan-line moves over the layout, the following different operations are performed:

1. The finite element mesh is constructed.
2. Device recognition (both MOS and bipolar) and characterization is realized.
3. The zeroth, first and second-order moments of each element are added to an intermediate admittance network, called the frontal network.

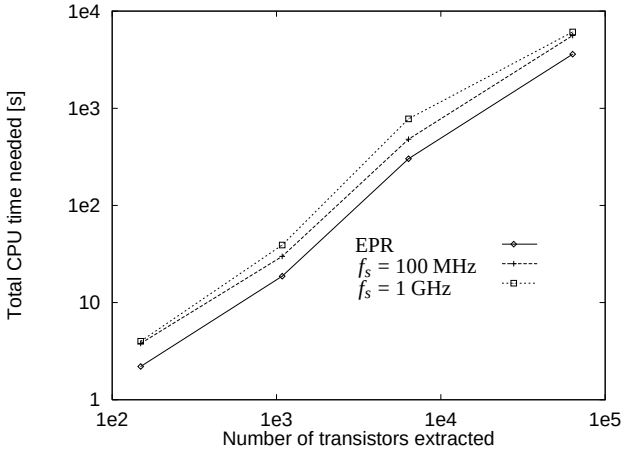

Figure 1: *Total extraction times on a HP 9000/735 of the test circuits as a function of number of transistors.*

4. After the scan-line has passed all elements that are connected to non-terminal node k, δ_k is calculated and node k is added to a priority queue with maximum length N_{max}. In this queue the nodes are ordered in increasing relative error. If the number of pending nodes exceeds N_{max}, the nodes with lowest δ_k are eliminated first. Nodes having (more or less) the same error are ordered in increasing degree in order to speedup the extraction considerably [7].
5. After the complete layout has been scanned, all remaining non-terminal nodes that have errors smaller than δ_{max} are eliminated.

To study the effect of higher order moments extraction on CPU time and memory use, four different circuits have been extracted from their layouts: a NMOS random counter (149 transistors), a full custom CMOS overflow detector (1,083 transistors), a writable logic array (6,360 transistors) and a Cordic processor (63,416 transistors). Of each, three different extractions have been executed and the CPU times and memory uses are plotted in Figures 1 and 2 respectively. These results include start-up, device and connectivity extraction and input/output. All resistive and (vertical) capacitive effects are extracted simultaneously. The time and memory use is plotted as a function of the number of transistors present.

The three different extractions are: (1) EPR extraction (up to first-order moments), (2-3) use error calculation with f_s =100 MHz and 1 GHz respectively, with δ_{max} = 5%. Due to the use of the priority queue and the optimalization of the order in which the nodes are eliminated, the cpu-time is nearly linear in the size of the layout (number of transistors). Figures 1 and 2 furthermore show that the second-order moments and the error functions are calculated at moderate extra costs—e.g. some 25% increase in memory and a doubling of cpu-time.

3.1 Examples

Two different structures are considered. The first example is a 3D structure consisting of a flat resistor, with a metal

Table 1: *Position of dominant poles as a function of the frequency parameter f_s. The number of poles equals the number of non-terminal nodes #N.*

	Exact	100 MHz	200 MHz	1 GHz	2 GHz
#N	107	1	2	4	5
1st pole	$8.734\ 10^8$	$7.618\ 10^8$	$8.385\ 10^8$	$8.676\ 10^8$	$8.740\ 10^8$
2nd pole	$9.894\ 10^9$		$8.386\ 10^9$	$1.045\ 10^{10}$	$9.975\ 10^9$
3rd pole	$2.440\ 10^{10}$			$2.113\ 10^{10}$	$2.529\ 10^{10}$
4rd pole	$4.317\ 10^{10}$			$4.370\ 10^{10}$	$4.560\ 10^{10}$

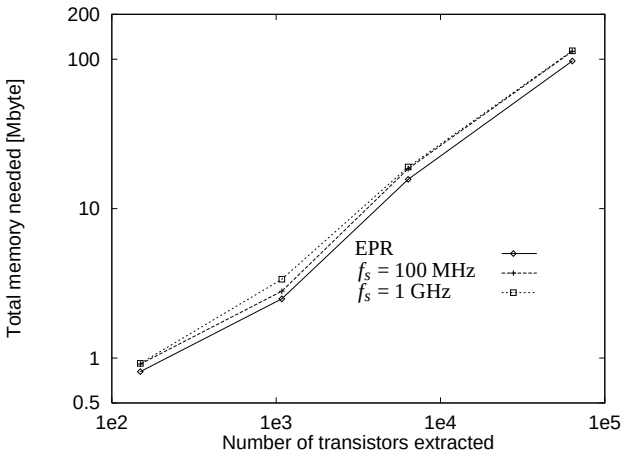

Figure 2: *Total memory use on a HP 9000/735 of the test circuits as a function of number of transistors.*

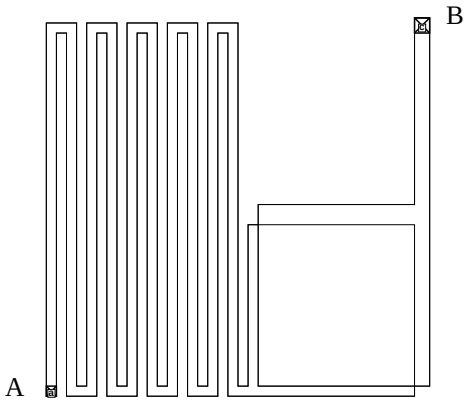

Figure 3: *Layout of a poly resistor (grey) on top of which a metal plate capacitor (black).*

plate capacitor on top of it. Secondly a 6-terminal line is considered. Resistors and Capacitances are extracted simultaneously, using a Finite Element Method for resistance extraction and a 3D Boundary Element Method for capacitance extraction. The grid size has been taken small enough to ensure that the distributed properties are accurately reflected by the extracted lumped RC mesh. In both examples the error functions are calculated with negligible extra time and some 2% extra memory requirements.

Of the first example the layout is given in Figure 3. This is a two terminal structure: terminal **A** at one end of the polysilicon spiral resistor, terminal **B** at the metal plate that is capacitively coupled to the resistor. The width of the resistor

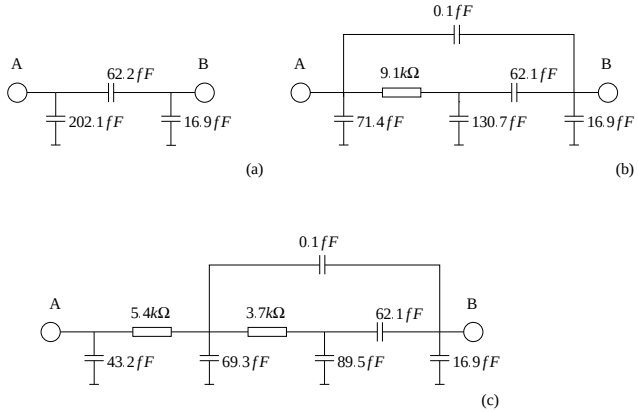

Figure 4: *Circuits extracted from layout of RC network: (a) EPR model (b) f_s = 100 MHz (c) f_s = 200 MHz*

equals 2 μm. Due to 3D capacitance extraction, lateral capacitive coupling between the resistor spirals is taken into account. This extraction yields an initial lumped RC mesh containing 109 nodes, 164 resistors and 720 capacitors. In EPR all non-terminal nodes are eliminated, and after elimination, the zeroth-order moments are represented by a resistor and the positive first-order moments by a capacitor. The resulting network is plotted in Figure 4a. Only capacitive couplings have survived the elimination process, and this minimal network is a low frequency representation of the original network. Now the elimination procedure is repeated, using the method described in this paper. The tolerance is valued δ_{max} = 5%. This value is fairly arbitrary, but experiments have proven this to be a very appropriate choice, for all cases studied. The signal frequency f_s is varied. It appears that if $f_s \leq 1$ MHz no extra non-terminal nodes are needed, in other words, if the signal frequency does not exceed 1 MHz, the circuit from Figure 4a is an accurate enough model. This is also demonstrated by Figures 5 and 6. In these plots the magnitude and phase of the trans-admittance Y_{AB} have been plotted for both the fully extracted network (109 nodes) and the EPR model. For $f < 1$ MHz the curves indeed coincide, as expected.

Further increase of f_s shows that, from f_s = 1 MHz up to 100 MHz, one non-terminal node is not eliminated. The resulting network has been plotted in Figure 4b. The total value of the spiral resistor exactly equals the value found if the resistor is considered without the extra capacitor. The total ground capacitance of the spiral resistor has been unchanged compared to the minimal EPR network. The same applies to the coupling capacitance and the metal ground capacitance. This means that the extra node is situated some-

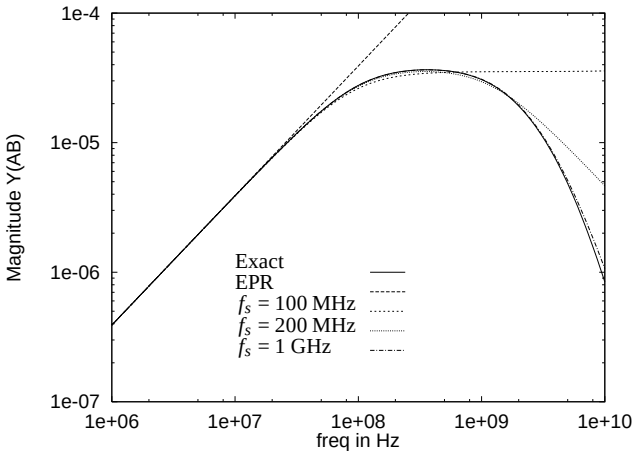

Figure 5: *Magnitude of the trans-admittance of the RC structure.*

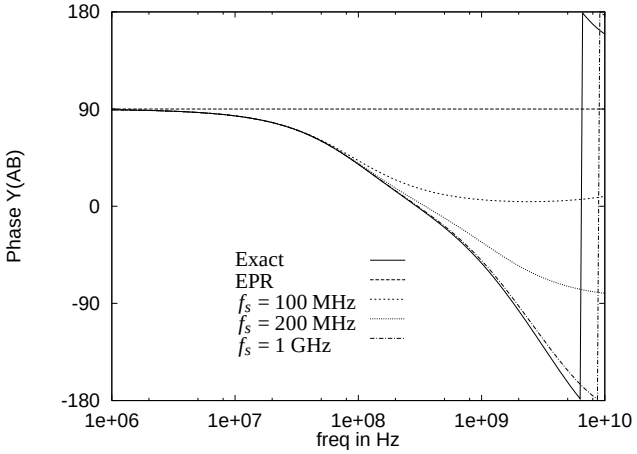

Figure 6: *Phase of the trans-admittance of the RC structure.*

where on the large pad at the end of the resistor, just below the metal pad. This shows that the method has successfully detected one of the crucial nodes. That is confirmed by Figures 5 and 6, where a very good agreement between the exact and the extracted network for f_s = 100 MHz is obtained for all $f < 100$MHz.

f_s is further increased, still with unchanged δ_{max}. At f_s = 200 MHz the method retains two extra nodes. After the elimination procedure all irrelevant, very large resistors (R $>$ 1 MΩ) and very small capacitors (C $<$ 0.1 fF) are neglected, yielding the network plotted in Figure 4c. The total resistance, coupling and ground capacitance are unchanged. At f_s = 1 GHz, three extra nodes are obtained. In all cases the fit with the exact results is good for $f < f_s$, as demonstrated by Figures 5 and 6.

In Table 1 the poles of the voltage transfer function of the exact network are compared to the poles obtained from the approximated networks, taking f_s = 100 MHz, 200 MHz, 1 GHz and 2 GHz respectively. These extractions leave 1, 2, 4 or 5 non-terminal nodes respectively. As can be seen from Table 1, the calculated poles converge to the exact poles.

The second example consists of a poly-silicon line connecting six terminals. The layout of the line is plotted in Figure 7. The width of the resistor equals 1 µm. A Finite Element Method has been used to calculate the capacitors and resistors. This extraction yields a lumped RC mesh containing 260 nodes, 433 resistors and 1366 capacitors. Taking f_s = 5 GHz, the number of non-terminal nodes that remain after extraction and elimination equals three. In principle the method yields a full graph between the 9 remaining nodes. However if all irrelevant resistors (R $>$ 1 MΩ) and capacitors (C $<$ 0.1 fF) are removed the simple network from Figure 8 is obtained. As can be seen, the original topological structure is perfectly reflected by the model. This example also shows that the introduction of extra nodes in the model can reduce the complexity of the total network by reducing the number of relevant components. This effect will be stronger if more terminals are connected to the line.

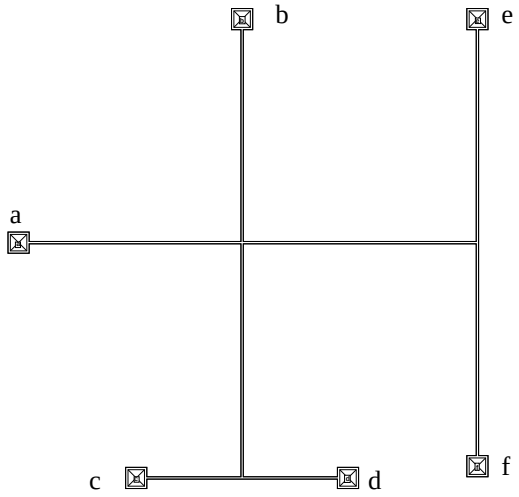

Figure 7: *Layout of a 6-terminal poly line.*

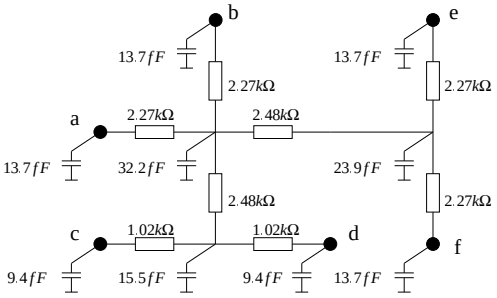

Figure 8: *Circuit extracted from layout of 6-terminal line with f_s = 5 GHz*

4 Conclusions

The method presented has some very important properties, that have not been encountered in any other method.

- It is capable of transforming large and complex RC circuits to much smaller, passive and physically realizable RC circuits, the transmission behavior of which accurately reflects the transmission behavior of the initial circuit up to a user defined maximal signal frequency.

- By increasing maximal signal frequency, the complexity of the extracted RC model also increases, such that the resulting model matches the original circuit up to the new frequency.

- Since the method uses the principle of selective elimination of the non-terminal nodes, in combination with linearized Gaussian elimination, the method always is numerically stable.

- The method is very suited to combine with scan-line oriented layout-to-circuit extractors.

- Implementation of the method in a layout-to-circuit extractor gives a moderate increase in cpu-time and memory consumption, if compared to much simpler extraction methods, such as EPR.

- Experiments show excellent results.

A Appendix

In this appendix we describe how to eliminate a single node from an initial linear and passive admittance network. After the elimination, an approximating admittance network results, while preserving the moments of the MacLaurin series expansion up to a given order $\mathcal{N}$. Applying this method iteratively, all non-terminal nodes of the network can be eliminated. Here, only a summary of the method is presented. For more detailed information we refer to [5].

The following definitions and assumptions are used:

- The initial network has N nodes. The relation between the node currents and the node potentials of the initial network is given by the admittance network $\mathbf{Y}(s)$. The admittance network after the elimination of non-terminal node k is given by $\overline{\mathbf{Y}}(s)$.

- Each element $Y_{ij}(s)$ of $\mathbf{Y}(s)$ is given as the first $\mathcal{N}$ terms of a truncated MacLaurin series expansion, yielding

$$Y_{ij}(s) = \begin{cases} \sum_{l=0}^{\mathcal{N}} M_{ij}^{(l)} s^l & (i \neq j) \\ \sum_{k=1, k\neq i}^{N} Y_{ik}(s) & (i = j) \end{cases} \tag{2}$$

The elimination procedure used is in fact Gaussian elimination. An important property of this is that, given the first $\mathcal{N}$ moments of the series expansion of the initial network $\mathbf{Y}(s)$, the first $\mathcal{N}$ moments of $\overline{\mathbf{Y}}(s)$ can be calculated exactly. This is also demonstrated by equation 4. Now define $\overline{\mathbf{Y}}^{\mathcal{N}}(s)$ as the first $\mathcal{N}$ terms of $\overline{\mathbf{Y}}(s)$. $\overline{\mathbf{Y}}^{\mathcal{N}}(s)$ is a good approximation of $\overline{\mathbf{Y}}(s)$ if $\mathcal{N}$ is sufficiently large. Then each element $\overline{Y}_{ij}^{\mathcal{N}}(s)$ with $i \neq j$ is defined by:

$$\overline{Y}_{ij}^{\mathcal{N}}(s) = \sum_{l=0}^{\mathcal{N}} (M_{ij}^{(l)} + \tilde{M}_{ij}^{(l)}) s^l \tag{3}$$

with, for all $\alpha_k \leq l \leq \mathcal{N}$:

$$\tilde{M}_{ij}^{(l)} = \frac{\sum_{p=\alpha_k}^{l} M_{kj}^{(p)} M_{ik}^{(l-p+\alpha_k)} - \sum_{p=\alpha_k}^{l-1} \left(\tilde{M}_{ij}^{(p)} \sum_{n=1, n\neq k}^{N} M_{kn}^{(l-p+\alpha_k)} \right)}{\sum_{n=1, n\neq k}^{N} M_{kn}^{(\alpha_k)}} \tag{4}$$

If the k-th node is coupled resistively to at least one other node, then $\alpha_k = 0$, else $\alpha_k = 1$.

Acknowledgement

We should acknowledge the sponsor: This research is sponsored in part by the Dutch Technology Foundation (STW) under grants nr. DEL 22.2810 and DEL 11.2450, and has been carried out in the context of DIMES, the Delft Institute of Microelectronics and Submicron Technology.

References

[1] A. J. van Genderen and N. P. van der Meijs, "Reduced RC Models for IC Interconnections with Coupling Capacitances," *Proc. IEEE 3rd European Design Automation Conference*, March 1992, pp. 132–136.

[2] C. L. Ratzlaff and L. T. Pillage, "RICE: Rapid Interconnect Circuit Evaluation Using AWE," *IEEE Trans. on Computer-Aided Design*, Vol. CAD-13, No. 6, pp. 763–776, 1994.

[3] S. Y. Kim, N. Gopal and L. T. Pillage, "Time-Domain Macromodels for VLSI Interconnect Analysis," *IEEE Trans. on Computer-Aided Design*, Vol. CAD-13, No. 10, pp. 1257–1270, 1994.

[4] M. Glez-Harbour and J. M. Drake, "Calculation of Signal Delay in Integrated Interconnections," *IEEE Trans. on Circuits and Systems*, Vol. CAS-36, No. 2, pp. 272–276, 1989.

[5] P. J. H. Elias and N. P. van der Meijs, "Including Higher-Order Moments of RC Interconnections in Layout-to-Circuit Extraction," *Proc. IEEE European Design and Test Conference*, March 1996.

[6] C. Y. Chu and M. A. Horowitz, "Charge-Sharing Models for Switch-Level Simulation," *IEEE Trans. on Computer-Aided-Design*, Vol. CAD-6, No. 6, pp. 1053–1061, 1987.

[7] N. P. van der Meijs and A. J. van Genderen, "Delayed Frontal Solution for Finite-Element based Resistance Extraction," *Proc. 32nd Design Automation Conf.*, pp. 273–278, 1995.