

Modeling of Polysilicide Gate Resistance Effect on Inverter Delay and Power Consumption Using Distributed RC Method and Branching Technique

Y. Koolivand
Department of Electrical and
Computer Engineering
University of Tehran
ykoolivand@ece.ut.ac.ir

A. Zahabi
Department of Electrical and
Computer Engineering
University of Tehran
azahabi@ut.ac.ir

N. Masoumi
Department of Electrical and
Computer Engineering
University of Tehran
nmasoumi@ut.ac.ir

ABSTRACT

Two parameters that contribute significantly in a CMOS inverter delay are the output load and propagation of the control signal across the gates of its transistors. The latter one which is due to the polysilicide gate resistance (*PGR*) is proportional to the gate width, W . The *PGR* effect causes the inverter transistors remain simultaneously on, for further time in the saturation region during transient instants, so that the short circuit power consumption increases largely. In this paper, we model the *PGR* resistance effect using the distributed RC approach based on a new proposed technique. Additionally, in order to reduce its negative impact on the circuit performance, we utilize a relevant optimized branching method. The results obtained from our model compared to the *HSPICE* simulation results verify a good agreement. Furthermore, our modeling technique can be implemented as a *CAD* tool for a primary estimation of the delay and the power consumption in complex circuits while retains an acceptable accuracy.

Categories and Subject Descriptors: B. [Hardware]

General Terms: Design, performance.

Keywords: Polysilicide gate resistance, propagation delay, short circuit power, performance degradation.

1. INTRODUCTION

With the technology scaling down and growing the demand for dense *VLSI* circuits, interconnect parasitic capacitances increase. This leads to the performance degradation because of the parasitic glitch and the increased delay. For modeling, extraction and evaluation of the parasitic capacitance effects, several articles have been published [4], [5], [6]. In conventional methods, buffers or boosters among the signal paths are employed in such a way that the total delay of the structure is made minimum [7], [8]. However, in all these approaches, inverters with a small inherent delay are needed. The main focus of this paper is the modeling of delay due to the transistor gate structure in inverters, which are used as repeaters for the interconnect wires. The gate of a

transistor is an important delay path, because any gate resistance along with its parasitic capacitance, cause an input time-constant, consequently a delay, in the signal propagation from the input to the output. The distributed *RC* model, which we have employed in this work for modeling the resistance effect of the gate, is a proper model for interconnect wires [3]. The analytical models presented in this work can be useful and remedial for high performance circuits where speed is very critical.

In section 2, the gate resistance effect on the circuit delay and the power consumption is explained. Section 3 presents a model for the *PGR*. To reduce the *PGR* effect, we propose a new optimized branching method in section 4. In section 5, we study effects of the output load on the proposed model. Finally, section 6 concludes this article.



Figure 1. Gate resistance modeling for *HSPICE* simulator.

2. EFFECTS OF *PGR* ON PROPAGATION DELAY AND POWER CONSUMPTION

In the recent *VLSI* technologies, the gate of a transistor is made up of the polysilicide which has a higher sheet resistance compared to the metalized interconnects. As such, for the delay propagation modeling in an inverter, the effect of the control signal slope as well as the output load effect must be considered.

The polysilicide-gate sheet-resistance (n^+ , p^+ polysilicide) is $5\Omega/\text{square}$ in $0.25\ \mu\text{m}$ CMOS technology [9]. The *PGR* effect can be significant for large channel-width transistors, so that the input time-constant becomes comparable with the output time-constant. This effect has not been considered in *HSPICE*¹ technology file. As Figure 1 illustrates, for including the gate resistance in our model, a transistor with a width $W\ \mu\text{m}$ (W is relatively a large integer value) is divided to W transistors with a $1\ \mu\text{m}$ channel length (for $W=1\ \mu\text{m}$ the *PGR* effect is negligible). Therefore, all the drains are connected together, and the same for the transistor sources. On the other hand, each gate is connected to the preceding

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

GLSVLSI'04, April 26-28, 2004, Boston, Massachusetts, USA
Copyright 2004 ACM 1-58113-853-9/04/0004...\$5.00.

1- www.mosis.org

gates via 20Ω resistances ($5 \Omega/\square \times (1 \mu\text{m}/0.25 \mu\text{m}) = 20 \Omega$). This method is based on the distributed RC model and indicates that each transistor with a channel width $W \mu\text{m}$ consists of many subtransistors with a smaller channel width. When we apply a signal to one transistor, it propagates across the channel width, so that the subtransistors sense the signal, gradually. Those subtransistors that are closer to the input line, receive the signal earlier than the others. Thus, the subtransistors that are farther to the input are much influenced; consequently, they get the slowest command signal on their gates. The PGR not only causes the input time-constant, thus increasing the delay, but also degrades the control signal slope at the inverter subtransistors gate. For this reason, the inverter transistors operate simultaneously long time in the saturation region which leads to the increased short-circuit power consumption.

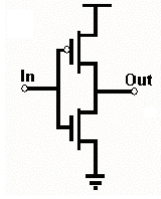


Figure 2. A simple inverter structure.

3. MODELING OF INVERTER PROPAGATION DELAY

As explained in [1] and [2], the propagation delays in a simple inverter shown in Figure 2 are comprised of two components.

$$t_{plh} = \alpha_1 t_{plhout} + \alpha_3 B t_{phlin} \quad (1)$$

$$t_{phl} = \alpha_2 t_{phlout} + \alpha_4 A t_{plhin} \quad (2)$$

Where indices “ lh ” and “ hl ” indicate the signal transition from low-to-high and high-to-low at the specified node, respectively. In (1) and (2), $A = I-2|V_{TP}|/V_{DD}$ and $B = I-2|V_{TN}|/V_{DD}$ where V_{TP} and V_{TN} are the threshold voltages of $PMOS$ and $NMOS$ transistors, respectively. Also, $\alpha_1, \dots, \alpha_4$ are fitting coefficients.

3.1 Input Time-Constant Computation

Applying the distributed RC model for the transistor gate, the time-constant at path A ($NMOS$ path) and for each node in Figure 3 according to [9] is as follows:

$$(\tau_{inN})_A = \frac{R_{gn} C_{gn}}{N^2} + 3 \frac{R_{gn} C_{gn}}{N^2} + \dots + N \frac{R_{gn} C_{gn}}{N^2} = \frac{N+1}{2N} R_{gn} C_{gn} \quad (3)$$

Where $(\tau_{inN})_A$ is the charging (discharging) time-constant of n 'th node in path A . Also, R_{gn} ($=R_{gp}=R_{\Omega\Box} \times (W_n/L)$) is the $NMOS$ ($PMOS$) gate resistance where $R_{\Omega\Box}$, W_n and L are the polysilicide sheet resistance, the width and the length of the $NMOS$ gate,

respectively. Also, C_{gn} ($=C_{gp}$) is the gate oxide capacitance of $NMOS$ ($PMOS$) transistors. Thus, the time-constant average in path A is determined as:

$$(\tau_{mean})_A = \frac{1}{N} \sum_{k=1}^N \frac{R_{gn} C_{gn}}{N^2} \frac{k(k+1)}{2} = \frac{R_{gn} C_{gn}}{2N^3} \left[\frac{N(N+1)}{2} + \frac{N(N+1)(2N+1)}{6} \right] \quad (4)$$

When N approaches to infinity:

$$\tau_A = \lim_{N \rightarrow \infty} (\tau_{mean})_A = \frac{R_{gn} C_{gn}}{6} \quad (5)$$

Similarly for path B ($PMOS$ path):

$$\tau_B = \lim_{N \rightarrow \infty} (\tau_{mean})_B = K^2 \frac{R_{gn} C_{gn}}{6} \quad (6)$$

Where $k=W_p/W_n$. We note that the input signal passes from both A and B paths and any input transition causes the charging of one path and discharging of the other path capacitors, simultaneously. Thus, similar to the technique presented in [1], the capacitors and resistors of both paths must be considered for the input delay (t_{plh_in} and t_{phl_in}) calculations:

$$t_{plh_in} = \ln 2 \frac{(R_{gp} + R_{gn})(C_{gp} + C_{gn})}{6} \left(\frac{C_{gp}}{C_{gp} + C_{gn}} \right) \quad (7)$$

$$t_{phl_in} = \ln 2 \frac{(R_{gp} + R_{gn})(C_{gp} + C_{gn})}{6} \left(\frac{C_{gn}}{C_{gp} + C_{gn}} \right) \quad (8)$$

The coefficient $C_{gp}/(C_{gp}+C_{gn})$ in (7) represents the fitting factor for matching with the HSPICE simulation results, and similarly for (8). The propagation delay of each node corresponds to the time interval for which, the voltage reaches to $V_{DD}/2$ (V_{DD} is supply voltage).

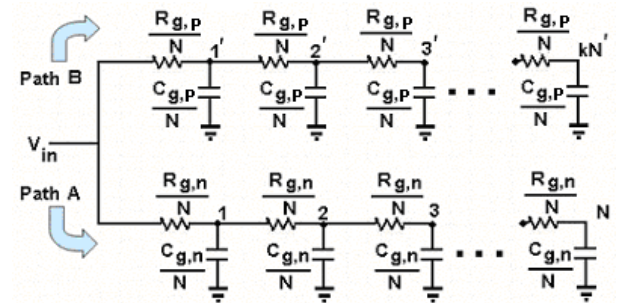


Figure 3. Analytical model for gate resistance presented in this paper.

3.2 Output Time-Constant Computation

Referring to [9], the equivalent output time-constant is given as:

$$\tau_{out} = R_{inv} C_l \quad (9)$$

where $R_{inv}(C_l)$ is the equivalent resistor (capacitor) related to the output node. Consequently, the delay due to the output time constant is:

$$t_{plh\ out} = \ln 2 (R_{invp} C_{lh\ out}) \quad (10)$$

$$t_{phl\ out} = \ln 2 (R_{invn} C_{hl\ out}) \quad (11)$$

where “n” (“p”) implies the NMOS (PMOS) transistor. Figure 4 illustrates high-to-low and low-to-high delays for several values of W versus K for two cases: *HSPICE* simulations and the proposed model.

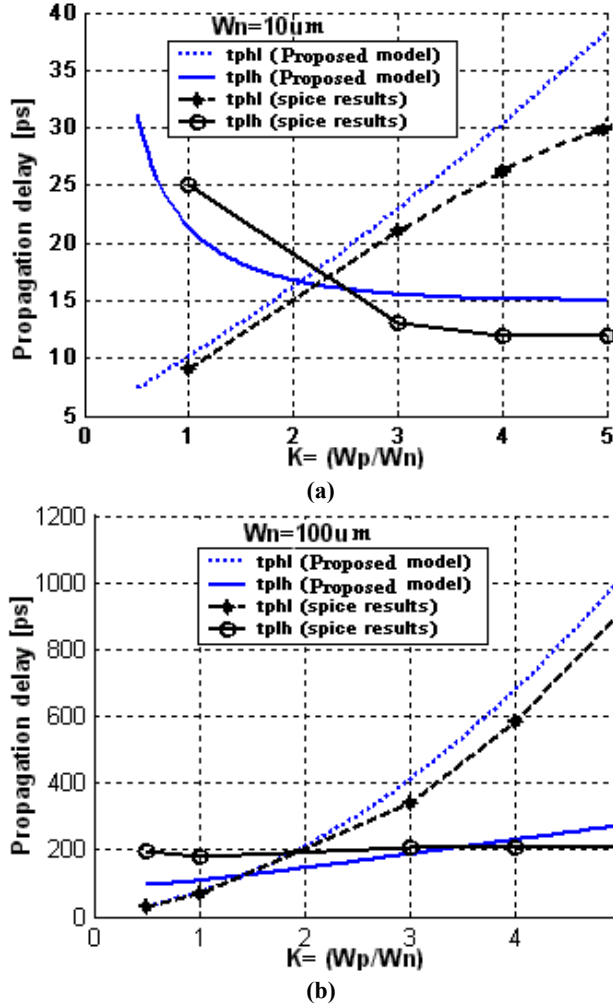


Figure 4. Delay propagations versus K for (a) $W_n=10\mu m$ (b) $W_n=100\mu m$.

As Figure 4 illustrates, at first, because of the constant NMOS output capacitor, t_{plh} is inversely proportional to K . Then, t_{plh} starts to rise, since the strong *PGR* effect exists. In contrast, t_{phl} is strictly proportional to K . For small values of W , the *PGR* effect is negligible, however, for large values the *PGR* effect becomes dominant. Results of *HSPICE* simulations agree with the proposed model and indicate a maximum error less than 27%.

4. REDUCING THE *PGR* EFFECT USING BRANCHING

As it is seen from the previous section, increasing the values of K and W_n causes a considerable performance degradation. To decrease the *PGR* effects, as depicted in Figure 5, we can use I transistors each with a channel width of (W/I) , instead of a transistor with a channel width of W . As a result, the input delay component is reduced with $(1/I^2)$ ratio. Each path in Figure 3 is divided into equal I paths. Thus, R_g and C_g are scaled as $(1/I)$. Finally, the input delays are modified as it follows:

$$(t_{plh\ in})_I = \ln 2 \frac{(R_{gp} + R_{gn})(C_{gp} + C_{gn})}{6I^2} \left(\frac{C_{gp}}{C_{gp} + C_{gn}} \right) \quad (12)$$

$$(t_{phl\ in})_I = \ln 2 \frac{(R_{gp} + R_{gn})(C_{gp} + C_{gn})}{6I^2} \left(\frac{C_{gn}}{C_{gp} + C_{gn}} \right) \quad (13)$$

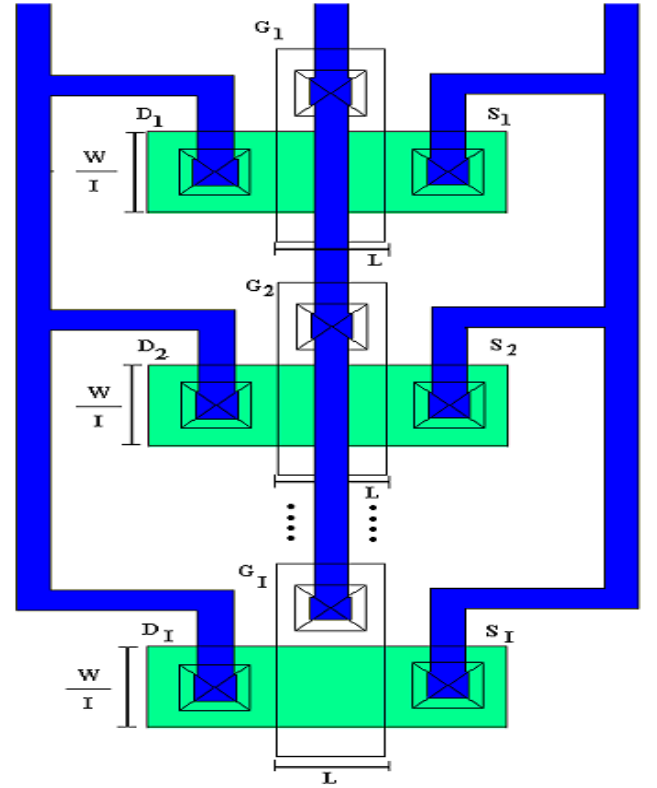


Figure 5. Branching in a transistor with channel width of W into I section.

The delay for $W_n=100\ \mu m$ with $K=5$ versus I are shown in Figure 6 for two cases: *HSPICE* simulation results and the proposed model. It is seen that increasing the number of branching reduces the delay, efficiently. The rate of performance improving versus I is very fast at the first, but it is almost not affected for large numbers of the branching process. More over, this rate is high for transistors of larger width with a small number of branching.

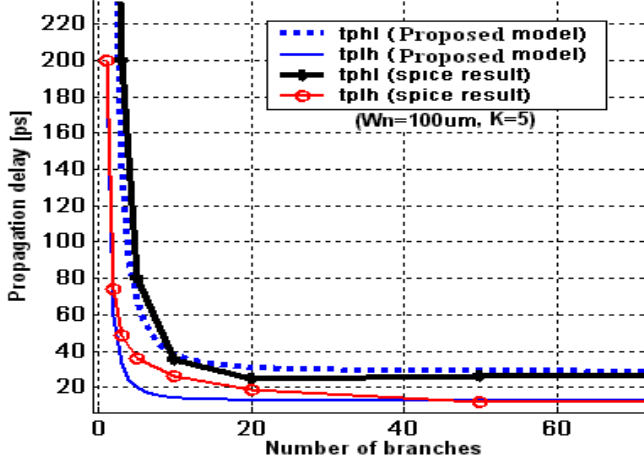


Figure 6. Delay versus number of branching for $W_n=100 \mu m$ with $K=5$.

Figure 7 shows the normalized mean power consumption for several values of W_n versus I for $K=5$ and the input signal frequency at 100 MHz . It is observed that the normalized power for larger values of W_n is high at the beginning, and then it decreases at the end for more branching. It means that branching shows desirable influence on wide channel transistors.

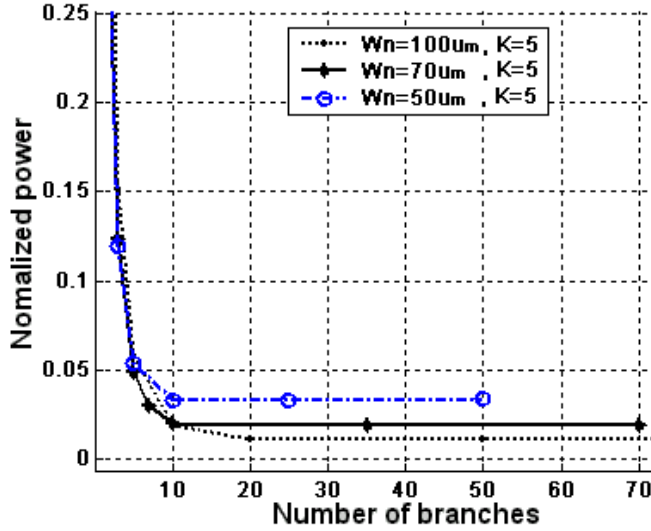


Figure 7. Normalized mean power consumption versus number of branching for some W_n .

5. EFFECT OF LOADING ON THE INVERTER CHAIN DELAY

Figure 8 depicts an inverter chain. Utilizing the expressions given in [1], the signal propagation delays in n 'th inverter are:

$$(t_{phl})_n = \alpha_2(t_{phs})_n + \alpha_4 A(t_{plhs})_n \quad (14)$$

$$(t_{plh})_n = \alpha_1(t_{plhs})_n + \alpha_3 B(t_{phls})_n \quad (15)$$

where the index "s" implies the delay when a step voltage is applied at the input of the inverter. So, for n 'th inverter:

$$(t_{hls})_n = \ln 2(R_{invn})_n C_{hl_n} \quad (17)$$

$$(t_{lhs})_n = \ln 2(R_{invp})_n C_{lh_n} \quad (18)$$

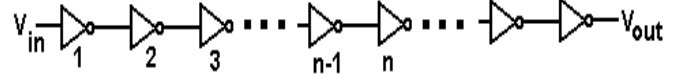


Figure 8. An inverter chain.

The input equivalent circuit of the n 'th inverter is shown in Figure 9. The delays in n 'th inverter input are:

$$(t_{hs})_{n-1} = \ln 2 \{ (R_{invp} C_{lh})_{n-1} + [(R_{invp})_{n-1} + \frac{(R_{gn} + R_{gp})_n}{I_n^2} \times \frac{(C_{gp})_n}{6(C_{gn} + C_{gp})_n}] (C_{gn} + C_{gp})_n \} \quad (18)$$

$$(t_{hs})_{n-1} = \ln 2 \{ (R_{invn} C_{hl})_{n-1} + [(R_{invn})_{n-1} + \frac{(R_{gn} + R_{gp})_n}{I_n^2} \times \frac{(C_{gn})_n}{6(C_{gn} + C_{gp})_n}] (C_{gn} + C_{gp})_n \} \quad (19)$$

where I_n is the number of branching in the n 'th inverter.

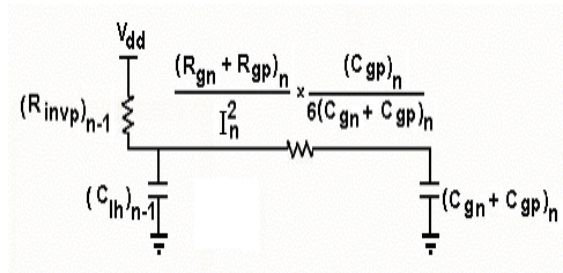


Figure 9. The equivalent circuit for the n th inverter input.

When it is desired that the delay introduced by the PGR effect in the n 'th inverter equals maximum x times of the delay of an inverter of no gate resistance, then, the number of branching for the n 'th stage is given by:

$$I_n = \max \{I_{n1}, I_{n2}\} \quad (20)$$

where I_{n1} and I_{n2} are given by the following expressions:

$$I_{1n} = \sqrt{\frac{(R_{gn} + R_{gp})_n (C_{gp})_n}{6 \left[\frac{x\alpha_2}{A\alpha_4} (R_{invn})_n (C_{hl})_n - (R_{invn})_{n-1} (C_{hl})_{n-1} - (R_{invn})_{n-1} (C_{gn} + C_{gp})_n \right]}} \quad (21)$$

$$I_{2n} = \sqrt{\frac{(R_{gn} + R_{gp})_n (C_{gn})_n}{6 \left[\frac{x\alpha_1}{B\alpha_3} (R_{invn})_n (C_{hl})_n - (R_{invn})_{n-1} (C_{hl})_{n-1} - (R_{invn})_{n-1} (C_{gn} + C_{gp})_n \right]}} \quad (22)$$

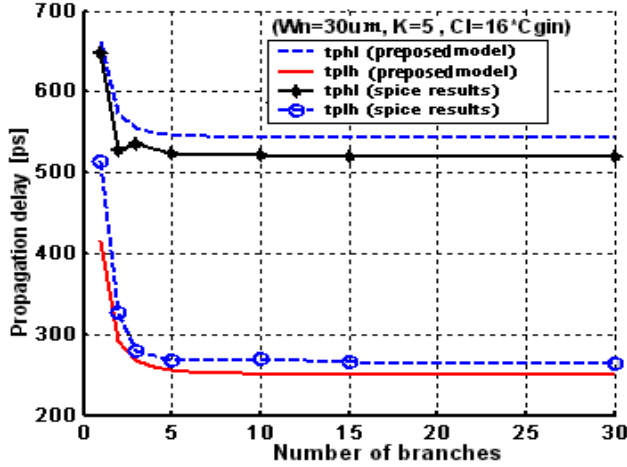


Figure 10. Delay in an inverter versus number of branching for $W_n=30 \mu m$, $K=5$ and $F=16$.

Figure 10 illustrates the propagation delays of an inverter versus I for $W_n=30 \mu m$, $K=5$ and $F=16$ ($F=C_{out}/C_{in}$) for two previous cases. The normalized mean power consumption for $W_n=100 \mu m$, $K=5$ and several values of F is depicted in Figure 11. Obviously, the PGR effect is much significant for small loads. In fact, the minimum number of branching required to reach a relatively steady-state performance is high for both large loads and small channel widths.

6. CONCLUSIONS

The PGR distributed resistance causes severe performance degradation in the circuits of wide channel transistors. To alleviate this effect, the branching technique is employed in fabrication process. We have shown that to obtain a relatively steady-state performance small loads need less branching. It is the same for wide channel transistors when the load is constant. In this work, we have employed the distributed RC modeling technique, to model the PGR , instead of using simple lumped models. The results obtained from our new proposed model show a good agreement with the $HSPICE$ simulation results, so that the maximum deviation is less than 27%.

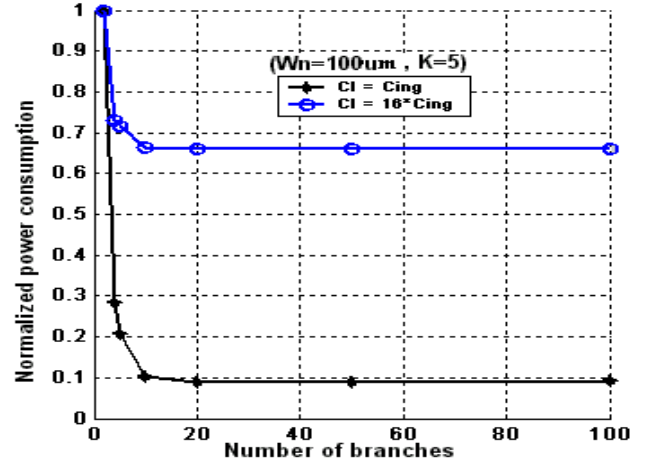


Figure 11. Normalized mean power consumption versus number of branching for $W_n=100 \mu m$, $K=5$ and some F .

7. REFERENCES

- [1] D. Deschacht, C. Dabrin, D. Auvergne, "Delay propagation effect in transistor gates," *IEEE J. Solid State Circuits*, vol. 31, no. 8, August 1996.
- [2] D. Deschacht, M. Robert, D. Auvergne, "Explicit formulation of delays in CMOS data paths," *IEEE J. Solid State Circuits*, vol. 23, no. 5, October 1988.
- [3] G. Servel, F. Huret, E. Poleczny, P. Kennis, D. Deschacht, "Impact of inductance on timing characteristics of VLSI interconnects," *International Caracas Conference on Devices, Circuits and Systems, Cancun*, C17.1 - C17.6, 2000.
- [4] E. Sicard, T. Demonchaux, J. I. Noullet, A. Rubio, "Cross-talk extraction from mask layout," *IEEE European Design Automation Conference, paris*, 1993.
- [5] K. Soumyanath, S. Borkar, C. Zhou, B. A. Bloechel, "Accurate on-chip interconnect evaluation: a time-domain technique," *IEEE J. Solid State Circuits*, vol. 34, no. 5, May 1999.
- [6] M. Lee "A multilevel parasitic interconnect capacitance modeling and extraction for reliable VLSI on-chip clock delay evaluation," *IEEE J. Solid State Circuits*, vol. 33, no. 4, April 1998.
- [7] K. Chaudhary, A. Onozawa, E. S. Kuh, "A spacing algorithm for performance enhancement and cross-talk reduction," *Proceedings of IEEE/ACM international conference on Computer-aided design*, p.697-702, 1993.
- [8] A. Nalamalpu, S. Srinivasan, "Boosters for driving long onchip interconnects design issues, interconnect synthesis, and comparison with repeaters," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 21, no. 1, January 2002.
- [9] J. Rabaey, "Digital integrated circuits: a new design perspective," *Prentice-Hall*, 1996.