



HAL
open science

Statistical Characterization of Library Timing Performance

Vincent Migairou, Robin P. Wilson, Sylvain Engels, Nadine Azemard,
Philippe Maurine

► **To cite this version:**

Vincent Migairou, Robin P. Wilson, Sylvain Engels, Nadine Azemard, Philippe Maurine. Statistical Characterization of Library Timing Performance. PATMOS: Power And Timing Modeling, Optimization and Simulation, Sep 2006, Montpellier, France. pp.468-476, 10.1007/11847083_45. lirmm-00093233

HAL Id: lirmm-00093233

<https://hal-lirmm.ccsd.cnrs.fr/lirmm-00093233>

Submitted on 13 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Statistical Characterization of Library Timing Performance

V. Migairou¹, R. Wilson¹, S. Engels¹, N. Azemard², and P. Maurine²

¹ STMicroelectronics Central CAD & Design Solution,
850 rue Monnet, 38926, Crolles, France

² LIRMM, UMR CNRS/Université de Montpellier II, (C5506),
161 rue Ada, 34392, Montpellier, France

Abstract. With the scaling of technology, the variability of timing performances of digital circuits is increasing. In this paper, we propose a first order analytical modeling of the standard deviations of basic CMOS cell timings. The proposed model is then used to define a statistical characterization protocol which is fully compliant with standard characterization flows. Validation of this protocol is given for a 90nm process.

1 Introduction

With the scaling of technology, digital circuits are being designed to operate at ever increasing clock frequencies. At the same time, process fabrication uncertainties are becoming more important resulting in larger parameter variations and thus in larger timing performance variations. Within this context, the verification before fabrication of the circuit timings, which has always been a critical design step, is becoming too pessimistic and inadequate.

The common way to verify the timings of a circuit, and thus to validate it, is to use the well known multiple corner (Process, Voltage, Temperature) based approach during the static timing analysis (STA). The main drawback of such an analysis lies in its conservatism [1]. If the resulting design margins guarantee obtaining high yield values, they may induce some convergence problem during the timing optimization step. The Statistical Static Timing Analysis (SSTA) [1-4] appears as a powerful alternative to reduce these design margins and also to take inter and intra-die process dispersions into account. Intra-die variability, which has become a substantial part of the overall process variability, is not modeled, resulting in an increasing loss of accuracy.

Like the traditional timing analysis STA, the main aim of SSTA is to provide timing information allowing to guarantee the functionality of a design. The main difference between STA and SSTA lies in the nature of the information reported. Indeed, statistical analysis propagates probability density function of timing performance rather than worst case timings. This constitutes a significant advantage allowing the estimation of the manufacturing yield.

However SSTA presents some drawbacks. Firstly, the timing analysis itself is, from a computation cost point of view, more expensive than the traditional corner

approach. Secondly, structural correlations introduced by the logic synthesis must be captured to obtain accurate *pdf* of the propagation delays [block based]. Thirdly, it requires, at the library level, an expensive characterization step.

This paper investigates the latter problem and a solution is proposed to accurately and quickly evaluate the dispersions of standard cell timing performances. In section III, first order analytical expressions of the standard deviations of CMOS cell timings are deduced from an analytical timing representation, which is briefly introduced in section II. Section IV is then dedicated to the validation of the resulting expressions and to the analysis of their properties. These properties are then exploited to define a statistical characterization protocol of CMOS cells. Finally a conclusion is drawn in section VI. Note that to facilitate the reading, table 1 lists all the notations used in the paper.

Table 1.

Symbol	Definition	Dim.
V _{dd}	Supply voltage	V
C _L	Total output load capacitance	fF
T _{OX}	Gate oxide thickness	μm
V _{tn/p}	Threshold voltage of N and P transistors	V
α _{n/p}	Velocity saturation index of N and P transistors	-
v _s	Saturation velocity	m/s
L	Transistor length	μm
W _{N/P}	N and P transistor width respectively	μm
DW _{HL/LH}	Digital weights of a logic gate [5]	-
τ _{N/P}	Metrics of the timing performance of a process	ps
C _{N/P}	Gate capacitance of N and P transistors	fF
T _{HL/LH}	Propagation delay of a cell	ps
τ _{OUTHL/LH}	Falling and rising output transition time	ps

2 Analytical Timing Representation

Modeling the transistor as a current generator [5], the output transition time of CMOS primitives can directly be obtained from the modeling of the (dis)charging current that flows during the switching process of the structure and from the amount of charge (C_L·V_{DD}) to be transferred from the output to the supply rails as :

$$\tau_{outHL} = C_L \cdot V_{DD} / I_{MAX} \quad (1)$$

where I_{MAX} is the maximum current available in the structure. The key point here is to evaluate this maximum current which depends on the input controlling condition and also on the output load value. For that, two domains have to be considered: the Fast input and the Slow input range.

In the *Fast input ramp domain*, the high slew rate of the incoming signal forces the structure to provide all the current it can deliver [5]. As a result, the switching current has a maximum and constant value which can easily be obtained from the alpha power law model [6]:

$$I_{MAX}^{Fast} = v_s \cdot C_{ox} \cdot W_N \cdot (V_{DD} - V_{in})^{\alpha_n} \quad (2)$$

a. Output transition time modeling

Combining then (1) and (2) finally leads to the output transition time expression associated to the Fast input ramp range

$$\tau_{outHL}^{Fast} = \tau_N \cdot \frac{DW_{HL} \cdot Cl}{C_N} = \frac{DW_{HL} \cdot Cl \cdot V_{DD}}{v_s \cdot C_{ox} \cdot W_N \cdot (V_{DD} - V_{in})^{\alpha_n}} \quad (3)$$

As shown, this expression captures all the sensitivity of the output transition time to process parameters such as V_T , C_{ox} but also on design parameters such as C_L and W .

In the *Slow input ramp domain*, the maximum switching current decreases with the input ramp duration. Extending the results of [5] to general value of the velocity saturation index, the maximum switching current flowing in a CMOS structure is

$$I_{NMAX}^{Slow} = \left\{ \frac{(\alpha_N \cdot v_s \cdot C_{ox} \cdot W_N)^{\frac{1}{\alpha_n}} \cdot Cl \cdot V_{DD}^2}{\tau_{IN}} \right\}^{\frac{\alpha_n}{1+\alpha_n}} \quad (4)$$

Where τ_{IN} is the input ramp duration (the output transition time of the controlling structure). It is usually measured between 80% and 20% of V_{DD} and extrapolated on the full voltage swing. Combining (1) and (3) with (5), we finally obtain a manageable transition time expression for a falling output edge in the Slow input ramp domain:

$$\tau_{outHL}^{Slow} = \left(\frac{DW_{HL} \cdot Cl \cdot \tau_{IN}^{\alpha_n} \cdot V_{DD}^{1-\alpha_n}}{\alpha \cdot v_s \cdot C_{ox} \cdot W_N} \right)^{\frac{1}{1+\alpha_n}} \quad (5)$$

with an equivalent expression for the rising edge. As expression (3), (5) constitutes an explicit expression of the output transition time in the slow input ramp domain. To conclude with the modeling of the output transition time, one can observe that in the Fast input range the transition time only depends on the output load while in the slow input range, it also depends on the input transition time duration, and is threshold voltage independent.

b. Propagation delay modeling

The delay of a basic CMOS structure is load, gate size and input slew dependent. Following [7-8], the input slope and the I/O coupling capacitance C_M can be introduced in the propagation delay as

$$T_{HL} = \frac{\tau_{in}}{\alpha_N + 1} \left(\frac{\alpha_N - 1}{2} + \frac{V_{m/p}}{V_{DD}} \right) + \left(1 + \frac{2C_M}{C_M + C_L} \right) \frac{\tau_{outHL}}{2} \quad (6)$$

This expression captures all the delay sensitivity of basic CMOS structures to its environment (τ_{IN} , τ_{out}), and also all the sensitivity to the main process parameters through the term τ_{out} .

3 Analytical Standard Deviation Model

Normal distributions of standard cell timing performance can be obtained from the analytical model introduced in section II. More precisely, analytical expressions of propagation delay and output transition time standard deviations can be derived from the statistical error propagation expression below:

$$\sigma^2(f) = \sum_i \left\{ \left(\frac{\partial f}{\partial p_i} \right)^2 \cdot \sigma^2_{p_i} \right\} + \sum_{i \neq j} \sum_j 2 \cdot \left(\frac{\partial f}{\partial p_i} \right) \left(\frac{\partial f}{\partial p_j} \right) \cdot \text{cov}(p_i, p_j) \quad (7)$$

where f stands for the timing performance under consideration (propagation delay or output transition time, p_i are the process parameters and $\text{cov}(p_i, p_j)$ is the co-variance between process parameters p_i and p_j .

a- Output transition time standard deviation

The output transition time model introduced in section II clearly distinguishes two kinds of behaviours for any CMOS cell: one for fast input ramp domain, and another one for slow input ramp domain.

In the *Fast input ramp domain*, it appears following (3) that output transition time does only depend on two geometrical process parameters (W and C_{ox}) and one electrical process parameter (V_t). Combining expressions (7) and (3) leads to the following normalized expression of the output transition time standard deviation:

$$\frac{\sigma(\tau_{out}^{Fast})}{\tau_{out}^{Fast}} = \left\{ \begin{array}{l} \left[\frac{\sigma^2(W)}{W^2} + \frac{\alpha^2 \cdot \sigma^2(V_t)}{(V_{DD} - V_t)^2} + \frac{\sigma^2(C_{ox})}{C_{ox}^2} \right]^{\frac{1}{2}} \\ + 2 \cdot \left[\begin{array}{l} - \frac{\alpha \cdot \text{Cor}(W, V_t) \cdot \sigma(W) \cdot \sigma(V_t)}{W \cdot (V_{DD} - V_t)} + \\ \frac{\text{Cor}(W, C_{ox}) \cdot \sigma(C_{ox}) \cdot \sigma(W)}{C_{ox} \cdot W} - \frac{\alpha \cdot \text{Cor}(C_{ox}, V_t) \cdot \sigma(C_{ox}) \cdot \sigma(V_t)}{C_{ox} \cdot (V_{DD} - V_t)} \end{array} \right] \end{array} \right\} \quad (8)$$

where $\sigma(V_t)$, $\sigma(1/T_{ox})$, $\sigma(W)$, are respectively the standard deviation of the threshold voltage, the oxide thickness and the width of the transistor, and $\text{Cor}(p_i, p_j)$ are the correlation factors between two process parameters p_i and p_j .

Following the same reasoning to determine the standard deviation of the output transition time in the *Slow input ramp domain*, we have obtained the following expression:

$$\frac{\sigma(\tau_{out}^{Slow})}{\tau_{out}^{Slow}} = \frac{1}{1 + \alpha} \cdot \left\{ \frac{\sigma^2(W)}{W^2} + \frac{\sigma^2(C_{ox})}{C_{ox}^2} + 2 \cdot \frac{\text{Cor}(W, C_{ox}) \cdot \sigma(C_{ox}) \cdot \sigma(W)}{C_{ox} \cdot W} \right\}^{\frac{1}{2}} \quad (9)$$

Note that expression (10) does not depend on the threshold voltage standard deviation resulting in a simpler expression than (9). This was expected from (6) which is threshold voltage independent.

b- Propagation delay standard deviation

Assuming, without loss of generality, that output load value is greater than the I/O coupling capacitance C_M (see (6)), we obtained the expression of the propagation delay standard deviation in the *Fast input ramp domain*:

$$\frac{\sigma(T)}{\tau_{out}^{Fast}} = \frac{1}{2} \cdot \left\{ \begin{aligned} & \left[\frac{\sigma^2(W)}{W^2} + \frac{\sigma^2(C_{ox})}{C_{ox}^2} + \left[\frac{\tau_{IN}}{V_{DD} \cdot \tau_{out}^{Fast}} + \left(\frac{\alpha}{V_{DD} - V_t} \right) \right]^2 \cdot \sigma^2(V_t) \right. \\ & \left. + \frac{2 \cdot Cor(W, C_{ox}) \cdot \sigma(W) \cdot \sigma(C_{ox})}{W \cdot C_{ox}} \right]^{\frac{1}{2}} \\ & - \left[\frac{\tau_{IN}}{V_{DD} \cdot \tau_{out}^{Fast}} + \left(\frac{\alpha}{V_{DD} - V_t} \right) \right] \cdot \sigma(V_t) \cdot \\ & \left(\frac{2 \cdot Cor(W, V_t) \cdot \sigma(W)}{W} + \frac{2 \cdot Cor(V_t, C_{ox}) \sigma(C_{ox})}{C_{ox}} \right) \end{aligned} \right\} \quad (10)$$

and in the *Slow one*:

$$\frac{\sigma(T)}{\tau_{out}^{Slow}} = \frac{1}{2} \cdot \left\{ \begin{aligned} & \left(\frac{1}{1 + \alpha} \right)^2 \cdot \left(\frac{\sigma^2(W)}{W^2} + \frac{\sigma^2(C_{ox})}{C_{ox}^2} + \frac{2 \cdot Cor(W, C_{ox}) \cdot \sigma(W) \cdot \sigma(C_{ox})}{W \cdot C_{ox}} \right)^{\frac{1}{2}} \\ & + \left[\frac{\tau_{IN}}{V_{DD} \cdot \tau_{out}^{Slow}} \right]^2 \cdot \sigma^2(V_t) \\ & - \frac{\tau_{IN}}{V_{DD} \cdot \tau_{out}^{Slow}} \cdot \sigma(V_t) \cdot \\ & \left(\frac{2 \cdot Cor(W, V_t) \cdot \sigma(W)}{(1 + \alpha) \cdot W} + \frac{2 \cdot Cor(V_t, C_{ox}) \sigma(C_{ox})}{(1 + \alpha) \cdot C_{ox}} \right) \end{aligned} \right\} \quad (11)$$

c- Discussion

In the two preceding paragraphs, we have deduced, from a first order timing representation, analytical expressions of the normalized (with respect to the nominal output transition time of the considered cell) standard deviation of both propagation delay and output transition time. If this choice may appear arbitrary, it leads to an interesting result.

Indeed, eq. (9) and (10) show that the normalized standard deviation of the output transition times does not depend on the output load and input ramp values. Similarly the normalized standard deviation of the propagation delay only depends on process parameters and on the input to output ramp duration ratio (τ_{IN}/τ_{OUT}). In other words, it is possible to extract from electrical simulations, unique normalized standard deviation curves of both propagation delay and output transition time. Since these curves are representative of the whole design space (defined by the max slew and load values) they can be efficiently exploited to speed up the statistical characterization

step of library cells. Before discussing in detail the implications of this result, let us first validate it.

4 Validation

In order to validate the existence of these unique and representative characteristics, we extracted from electrical simulations, the standard deviations of the propagation delay and output transition time of various basic cells designed with 90nm and 65nm processes. Statistical BSIM4 model cards, in which inter-die and intra-die variations are characterized separately, were used to perform these simulations.

The simulated values, obtained for a wide range of output load ($C_L=1\text{fF}$ to 660fF) and input ramp values ($\tau_{IN}= 1\text{ps}$ to 1ns), have allowed plotting the evolutions of the normalized standard deviations with respect to the input to output ramp duration ratio (τ_{IN}/τ_{OUT}). Fig.1 is typical illustration of the evolutions for various standard cells designed with 90nm process.

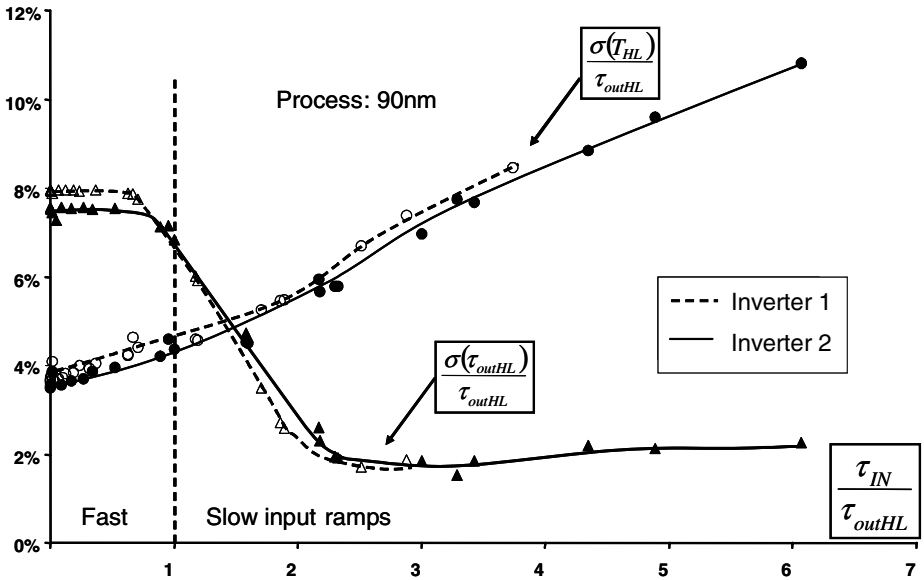


Fig. 1. $\sigma(T_{HL})/T_{HL}$ and $\sigma(\tau_{outHL})/\tau_{outHL}$ of two different inverters designed with a 90nm process

As expected from the preceding discussion, the normalized values of $\sigma(T_{HL})$ and $\sigma(\tau_{outHL})$ obtained, for a given inverter, and for various loading and controlling conditions belong to the same curves.

However the evolution of the output transition time standard deviation exhibits a linear behaviour (for τ_{IN}/τ_{OUT} values ranging between 1 and 2.5) which is not captured by expressions (9) and (10). This linear behaviour corresponds to the transition from the fast input ramp domain to the slow one. Despite this lack of the model, the

existence of a unique characteristic is verified by simulation validating thus the existence of unique standard deviation characteristics.

5 Statistical Characterization of Library Timing Performances

As mentioned in section III, the existence of these characteristic curves can be exploited to speed up the statistical characterization of library timing performances. Indeed rather than performing Monte Carlo simulations to evaluate the standard deviations of the timings for every (τ_{IN}, C_L) couple reported in the timing look up tables, one can run few Monte Carlo simulations to obtain the evolutions of the normalized standard deviations. Then, look up tables reporting the absolute values of the timing performances variability can be filled (by interpolation) from the mean values of the output transition time, usually reported in the tlf. Fig.2 illustrates the resulting characterization protocol. It goes off as explain below.

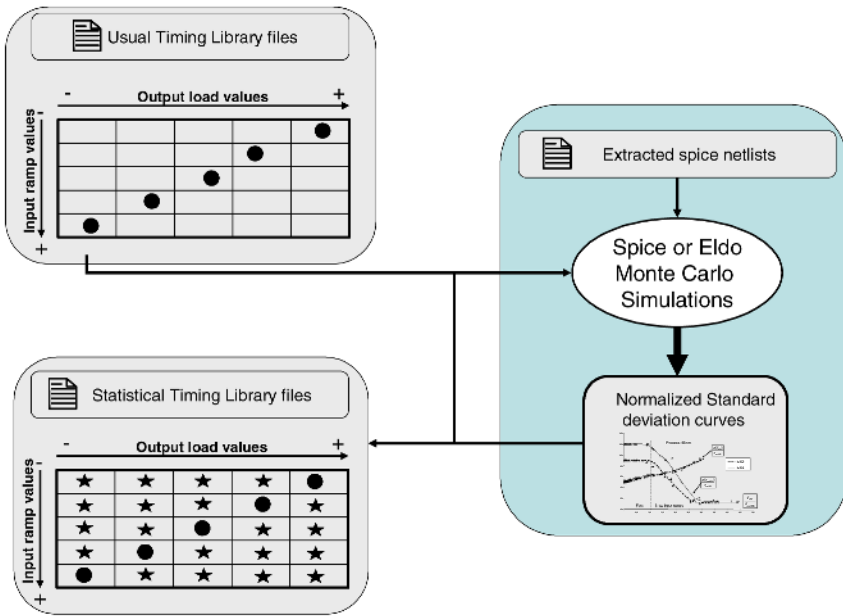


Fig. 2. Statistical timing performance characterization protocol

In a first step, the output load and input ramp values for which the standard cell under consideration has been characterized are extracted from the timing library format. In a second step, Monte Carlo analyses are performed in order to determine the standard deviation values for the few load values and all the input ramp values (i.e. for loading and controlling conditions identified by • in Fig.2). The choice of C_L values, for which these Monte Carlo simulations are performed, is done in order to capture the minimum and maximum (τ_{IN}/τ_{OUT}) values but also the (τ_{IN}/τ_{OUT}) values

ranging from 0.5 to 2.5 i.e. to properly sample the transition for the fast input ramp domain to the slow one. Note that the limit between these two domains is fully defined by the following expression [5]:

$$\tau_{IN} \geq \left(\frac{V_{DD}}{V_{DD} - V_t} \right) \cdot \tau_{out}^{Fast} \Rightarrow \frac{\tau_{IN}}{\tau_{out}^{Fast}} = \left(\frac{V_{DD}}{V_{DD} - V_t} \right) \quad (12)$$

The evolution of the normalized standard deviation is then plotted (i.e. reported in a one line look up table). Finally the statistical timing look up table is filled (by interpolation) considering the mean values of the output transition time provided by the tlf. As illustrated by Fig.2, the proposed characterization method requires 80% less Monte Carlo simulations for 5 by 5 look up tables while maintaining a high accuracy level.

The accuracy of the method is illustrated by tables 2 and 3. They give, for all inverters of 65nm library, the relative discrepancies obtained for the transition time and the delay with respect to a brute force method for which Monte Carlo simulations have been performed for all (τ_{IN}, C_L) couples. As shown, the relative discrepancies are lower than 5% validating the proposed characterization protocol.

Table 2. % of error between simulated and calculated $\sigma(\tau_{OUT\,RISE})$ values

$\tau_{IN}(ps) / C_L (fF)$	7	27	84	166,6	331,2
1	2%	1%	0%	0%	0%
50	3%	1%	1%	0%	0%
150	0%	0%	3%	2%	1%
500	0%	0%	3%	0%	3%
1000	0%	1%	0%	2%	0%

Table 3. % of error between simulated and calculated $\sigma(\tau_{OUT\,FALL})$ values

$\tau_{IN}(ps) / C_L (fF)$	7	27	84	166,6	331,2
1	1%	0%	0%	0%	0%
50	2%	1%	0%	0%	0%
150	1%	0%	1%	1%	0%
500	5%	0%	0%	1%	1%
1000	0%	0%	0%	0%	1%

6 Conclusion

In this paper, we have derived from a first order analytical modeling of timing performance, a method allowing the statistical characterization of library timing performance. The proposed method presents two main advantages. Firstly, it is fully compliant with usual characterization methods and can easily be automated. Secondly, it requires a number of Monte Carlo simulations which is, in average, 80% less important than with a brute force approach, while keeping an equivalent level of accuracy (95%).

References

- [1] C. Visweswariah. "Statistical timing of digital integrated circuits". IEEE International Solid-State Circuits Conference, CA, 2004.
- [2] A. Agarwal and al, "Statistical timing analysis for intra-die process variations with spatial correlations", ICCAD, 2003
- [3] J.Y. Le and al, "STAC: Statistical Timing Analysis with Correlation", the Design Automation Conference, June 2004.
- [4] M. Orshansky and al, "A general probabilistic framework for worst case timing analysis," DAC 2002, pp. 556-561, 2002.
- [5] P. Maurine and al, "Transition time modeling in deep submicron CMOS" IEEE Trans. on CAD, vol.21, n11, pp.1352-1363, 2002.
- [6] T. Sakurai and A.R. Newton, "Alpha-power model, and its application to CMOS inverter delay and other formulas", J. Solid State Circuits vol. 25, pp. 584-594, April 1990.
- [7] K.O. Jeppson, "Modeling the Influence of the Transistor Gain Ratio and the Input-to-Output Coupling Capacitance on the CMOS Inverter Delay", IEEE JSSC, Vol. 29, pp. 646-654, 1994.
- [8] J.M. Daga and al "Temperature effect on delay for low voltage applications", DATE, pp. 680-685, 1998, Paris.