A New Efficient Approach to Statistical Delay Modeling of CMOS Digital Combinational Circuits

Syed A. Aftab  
Motorola Strategic Systems Tech.  
2100 E. Elliot Rd, MD EL612  
Tempe, AZ 85284.

M. A. Styblinski  
Dept. of Electrical Engineering  
Texas A&M University  
College Station, TX 77843.

Abstract

This paper presents one of the first attempts to statistically characterize signal delays of basic CMOS digital combinational circuits using the transistor level approach. Hybrid analytical/iterative delay expressions in terms of the transistor geometries and technological process variations are created for basic building blocks. Local delays of blocks along specific signal paths are combined together for the analysis of complex combinational VLSI circuits. The speed of analysis is increased by 2 to 4 orders of magnitude relative to SPICE, with about 5-10% accuracy. The proposed approach shows good accuracy in modeling the influence of the “noise” parameters on circuit delay relative to direct SPICE-based Monte Carlo analysis. Examples of statistical delay characterization are shown. The important impact of the proposed approach is that statistical evaluation and optimization of delays in much larger VLSI circuits will become possible.

1 Introduction

Statistical analysis, optimization and general Design for Quality (DFQ) of large CMOS digital circuits is prohibitively expensive, due to long simulation times (e.g., using SPICE) caused by the large number of circuit analyses required. Switch-level simulators [1]-[5] reduce costs by replacing the transistors with a linear resistive switches and solving the equivalent RC network. Alternatively, the appropriate circuit delay equations can be directly solved [6, 7, 8] to improve accuracy. Unfortunately, these methodologies have so far proved to be inadequate for statistical analysis, because they do not take the technological process parameters into account. Modifications (to include the effects of the “noise” parameters) to the RC network based simulators is possible for simple structures only such as inverters [9]. In addition, statistical analysis requires the accurate estimation of the delay functions and its derivative, which requires greater simulator accuracy than for the nominal analysis.

The present paper is believed to be one of the first attempts to statistically characterize complex VLSI building blocks (NAND/NOR), using the transistor level approach.

2 The Statistical Modeling Methodology

The proposed statistical modeling approach is novel in several ways. The major objective is to generate efficient C-code relating individual transistor delays to the geometric (X) and noise (θ) parameters, using analytical (i.e., “symbolic”) formulas mixed with calls to iterative algorithms, as needed. The pre-simulation stage (Fig. 1(a)) is performed once. First,
a statistical model is selected that can accurately predict the variation of the transistor model parameters (as in SPICE), in terms of a set of independent (uncorrelated) noise parameters \( \theta \). Next, the device models for the transistors and capacitors are selected. The transistor model proposed in [8] is accurate (even for submicron MOSFETs) and analytically simple. In addition, it’s empirical nature allows the modeling of MOS structures, such as short chains, with the non-physical model parameters extracted directly from I-V curves (simulation or measurements). The capacitance model approximates the non-linear MOS parasitic capacitances with average capacitors (additionally parameterized by tuning parameters), thus greatly simplifying circuit analysis. These device models are then statistically characterized, by determining the relationship between the device model parameters \( p \) and the geometric \( X \) and noise parameters \( \theta \). An advanced interpolation method [10] is used to link the designable and noise parameters to the non-physical transistor model parameters \( p \) from [8], i.e., \( p \approx p(X, \theta) \), which is computationally inexpensive (since each transistor type is individually characterized). The capacitor model parameters \( p \) are obtained directly from the selected statistical model [11], so no approximation is needed. Each basic building block is characterized in terms of the device model parameters and the shape of the input waveform, by explicitly solving the generic equation:

\[
dV_i(t) = I_D(V_{in}(t), p_c(X, \theta), p(X, \theta), V_{DC}) dt.
\]

Several cases may arise, with a different expression for the current \( I_D(t) \) in each case. The input waveform is represented by a multi-segment piece-wise linear approximation, although usually a two-segment waveform is selected for efficiency reasons. The proposed solutions are combinations of analytical and simple iterative methods, because it is practically impossible to always be able to obtain closed form delay formulas of sufficient statistical accuracy (see Section 3). New approximations are made to reduce the complexity of the circuit solutions, while maintaining high statistical accuracy. Wherever possible, more complex building blocks are reduced to simpler blocks, whose equations are easier to manage. For example, series connected transistors are replaced by an equivalent composite transistor, whose model parameters are determined using a simple iterative procedure [12]. For more details, regarding the actual delay models implemented, refer to [12].

The second stage of the proposed methodology is used to generate delay models based on the specific circuit topology (Fig. 1(b)). The circuit is first decomposed into cells along critical paths. Each cell is further decomposed into basic building blocks. The local delay expressions of each building block, in a selected path, are combined and C-code corresponding to this delay is generated. As demonstrated in [9], this can result in a speed increase of up to 2 orders of magnitude relative to a numerical simulator performing exactly the same operations.

3 Nominal Analysis vs Statistical Analysis

As mentioned earlier, good nominal accuracy is not equivalent to good statistical accuracy. Statistical accuracy requires that both the function and its derivative be accurately determined. Therefore, delay formulas with simplified approximations (based on empirical or theoretical considerations) that are accurate for nominal analysis, may not be sufficient for estimating \( \sigma \)'s of delays. In [9], the RC delay models from [2] were modified to include the effects of the noise parameters. The effort was successful for simple cases only (such as inverters). For more complicated structures (such as NAND/NOR gates), errors in estimating \( \sigma \) approached 50%, even if the error in estimating the mean was within 5%.

Fig. 2 is used to illustrate the difficulty in achieving good statistical accuracy relative to nominal analysis. In this case, simplified formulas (with high nominal accuracy) than those proposed were used to solve the circuit equations, in order to estimate the delay of four 2-input NAND gates in series. The output voltage waveforms \( V_{out}(t) \) from the simplified formulas were compared with the actual waveforms \( V_i(t) \) from SPICE3 for stages, \( i = 1, \cdots, 4 \). For such a configuration, the nominal analysis showed an error of about 4% when comparing the time when the output waveform reaches \( V_{dd}/2 \), i.e., \( t = t|_{V_{out}=V_{dd}/2} \). The mean of a 100 point Monte-Carlo sample data set showed similar accuracy. However, \( \sigma \) had an error of about 30%, even though the statistical accuracy of each 2-input NAND gate in isolation was within 5%. The figure shows how the error propagated from one logic gate to the next for a typical “bad” point in the Monte-Carlo sample. Even though, the number of “bad” points was a small fraction of the total number of points, they had a tendency to affect the extreme boundaries of the distribution of \( t \) (close to the minimum and maximum values), causing a significant error in estimating \( \sigma \). The output waveforms estimated with the strategy proposed in this paper, at the same data point is represented as \( V_{out}(t) \) in Fig. 2. Notice that the error is dramatically reduced, and consequently, estimation of standard deviation \( \sigma \) is also more accurate.

In the following section, the practical applications of the proposed methodology are illustrated on some examples.

4 Examples

Example 1: Inverter and NAND gates

The methodology was initially tested for a CMOS inverter, a 2 input and a 5 input NAND gate, with a loading capacitance of \( 1pF \). The NMOS and
PMOS transistors of the inverter were 3μ/3μ, while all NMOS and PMOS transistors of the NAND gates were 12μ/3μ and 8μ/3μ, respectively. For both NAND gates, the gate voltage $V_g(t)$ was applied to the NMOS transistor closest to the output node, while $V_{dd}$ was applied to the remaining NMOS transistors. A 100 random Monte Carlo samples were generated and the output waveform from the model was compared with the actual waveform. In Table 1, the time $t = t|_{V_a(t)=5V_{dd}}$ calculated from the analytical model and form SPICE are compared. The statistical accuracy of the model is very high ($\sim 5\%$ error in estimating standard deviation’s). The model showed similar accuracy for different sets of widths, lengths, and input waveform shapes. The reduction in analysis times for the inverter and the NAND gates was about 100 and 400, respectively. Next, a more complex example is shown with loading inverters instead of large loading capacitances.

Example 2: 2 bit Adder
Our statistical model was applied to a 2 bit adder circuit (Fig. 3), consisting of 18 two-input NAND gates (so NFETS are connected in short chains), with each output node loaded with an inverter, so the accuracy of the parasitic capacitance models is critical.

![Fig. 2: Error in delay estimation for four 2-input NAND gates in series using simplified formulas and proposed formulas.](image)

![Fig. 3: Example 2: Delay model for a 2 bit adder.](image)

Table 1: Accuracy of delays models for inverter and NAND gates

<table>
<thead>
<tr>
<th></th>
<th>Inverter</th>
<th>NAND2</th>
<th>NAND5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean ($t$)</strong></td>
<td>SPICE</td>
<td>11.1ns</td>
<td>8.40ns</td>
</tr>
<tr>
<td><strong>Model</strong></td>
<td>11.32ns</td>
<td>8.63ns</td>
<td>12.46ns</td>
</tr>
<tr>
<td><strong>Error</strong></td>
<td>1.7%</td>
<td>1.6%</td>
<td>3.6%</td>
</tr>
<tr>
<td><strong>Sigma($\sigma$)</strong></td>
<td>SPICE</td>
<td>0.51ns</td>
<td>0.45ns</td>
</tr>
<tr>
<td><strong>Model</strong></td>
<td>0.60ns</td>
<td>0.51ns</td>
<td>0.51ns</td>
</tr>
<tr>
<td><strong>Error</strong></td>
<td>5.3%</td>
<td>6.8%</td>
<td>3.8%</td>
</tr>
</tbody>
</table>

$V_{th}(t = t|_{V_a(t)=5V_{dd}} = 0.5V/\text{ns}$ and $V_{th}(t > 10\text{ns}) = V_{dd}$.

Example 3: 5 × 5 Baugh-Wooley Multiplier
In order to test the accuracy of the models for a large circuit, a 5 × 5 Baugh-Wooley multiplier (Fig. 4) was selected as an example. The circuit contains approximately 1200 transistors. It has a very regular structure with one bit full adders arranged in the form of rows and columns. The size of the circuit prevented the determination of the standard deviation from SPICE, because of long simulation times involved. In fact, SPICE3 ran into convergence problems and could not simulate this circuit. Therefore, HSPIE was used to determine the delay from node $a_{3b_0}$ to the output node $P_4$. Using a modified version of the program rc_ko [9], the signal path was automatically generated between

![Fig. 4: Example 3: The 5 × 5 Baugh-Wooley multiplier.](image)
these nodes. Since it was not possible to determine \( \sigma \) from the simulator, the accuracy of the model was tested at three selected points only. The nominal point showed very good accuracy (6.5\%) when estimating the time \( t \) when the output waveform reaches \( V_{dd}/2 \). The two data points from the Monte Carlo test set from the previous example (adder) that resulted in the maximum and minimum values of \( t \) were assumed to be critical for this example as well, since both use a very similar structure. The errors at these “estimated” worst-cases, also showed good modeling accuracy (9.9\% and 11.1\%). The analysis was more than 6000 times faster than HSPICE. The effects of the systematic errors from the capacitor model is apparent as the model errors showed systematic increase (from about 6\% at the output of the first full adder B to a maximum of 11\% at the node \( P_b \)) from one node to the next in the signal path selected. Since the errors are still acceptable, no tuning was performed. For much larger circuits, local tuning may have to be performed for some subcircuits (for example, for the full adder), to increase statistical accuracy. Note that the proposed scheme can model significantly larger circuits than this, but there is really no way to check statistical model accuracy for such circuits (SPICE is too slow and runs into convergence problems). The importance of the high modeling accuracy was illustrated, when a race condition was created within an adder block for some specific combinations of input waveform shape and the noise factors. This phenomenon was entirely missed by the RC modeling strategy [2], but was correctly detected by the proposed methodology.

5 Conclusions

In this paper, an efficient methodology was presented to statistically characterize signal delays of CMOS VLSI combinational circuits. Newly developed hybrid analytical/iterative delay formulas are independently dependent on the geometrical (widths and lengths) and noise parameters, through the device (transistor and capacitor) model parameters. The analytical models of the delay of basic building blocks are combined together for analysis of complex combinational circuits. To increase speed of analysis, C-code is generated for specific signal paths. As seen from the examples, the influence of the “noise” parameters is modeled very well. The resulting models increase the speed of analysis by 2–4 orders of magnitude relative to SPICE. Due to this high efficiency, statistical delay characterization of large combinational VLSI circuits has become possible.

Acknowledgements

The authors would like to thank Dr. L. J. Opalski and K. M. Opalska for the code generation program \texttt{rc.ko} [13], used by the authors with modifications. \texttt{rc.ko} was based on the RC delay model program by A. C. Deng and J. F. Tuan.

References