An Enhanced Analytical Electrical Masking Model for Multiple Event Transients

Adam Watkins and Spyros Tragoudas
Department of Electrical and Computer Engineering
Southern Illinois University, Carbondale, IL 62901
watkins@engr.siu.edu, spyros@engr.siu.edu

ABSTRACT

Due to the reducing transistor feature size, the susceptibility of modern circuits to radiation induced errors has increased. This, as a result, has increased the likelihood of multiple transients affecting a circuit. An important aspect when modeling convergent pulses is the approximation of the gate output. Thus, in this paper, a model that approximates the output pulse shape for convergent inputs is proposed. Extensive simulations showed that the proposed model matched closely with HSPICE and provides a speed-up of 15X.

Keywords

Soft Errors; Transient Pulses; Transient Errors; Radiation Induced Errors; Multiple Event Transients; Multiple Event Upsets

1. INTRODUCTION

A soft error occurs when a neutron particle from cosmic radiation or an alpha particle from packaging hits the diffusion region of a transistor. This strike causes a temporary flow of current through the transistor causing a change on the gate output voltage. If the strike is sufficiently severe, the change will manifest into a voltage pulse commonly referred to as a single event upset (SEU). This SEU may propagate through the circuit and cause an observable error at the output.

As the transistor size has continued to decrease, the probability of a SEU occurring has continued to greatly increase. This, in effect, has created a need for accurate and efficient SEU simulators. The SEU problem has been investigated deeply [1], [2], [3]. However, due to the close proximity of transistors and the increase in transistor error sensitivity, the probability of a single or multiple radiation strikes producing multiple voltage pulses has become a concern. The authors in [4] investigated a 65 nm 3x3 inverter matrix and found that 40% of all single radiation strikes resulted in the generation multiple significant voltage pulses. Considering the possibility that a modern circuit is also vulnerable to multiple simultaneous strike, there is a significant concern for multiple transients. This creates the need for either the enhancement of existing SEU tools or the development of new tools which can consider the single event multiple transient (SEMT) and multiple even multiple transient (MEMT) effect.

Previous work in SEU's have shown that combinational circuits are vulnerable to 3 types of error masking effects: logical masking, electrical masking and temporal masking. Accurate estimation of the soft error rate requires consideration of all masking factors for not only SEU's but also the SEMT and MEMT cases. In this paper, the focus is on the consideration of the electrical masking effect. There have been many approaches to the electrical masking problem for the SEU case [5], [6], [7]. However, a major drawback with these approaches is that they do not accurately consider the effect of multiple pulses occurring on the gate input (referred to as convergent pulses). Since, in the SEMT and MEMT cases, there are many pulses generated, the occurrence of convergent pulses is much higher. To model these cases, an accurate electrical masking model must consider convergent pulses.

There has been a few attempts at modeling the effect multiple transient pules, however all approaches suffer from inaccurate estimation of electrical masking. The authors in [8], [9] both proposed a simulation tool to determine the soft error rate (SER) of a combinational circuit in the presence of multiple transients, however, they calculate the output transient of a gate with a convergent input through the simple superposition of the pulse widths. As stated in [6], this is insufficient since accurate electrical masking modeling requires the inclusion of the pulse amplitude and non-linear portions of the glitch.

An additional possible solution to the pulse convergence problem is the simple extension of [6] for convergent pulses. In [6] an accurate iterative model for the single pulse case is proposed. Their model can easily be extended by generating larger look-up tables. However, since they characterize stacked transistors using look-up tables, gates with more than 2-inputs require tables with 125,000 entries for 3-input gates and over 7 million for 4-input gates. Thus our model provides an enhancement to [6] by allowing the characterization of stacked transistors using a look-up tables for the transistor and the Miller capacitance between the gate and drain terminals.

Based on the above discussion, there is a lack of accurate and fast methods to consider convergent pulses. This provides the motivation of this work. To the author's knowledge, the proposed model is the first masking model that accurately and efficiently approximates the entire output pulse shape for convergent input pulses. Through extensive experimentation, the proposed model has been shown to be robust for various input shapes and process technologies while providing a 15X speed-up over HSPICE.

The rest of this paper is organized as follows. In Section II, a background discussing the premise for the model is given. Section III details the derivation of the model's equations. Section IV provides the results and Section V concludes the paper.

2. PULSE APPROXIMATION MODEL

As in [6], an important aspect of the proposed model is the accurate approximation of the transistor drain to source current (I_{DS}). To consider the I_{DS} current, a look-up table for each transistor is created by sweeping both the gate-source voltage (V_{GS}) and the drain-source voltage (V_{DS}).

Additionally, as in [6], the effects of pulse overshoot and undershoot are considered through the use of a Miller capacitor. In contrast to [6],

This research has been supported in part by grants NSF IIP 1432026, and NSF IIP 1361847 from the NSF I/UCRC for Embedded Systems at SIUC. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation

GLSVLSI '16, May 18-20, 2016, Boston, MA, USA
© 2016 ACM. ISBN 978-1-4503-4274-2/16/05…$15.00
DOI: http://dx.doi.org/10.1145/2902961.2903007

however, we consider the MOS parasitic effects on each individual transistor through the inclusion of a Miller capacitance between the gate and drain terminals. Fig. 1 demonstrates the proposed model for a NMOS and a PMOS transistor.

As stated in [6], it was observed that the value of the Miller capacitance varied over the range of drain-source (V_{DS}) and gate-source (V_{GS}) voltages. To consider this effect, it is proposed to generate a look-up table with each entry pertaining to the Miller capacitance at each specific V_{DS} and V_{GS} bias. The Miller capacitance look-up tables can be characterized using HSPICE by sweeping the V_{DS} and V_{GS} voltages and measuring the current flowing through the gate node of the transistor (i_m). Using the current i_m, the derivative of the difference between the gate and drain voltages and equation (1), the capacitance C_{Miller} can be calculated.

$$C_{Miller} = \frac{i_m}{\frac{d(V_{DS} - V_{GS})}{dt}} \tag{1}$$

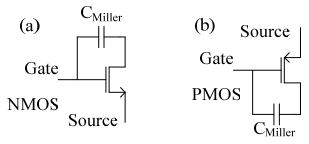

Fig. 1. (a) The transistor model used for a NMOS. (b) The transistor model for a PMOS.

As stated before, the I_{DS} current and the Miller capacitance for both the PMOS and NMOS transistors are characterized. Using this data, each transistor is modeled as a current source and its Miller capacitor. Any given gate can then be represented by replacing each transistor with a current source and a Miller capacitor connected between the drain and gate terminals. At the output of the gate, a load capacitance is added to consider the gate loading effects and to calibrate the model. Fig. 2. demonstrates the conversion to the proposed model on a 2-input NAND gate.

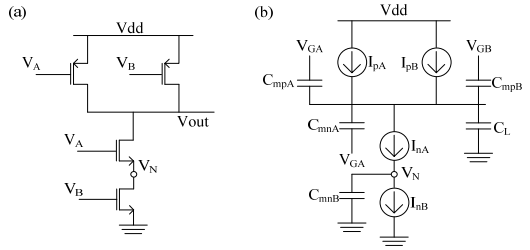

Fig. 2. (a) A standard 2-input CMOS NAND gate. (b) The NAND gate converted to the proposed model.

Before beginning the discussion on the equation derivation it is important to note that this model is general for all types of gates but for the sake of this discussion, the derivation will focus on CMOS NAND and NOR gates.

The proposed gate modeling for a given gate, as in Fig.2b, can be generalized for an *i* number of inputs as in Fig. 3. where a single current source represents the current for the pull-up network and a second current source represents the pull-down network. This is based on the assumption that the gate is a CMOS gate in which one network consists of parallel current sources and the equivalent current is equal to the sum of each parallel transistor current. The other current source represents the current flowing through series connected transistors which ideally have a single constant current. However, in practice, due to parasitic effects, the current in each stacked transistor varies. It was found through many simulations that the best approximation for this current was by taking the average between each transistor current in the stack. For example, if we have the circuit in Fig. 2b, the current on the stacked transistors can be calculated using the following equation:

$$I_{pd} = \frac{I_{nA} + I_{nB}}{2} \tag{2}$$

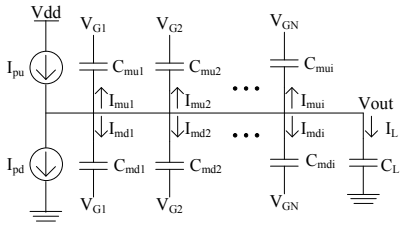

Fig. 3. A generalized representation for the gate model.

To derive the output equation for the node V_{out} in Fig. 3., we first use Kirchoff's Current Law (KCL) at the node V_{out}. Assuming that $I_{mu_i}$and I_{md_i} are the currents through the *i*-th Miller capacitor, I_L is the current through the load capacitor and I_{pu} and I_{pd} represent the pull-up and pull-down current respectively, KCL will result in the following:

$$I_{pu} = I_{pd} + I_L + \sum_{i=1}^{N} I_{mu_i} + \sum_{i=1}^{N} I_{md_i} \tag{3}$$

For our model, we discretize the time into time steps with the variable t_{step}. Additionally, we represent each variable in the current time instant with the subscript *i* and each variable in the previous time instant with the subscript *i-1*. Lastly, we represent the change in a value between two time instances using Δ (i.e. $V_i - V_{i-1} = \Delta V$).

According to the charge conservation law, we get equation (4) for each Miller capacitor C_{m_i} connected between V_{out} and the input voltage V_{G_i}.

$$I_{m_i} = \frac{[\Delta V_{out} - \Delta V_{G_i}]C_{m_i}}{t_{step}} \tag{4}$$

Next, consideration of the load capacitance C_L gives the following equation:

$$I_L = \frac{(\Delta V_{out})C_L}{t_{step}} \tag{5}$$

By substituting into equation (3) the current I_{m_i} from (4) for the currents $I_{mu_i}$and I_{md_i} and replacing the current I_L with equation (5), we get the following for a N_u number of Miller capacitors connected between V_{out} and the *i*-th input in the pull-up network (denoted by C_{mu_i}) and a N_d number of Miller capacitors connected to V_{out} and the *i*-th input on the pull-down network (denoted by C_{md_i}):

$$I_{pu} = I_{pd} + \frac{\Delta V_{out} C_L}{t_{step}} + \sum_{i=1}^{N_u} \frac{[\Delta V_{out} - \Delta V_{G_i}]C_{mu_i}}{t_{step}} + \sum_{i=1}^{N_d} \frac{[\Delta V_{out} - \Delta V_{G_i}]C_{md_i}}{t_{step}} \tag{6}$$

Rearranging (6) and solving for ΔV_{out} gives the following:

$$\Delta V_{out} = \frac{(I_{pu} - I_{pd})t_{step} + \sum_{i=1}^{N_u} \Delta V_{G_i} C_{mu_i} + \sum_{i=1}^{N_d} \Delta V_{G_i} C_{md_i}}{C_L + \sum_{i=1}^{N_u} C_{mu_i} + \sum_{i=1}^{N_d} C_{md_i}} \tag{7}$$

Assuming that ΔV_{out_i} equals $V_{out_i} - V_{out_{i-1}}$we get

$$V_{out_i} = \Delta V_{out} + V_{out_{i-1}} \tag{8}$$

As can be observed in equations (7) and (8) an important aspect of the proposed model is the accurate consideration of both the pull-up

and pull-down currents. As stated before, cases where the pull-up or pull-down transistors are composed of parallel connected transistors the current can simply be found through the sum of the drain to source currents. However, in cases where the transistors are in series, the determination of the equivalent current is more complicated.

In Fig. 2a. the pull-down branch consists of series connected NMOS transistors which will be used to demonstrate the derivation. Note that using the same method, the equations for stacked PMOS devices can be easily derived. Within the stack there is an intermediate node voltage *Vn* which is crucial for the determination of the drain to source currents in the transistors. Through many simulations using the transistor drain to source current look-up tables, it was found that accurate approximation of the node voltages provided accurate current values for each transistor in the stack.

To calculate the intermediate voltage, it is proposed to first convert each transistor to the representations in Fig. 1a and 1b. which results in the circuit in Fig. 2b. In Fig. 2b. the intermediate node voltage *Vn* is connected to a static capacitor representing the loading effect of the subsequent transistor stack. Next, we will look at how to derive the node voltage equation.

Similar to equation (7), the node voltage *Vn* is modeled using discrete time steps. Using KCL at *Vn* in Fig. 2b we get the following equation:

$$I_{D1} = I_{D2} + I_m \tag{9}$$

As in equation (4), the current through the static capacitor can be determined using the charge conservation law resulting in the following for the current I_m.

$$I_m = \frac{C_{m1}\left(V_{n_i} - V_{n_{i-1}}\right)}{t_{step}} \tag{10}$$

By substituting (10) into (9) and solving for V_{n_i}, we get the equation for the node voltage *Vn* at the time instant *i*.

$$V_{n_i} = \frac{(I_{D1} - I_{D2})t_{step}}{C_{m1}} + V_{n_{i-1}} \tag{11}$$

In cases where there are more than 2 stacked transistors, the described model is directly extendable for each intermediate node. Fig. 4a and b provides the conversion for 3 stacked NMOS transistors to our model. The node voltages V_{N1} and V_{N2} can be found in any order using equation (11) and by considering the currents flowing into the node. For example, if we wanted to solve for the node voltage V_{N2} the currents I_{D2} and I_{D3} would be used in equation (11) since they are directly connected to the node. Additionally, the order for solving the node equations does not matter since the node voltages from the previous step (step *i-1*) are used.

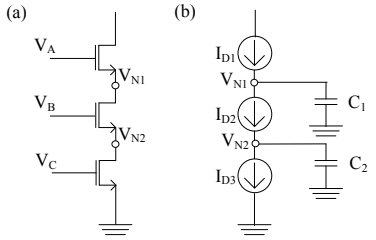

Fig. 4. (a) Three stacked NMOS transistors with two intermediate node *Vn1* and *Vn2*. (b) The transistors converted to the proposed circuit model.

3. EXPERIMENTAL RESULTS

The proposed algorithm was implemented in Matlab and characterized using the 32 nm and 45 nm PTM library and the 65 nm IBM library. To test the proposed model, we created 3 benchmark circuits for each process library to test two convergent pulses on a 2 and 3 input NAND gate and a 2 input NOR gate. The first two benchmark circuits denoted by NAND2R and NAND2F in column 1 of Table 1, consists of two NAND gate chains with each off-input held to a non-controlling value. The chains contain two gates and the R and F in the benchmark names represent either a rising (R) or falling (F) pulse on the convergent gate's inputs. Each chain is then connected to an input of the NAND gate. The benchmark circuits NOR2R and NOR2F are constructed similarly but the chains consist of NOR gates and are connected to the input of a 2 input NOR gate. The last benchmark circuits constructed, NAND3R and NAND3F, were created with 2 2 gate chains connected to the first and second input of the 3-input NAND gate. The remaining off-input was held to a non-controlling value.

Next, to accurately model the injected pulse shape, we represented the injection current as a current source located at the output of the first gate in the chain. We then used the function for the pulse shape given in [10] to relate the strike current to a given charge. The equation is given below with Q_o representing the charge resulting from a strike, τ representing a technology dependent pulse shaping parameter and *t* being the variable for time.

$$I(t) = \frac{2Q_o}{\tau\sqrt{\pi}}\sqrt{\frac{t}{\tau}}e^{\frac{-t}{\tau}} \tag{12}$$

Using equation (12), Q_o was varied over a range of charge values that results in an observable error. For our test cases, we found that a charge ranging from 3-16 fC for a τ of $90\text{x}10^{-12}$ provided an observable output shape for all processes tested. We then propagated the pulses through the gate chains to gate of interest and compared the generated outputs using the proposed model and the model in [6] to HSPICE. Table 1 details the average error for 30 cases per gate and transition type for the 32, 45 and 65 nm processes using 2400 data points over a range of 0 to 1.2 pico-seconds. The first column gives the benchmark name. All subsequent columns provide the MSE error for each test case for the proposed method and the method in [6]. Assuming that N represents the number of calculated points, V_i^H represents the pulse voltage at point *i* on the signal from HSPICE and V_i^C is the voltage calculated by the proposed model, the MSE was calculated using the following equation:

$$MSE = \frac{\sum_{i=0}^{N}(V_i^H - V_i^C)^2}{N} \tag{13}$$

As can be seen from Table 1, the proposed model is much more accurate over the model proposed in [6]. In all tested cases, the method in [6] predicted there would be an output pulse. However, they only propagate the pulse with the largest width which fails to consider the effect of multiple pulses and provides a pulse with a much smaller width. For this reason, their method tends to underestimate the severity of the pulse in the presence of convergent pulses. According to Table 1, our method closely predicts the output pulse shape.

Next, we run a simulation to observe how the number of points selected relates to the speed and accuracy. In this simulation, we used the NAND2F and NAND3F benchmark and applied a pulse to each input of the terminating NAND gate. We then varied the number of output data points for the terminating NAND gate in both HSPICE and the proposed model. Table 2 provides the results of this simulation.

The first column provides the number of simulation points used to calculate the result. Columns 2 and 3 provide the error of the proposed model and [6] for the given number of points. Columns 4, 5 and 6 gives the execution time of the proposed method, the method in [6] and HSPICE respectively. As can be seen from Table 2, the reduction in simulation time scales linearly to the reduction in simulation points.

In Fig. 5, we provide the output waveform of the simulation in Table 2 for 900 and 300 points respectively. As can be observed in Fig. 5a, the result for the proposed model closely matches the result obtained in HSPICE when 900 simulation points are used. If a higher accuracy is desired, the proposed model provides a near perfect match with 2400 points. For both the 900 and 300 point cases, it can be observed that the method in [6] greatly underestimates the resulting pulse thus providing a high MSE. The proposed model provides an accurate result with an MSE of 7.40×10^{-3} while still providing a speed-up of 15X compared to HSPICE for this benchmark. In general we have observed a speed-up of 15X for any 2-input gate using 900 points while maintaining a high accuracy

In Fig. 6a and 6b, we provide the output waveforms for the NAND3F simulation for 900 and 300 points respectively. Similar to the NAND2F benchmark, a simulation with 900 points provides a close waveform approximation. However, if a lower accuracy is tolerable, the number of points can be reduced further providing a faster simulation time as demonstrated in Table 2 and Fig. 6b. As can be observed in Table 2, proposed model provides a speed-up of 12X for the 900 point case. In general for any 3-input gate, we have observed a speed-up of 12X using 900 points while maintaining a high accuracy

Table 1. Proposed and [6] compared to HSPICE for 2400 points

Circuit	MSE [65nm] ($x10^{-3}$)		MSE [45nm] ($x10^{-3}$)		MSE [32nm] ($x10^{-3}$)	
	Proposed	[6]	Proposed	[6]	Proposed	[6]
NAND2R	0.760	108	0.453	35.6	0.190	45.6
NAND2F	1.000	110	0.550	67.8	0.467	50.9
NOR2R	0.220	23.0	0.345	133.0	0.402	146.2
NOR2F	0.986	103	0.678	42.5	0.493	33.9
NAND3R	0.887	115	0.785	55.6	0.352	39.2
NAND3F	0.998	98.9	0.467	39.8	0.465	30.8

Table 2. Execution Time vs Accuracy ($x10^{-3}$)

Points	Error Proposed	Error [6]	Execution Time Proposed	Execution Time [6]	Execution Time HSPICE
NAND2F					
2400	0.486	85.7	8.39	8.03	100
1200	3.90	92.0	4.34	4.01	60
900	7.40	97.8	3.43	3.05	50
300	41.1	115.9	1.23	1.18	20
NAND3F					
2400	0.529	37.5	12.1	11.4	120
1200	6.40	38.7	6.40	5.90	70
900	13.0	42.8	4.90	4.50	60
300	60.9	85.0	1.91	1.50	30

4. CONCLUSION

In this paper, we proposed a model that improves SEMT and MEMT error simulation by calculating the output pulse shape in the presence of convergence with an MSE of less than 7.40×10^{-3} for 2-input gates and 13×10^{-3} for 3-input gates using 900 points and a speed-up of 15X and 12X respectively compared to HSPICE. Future work with our model includes the study of the number of points effect the simulation result on large benchmarks and the study of SEMT and MEMT on large standard benchmark circuits using the proposed model.

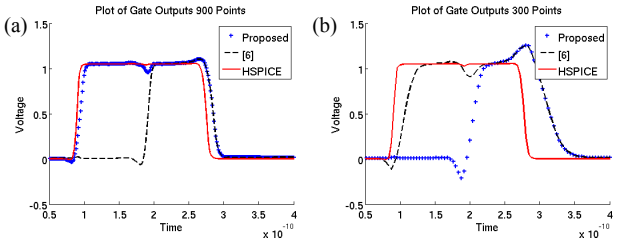

Fig. 5. (a) Output of NAND2F with 900 points. (b) Output of NAND2F with 300 points.

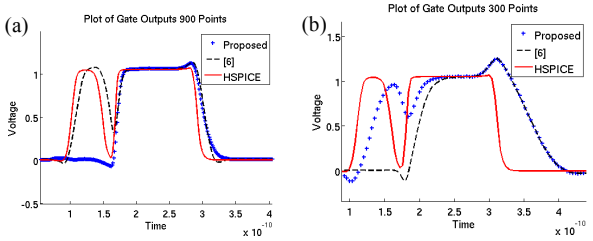

Fig. 6. (a) Output of NAND3F with 900 points. (b) Output of NAND3F with 300 points.

5. REFERENCES

[1] Miskov-Zivanov, N.; Marculescu, D.; , "MARS-C: modeling and reduction of soft errors in combinational circuits," Design Automation Conference, 2006 43rd ACM/IEEE , vol., no., pp.767-772, 0-0 0

[2] Ming Zhang; Shanbhag, N.R.; , "Soft-Error-Rate-Analysis (SERA) Methodology," Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on , vol.25, no.10, pp.2140-2155, Oct. 2006

[3] Bin Zhang; Wei-Shen Wang; Orshansky, M.; , "FASER: fast analysis of soft error susceptibility for cell-based designs," Quality Electronic Design, 2006. ISQED '06. 7th International Symposium on , vol., no., pp.6 pp.-760, 27-29 March 2006

[4] Harada, R.; Mitsuyama, Y.; Hashimoto, M.; Onoye, T., "Neutron induced single event multiple transients with voltage scaling and body biasing," in *Reliability Physics Symposium (IRPS), 2011 IEEE International* , vol., no., pp.3C.4.1-3C.4.5, 10-14 April 2011

[5] Omana, M.; Papasso, G.; Rossi, D.; Metra, C.; , "A model for transient fault propagation in combinatorial logic," On-Line Testing Symposium, 2003. IOLTS 2003. 9th IEEE , vol., no., pp. 111- 115, 7-9 July 2003

[6] Feng Wang; Yuan Xie; , "Soft Error Rate Analysis for Combinational Logic Using an Accurate Electrical Masking Model," Dependable and Secure Computing, IEEE Transactions on , vol.8, no.1, pp.137-146, Jan.-Feb. 2011

[7] Watkins, A.; Tragoudas, S., "Transient pulse propagation using the Weibull distribution function," in *Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFT), 2012 IEEE International Symposium on* , vol., no., pp.109-114, 3-5 Oct. 2012

[8] Miskov-Zivanov, Natasa; Marculescu, D., "Multiple Transient Faults in Combinational and Sequential Circuits: A Systematic Approach," in *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on* , vol.29, no.10, pp.1614-1627, Oct. 2010

[9] Fazeli, M.; Ahmadian, S.N.; Miremadi, S.G.; Asadi, H.; Tahoori, M.B., "Soft error rate estimation of digital circuits in the presence of Multiple Event Transients (METs)," in *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2011* , vol., no., pp.1-6, 14-18 March 2011

[10] J. Ziegler, "Terrestrial cosmic rays," IBM J. Res. Develop., vol. 40, no. 1, pp. 19–39, Jan 1996