A Design-Oriented Soft Error Rate Variation Model Accounting for Both Die-to-Die and Within-Die Variations in Submicrometer CMOS SRAM Cells

Hassan Mostafa, *Student Member, IEEE*, Mohab Anis, *Member, IEEE*, and Mohamed Elmasry, *Fellow, IEEE*

***Abstract*—Submicrometer static random access memory cells are more susceptible to particle strike soft errors and increased statistical process variations, in advanced nanometer CMOS technologies. In this paper, analytical models for the critical charge variations accounting for both die-to-die and within-die variations are proposed. The derived models are verified and compared to Monte Carlo simulations by using industrial 65-nm CMOS technology. This paper provides new design insights such as the impact of the coupling capacitor, one of the most common soft error mitigation techniques, on the critical charge variability, especially, at lower supply voltages. It demonstrates that two extreme values of this coupling capacitor exist. The first value results in maximum relative variations and the other results in minimum relative variations. Therefore, the circuit designers can utilize these results to design the coupling capacitor to limit the variations under power and performance constraints in early design cycles. The derived analytical models account for the impact of the supply voltage and different particle strike conditions. These results are particularly important for soft error tolerant and variation tolerant designs in submicrometer technologies, especially, for low power operations.**

***Index Terms*—Deep submicrometer, process variations, reliability, soft errors, static random access memory (SRAM).**

I. Introduction

RELIABILITY is one of the major design challenges for submicrometer CMOS technology. Shrinking geometries, lower power supply, higher clock frequencies, and higher density circuits all have a great impact on reliability [1]–[7]. As CMOS technology further scales, soft errors become one of the major reliability concerns. Soft errors are caused by two types of radiation: 1) alpha particles emitted by radioactive impurities in integrated circuits (ICs) and package materials and 2) high energy neutrons resulting from the interaction between cosmic rays and the earth atmosphere [3], [4]. When an alpha particle hits a silicon substrate, the particle generates electron-hole pairs, as it passes through p-n junctions. Although a neutron does not ionize the material directly, it does collide with atoms, resulting in products capable of inducing electron-hole pairs. The generated charges are transported to circuit nodes by drift and diffusion mechanisms, causing a current pulse that disturbs the node voltage and can lead to soft errors [2].

Manuscript received April 10, 2009; revised June 26, 2009; accepted August 24, 2009. First published January 22, 2010; current version published June 09, 2010. This paper was recommended by Associate Editor V. Kursun.

The authors are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada N2L3G1 (e-mail: hmostafa@uwaterloo.ca; manis@vlsi.uwaterloo.ca; elmasry@uwaterloo.ca).

Digital Object Identifier 10.1109/TCSI.2009.2033528

In memory elements, this disturbance can cause bit flips (a 0-to-1 flip or a 1-to-0 flip) which may corrupt the logic state of the circuit. However, in combinational circuits, it may cause a temporary change in the output node voltage. This temporary change can be tolerated, unless it is latched by a succeeding memory element.

For memory elements such as static random access memory (SRAM) and flip-flops, if the charge collected by the particle strike at the storage node, is more than a minimum value, the node is flipped and a soft error occurs. This minimum value is called a critical charge (Q_{critical}), which can be used as a measure of the memory element vulnerability to soft errors [2], [5], [7]–[10]. This critical charge exhibits an exponential relationship with the soft error rate (SER) [2], and consequently, this critical charge should be designed high enough, to limit the SER. SRAM cells are more vulnerable to soft errors due to their lower node capacitance. Moreover, since SRAM occupies the majority of the die area in system-on-chips and microprocessors, different leakage reduction techniques such as supply voltage reduction and dynamic voltage scaling, are applied to SRAMs to limit the overall chip leakage. These techniques increase the SRAM soft error vulnerability by reducing the critical charge.

Process variations are expected to worsen in future technologies, due to difficulties with printing nanometer scale geometries in standard lithography. Therefore, these variations are considered another main challenge in CMOS technology scaling [11]–[15]. They can be classified as die-to-die (D2D) variations and within-die (WID) variations. In D2D variations, all the devices on the same die are assumed to have the same parameters values. However, the devices on the same die are assumed to behave differently in WID variations [11]. Although D2D variations were originally considered the main source of process variations, WID variations are posing the major design challenge as technology scales [12], [13]. The D2D variations can be easily modeled by using corner-based models, which assume that all devices on a given die have the same parameter value, that is shifted away from the mean by a fixed amount. However, WID variations modeling requires representing each device parameter, within the same die, by a separate random variable. These WID variations random variables should be treated statistically which makes the WID variations modeling much more complex and difficult than D2D variations modeling.

Due to the existence of process variations, the critical charge has variations around its nominal value which can result

in SRAM failure to meet robustness constraints. Recently, researchers have attempted to calculate the critical charge nominal value as well as addressing the impact of process variations on the critical charge in memory elements such as SRAM cells and flip-flops. However, most of this research is conducted by using Monte Carlo analysis tools [1], [16]–[18], which are time consuming and provide little design insights. Moreover, these Monte Carlo analysis tools are not scalable with technology. From a design perspective, few articles have been published on modeling the critical charge and its variations. In [19]–[21], different models for the critical charge are proposed, however, these models overestimate the critical charge value and provide little insights to circuit designers. In [22], an analytical model to estimate the critical charge is presented. Despite its accuracy in modeling the critical charge, this model depends mainly on SPICE simulations. Thus, this model can be used only when dealing with D2D variations. These D2D variations are estimated by applying corner-based analysis that have been already performed in [22]. These techniques tend to be inefficient, and completely pessimistic in the presence of relatively large variations. Therefore, statistical design-oriented techniques are required, especially, when dealing with the WID variations [23].

In this paper, an accurate analytical model of the critical charge, accounting for both D2D and WID variations, is proposed. This model is further simplified to provide more design insights on the impact of process variations on the critical charge. The derived model is simple, scalable in terms of technology scaling. Moreover, it shows explicit dependence on design parameters such as node capacitance, transistors sizing, transistor parameters, and supply voltage. This is a very essential step since supply voltage reduction is one of the most common techniques for low power applications. The results are verified by using SPICE transient and Monte Carlo simulations and an industrial 65-nm CMOS technology transistor model. These results are particularly important for the design of nanometer technology, when WID variations dominate the process variations [13].

The rest of this paper is organized as follows. In Section II, the proposed models and the previous critical charge models are compared qualitatively to show the advantages of the proposed models, especially, in accounting for WID variations. The exact model assumptions and derivations for both the nominal critical charge value and its variability are proposed in Section III. This exact model is further simplified in Section IV to provide more design insights to circuit designers. The proposed models are compared with SPICE transient and Monte Carlo simulations in Section V. In Section VI, the design insights extracted from the proposed models are discussed. Finally, some conclusions are drawn in Section VII.

II. Review of the Previous Critical Charge Models

The previous critical charge models, introduced in [19]–[22], exhibit some limitations, that make them incapable of modeling the WID variations. For example, the model introduced in [19] modeled Q_{critical} as follows:

$$Q_{\text{critical}} = C_1 V_{\text{DD}} + i_{p1\,\max} t_f \tag{1}$$

where $i_{p1\,\max}$ is the maximum restoring current of the transistor M_{p1}. The critical charge obtained from this model is overestimated, because of the following two reasons: 1) the flipping threshold voltage of an inverter is less than V_{DD} (around $V_{\text{DD}}/2$) and 2) the restoring current term ($i_{p1\,\max} t_f$) considers only the maximum current value which is not a valid assumption for the time varying restoring current. These issues have been refined to some extent in [20], by defining the critical charge as

$$Q_{\text{critical}} = \int_0^{V_{\text{trip}}} C_1 dV + \rho i_{p1} t_{\text{pulse}} \tag{2}$$

where V_{trip} is the tripping point of the SRAM cell, ρ is a correction factor, and t_{pulse} is the duration of the particle induced current pulse. This model provides a better estimation of Q_{critical}. However, both models in [19] and [20] cannot be used to model the variations (D2D or WID variations), since they account only for M_{p1} current and ignore the currents of M_{n2} and M_{p2} which can have a significant contribution to the critical charge variability. The work in [21] presents an analytical method to calculate Q_{critical} in terms of the transistor parameters and the injected current pulse magnitude and duration. This model uses a rectangular current pulse, instead of using an exponential current pulse, to model the particle strike induced current pulse, which makes its accuracy in calculating Q_{critical} very poor. If an exponential current pulse is to be used, the model becomes complex and provides little insights. In addition, the model ignores the nMOS transistors current (i.e., M_{n2}), and does not show its effectiveness in calculating Q_{critical}, when different transistor parameters vary.

Finally, the work in [22] introduces a very accurate model in calculating Q_{critical}. However, the value of the injected current pulse charge Q is obtained via iterative transient simulations by increasing Q by a small amount (~ 0.001 fC) in SPICE till flipping occurs. Although this method can be used in calculating D2D variations by using corner-based or worst-case methods, in which the value of Q can be obtained by using SPICE simulations. This technique can not be used for the WID statistical variations, since Q must be calculated for each statistical run. Consequently, this model accounts only for D2D variations, which have been already performed in [22].

The proposed exact model overcomes all the previous limitations, and introduces analytical formulas for Q_{critical} which can be employed without SPICE simulations (assuming that the transistor parameters such as α and V_t are known). Moreover, the developed exact model accounts for both D2D and WID variations. The disadvantage of this exact model is its complexity in the WID variations modeling, which is refined by using the simplified model. The simplified model introduces only three equations (25), (26), and (27), that provide useful design insights reported later in Section VI.

III. Exact Model Assumptions and Derivations

Fig. 1 shows a typical six transistor (6T) SRAM cell. It consists of two cross-coupled inverters, that store two complementary logic values ("1" and "0") at their output nodes. These output nodes are denoted by V_1 and V_2. The SRAM cell has its highest susceptibility to particle strikes in the standby mode, since, in the standby mode, the storage nodes are disconnected

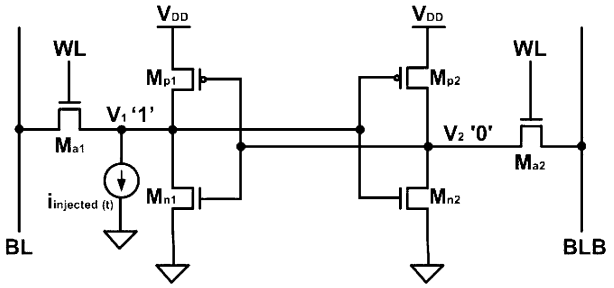


Fig. 1. SRAM cell with the particle strike induced current pulse ($i_{\text{injected}}(t)$). Node V_1 is assumed to be at logic "1" and node V_2 is assumed to be at logic "0".

from the highly capacitive bitlines. Therefore, their critical charge is smaller than that when the SRAM cell is operating in the read mode. In addition, the SRAM cell is most likely in the standby mode during its operating time. Thus, the access transistors M_{a1} and M_{a2} are excluded from the analysis. For the proper operation of the SRAM cell, the pMOS pull-up transistors are sized to be weaker than the nMOS pull-down transistors. Consequently, the data node storing logic "1" is the most susceptible to particle strikes. It has been reported that Q_{critical} of a 0-to-1 flip in SRAM is about 22× larger than that for a 1-to-0 flip [24]. Therefore, the proposed critical charge models account for the 1-to-0 flip case only. Assume that node V_1 stores logic "1" and accordingly node V_2 stores logic "0". Hence, only transistors M_{p1} and M_{n2} are "ON".

A. Critical Charge Model

In order to determine the critical charge model at node V_1, which is more susceptible to soft errors, the particle strike is modeled by a double exponential current pulse given by [25]

$$i_{\text{injected}}(t) = \frac{Q}{\tau_f - \tau_r} \times [\exp(-t/\tau_f) - \exp(-t/\tau_r)] \quad (3)$$

where Q is the total charge deposited by this current pulse at the struck node, and τ_f and τ_r are the falling time and the rising time constants, respectively [25]. Although different current pulse waveforms are reported in [2], the current pulse waveform in (3) has the advantage of being accurate, as well as simple for the proposed analytical model. Typically, for a particle induced current pulse, τ_f is much larger than τ_r [2], [22]. Based on this fact, and for model simplicity, we further approximate (3) as a single exponential current pulse, as given in the following equation:

$$i_{\text{injected}}(t) \approx \frac{Q}{\tau} \times \exp(-t/\tau) \quad (4)$$

where τ is equal to τ_f in (3). The nodal current equation at node $\mathrm{V_1}$ is written as

$$C_1 \frac{dV_1}{dt} = i_{p1}(t) - i_{\text{injected}}(t) \quad (5)$$

where C_1 is node V_1 capacitance; $i_{p1}(t)$ is the pMOS transistor, M_{p1}, restoring current, which tries to pull-up node V_1; and $i_{\text{injected}}(t)$ is the injected current pulse given in (4). It should be noted that transistor M_{n1} subthreshold current is ignored in this analysis [22].

From (5), the values of Q and τ, that equalize $i_{p1}(t)$ and $i_{\text{injected}}(t)$ currents, can be obtained. Hence, node V_1 voltage attains a certain minimum value, V_{min}, which can be obtained by equating these two currents. Since transistor M_{p1} is in the linear region, M_{p1} can be modeled by a resistor R_{p1}. As a result, (5) is rewritten as follows:

$$C_1 \frac{dV_1}{dt} = \frac{V_{\text{DD}} - V_1}{R_{p1}} - \frac{Q}{\tau} \times \exp(-t/\tau) \quad (6)$$

where V_{DD} is the supply voltage. The minimum voltage V_{min} is computed by equating the two currents and the time at which this V_{min} occurs, t_{min}, is obtained by solving the differential equation in (6) and finding the time at which $V_1 = V_{\text{min}} \cdot t_{\text{min}}$ and V_{min} are expressed as [22]

$$t_{\text{min}} = \frac{\tau R_{p1} C_1}{\tau - R_{p1} C_1} \times \ln\left(\frac{\tau}{R_{p1} C_1}\right) \quad (7)$$

$$V_{\text{min}} = V_{\text{DD}} - \frac{Q R_{p1}}{\tau} \times \left(\frac{R_{p1} C_1}{\tau}\right)^{\frac{R_{p1} C_1}{\tau - R_{p1} C_1}}. \quad (8)$$

The work in [22] finds Q by using transient SPICE simulations. Therefore, if the model in [22] is to be used for statistical WID variations modeling, this value of Q must be found for each run, which turns out to be completely inefficient. This is the reason why this model can only be used for the D2D variations modeling, which has been already performed in [22].

In the proposed model, we assume that once node V_1 voltage hits its minimum value, V_{min}, the pMOS transistor, M_{p1}, restoring current causes V_1 voltage to either recover to logic "1" and no flipping occurs, or flip to logic "0" and flipping occurs. This assumption is justified by noting that after the time t_{min}, the injected current $i_{\text{injected}}(t)$ continues decaying exponentially according to (4). Therefore, the goal is to find the condition on the restoring current, $i_{p1}(t)$, that causes node V_1 to flip. This restoring current is controlled by its gate voltage, V_2. Accordingly, if V_2 is rising, the source to gate voltage of M_{p1} decreases, and correspondingly, the restoring current decreases resulting in a soft error. On the other hand, if V_2 is falling, the restoring current increases, and correspondingly, node V_1 voltage recovers and no flipping occurs.

Due to the fact that the inverter switching voltage V_M is defined as, the threshold between logic "1" and logic "0" (i.e., when the inverter input slightly exceeds V_M, the inverter output is assumed to be at logic "0", and vice versa). If V_{min} is slightly below the switching voltage of the second inverter, V_{M2}, V_2 rises to logic "1" decreasing the restoring current, and resulting in a soft error.

Consider the flipping case (i.e., $V_{\text{min}} < V_{M2}$), node V_2 voltage stays around 0 V, for the time interval over which V_1 is approaching V_{min} (i.e., t_{min}), and then starts to rise. Furthermore, V_1 is assumed to remain constant at V_{min}, until V_2 rises and exceeds the switching threshold of the first inverter V_{M1}. The time at which V_2 hits (V_{M1}) is denoted by t_f, which refers to the SRAM cell flipping time. These assumptions are validated by noticing that once V_2 hits V_{M1}, the positive feedback of the cell becomes strong enough to continue flipping the cell state. Moreover, these assumptions allow us to decouple the cross-coupled inverters of the SRAM cell, as proposed in [22].

From (8), and for a given τ, the value of Q, that just cause V_1 to flip, is obtained by equating $V_{\min}$ to V_{M2}. Correspondingly, Q is determined by

$$Q = \frac{(V_{\rm DD} - V_{M2})\tau}{R_{p1}\beta} \text{ where } \beta = \left(\frac{R_{p1}C_1}{\tau}\right)^{\frac{R_{p1}C_1}{\tau - R_{p1}C_1}}. \quad (9)$$

From (9), Q is obtained without SPICE simulations. Therefore, the main limitation in [22] for WID variations modeling is refined.

Now, the objective is to find the flipping time, t_f. The flipping time, t_f, is the sum of $t_{\min}$, and the time delay that V_2 takes to rise from 0 V to V_{M1} (this time is denoted by $t_{\rm up}$). This delay is driven by transistors M_{p2} and M_{n2}, where their gate voltage V_1 is constant at V_{M2}. Transistor M_{p2} is in the saturation region. However, transistor M_{n2} is in the linear region, when V_2 rises from 0 V to $(V_{M2} - V_{tn2})$, where V_{tn2} is the threshold voltage of M_{n2}. When V_2 exceeds $(V_{M2} - V_{tn2})$, transistor M_{n2} is in the saturation region. The currents of these two transistors are given by

$$i_{n2} = \begin{cases} \frac{V_2}{R_{n2}} & 0 \le V_2 \le (V_{M2} - V_{tn2}) \\ i_{n2\rm sat} & (V_{M2} - V_{tn2}) \le V_2 \le V_{M1} \end{cases}$$
$$i_{p2} = \{\, i_{p2\rm sat} \quad 0 \le V_2 \le V_{M1} \quad (10)$$

where i_{p2} and i_{n2} are the currents of transistors, M_{p2} and M_{n2}, respectively, $i_{p2\rm sat}$ and $i_{n2\rm sat}$ are the saturation currents of transistors M_{p2} and M_{n2}, respectively, and R_{n2} is the linear region equivalent resistance of transistor M_{n2}. The nodal current equation at node V_2 is given by

$$C_2\frac{dV_2}{dt} = i_{p2} - i_{n2} \quad (11)$$

where C_2 is the node capacitance of node V_2. From (10) and (11), it is obvious that $t_{\rm up}$ can be divided into two time delays. The first time delay $t_{\rm up1}$ is the time delay taken when V_2 rises from 0 V to $(V_{M2} - V_{tn2})$, while transistor M_{n2} is in the linear region. The other time delay $t_{\rm up2}$ is the time elapsed when V_2 rises from $(V_{M2} - V_{tn2})$ to V_{M1}, while M_{n2} is in the saturation region. These assumptions are justified by noticing that the velocity saturation voltage value $V_{\rm DSAT}$ is close to $(V_{M2} - V_{tn2})$, as given in [26] for deep submicrometer technologies. Following that, the differential equation in (11) is solved in two time intervals with the following boundary conditions, $V_2(t_{\min}) = 0$ V, $V_2(t_{\min} + t_{\rm up1}) = (V_{M2} - V_{tn2})$, and $V_2(t_f) = V_{M1}$, yielding

$$t_{\rm up1} = C_2R_{n2}\ln\left(\frac{i_{p2\rm sat}R_{n2}}{i_{p2\rm sat}R_{n2} - (V_{M2} - V_{tn2})}\right)$$
$$t_{\rm up2} = C_2\frac{V_{M1} - (V_{M2} - V_{tn2})}{i_{p2\rm sat} - i_{n2\rm sat}}. \quad (12)$$

It should be noted that the above assumptions are valid only if V_{M1} is larger than $(V_{M2} - V_{tn2})$, which is usually the case. However, if V_{M1} is smaller, the transistor M_{n2} does not enter the saturation region. As a result, $t_{\rm up}$ becomes the time elapsed when V_2 rises from 0 V to V_{M1} with transistor M_{n2} in the linear region. This time has the same formula as $t_{\rm up1}$ by replacing $(V_{M2} - V_{tn2})$ with V_{M1}. By using (7), (8), and (12), the flipping time t_f is expressed as (13), shown at the bottom of the page. Thus, the critical charge $Q_{\rm critical}$ is obtained as follows [19]–[22], [27]:

$$Q_{\rm critical} = Q(1 - \exp(-t_f/\tau)). \quad (14)$$

In this derivation, the focus is on the supply voltage range covering the super-threshold region, without accounting for the subthreshold operation. To simplify the analysis, the well-known alpha-power law model for the transistors current [28], is adopted. In [28], the transistor current in the saturation region is modeled by

$$i_n = K_{n'}(W/L)(V_{\rm GS} - V_{tn})^{\alpha_n} \quad (15)$$

where V_{tn} is the threshold voltage, $K_{n'}$ is a technological parameter, α_n is the velocity saturation exponent ranging from 1 to 2, depending on whether the transistor is in deep velocity or pinch-off saturation, and W and L are the width and length of the transistor channel, respectively.

According to this model, the inverter switching voltage V_M is given by [28]

$$V_M = \frac{r(V_{\rm DD} - |V_{tp}|) + V_{tn}}{1 + r}$$

where

$$r = \left(\frac{K_{p'}(W/L)_p}{K_{n'}(W/L)_n}\right)^{1/\alpha} \qquad \alpha = \alpha_n = \alpha_p \quad (16)$$

where V_{tn} and V_{tp} are the threshold voltages, α_n and α_p are the velocity saturation exponents, $K_{n'}$ and $K_{p'}$ are the technology parameters, and $(W/L)_n$ and $(W/L)_p$ are the aspect ratios of the nMOS and pMOS transistors, respectively.

In addition, the currents $i_{p2\rm sat}$ and $i_{n2\rm sat}$ are given by

$$i_{p2\rm sat} = K_{p'}(W/L)_{p2}(V_{\rm DD} - V_{M2} - |V_{tp2}|)^{\alpha}$$
$$i_{n2\rm sat} = K_{n'}(W/L)_{n2}(V_{M2} - V_{tn2})^{\alpha} \quad (17)$$

and the resistances R_{p1} and R_{n2} are computed by

$$R_{p1} = \frac{1}{K_{p'}(W/L)_{p1}(V_{\rm DD} - |V_{tp1}|)}$$
$$R_{n2} = \frac{1}{K_{n'}(W/L)_{n2}(V_{M2} - V_{tn2})}. \quad (18)$$

Using (7)–(9), (12)–(14), and (16)–(18), The critical charge, $Q_{\rm critical}$, can be obtained without doing any SPICE simulations.

$$t_f = \begin{cases} \tau\ln\left(\frac{1}{\beta}\right) + t_{\rm up1} + t_{\rm up2} & V_{M1} > (V_{M2} - V_{tn2}), \\ \tau\ln(\frac{1}{\beta}) + C_2R_{n2}\ln\left(\frac{i_{p2\rm sat}R_{n2}}{i_{p2\rm sat}R_{n2} - V_{M1}}\right), & V_{M1} < (V_{M2} - V_{tn2}) \end{cases} \quad (13)$$

B. Statistical Critical Charge Variation Model

Process variations affect device parameters, resulting in fluctuations in the critical charge. The primary sources of process variations, that affect the device parameters, are as follows.

1) *Random Dopant Fluctuations (RDF).* The number of dopants in the MOSFET depletion region decreases, as technology scales. Due to the discreteness of the dopant atoms, there is a statistical random fluctuation of the number of dopants, within a given volume, around their average value [23], [29]. This fluctuation in the number of dopants in the transistor channel results in device threshold voltage variations. It has been shown that the threshold voltage variation, due to RDF, is normally distributed, and its standard deviation $\sigma_{V_t,\mathrm{RDF}}$ is inversely proportional to the square root of the transistor active area (WL). Therefore, these variations can be mitigated by sizing the transistors up, at the expense of more power consumption, and area overhead [23], [29], [30].
2) *Channel Length Variations.* For sub-90-nm nodes, optical lithography requires light sources with wavelengths much larger than the minimum feature sizes for the technology [14]. Therefore, controlling the critical dimension (CD) at these technology nodes becomes so difficult. The variation in CD (i.e., channel length of the transistor) impacts directly the transistor V_t. In short channel devices, the threshold voltage, V_t, has an exponential dependence on the channel length, L, due to charge sharing and drain-induced barrier lowering (DIBL) effects [23], [26], [29]. Thus, a slight variation in L introduces large variation in V_t due to this exponential dependence.

Although the RDF and channel length variations are considered the dominant sources of device variations [13], there are many other sources such as line edge roughness (LER), oxide charge variations, mobility fluctuations, gate oxide thickness variations, channel width variation, and aging effects that affect the device threshold voltage variations [29].

From a circuit modeling perspective, the total variation in V_t, due to RDF, channel length variation, as well as other sources of variation, is modeled as [23]

$$\sigma_{V_t} = \sqrt{\sigma^2_{V_t,\mathrm{RDF}} + \sigma^2_{V_t,L} + \sigma^2_{V_t,\mathrm{Other}}}. \tag{19}$$

Throughout this paper, we are dealing with the total variation in threshold voltage (σ_{Vt}), as modeled in (19).

From the equations derived in Section III-A, it is evident that the critical charge Q_{critical} is dependent on the threshold voltages of transistors M_{p1}, M_{p2}, M_{n1}, and M_{n2}, which are denoted by $V_{tp1}, V_{tp2}, V_{tn1}$, and V_{tn2}, respectively. A small change in these threshold voltages results in an incremental change in the critical charge $(\Delta(Q_{\mathrm{critical}}))$ that is calculated by using Taylor expansion around the nominal value as follows:

$$\Delta Q_{\mathrm{critical}} = \frac{\partial Q_{\mathrm{critical}}}{\partial V_{tp1}}\Delta V_{tp1} + \frac{\partial Q_{\mathrm{critical}}}{\partial V_{tp2}}\Delta V_{tp2} + \frac{\partial Q_{\mathrm{critical}}}{\partial V_{tn1}}\Delta V_{tn1} + \frac{\partial Q_{\mathrm{critical}}}{\partial V_{tn2}}\Delta V_{tn2} \tag{20}$$

where $\Delta Vt_{p1}, \Delta Vt_{p2}, \Delta Vt_{n1}$, and ΔVt_{n2} are the variations of the threshold voltages. The partial derivative terms in (20) can be computed numerically at the mean threshold voltages. Therefore, the standard deviation of the critical charge variations is calculated as follows:

$$\sigma_{Q_{\mathrm{critical}}} = \left\{ \left(\frac{\partial Q_{\mathrm{critical}}}{\partial V_{tp1}}\right)^2 \sigma^2_{V_{tp1}} + \left(\frac{\partial Q_{\mathrm{critical}}}{\partial V_{tp2}}\right)^2 \sigma^2_{V_{tp2}} + \left(\frac{\partial Q_{\mathrm{critical}}}{\partial V_{tn1}}\right)^2 \sigma^2_{V_{tn1}} + \left(\frac{\partial Q_{\mathrm{critical}}}{\partial V_{tn2}}\right)^2 \sigma^2_{V_{tn2}} \right\}^{0.5} \tag{21}$$

where $\sigma_{Vtp1}, \sigma_{Vtp2}, \sigma_{Vtn1}$, and σ_{Vtn2} are the standard deviations of the threshold voltages $V_{tp1}, V_{tp2}, V_{tn1}$, and V_{tn2}, respectively.

This model is valid under the following assumptions.

1) The dominant source of variations is the transistor V_t variations. The channel length variations are assumed to affect only V_t through short channel effects. While the variations in the channel length introduce also fluctuations in the input gate capacitance, nevertheless, this contribution is much smaller than the variation in the threshold voltage variations [23], [31].
2) The impact of process variations on the critical charge variations is computed by using a linear approximation. This assumption is accurate, since, WID variations are usually small and can be linearized around the nominal value [31]–[38]. Under this linear approximation, the critical charge mean value is assumed to be equal to its deterministic value, when no variations are introduced. Therefore, process variations affect only the variance of the critical charge (i.e., the critical charge spread around its nominal value).
3) According to [39], the correlation between the different transistors threshold voltages can be neglected for WID variations. This is due to the fact that the RDF is random, and therefore, V_t of the four transistors, in consideration, are identified as four independent and uncorrelated Gaussian random variables [40]. This assumption simplifies the derivation of (21).

IV. Simplified Model for Statistical Design-Oriented Critical Charge Variation

A. Simplified Model Assumptions and Derivations

The model, which is introduced in Section III, for the critical charge variations, is calculated numerically. Therefore, it does not present obvious design insights for WID variations due to its complexity. However, it can be used for the D2D variations by adopting corner-based (or worst-case) analysis methods. In this section, this complex model is simplified for the case of a symmetric 6T SRAM, to account for the critical charge variations from a design perspective. The following assumptions are made to derive this simplified model.

1) The inverters switching voltages are equal to half the supply voltage (i.e., $V_{M1} = V_{M2} = 0.5V_{\mathrm{DD}}$). Thus, the variations in V_{M1} and V_{M2} are ignored.
2) The variation of the factor β, expressed in (9), which is dependent only on V_{tp1} through R_{p1} is calculated to be less

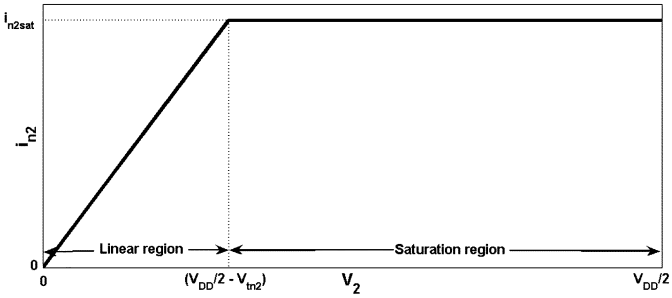


Fig. 2. Transistor M_{n2} current approximation. This current is assumed linear as V_2 changes from 0 to then it saturates at $i_{n2\text{sat}}$ when V_2 changes from $(V_{\text{DD}}/2 - V_{tn2})$ to $V_{\text{DD}}/2$.

than 0.8%, relative to its mean value. As a result, the variations in this factor are ignored, and this factor is assumed constant from the variability perspective.

3) The time delay t_{up} is obtained simply by using a first order approximation of the low to high propagation delay of an inverter, which can be modeled as follows:

$$t_{\text{up}} = \frac{C_2 \Delta V}{i_{\text{average}}} \tag{22}$$

where ΔV is the output voltage swing, that is usually assumed to be $0.5V_{\text{DD}}$, and i_{average} is the average charging current, that is the difference between transistor M_{p2} current and transistor M_{n2} current. Since M_{p2} is in the saturation region during the entire charging process time, hence, its average current is $i_{p2\text{sat}}$. While, transistor M_{n2} current rises from 0A, when the output voltage V_2 (V_{DS} of transistor M_{n2}) equals 0 V, up to $i_{n2\text{sat}}$, when the transistor enters the saturation region. This current is assumed linear with V_2 in the linear region, as depicted in Fig. 2. The average of this current is obtained from Fig. 2 as follows:

$$i_{n2\text{average}} = i_{n2\text{sat}}(0.5 + V_{tn2}/V_{\text{DD}}). \tag{23}$$

The relative variations of this current are given by

$$\frac{\Delta i_{n2\text{average}}}{i_{n2\text{average}}} = \left[\frac{-\alpha}{(V_{\text{DD}}/2) - V_{tn2}} + \frac{1}{V_{\text{DD}}(0.5 + V_{tn2}/V_{\text{DD}})}\right] \Delta V_{tn2}. \tag{24}$$

The variations due to the first term in (24) dominates the second term (as a numeric example, when $V_{\text{DD}} = 1$ V, $\alpha = 1.2$, and $V_{tn2} = 0.342$ V, the first term is 7× higher than the second term). Therefore, in the following derivations, $i_{n2\text{average}}$ is assumed to be equal $i_{n2\text{sat}}(0.5 + V_{tn2}/V_{\text{DD}})$, while the variations of the term $(0.5 + V_{tn2}/V_{\text{DD}})$ are not considered, and this factor is assumed constant, from the variability perspective.

B. Statistical Design-Oriented Critical Charge Variation Model Accounting for WID Variation

By using the simplified model formulas in Section IV-A, and the previous simplified model assumptions, the partial derivatives, defined in (21), are calculated analytically and normalized to the mean value of Q_{critical} as follows:

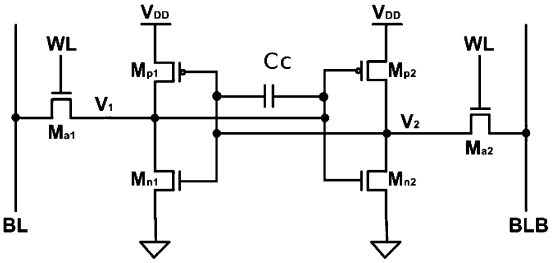


Fig. 3. SRAM cell with the coupling capacitor C_c which increases the critical charge value of its storage nodes (V_1 and V_2).

$$\frac{\frac{\partial Q_{\text{critical}}}{\partial |V_{tp1}|}}{Q_{\text{critical}}} = \frac{-1}{(V_{\text{DD}} - |V_{tp1}|)} \tag{25}$$

$$\frac{\frac{\partial Q_{\text{critical}}}{\partial |V_{tp2}|}}{Q_{\text{critical}}} = \frac{(\alpha/\tau) i_{p2\text{avergae}}}{C_2\left(\frac{V_{\text{DD}}}{2}\right)\left(\frac{V_{\text{DD}}}{2} - |V_{tp2}|\right)} \frac{\beta t_{\text{up}}^2 \exp\left(\frac{-t_{\text{up}}}{\tau}\right)}{1 - \beta \exp\left(\frac{-t_{\text{up}}}{\tau}\right)} \tag{26}$$

$$\frac{\frac{\partial Q_{\text{critical}}}{\partial V_{tn2}}}{Q_{\text{critical}}} = \frac{-(\alpha/\tau) i_{n2\text{avergae}}}{C_2\left(\frac{V_{\text{DD}}}{2}\right)\left(\frac{V_{\text{DD}}}{2} - V_{tn2}\right)} \frac{\beta t_{\text{up}}^2 \exp\left(\frac{-t_{\text{up}}}{\tau}\right)}{1 - \beta \exp\left(\frac{-t_{\text{up}}}{\tau}\right)}. \tag{27}$$

From (25), it is clear that reducing $|V_{tp1}|$ results in reducing the relative variations. Accordingly, it is recommended that transistor M_{p1} is used as a low-V_t device, if the dual-V_t technique is to be used (the same for M_{p2}, when the hit occurs at the other node). Moreover, as the supply voltage V_{DD} is reduced, the variations due to V_{tp1} are increased.

Since increasing the node capacitance is one of the most common techniques to mitigate soft errors in SRAM cells, it is important to see the impact of increasing the node capacitance on the relative critical charge variations. Usually, a coupling capacitor C_c is employed between the storage nodes (V_1 and V_2) as shown in Fig. 3. This coupling capacitor, C_c, increases the nodal capacitances of the SRAM cell storage nodes, and therefore, their critical charge is increased significantly. This C_c is stacked on top of the SRAM cell [metal-insulator-metal (MIM) capacitor] to minimize the required area overhead. However, its value can not be too large, since it depends on the inter-metal dielectric and the cell area. A typical 1 μm^2 C_c has a value of the order of 1 fF [22]. The model capacitances C_1 and C_2, have to be modified to account for C_c, by applying the Miller effect as follows [22]:

$$C_1' = C_1 + 2C_c \qquad C_2' = C_2 + 2C_c. \tag{28}$$

From (26) and (27), and by using t_{up} and β formulas derived in Section IV-A, the relative critical charge variations $(((\partial Q_{\text{critical}})/(\partial |V_{tp2}|))/(Q_{\text{critical}})$ and $((\partial Q_{\text{critical}})/(\partial V_{tn2}))/(Q_{\text{critical}}))$ have the same dependence on the node capacitance, C' (assuming $C_1' = C_2' = C'$ for a symmetric SRAM). This dependence is in the form $(\beta C' \exp(-\gamma C'))/((1 - \beta \exp(-\gamma C')))$, where $\gamma = (((V_{\text{DD}})/(2)))/(\tau(i_{p2\text{avergae}} - i_{n2\text{average}}))$. Therefore, it is possible to obtain the value of the node capacitance, C',

that maximizes these relative variations, by differentiating with respect to C', and equating the result to zero. After some simplifications, the condition on C' for the maximum possible relative variations is given by

$$1-\beta\exp(-\gamma C')=C'(\gamma-\theta)$$

where

$$\gamma=\frac{\left(\frac{V_{\rm DD}}{2}\right)}{\tau\left(i_{p2\rm avergae}-i_{n2\rm average}\right)}$$

$$\theta=\left(\frac{\frac{\partial\beta}{\partial C'}}{\beta}\right)=\frac{\frac{R_{p1}}{\tau}}{1-\frac{R_{p1}C'}{\tau}}\left(1+\frac{\ln\left(\frac{R_{p1}C'}{\tau}\right)}{1-\frac{R_{p1}C'}{\tau}}\right). \quad (29)$$

From (29), the value of C' that maximizes the relative variations, denoted by C'_m, is obtained for a given value of $V_{\rm DD},\tau$, and average currents ($i_{p2\rm averge}$ and $i_{n2\rm average}$). These average currents are dependent on transistors M_{p2} and M_{n2} parameters (W/L and V_t). Since C'_m results in the maximum relative variations, it is essential at the design level to avoid the satisfaction of this condition reported in (29). Otherwise, the SRAM cell will exhibit the maximum possible relative critical charge variations. These maximum variations are calculated by substituting the condition in (29) in (26) and (27) and are given by

$$\left(\frac{\frac{\partial Q_{\rm critical}}{\partial|V_{tp2}|}}{Q_{\rm critical}}\right)|_{\rm max}=\frac{\alpha\beta\tau\left(\frac{\gamma^2}{(\gamma-\theta)}\right)i_{p2\rm average}}{\left(\frac{V_{\rm DD}}{2}\right)\left(\frac{V_{\rm DD}}{2}-|V_{tp2}|\right)}\exp(-\gamma C'_m) \quad (30)$$

$$\left(\frac{\frac{\partial Q_{\rm critical}}{\partial V_{tp2}}}{Q_{\rm critical}}\right)|_{max}=\frac{\alpha\beta\tau\left(\frac{\gamma^2}{(\gamma-\theta)}\right)i_{n2\rm average}}{\left(\frac{V_{\rm DD}}{2}\right)\left(\frac{V_{\rm DD}}{2}-V_{tn2}\right)}\exp(-\gamma C'_m). \quad (31)$$

By using (25) with (30) and (31), the maximum possible relative critical charge variations, for a give SRAM cell design with respect to C', are estimated.

It should be mentioned that for a given C',τ, and $V_{\rm DD}$, the condition on the saturation currents to achieve the maximum possible relative critical charge variations is known. Therefore, this condition should be avoided by designing the transistors currents to be far from this maximum relative variations condition. In addition, (26) and (27) indicate that the relative variations, due to V_{tn2} and V_{tp2}, are decaying exponentially with $(t_{\rm up}/\tau)$. From (22), $t_{\rm up}$ is dependent on C', therefore, there exists a certain value of C' for a given τ that makes the relative variations contributions of V_{tn2} and V_{tp2} smaller than that of V_{tp1}. In this situation, the variations of V_{tp1} dominate, and further increasing C' does not reduce the overall relative variations which are at a minimum value. The knowledge of C', which results in maximum and minimum relative variations, provides a vital design insight for circuit designers, who target at mitigating the soft errors, while keeping the variability at a certain level.

Finally, the proposed models can be used for future CMOS technology nodes (i.e., 45, 32, and 22 nm), since, the transistor model parameters such as the technology parameters and the threshold voltage standard deviation σ_{V_t} can be easily obtained. Therefore, the proposed models are scalable in terms of technology scaling and can be used to predict the critical charge variability for future technology nodes as long as the models assumptions are satisfied.

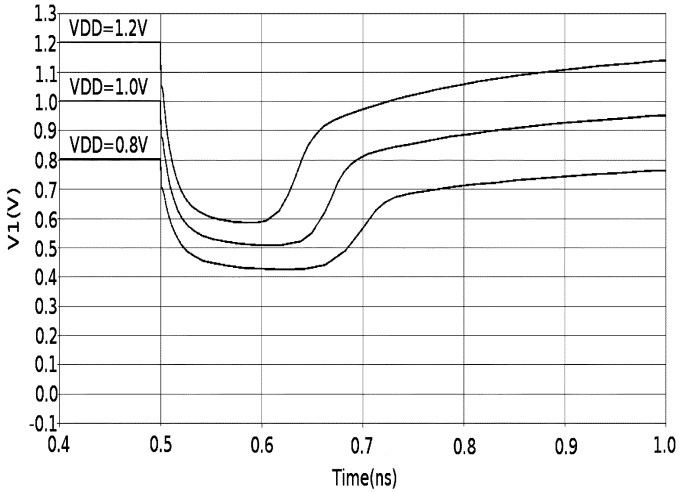


Fig. 4. Nonflipping case when the SRAM cell recovers for different values of $V_{\rm DD}$. Node V_1 voltage falls down till it hits $V_{\rm min}$ then it recovers back to $V_{\rm DD}$. This $V_{\rm min}$ is close to $V_{\rm DD}/2$ which validates the assumptions used in the simplified model.

V. Results and Discussion

In all the following simulations, an industrial 65-nm technology, with technological parameters shown in Table I, is employed. The SRAM cell is sized such that its stability is maintained, as reported in [36].

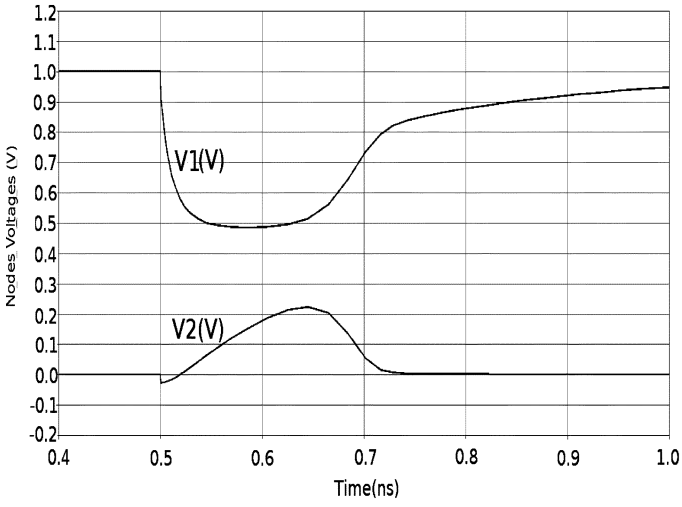
A. Verification of the Models Assumptions

First, the assumptions, used in deriving (7) and (8), are verified. Fig. 4 illustrates the nonflipping case, where the SRAM cell recovers for different values of $V_{\rm DD}$. Node V_1 voltage falls down till it hits a minimum voltage [which is called $V_{\rm min}$, and given in (8)] then recovers back to $V_{\rm DD}$. From Fig. 4, this minimum voltage $V_{\rm min}$ is close to $V_{\rm DD}/2$ justifying the assumptions used in the simplified model. Fig. 5(a) shows the two nodes V_1 and V_2 voltages in the nonflipping case. It is clear that, since V_2 voltage can not hit V_{M1}, the SRAM cell is recovered. However, in Fig. 5(b), the V_2 node voltage hits V_{M1}, and hence, the SRAM cell exhibits a soft error. Moreover, Fig. 5(b) shows that node V_2 voltage is around 0 V as long as node V_1 voltage is falling. Once node V_1 voltage hits $V_{\rm min}$, V_1 stays constant at $V_{\rm min}$, whereas, node V2 voltage rises to V_{M1}. It should be mentioned that the minimum voltage, $V_{\rm min}$, shown in Fig. 5(b), at which V_1 stays constant before flipping to 0 V, is slightly less than $V_{\rm min}$ shown in Fig. 5(a) for the nonflipping case. The difference between these two minima is approximately 10–20 mV, which demonstrates that the flipping occurs, when $V_{\rm min}$ is less than V_{M2}.

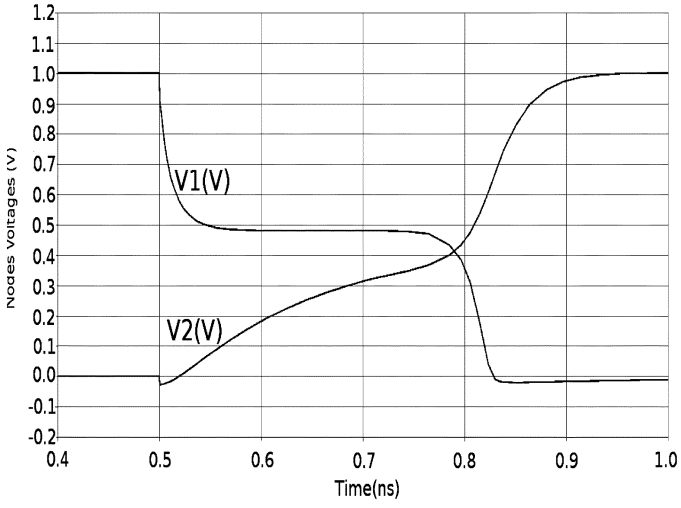
B. Verification of the Models Estimated Critical Charge

To verify the critical charge nominal value, and the critical charge variations models, the analytical models are compared to the simulation results using SPICE transient and Monte Carlo simulations. These simulations are performed to validate the nominal critical charge, and the critical charge variability models, respectively, for both the exact and the simplified models.

In the following, the validation results for these models are presented. A large number of Monte Carlo runs (4000 runs) are used to provide a good accuracy in determining the critical charge mean and standard deviation. For each Monte Carlo run, the value of the current pulse charge Q that causes the cell to



(a)



(b)

Fig. 5. Two nodes V_1 and V_2 voltages in: (a) the nonflipping case when V_2 voltage can not hit V_{M1}, and hence, the SRAM cell is recovered; (b) the flipping case when V_2 voltage hits V_{M1}, and hence, the SRAM cell exhibits a soft error. The minimum voltage $V_{\min}$ shown in the flipping case (b) at which V_1 stays constant before flipping to 0 V is slightly less than $V_{\min}$ shown in the nonflipping case (a).

TABLE I
65-nm TECHNOLOGY INFORMATION AND SRAM SIZING [29]

	NMOS	PMOS
Nominal V_{DD}	1.0-1.2 V	
W/L ($\mu m/\mu m$)	0.195/0.065	0.11/0.065
V_{to} (mV)	342	-204
$\sigma_{V_{to}}$ (mV)	25.8	34.3

flip is determined. Then, the simulations are repeated for different $V_{\rm DD}$ (from 0.7 to 1.2 V), to find the effect of reducing $V_{\rm DD}$ on the critical charge mean and variations. The SRAM sizing, shown in Table I, is used in the simulation setups. Hardware-calibrated statistical models are used to account for V_t variations.

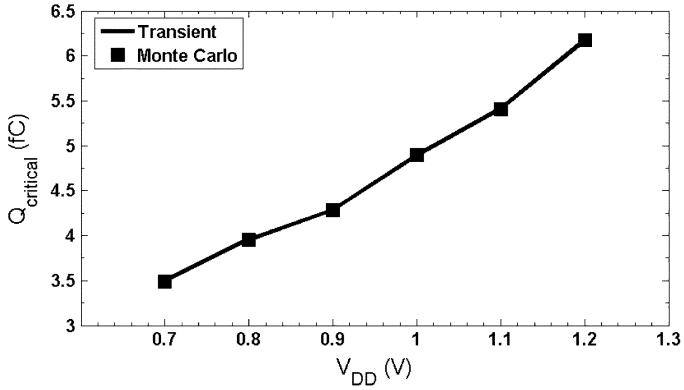


Fig. 6. $Q_{\rm critical}$ versus $V_{\rm DD}$ for $\tau = 250$ ps from the transient simulations (when no variations are introduced) and from Monte Carlo simulations. Clear agreement between $Q_{\rm critical}$ (obtained from transient simulations) and $\mu_{Q_{\rm critical}}$ (obtained from Monte Carlo simulations) justifies the linearity approximation assumption which states that process variations affect only on the critical charge variance (spread) and have no effect on its mean.

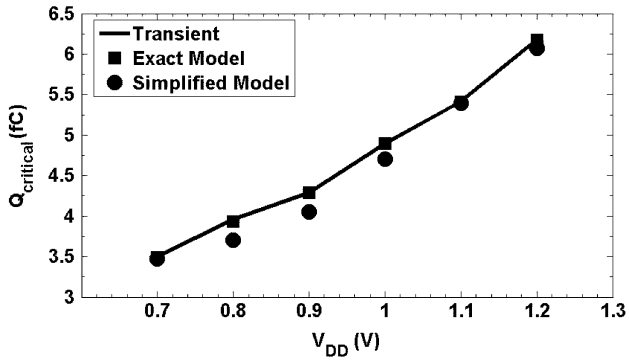


Fig. 7. $Q_{\rm critical}$ versus $V_{\rm DD}$ for $\tau = 250$ ps from Monte Carlo simulations. Also shown the results from the proposed exact and simplified models.

Typically, random variations are inversely proportional to the square root of the gate area (WL), as explained in Section III-B, [29], [30]. Therefore, the pMOS transistors have higher V_t variations than the nMOS transistors, since the pMOS transistors exhibit lower driving strength (weaker) than the nMOS transistors in the SRAM cell.

1) Nominal Critical Charge:

Fig. 6 displays the nominal critical charge, which is obtained by using the transient simulations $(Q_{\rm critical})$ and Monte Carlo simulations $(\mu_{Q_{\rm critical}})$. Clear agreement between $Q_{\rm critical}$ and $\mu_{Q_{\rm critical}}$ justifies the linearity approximation assumption used in Section III-B, down to $V_{\rm DD} = 0.7$ V (i.e., process variations affect only on the critical charge variance (spread) and have no effect on its mean).

Fig. 7 shows the nominal critical charge value calculated from the proposed exact and simplified model versions, and compared to the transient simulations results for different supply voltage values. It should be highlighted that the simplified model is proposed only for the WID variations estimation, although it still shows an acceptable match for the nominal critical charge value. These results are obtained by using $\tau = 250$ ps to ensure that the primary assumption used in (4) is satisfied ($\tau_r = 1$ ps and $\tau_f = 250$ ps). It is obvious from Figs. 6 and 7 that reducing the supply voltage decreases the critical charge, which is expected.

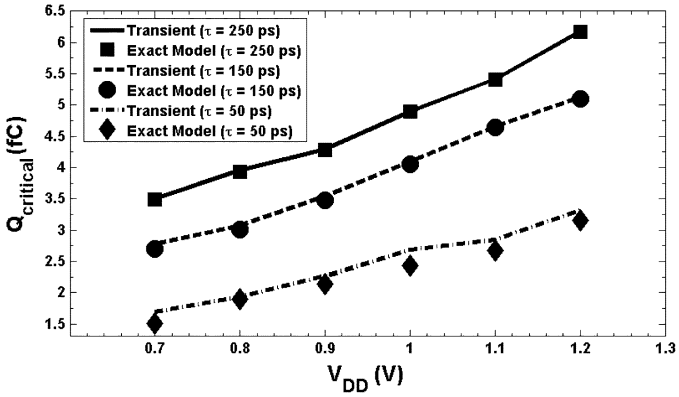


Fig. 8. Q_{critical} versus V_{DD} for different values of τ (50 to 250 ps) from the transient simulations and from the proposed exact model.

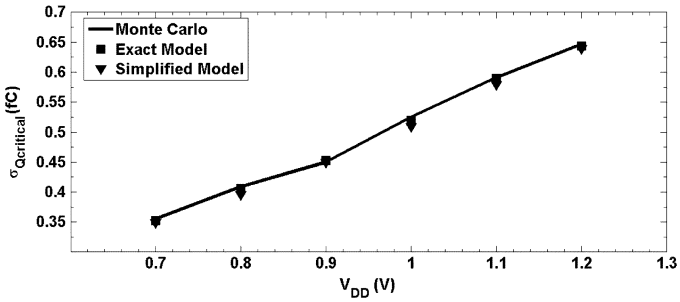


Fig. 9. Critical charge variations $\sigma_{Q_{\mathrm{critical}}}$ versus V_{DD} for $\tau = 250$ ps from Monte Carlo simulation and from the proposed exact and simplified model.

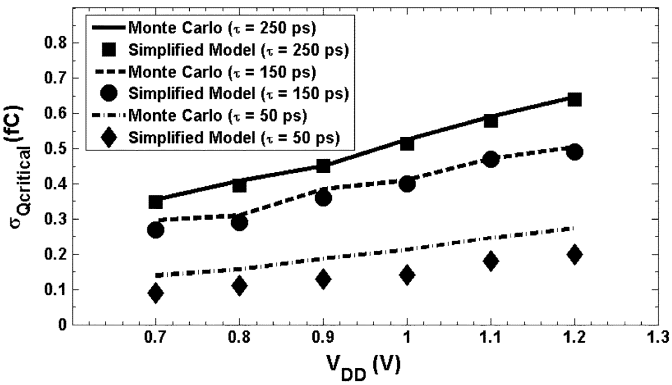


Fig. 10. Critical charge variations $\sigma_{Q_{\mathrm{critical}}}$ versus V_{DD} for different values of τ (50–250 ps) from Monte Carlo simulation and from the proposed simplified model.

According to [2] and [35], the current pulse, used in circuit level modeling of soft errors, might have a varying width from a few picoseconds to hundreds of picoseconds. The narrow current pulse represents the worst-case situation, because the critical charge, Q_{critical}, is minimal. This narrow current pulse corresponds to an event, in which the track of an ionized particle intersects the drain of the nMOS transistor in the OFF-state (like M_{n1} in the analyzed case). This means that the charge collection mechanism is dominated by the drift current (due to local electric fields) in a very short time. On the other hand, the charge collection mechanism is dominated by diffusion current in the events in which the ion track does not intersect the drain [2]. Theoretical studies showed that, typically, 80%–90% of the neutron induced SER is represented by the latter events in which the current pulse is relatively wide [41], [42]. Such a discussion demonstrates that both narrow and wide current pulses must be considered in Q_{critical} calculations.

Therefore, the values of Q_{critical}, calculated from the proposed exact model and from SPICE transient simulations for different current pulse widths (by varying τ from 50 to 250 ps), are shown in Fig. 8. The simplified model results are not shown in this figure as the simplified model is mainly introduced for WID variations estimation. In Fig. 8, it is shown that as the current pulse width increases (i.e., diffusion current dominates), the critical charge increases. In addition, the proposed model accuracy degrades as τ decreased because this contradicts the primary assumption used in deriving (4).

2) Critical Charge Variations: In Sections III-B and IV-B, the derivation of the critical charge standard deviation using the exact model and the simplified model is described. Fig. 9 shows the simulation result for $\sigma_{Q_{\mathrm{critical}}}$ for different V_{DD} values. Note that each data point represents $\sigma_{Q_{\mathrm{critical}}}$ calculated from 4000 Monte Carlo runs. Also, Fig. 9 shows the results from the proposed models. Both models results exhibit a good match with the simulation results. Fig. 10 shows $\sigma_{Q_{\mathrm{critical}}}$ obtained from Monte Carlo simulations and from the simplified model for different values of τ which demonstrates that, as τ is reduced, the critical charge variations are reduced as well. The simplified model accuracy is reduced, as τ is decreased. Therefore, it is recommended to use the proposed models with caution for small values of τ (Actually, the proposed models can account for small values of τ, by deriving the models again by taking τ_r and τ_f into account, however, this complicates the models, and hides some design insights).

For all the succeeding discussions, only the simplified model is used and, therefore, larger values of τ is used ($\tau \geq 50$ ps). It is important to show that as V_{DD} is reduced for low power applications, $\sigma_{Q_{\mathrm{critical}}}$ is decreased, which is a promising result for low power SRAM cells.

C. Effect of the Coupling Capacitor on the Critical Charge Relative Variability

In Section IV-B, it has been shown that the capacitance C'_m, which results in the maximum relative variations, can be obtained from the condition given in (29). For $V_{\mathrm{DD}} = 1$ V, $\alpha = 1.2$ (extracted from fitting Log i_D-Log V_{GS} characteristics to the alpha-power model), $i_{p2\mathrm{sat}} = 12.9$ μA, $i_{2n\mathrm{sat}} = 11.2$ μA, and $\tau = 250$ ps. By using (29), $\gamma = 0.6 * 10^{15}\mathrm{F}^{-1}$, and solving this equation yields that $C'_m = 0.143$ fF. The node capacitance C equals 0.93 fF (extracted from simulations), therefore, the condition for the maximum relative variations is not met in this case, since C is already larger than C'_m. Fig. 11 shows how the relative variations in (26) and (27) vary with the capacitance C'. For a given relative variations specifications, the value of the capacitance C', that results in these relative variations, can be obtained from this figure. For example, the value of C', that results in 50% of the maximum relative variations value, equals 1.9 fF. Consequently, the coupling capacitor, that results in half maximum relative variations, is $C_c = (1.9 - 0.93)/2 = 0.485$ fF.

One important design insight from this discussion is that the proposed model can aid circuit designers to choose the value of C_c that enhances the critical charge nominal value, while

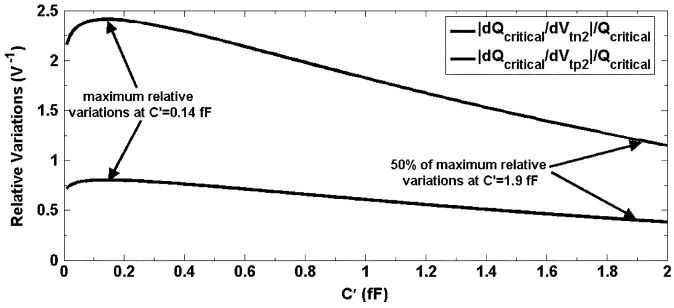


Fig. 11. Relative variations $(\partial Q_{\rm critical}/\partial|V_{tp2}|)/Q_{\rm critical}$ and $(\partial Q_{\rm critical}/\partial V_{tn2})/Q_{\rm critical}$ versus C' showing that the maximum relative variations occurs at $C' = 0.14$ fF and 50% of the maximum variations occurs at $C' = 1.9$ fF.

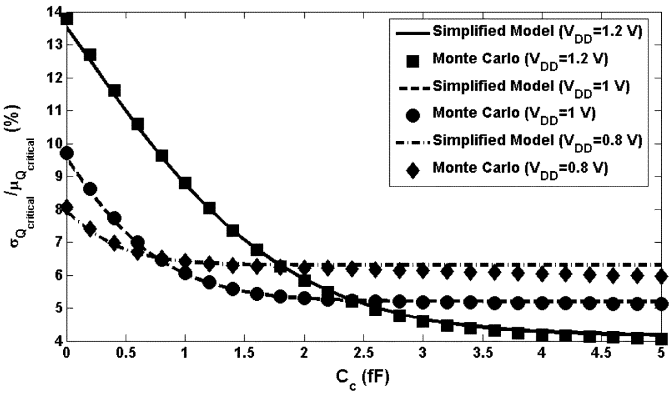


Fig. 12. Overall relative variations $(\sigma_{Q_{\rm critical}}/\mu_{Q_{\rm critical}})$ versus C_c obtained from Monte Carlo simulations and from the proposed simplified model for different values of $V_{\rm DD}$ when $\tau = 250$ ps (which represents the drain nonintersecting particle strike event).

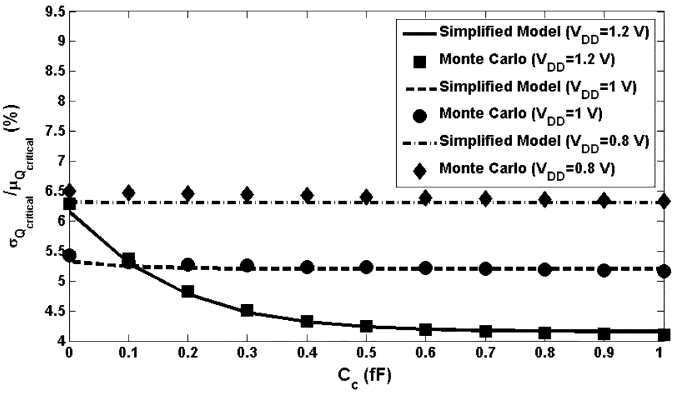


Fig. 13. Overall relative variations $(\sigma_{Q_{\rm critical}}/\mu_{Q_{\rm critical}})$ versus C_c obtained from Monte Carlo simulations and from the proposed simplified model for different values of $V_{\rm DD}$ when $\tau = 50$ ps (which represents the drain intersecting particle strike event).

keeping the relative variations at the required level (50% of the maximum relative variations is just an example).

Figs. 12 and 13 portray the overall relative variations $(\sigma_{Q_{\rm critical}}/\mu_{Q_{\rm critical}})$ versus C_c obtained from Monte Carlo simulations, and from the proposed model, for different values of $V_{\rm DD}$, when $\tau = 250$ ps (which represents the drain non-intersecting event), and $\tau = 50$ ps (which represents the drain intersecting event). The proposed model is in good agreement with the simulation results.

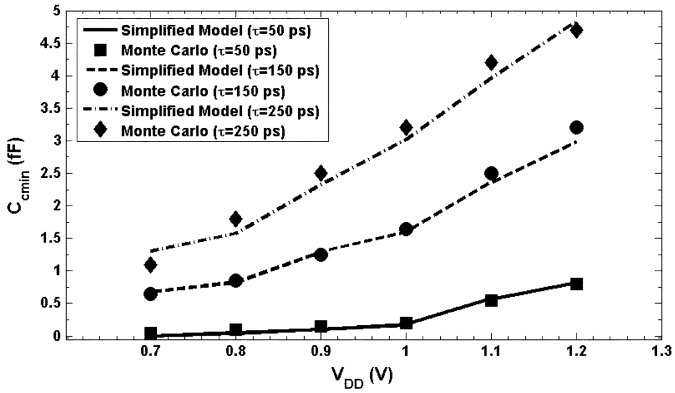


Fig. 14. Coupling capacitor that results in minimum relative critical charge variations $C_{c\,\rm min}$ versus $V_{\rm DD}$ for different values of τ which shows that when $V_{\rm DD}$ is reduced, the value of $C_{c\,\rm min}$ that results in minimum relative variations is decreased. These results are obtained from the proposed simplified model and from Monte Carlo simulations.

It is obvious from Figs. 12 and 13 that, as C_c increased, $\sigma_{Q_{\rm critical}}/\mu_{Q_{\rm critical}}$ decreases, till reaching a minimum value at which increasing C_c has no effect on $\sigma_{Q_{\rm critical}}/\mu_{Q_{\rm critical}}$. The reason for this is readily explained by recalling (26) and (27), which show that for large values of C_c, the variations from V_{tn2} and V_{tp2} are vanished (since increasing C_c increases $t_{\rm up}$) and, hence, the variations from V_{tp1} dominate the overall variations. Therefore, $\sigma_{Q_{\rm critical}}/\mu_{Q_{\rm critical}}$ is proportional to $(1/(V_{\rm DD} - |V_{tp1}|)$. This latter observation explains why $\sigma_{Q_{\rm critical}}/\mu_{Q_{\rm critical}}$ saturates at the highest value for the case $(V_{\rm DD} = 0.8$ V). Figs. 12 and 13 show also that $\sigma_{Q_{\rm critical}}/\mu_{Q_{\rm critical}}$ decreases, as $V_{\rm DD}$ is reduced, before reaching its minimum level. However, $\sigma_{Q_{\rm critical}}/\mu_{Q_{\rm critical}}$ decreases, as $V_{\rm DD}$ increases, when V_{tp1} variations dominate (at large values of C_c).

Finally, as shown in these two figures, $\sigma_{Q_{\rm critical}}/\mu_{Q_{\rm critical}}$ reaches a minimum value, at smaller values of C_c, for smaller τ values. Hence, for the drain intersecting event case (small τ values), C_c, that results in the minimum $\sigma_{Q_{\rm critical}}/\mu_{Q_{\rm critical}}$, is smaller than that for the drain non-intersecting case. The value of C_c, that causes $\sigma_{Q_{\rm critical}}/\mu_{Q_{\rm critical}}$ to reach its minimum value, is denoted by $C_{c\,\rm min}$, and is obtained from $\sigma_{Q_{\rm critical}}/\mu_{Q_{\rm critical}}$ plots. It might be beneficial for designers to know, in advance, the value of $C_{c\,\rm min}$, and the impact of $V_{\rm DD}$ and τ on it. Fig. 14 shows how $V_{\rm DD}$ and τ affect on $C_{c\,\rm min}$, as obtained from the proposed simplified model and from Monte Carlo simulations. According to Fig. 14, it is clear that $C_{c\,\rm min}$ increases when $V_{\rm DD}$ increases, and also when τ increases. This result is promising for low power SRAM cells, since a smaller coupling capacitor is required to have the minimum relative critical charge variations.

Now, the values of C_c, that result in maximum and minimum $\sigma_{Q_{\rm critical}}/\mu_{Q_{\rm critical}}$, are calculated. Thus, a good design insight is to use a coupling capacitor between these two extremes, to enhance the critical charge mean, and minimize the relative critical charge variations, under certain power and performance constraints.

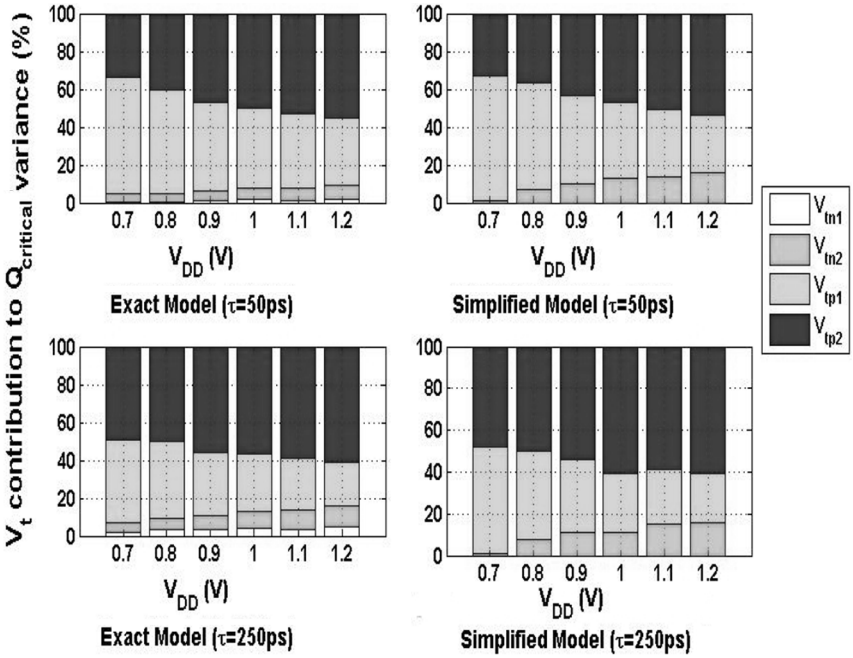


Fig. 15. Percentage contribution of each transistor threshold voltage variations for different values of $V_{\rm DD}$ when $\tau = 50$ and 250 ps obtained from the two proposed models. The contribution of V_{tp1} increases as the supply voltage is reduced which is well explained from (25) (inversely proportional to $(V_{\rm DD} - |V_{tp1}|)$).

D. SRAM Cell Transistors Contribution to the Overall Critical Charge Variability

The overall critical charge standard deviation $(\sigma_{Q_{\rm critical}})$ has contributions from different transistors threshold voltages variations (i.e., $V_{tn1}, V_{tn2}, V_{tp1}$, and V_{tp2}). Fig. 15 shows the percentage contribution of each transistor threshold voltage variations for different values of $V_{\rm DD}$, when $\tau = 50$ and 250 ps obtained from the two proposed models. It is evident that the contribution of V_{tn1} in the exact model is very small (less than 6%). This justifies the assumptions used in deriving the simplified model, which ignores its variations contribution (when we assume that $V_{M1} = V_{\rm DD}/2$). According to Fig. 15, the contribution of V_{tp1} increases, as the supply voltage is reduced which is well explained by (25) (inversely proportional to $(V_{\rm DD} - |V_{tp1}|)$). At $V_{\rm DD} = 0.7$ V, the transistor M_{p1} dominates the variations (62%) for the case $\tau = 50$ ps.

Moreover, when τ increases, V_{tn2} and V_{tp2} contributions to the critical charge variance are increased, and V_{tp1} contribution is decreased. These results agree with (26) and (27). In addition, Fig. 15 shows that the contributions of the pMOS transistors, M_{p1} and M_{p2}, dominate the variations, because their percentage contributions is larger than 84% in all cases. This fact can be justified by noting that the gate area of the pMOS transistors is smaller than that of the nMOS transistors (as reported in Table I). Since the threshold voltage variations are inversely proportional to the square root of the gate area (WL), the pMOS transistors dominate the variations.

E. Accuracy of the Proposed Models

In Fig. 16, $Q_{\rm critical}$ from the proposed exact model is plotted versus the transient simulations results for different values of $\tau, V_{\rm DD}$, and C_c. The maximum error is 6.2%, and the average error is 1.8%. Fig. 17 shows $\sigma_{\rm Qcritical}$ from the simplified model plotted versus Monte Carlo simulation results for different values of $\tau, V_{\rm DD}$, and C_c. The maximum error is 9.2%, and the average error is 4%. Good agreement between the proposed models and the simulation results justifies all

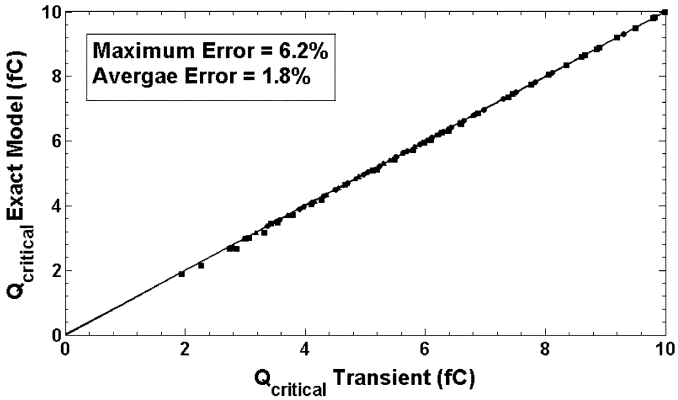


Fig. 16. $Q_{\rm critical}$ from the proposed exact model is plotted versus the transient simulations results for different values of $\tau, V_{\rm DD}$ and C_c.

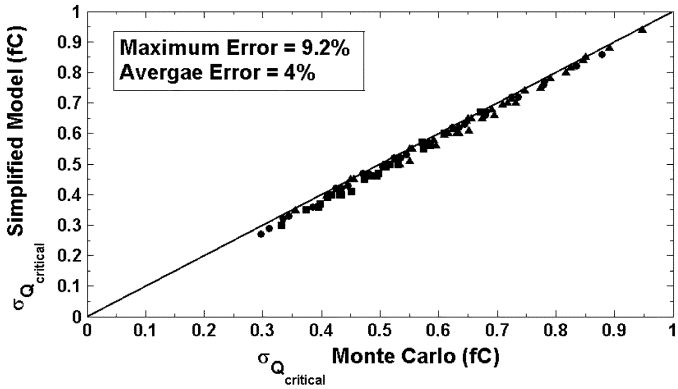


Fig. 17. $\sigma_{Q\rm critical}$ from our simplified model plotted versus Monte Carlo simulation results for the same ranges of $\tau, V_{\rm DD}$ and C_c.

the assumptions used to derive the models, as explained in Sections III and IV.

As shown in the previous discussions, the proposed models are based on easily measurable parameters, which can be directly extracted from the measurements or technology information (i.e., C, σ_{V_t}, $\mathrm{V_{to}}$, and α). In addition, the proposed models are very efficient when compared to the computationally expensive, and time consuming Monte Carlo simulations. The models can be used to explore design tradeoffs to increase the critical charge or control its variability. The proposed model shows how the coupling capacitor, one of the most common soft error mitigation techniques in SRAM cells, affect on the critical charge relative variability. Moreover, the proposed model provides a certain range for this coupling capacitor C_c to keep the variability within an acceptable limit.

VI. Design Insights

In this section, some design insights, extracted from the proposed models in this paper, are reported. The proposed models provide the following design insights.

1) Increasing the supply voltage V_{DD} results in increasing both Q_{critical} and $\sigma_{Q_{\mathrm{critical}}}$. Therefore, the choice of V_{DD}, that yields acceptable values of $\sigma_{Q_{\mathrm{critical}}}/\mu_{Q_{\mathrm{critical}}}$, is essential as explained in the proposed models.
2) From the formulas derived in Section III-A, the critical charge nominal value for the SRAM cell is estimated accurately without time consuming transient simulations. For a target SER, the critical charge value can be calculated by the following empirical equation:

$$\mathrm{SER}\alpha N_{\mathrm{flux}} \times \mathrm{CS} \times \exp(-Q_{\mathrm{critical}}/Q_s) \tag{32}$$

where N_{flux} refers to the intensity of the neutron flux, CS is the cross section area of the struck node, and Q_s is the charge collection efficiency. These parameters depend mainly on the SRAM cell technology and layout. Once the required critical charge is known, the circuit parameters are designed to achieve it without doing any SPICE simulations.
3) The coupling capacitor, C_c, can result in a maximum $\sigma_{Q_{\mathrm{critical}}}/\mu_{Q_{\mathrm{critical}}}$, as depicted in (29). Although, this occurs in the designed SRAM, proposed in this work, only when τ exceeds 1000 ps, it can occur at lower τ values for a different SRAM design, when the condition, in (29), is satisfied. Therefore, the circuit designer must be aware, at the design level, of this condition and avoid it.
4) For $C_c = C_{c\,\mathrm{min}}$, V_{tp1} variations dominate the overall critical charge variations. Thus, $\sigma_{Q_{\mathrm{critical}}}/\mu_{Q_{\mathrm{critical}}}$ is at its minimum value and inversely proportional to $(V_{\mathrm{DD}}\text{-}|V_{tp1}|)$. Therefore, a further increase in C_c results in increasing Q_{critical}, while keeping $\sigma_{Q_{\mathrm{critical}}}/\mu_{Q_{\mathrm{critical}}}$ constant. If it is required to further reduce $\sigma_{Q_{\mathrm{critical}}}/\mu_{Q_{\mathrm{critical}}}$, either V_{DD} can be increased or low-V_t pMOS transistors can be used.
5) For $C_c < C_{c\,\mathrm{min}}$, the variations of both V_{tn2} and V_{tp2} dominate $\sigma_{Q_{\mathrm{critical}}}/\mu_{Q_{\mathrm{critical}}}$. These variations decay exponentially with (t_{up}/τ). Therefore, to reduce $\sigma_{Q_{\mathrm{critical}}}/\mu_{Q_{\mathrm{critical}}}$ in this case, either increasing t_{up} (by increasing C_c or V_{DD}), or reducing the average charging current, or reduce τ. Since a small τ represents only 10%–20% of the neutron induced SER events, the latter condition is out of control.
6) Since the two extremes of C_c, that result in maximum and minimum $\sigma_{Q_{\mathrm{critical}}}/\mu_{Q_{\mathrm{critical}}}$, can be obtained from the proposed models, the circuit designer can determine C_c that results in a certain Q_{critical} and $\sigma_{Q_{\mathrm{critical}}}/\mu_{Q_{\mathrm{critical}}}$, while satisfying the power and performance constraints at the design level.

Although this paper has focused on the critical charge and its variability modeling for the SRAM cell, it can be extended to model them in flip-flops circuits. This is possible, because all the flip-flops topologies consist of an embedded cross-coupled inverters as those in the SRAM cell. However, these inverters are not symmetric like those in the SRAM cell. The proposed models can be extended to account for asymmetrical inverters by simply assuming that $V_{M1} \neq V_{M2}$.

VII. Conclusion

In this paper, analytical models accounting for both D2D and WID variations, are proposed. The proposed models deal with the D2D variations, by using corner-based methods. Moreover, they deal with the WID variations, by using statistical techniques. The accuracy of the proposed models is validated by transient and Monte Carlo SPICE simulation results, for an industrial 65-nm technology, over a wide range of supply voltages, particle strike induced current pulse widths, and coupling capacitors. The proposed models show that, the use of the coupling capacitor in the SRAM cell, as a soft error mitigation technique, is limited by the relative variations. The proposed models provide an analytical equation, to calculate the value of the coupling capacitor, that results in minimum relative variations. Finally, the proposed models show that, the pMOS transistors in the SRAM cell, are dominating the variations, and hence, the pMOS transistors must be designed, while taking the critical charge variations into account.

The derived statistical models are scalable, bias dependent, and require only the knowledge of easily measurable parameters. Moreover, the models are very efficient, compared to Monte Carlo simulations. This makes them very useful in early design cycles, SRAM design optimization, and technology prediction. Finally, the proposed models can be extended for the flip-flops critical charge variability as well.

References

[1] Q. Ding, R. Luo, and Y. Xie, "Impact of process variation on soft error vulnerability for nanometer VLSI circuits," in *Proc. ASICON*, 2005, pp. 1023–1026.
[2] T. Heijmen, D. Giot, and P. Roche, "Factors that impact the critical charge of memory elements," in *Proc. 12th IEEE Int. On-Line Test. Symp. (IOLTS)*, 2006.
[3] T. P. Ma and P. V. Dressendorfer, *Inonizing Radiation Effects in MOS Devices and Circuits*. New York: Wiley, 1989.
[4] T. Nakamura, M. Baba, E. Ibe, Y. Yahag, and H. Kameyama, *Terrestrial Neutron-Induced Soft Errors in Advanced Memory Devices*. Singapore: World Scientific, 2008.
[5] R. Ramanarayanan, V. Degalahal, N. Vijaykrishnan, M. J. Irwin, and D. Duarte, "Analysis of soft error rate in flip-flops and scannable latches," in *Proc. ASIC*, 2003, pp. 231–234.
[6] K. Ramakrishnan, R. Rajaraman, S. Suresh, N. Vijaykrishnan, Y. Xie, and M. J. Irwin, "Variation impact on ser of combinational circuits," in *Proc. Int. Symp. Quality Electron. Des. (ISQED)*, 2007, pp. 911–916.
[7] T. Heijmen, "Soft error vulnerability of sub-100-nm flip-flops," in *Proc. 14th IEEE Int. On-Line Test. Symp. (IOLTS)*, 2008, pp. 247–252.
[8] P. Hazucha and C. Svensson, "Impact of CMOS technology scaling on the atmospheric neutron soft error rate," *IEEE Trans. Nucl. Sci.*, vol. 47, no. 6, pp. 2586–2594, 2000.

[9] P. Shivakumar, S. W. Keckler, D. Burger, M. Kistler, and L. Alvisi, "Modeling the effect of technology trends on the soft error rate of combinational logic," in *Proc. Int. Conf. Dependable Syst. Netw.*, 2002, pp. 389–398.
[10] R. C. Baumanm, "Soft errors in advanced semi-conductor devices-Part I: The three radiation sources," *IEEE Trans. Device Mater. Reliab.*, vol. 1, no. 1, pp. 17–22, 2001.
[11] S. Borkar, T. Karnik, S. Narendra, J. Tschanz, A. Keshavarzi, and V. De, "Parameter variations and impact on circuits and microarchitecture," in *Proc. 40th Conf. Des. Autom. (DAC)*, 2003, pp. 338–342.
[12] K. Bowman, S. Duvall, and J. Meindl, "Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution for gigascale integration," *IEEE J. Solid-State Circuits*, vol. 37, no. 2, pp. 183–190, 2002.
[13] H. Masuda, S. Ohkawa, A. Kurokawa, and M. Aoki, "Challenge: Variability characterization and modeling for 65-nm to 90-nm processes," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, 2005, pp. 593–599.
[14] J. Tschanz, K. Bowman, and V. De, "Variation tolerant circuits: Circuit solutions and techniques," in *Proc. Des. Autom. Conf. (DAC)*, 2005, pp. 762–763.
[15] ITRS Web-Site, "The international technology roadmap for semiconductors," [Online]. Available: http://public.itrs.net
[16] Q. Ding, R. Luo, H. Wang, H. Yang, and Y. Xie, "Modeling the impact of process variation on critical charge distribution," in *Proc. Syst. Chip Conf. (SOC)*, 2006, pp. 243–246.
[17] E. H. Cannon, A. J. KleinOsowski, R. Kanj, D. D. Reinhardt, and R. V. Joshi, "The impact of aging effects and manufacturing variation on SRAM soft-error rate," *IEEE Trans. Device Mater. Reliab.*, vol. 8, no. 1, pp. 145–152, 2008.
[18] T. Heijmen and B. Kruseman, "Alpha-particle-induced SER of embedded SRAMs affected by variations in process parameters and by the use of process options," *Solid-State Electron.*, vol. 49, pp. 1783–1790, 2005.
[19] J. M. Palau, G. Hubert, K. Coulie, B. Sagnes, M.-C. Calvet, and S. Fourtine, "Device simulation study of the SEU sensitivity of SRAMs to internal ion tracks generated by nuclear reactions," *IEEE Trans. Nucl. Sci.*, vol. 48, no. 2, pp. 225–231, 2001.
[20] Y. Z. Xu, H. Puchner, A. Chatila, O. Pohland, B. Bruggeman, B. Jin, D. Radaelli, and S. Daniel, "Process impact on SRAM alpha-particle SEU performance," in *Proc. IEEE Int. Reliab. Phys. Symp.*, Phoenix, AZ, 2004, pp. 294–299.
[21] B. Zhang, A. Arapostathis, S. Nassif, and M. Orshansky, "Analytical modeling of SRAM dynamic stability," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Des.*, San Jose, CA, 2006, pp. 315–322.
[22] S. M. Jahinuzzaman, M. Sharifkhani, and M. Sachdev, "Investigation of process impact on soft error susceptibility of nanometric SRAMs using a compact critical charge model," in *Proc. Int. Symp. Quality Electron. Des. (ISQED)*, 2008, pp. 207–212.
[23] M. H. Abu-Rahma and M. Anis, "A statistical design-oriented delay variation model accounting for within-die variations," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 27, no. 11, pp. 1983–1995, 2008.
[24] V. Degalahal, N. Vijaykrishnan, and M. Irwin, "Analyzing soft errors in leakage optimized SRAM design," in *Proc. IEEE Int. Conf. VSLI Des.*, 2003, pp. 227–233.
[25] G. R. Srinivasan, P. C. Murley, and H. K. Tang, "Accurate, predictive modeling of soft error rate due to cosmic rays and chip alpha radiation," in *Proc. IEEE Int. Reliab. Phys. Symp.*, 1994, pp. 12–16.
[26] W. Liu, *MOSFET Models for SPICE Simulation Including BSIM3v3 and BSIM4*. New York: Wiley, 2001.
[27] R. C. Jaeger, R. M. Fox, and S. E. Diehl, "Analytic expressions for the critical charge in CMOS static RAM cells," *IEEE Trans. Nucl. Sci.*, vol. 30, no. 6, pp. 4616–4619, 1983.
[28] T. Sakurai and A. Newton, "Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas," *IEEE J. Solid-State Circuits*, vol. 25, no. 2, pp. 584–594, 1990.
[29] Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI Devices*. New York: Cambridge University, 1998.
[30] K. Takeuchi, T. Tatsumi, and A. Furukawa, "Channel engineering for the reduction of random dopant placement induced threshold voltage fluctuations," in *Int. Electron Devices Meet., IEDM, Techn. Dig.*, 1996, pp. 841–844.
[31] H. Masuda, S. Okawa, and M. Aoki, "Approach for physical design in sub-100 nm era," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, 2005, pp. 5934–5937.
[32] M. Eisele, J. Berthold, D. Schmitt-Landsiedel, and R. Mahnkopf, "The impact of intra-die device parameter variations on path delays and on the design for yield of low voltage digital circuits," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 4, pp. 360–368, 1997.
[33] L. Brusamarello, R. da Silva, G. I. Wirth, and R. A. L. Reis, "Probabilistic approach for yield analysis of dynamic logic circuits," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 8, pp. 2238–2248, 2008.
[34] Y. Cao and L. T. Clark, "Mapping statistical process variations towards circuit performance variability: An analytical modeling approach," in *Proc. Des. Autom. Conf. (DAC)*, 2005, pp. 658–663.
[35] H. Nho, S. Yoon, S. S. Wong, and S. Jung, "Numerical estimation of yield in sub-100-nm SRAM design using Monte Carlo simulation," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 55, no. 9, pp. 907–911, 2008.
[36] M. R. de Alba Rosano and A. D. Garcia-Garcia, "Measuring leakage power in nanometer CMOS 6T SRAM cells," in *Proc. IEEE Int. Conf. Reconfigurable Comput. FPGA's*, 2006.
[37] C. Wang, C. Lee, and W. Lin, "A 4-kb low-power SRAM design with negative word-line scheme," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 54, no. 5, pp. 1069–1076, 2007.
[38] S. V. Walstra and C. Dai, "Circuit-level modeling of soft errors in integrated circuits," *IEEE Trans. Device Mater. Reliab.*, vol. 5, no. 3, 2005.
[39] J. T. Horstmann, U. Hilleringmann, and K. Goser, "Correlation analysis of the statistical electrical parameter fluctuations in 50 nm MOS transistors," in *Proc. 28th Eur. Solid-State Devices Conf.*, 1998, pp. 512–515.
[40] S. Mukhopadhyay, H. Mahmoodi, and K. Roy, "Modeling of failure probability and statistical design of SRAM array for yield enhancement in nanoscaled CMOS," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, pp. 1859–1879, 2005.
[41] G. Hubert, N. Buard, C. Weulersse, T. Carriere, M.-C. Palau, J.-M. Palau, D. Lambert, J. Baggio, F. Wrobel, F. Saigne, and R. Gaillard, "A review of DASIE code family contribution to SEU/MBU understanding," in *Proc. Int. Online Test Symp. (IOLTS)*, 2005, pp. 87–94.
[42] P. R. Fleming, B. D. Olson, W. T. Holman, B. L. Bhuva, and L. W. Massengill, "Design technique for mitigation of soft errors in differential switched-capacitor circuits," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 55, no. 9, pp. 838–842, 2008.

Hassan Mostafa (S'01) received the B.Sc. and M.Sc. degrees (with honors) in electronics from Cairo University, Cairo, Egypt, in 2001 and 2005, respectively. He is currently working toward the Ph.D. degree in electrical and computer engineering in the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada.

He was working in a project with Imec, Leuven, Belgium, in 2000. This project includes modeling and fabricating the ISFET transistor. His research interests include low-power circuits, variation-tolerant design, soft errror tolerant design, and statistical design methodologies.

Mohab Anis (S'98-M'03) received the B.Sc. degree (with honors) in electronics and communication engineering from Cairo University, Cairo, Egypt, in 1997 and the M.A.Sc. and Ph.D. degrees in electrical engineering from the University of Waterloo, Waterloo, ON, Canada, in 1999 and 2003, respectively.

He is currently an Associate Professor and the Codirector of the VLSI Research Group, Department of Electrical and Computer Engineering, University of Waterloo. He has authored/coauthored over 90 papers in international journals and conferences and is the author of the following two books: *Multi-Threshold CMOS Digital Circuits-Managing Leakage Power* (Kluwer, 2003) and *Low-Power Design of Nanometer FPGAs: Architecture and EDA* (Morgan Kaufmann: 2009). His research interests include integrated circuit design and design automation for VLSI systems in the deep submicrometer regime. He is the Cofounder of Spry Design Automation.

Dr. Anis is an Associate Editor of the *Journal of Circuits, Systems and Computers, ASP Journal of Low Power Electronics,*, and *VLSI Design*. He is an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—II: EXPRESS BRIEFS. He is also a member of the program committee for several IEEE conferences. He was the recipient of the 2009 Early Research Award from Ontario's Ministry of Research and Innovation, and the 2004 Douglas R. Colton Medal for Research Excellence in recognition of his excellence in research, leading to new understanding and novel developments in microsystems in Canada. He won the 2002 International Low-Power Design Contest.

Mohamed Elmasry (S'69-M'73-SM'79-F'88) was born in Cairo, Egypt, on December 24, 1943. He received the B.Sc. degree from Cairo University, Cairo, Egypt, in 1965, and the M.A.Sc. and Ph.D. degrees from the University of Ottawa, Ottawa, ON, Canada, in 1970 and 1974, respectively, all in electrical engineering.

He has worked in the area of digital integrated circuits and system design for the last 35 years. From 1965 to 1968, he was with Cairo University, and from 1972 to 1974, he was with Bell-Northern Research, Ottawa. Since 1974, he has been with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, where, from 1986 to 1991, he held the NSERC/BNR Research Chair in VLSI design, and where he is currently a Professor and founding Director of the VLSI Research Group. He has served as a Consultant to research laboratories in Canada, Japan, and the United States. He has authored or coauthored over 400 papers and 14 books on integrated circuit design and design automation. He is the holder of several patents. He is the founding President of Pico Electronics Inc., Waterloo, ON, Canada.

Dr. Elmasry has served in many professional organizations in different positions and received many Canadian and international awards. He is a Founding Member of the Canadian Conference on VLSI, the Canadian Microelectronics Corporation (CMC), the International Conference on Microelectronics (ICM), MICRONET, and Canadian Institute for Teaching Overseas (CITO). He is a Fellow of the Royal Society of Canada and a Fellow of the Canadian Academy of Engineers.