Analytical Modeling of SRAM Dynamic Stability

Bin Zhang, Ari Arapostathis, #Sani Nassif, and Michael Orshansky
Department of Electrical and Computer Engineering, University of Texas at Austin
#IBM Austin Research Labs
bzhang@ece.utexas.edu

ABSTRACT
In this paper, for the first time, a theory for evaluating dynamic noise margins of SRAM cells is developed analytically. The results allow predicting the transient error susceptibility of an SRAM cell using a closed-form expression. The key innovation involves using the methods of nonlinear system theory in developing the model. It is shown that when a transient noise of given magnitude affects a sensitive node of a cell, the bi-stable, feedback-driven nature of the cell determines whether the noise will be suppressed or will evolve to eventually flip state. The specific formal and quantitative result is a closed-form expression that can be used to predict whether a cell flip will occur for a noise signal with specific characteristics, and for a given SRAM cell design. Experiments show excellent match between the analytical prediction and the SPICE simulation results.

1. INTRODUCTION
Verifying stability of SRAM-based memory arrays is an essential design task. For memory design in nanometer scale technologies, the traditionally-used static stability analysis may no longer be sufficient. The overriding reason is that the power-minimization driven supply voltage scaling dangerously reduces design options. This is especially so when static noise margins (SNM) [1] are used. Checking stability using dynamic noise margins requires complex time-dependent stability analysis but permits a more aggressive flexible design, without jeopardizing reliability. Dynamic noise margins (DNM) take into account spectral and time-dependent properties of the specific noise patterns. These include the on-chip noise sources such as power and ground network noise, substrate injection noise, capacitive coupling, as well as the extrinsic transient noise such as those due to single event upsets (SEU) [3-6]. Single event upsets are becoming especially troublesome for memory arrays (much more than for combinational logic with its natural masking mechanisms) and dynamic stability analysis is essential for predicting bit error rates accurately. An additional factor for the growing importance of DNM analysis is the adoption of SOI devices for SRAM cells because the floating-body effect leads to a dynamic variation of transistor strengths during read and write operations [7].

ICCAD'06, November 5-9, 2006, San Jose, CA

Stability analysis is concerned with identifying the maximum possible unintended violations that will not cause the stored state to be lost. Traditional static stability analysis finds the noise margins by identifying the maximum *amplitude* of a voltage deviation on an input node that can be tolerated. Static analysis ignores the fact that not all transient noise affecting a sensitive node of an SRAM cell will cause the state to flip. There is a range of noise patterns that will cause only *a temporary* disturbance of the internal node voltage. Whether a transient noise signal is benign (will not cause a flip) or malignant (will cause a flip) depends not only on the amplitude of the signal but also on *its duration* (or pulse width). In addition, the outcome of the temporary disturbance is also determined by the electrical and geometric parameters of the cell. *This paper demonstrates how nonlinear study can be used to develop an analytical theory of the bi-stable cross-coupled inverter system affected by transient noise.*

A well-developed and diverse theory exists for analysis of SRAM cell stability based on the notion of SNM [1]. SNM analysis ignores the transient aspect of noise and posits a constant noise signal is present at a node. Despite the wealth of methods, it has been shown the various criteria are equivalent [2], including coincidence of equilibrium states, the small signal closed-loop gain of the SRAM being unity, the setting of the Jacobian of the Kirchhoff equations of the SRAM to zero. The equivalency extends to the most "famous" criterion that is familiar to many SRAM circuit designers and is based on the graphical method of inscribing the maximum square between normal and mirrored voltage transfer characteristics [2].

While there exists rich theory for static stability analysis of SRAM cells, extending it to dynamic analysis has proven difficult. The importance of transient noise susceptibility of SRAM cells has been of course realized [6]. However, all of the prior works rely on simulation techniques at the device level [8], transistor level [9], or in a mixed mode of device and transistor levels [10]. Simulation techniques search for the minimum strength of a malignant transient noise (for instance, in the form of critical charge deposited by an SEU). Although accurate, the simulation results can only provide limited insight into the problem. A somewhat similar problem is encountered in the characterization of combinational logic gates' susceptibility to transient errors [11-13]. Dynamic response of a gate is obtained by solving the simplified differential equations describing its dynamic behavior. However, because of the nonlinear cross-coupling between the nodal voltages of an SRAM cell, it is difficult to extend this literature to the problem at hand. Thus, a closed-form *analytical* formulation that reveals the relationship between the minimum strength of a malignant transient noise and the configuration parameters of the SRAM cell is sought. In addition to giving invaluable intuition about SRAM dynamic stability, it can be directly used for reliability, yield and optimization to avoid lengthy iterative SPICE simulations [14].

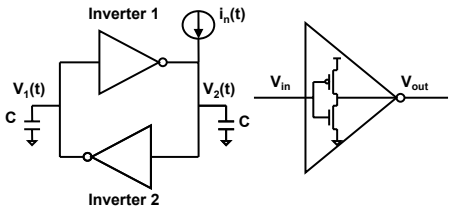
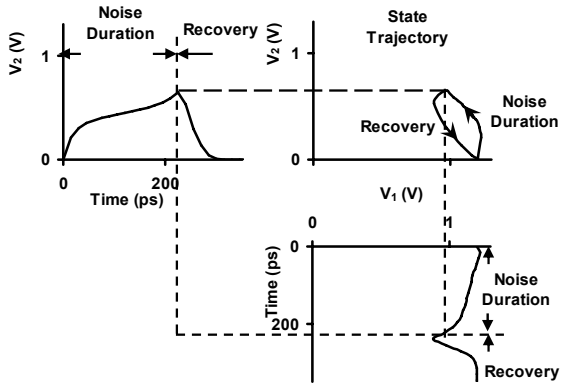
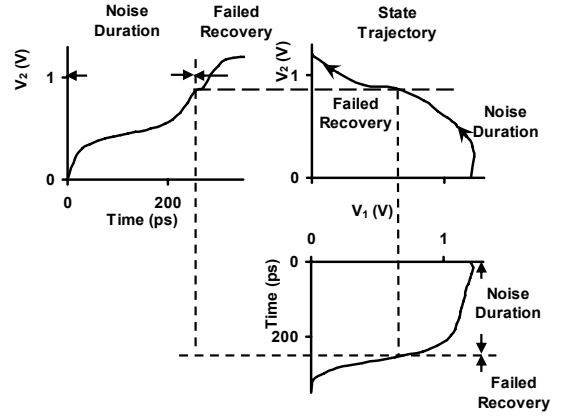


Figure 1. A 6T SRAM cell with a transient current noise being injected (access transistors not shown).



(a)



(b)

Figure 2. State space and time-domain plots of SRAM nodal voltages with injected noise. Depending on noise duration (a) cell dynamically recovers, or (b) state-flip occurs.

In this paper, for the first time, a method for evaluating dynamic noise margins is developed analytically. The results allow predicting the transient error susceptibility of an SRAM cell using a closed-form expression. The key innovation involves using the methods of nonlinear system theory in developing this model. It is shown that when a transient noise of a given magnitude affects a sensitive node of a cell, the bi-stable, feedback-driven nature of the cell determine whether the noise will be suppressed or will evolve to eventually flip state. It is shown that for the flip to take place, the noise needs to exceed specific threshold amplitude, which is shown to be equal to the SNM, and be sustained longer than a minimum critical duration.

The inverters of the cross-coupled pair in an SRAM cell are described as operating in two modes: high gain and low gain, such that in each mode, the driving current of the inverter is a linear function of either the input voltage or the output voltage but not both [15]. This permits decoupling of the input and output voltage dependences. Using a piece-wise linear MOSFET model, rigorous analysis of the coupled nonlinear feedback system is performed. The specific formal and quantitative result is a closed-form expression that can be used to predict whether a cell flip will occur for a noise signal with specific characteristics. For a given SRAM cell design, the expression evaluates the critical pulse width for various noise amplitudes. Experiments show excellent match between the analytical prediction and the SPICE simulation results.

This paper is organized as follows. In Section 2, we set up differential equations to describe the nonlinear dynamics of the SRAM cell. Based on the mathematical formulation of Section 2, we conduct the analytical study on the dynamic stability of the SRAM cell in Sections 3 and 4. We present the experimental results in Section 5. Finally we draw conclusions in Section 6.

2. SYSTEM MODELING SETUP

In this paper, we focus on the stability of a 6T SRAM cell with respect to an SEU. An SEU is a rare and random event. At the moment when it occurs, the SRAM cell is most likely in the standby mode. Therefore, we will only evaluate the SRAM stability in the standby mode. For other noise sources, the most interested modes might be the read or write mode. However, the framework built in this paper can be easily extended to those cases.

A 6T SRAM cell consists of two identical cross-coupled inverters and two access transistors. For notational simplicity, the two access transistors are not included in the analysis. Without loss of generality, the noise source in this paper is modeled as a current pulse injected into the internal node where the initial voltage is zero, Figure 1. If the noise source flows in the opposite direction, i.e. drawing positive charge from the internal node where the initial voltage is V_{dd}, the following discussion still applies, but with the roles of the NMOS and PMOS being switched. Capacitance, C, is the total lumped capacitance at the output node of each inverter. We describe the *state* of the cell by the *state vector* $\vec{V} = (V_1,\ V_2)$, where V_1 and V_2 are the two nodal voltages of the cell. For a given state vector value, the currents flowing through the four transistors can be determined. The set of $\vec{V}$ of all possible values forms the *state space*. The transient noise current is modeled as a square pulse with amplitude, I_n, and pulse width, pw:

$$i_n(t) = \begin{cases} I_n & (0 < t < pw) \\ 0 & (t \le 0 \text{ or } t \ge pw) \end{cases} \quad (1)$$

The square pulse model of the noise signal is an obviously simplified representation of the actual noise. For many noise source patterns, such a model may be sufficient. Although it ignores the details of the pulse's leading and trailing edges, the basic features of a transient noise are captured. Specifically, the model captures the average (or maximum) value of the noise through the amplitude and the transient nature of the noise through the pulse width.

The question of interest is whether under the impact of the noise signal the cell will undergo a *state-flip*, that is, will change the state. In SNM analysis [1], the voltage transfer curves are used to

identify the minimum voltage/current noise amplitude that will lead to a state-flip. Figure 2 shows the time behavior of the two nodal voltages under the impact of two separate transient noises with the same amplitude but different pulse widths. In Figure 2 (a), the SRAM cell successfully recovers to its original state after the short noise vanishes, while in Figure 2 (b), the SRAM cell fails to recover due to longer noise duration and undergoes a state-flip. Figure 2 also shows the *state trajectories* for both cases. The state trajectories are obtained by plotting the state vectors at various times in the state space. It can be seen that the state trajectory is a closed loop if no state-flip occurs and an open curve otherwise.

We start by setting up a general framework. It can be later shown how SNM analysis can be derived from it. The dynamic stability analysis is based on a system of coupled differential equations (2) and (3) that describe the dynamics of the SRAM cell - the evolution of the state vector driven by the charging/discharging of the nodal capacitances:

$$dV_1(t)/dt = -I_{inv}(V_2(t),V_1(t))/C \quad (2)$$

$$dV_2(t)/dt = [-I_{inv}(V_1(t),V_2(t)) + i_n(t)]/C \quad (3)$$

In these equations, $I_{inv}(V_{in},V_{out})$ is the *driving current* of an inverter, i.e. the current flowing into the inverter through its output node. The positive current direction is chosen to be discharging any nodal capacitor. Since I_{inv} is a nonlinear function of the inverter's input and output voltages, equations (2) and (3) describe a *nonlinear system*. The driving current I_{inv} is the sum of the drain-to-source currents of the NMOS and PMOS transistors, I_{nmos} and I_{pmos} :

$$I_{inv}(V_{in},V_{out}) = I_{nmos}(V_{in},V_{out}) + I_{pmos}(V_{in},V_{out}) \quad (4)$$

For both transistors, the driving currents are completely characterized by their gate-to-source voltage (V_{gs}) and drain-to-source voltage (V_{ds}). In particular, for NMOS, $V_{gs} = V_{in}$ and $V_{ds} = V_{out}$.

The essence of the inverter's driving current is best captured through the *drive curves* [15]. The drive curves are a contour map of the inverter's driving current versus its input and output voltages, Figure 3. The drive curves can be characterized by SPICE simulations. The characterization procedure is similar to that of the dc transfer curve, except that the output node of the inverter is loaded with a constant current source which the driving current is equal to. The dc voltage transfer curve is just one in the set of the drive curves, with:

$$I_{inv}(V_{in},V_{out}) = 0 \quad (5)$$

Drive curves are described by $I_{inv}(V_{in},V_{out}) = \kappa$, where κ is a constant. Each curve divides the state space into two regions, with one region $I_{inv}(V_{in},V_{out}) > \kappa$ and the other $I_{inv}(V_{in},V_{out}) < \kappa$.

Analytical study of the dynamic stability requires analytical modeling of the dependence of the driving currents on circuit parameters. Solvability of a nonlinear system problem is critical because known mathematical methods do not provide powerful enough means to analytically solve many nonlinear system problems [16]. Therefore, in addition to accuracy, the model of driving current must allow us to solve the SRAM equations (2) and (3) *analytically*.

The above requirements mean that a *linear* and *separable* model of driving current needs to be employed. We adopt the linear gate model of [15]. It is based on functional variable decoupling and piece-wise linearization. Specifically, an inverter is modeled as operating in two modes - a high gain and a low gain mode. In each mode, the driving current of the inverter is either a linear function of the input voltage or the output voltage, but never both.

Additionally, the model assumes that the short-circuit current is negligibly small [17]. This means that the NMOS and PMOS transistors never conduct simultaneously. (This assumption holds if the rise/fall time of the output voltage of the inverter is longer or comparable to that of the input voltage [17]. The assumption is verified and appears justified in this work). When the NMOS transistor conducts, the driving current of the inverter, I_{inv}, can be written as:

$$I_{inv}(V_{in},V_{out}) = \begin{cases} 0 & \text{(cutoff)} \\ g_{mn}(V_{in} - V_{thn}) & \text{(saturation)} \\ V_{out}/R_n & \text{(linear)} \end{cases} \quad (6)$$

where g_{mn}, V_{thn}, R_n are respectively the transconductance, threshold voltage, and linear-region resistance of the NMOS transistor. We refer to (6) as the "linear gate model". The key to the linear gate model is that these parameters are independent of the input and output voltages. When the PMOS transistor conducts, I_{inv} can be modeled similarly. With this approximation, the number of piece-wise linear regions that need to be considered is greatly reduced. This enables us to obtain a concise analytical solution without significant loss of accuracy.

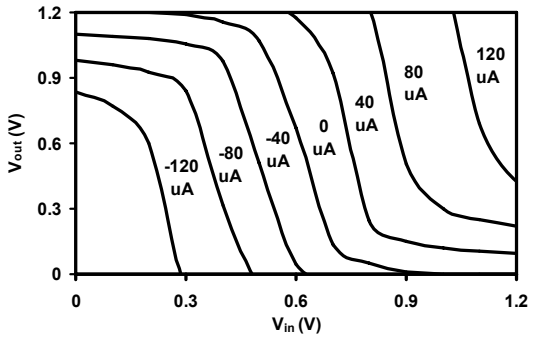
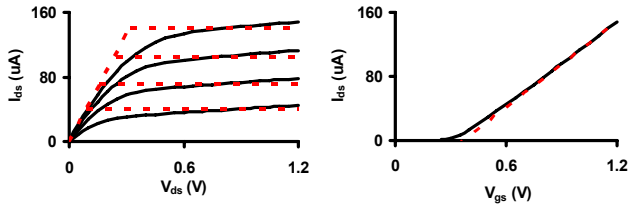


Figure 3. Equi-current drive curves are used to capture I-V characteristics of an inverter.



(a) **(b)**

Figure 4. To enable analytical solution a linear decoupled MOSFET model is used. I-V plots of a 100nm NMOS show that the approximation (dashed lines) is reasonable: (a) I_{ds}-V_{ds}, and (b) I_{ds}-V_{gs} for V_{ds} = 1.2V.

Characterizing these parameters is based on the I-V characteristics of the conducting MOS transistor, Figure 4. MOS transistors in the deep sub-micron domain follow the α-power law [18]. As transistor feature size scales down, α approaches one (α = 1.27 for NMOS and 1.46 for PMOS in 100nm technology). As a result, the transistor's drain-to-source current, I_{ds}, in saturation mode will be closer to a linear function of V_{gs} [18]. Thus, we can expect a better match between the piece-wise linear model and the actual I-V characteristic in future technologies.

3. DYNAMIC STATE SPACE ANALYSIS

According to the SRAM equations (2) and (3), a dynamic process of charging/discharging of the two capacitors that causes the evolution of the state vector $\vec{V}$ will continue until it reaches the *steady state* ($d\vec{V}/dt = 0$).The values of state vectors over time form the state trajectory, a curve linking the initial and steady states. If the currents are finite, the state trajectory is continuous. In the language of nonlinear system theory [16], any state that satisfies the condition $d\vec{V}/dt = 0$ is an *equilibrium state*. Thus, the steady state of a dynamic process is an equilibrium state. When there are multiple equilibrium states, the steady state is also determined by the initial state of the dynamic process. Evaluating (2) and (3) for $d\vec{V}/dt = 0$, it can be seen that the equilibrium states of an SRAM cell are given by the roots of the following equations:

$$-I_{inv}(V_2,V_1)/C = 0 \tag{7}$$

$$[-I_{inv}(V_1,V_2) + i_n(t)]/C = 0 \tag{8}$$

The above equations suggest that in the presence of transient noise, equilibrium states change over time. Under the noise model in (1), the change is not continuous and only occurs at $t = 0$ and $t = pw$. Specifically, equations (7) and (8) simplify to

$$I_{inv}(V_2,V_1) = 0 \tag{9}$$

$$I_{inv}(V_1,V_2) = \begin{cases} I_n & (0 < t < pw) \\ 0 & (t \le 0 \text{ or } t \ge pw) \end{cases} \tag{10}$$

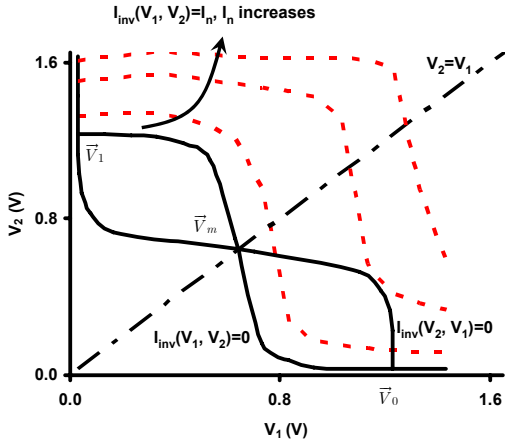


Figure 5. Stability analysis is done using superposition of mirrored transfer curves (drive curves). Injection of noise dynamically shifts the curves changing equilibrium conditions.

The traditional way of showing the roots of these equations is to plot the functions in (9) and (10) in the state space [1], Figure 5. The function described by $I_{inv}(V_1,V_2) = I_n$ is plotted for different I_n. These curves are in fact the drive curves of inverters discussed earlier.

The three crossing points between the drive curves $I_{inv}(V_2,V_1) = 0$ and $I_{inv}(V_1,V_2) = 0$ form the equilibrium states for $t \le 0$ or $t \ge pw$: $\vec{V}_0$, $\vec{V}_1$, and $\vec{V}_m$, referring to the states "0", "1", and "meta-stable". Importantly, these are also the equilibrium states for the noise-free case. The crossing points between the curve $I_{inv}(V_2,V_1) = 0$ and a dashed line are the equilibrium states for a particular I_n during $0 < t < pw$. It can be seen from Figure 5 that as I_n increases, the number of equilibrium states for $0 < t < pw$ reduces from three to one.

It is also possible to relate the current discussion to the traditional SNM analysis. Several different criteria have been shown equivalent, including the one based on coincidence of equilibrium states. This criterion is clearly identical to the case when the number of equilibrium states is two, Figure 5. The noise amplitude that will result in this case is then precisely the SNM of the cell (I_{snm}). Any higher noise amplitude will result in there remaining a single equilibrium state, making the SRAM cell unable to store either "0" or "1". Specifically, if the remaining equilibrium state represents a "1", the charging/discharging process on the two capacitors will pull any initial state, even if it is a "0", to the only equilibrium state "1". Effectively, SNM analysis ignores the fact that the change of the equilibrium states is only temporary, due to the transient nature of the noise. After the transient noise vanishes ($t > pw$), the equilibrium states will return to their original locations. If during the time when the noise is present ($0 \le t \le pw$) the state trajectory has not moved sufficiently far away from the initial state, the driving currents of the inverters are able to pull the state vector back to the correct equilibrium state (the initial state) after the noise vanishes.

After the transient noise vanishes, any initial state in this recovering stage, which is also the final state in the noise duration, Figure 2, will eventually be driven to one of the three equilibrium states as $t \to \infty$. Figure 6 shows three such state trajectories in the recovering stage. The equilibrium states behave as attractors, with each equilibrium state $\vec{V}_e$ having its *attraction region* $\mathbf{A}(\vec{V}_e)$ [16]. The attraction region is defined by the following: for any initial state vector in $\mathbf{A}(\vec{V}_e)$, its state trajectory will terminate in $\vec{V}_e$ for $t \to \infty$. Because of the assumed symmetry of the SRAM cell, the two cross-coupled inverters are identical. By symmetry, the boundary of attraction regions in the recovering stage is:

$$V_2 = V_1 \tag{11}$$

Any point in the state space that is below (above) the boundary belongs to the attraction region of $\vec{V}_0$ ($\vec{V}_1$). Any point on the line is attracted to $\vec{V}_m$. The transient noise modifies (moves) the equilibrium states of the SRAM cell and their attraction regions only during $0 < t < pw$. During this period, the attraction regions can be approximately identified from the locations of their associated equilibrium states in the state space. However, the exact location of the boundary of the attraction regions for $0 < t < pw$ is not required in this work.

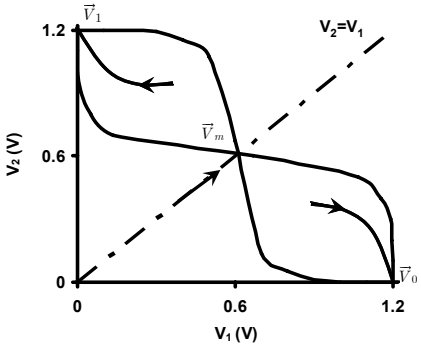
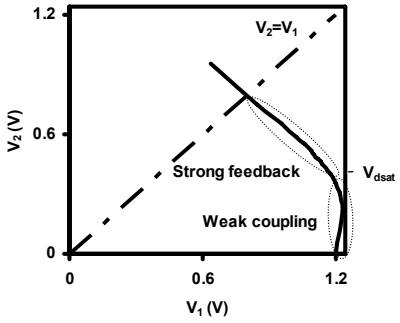
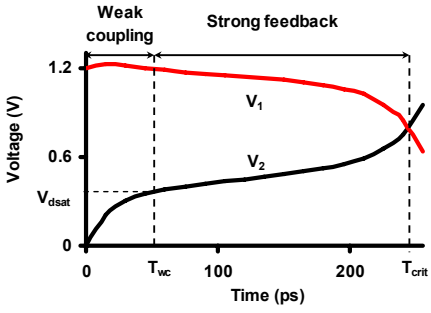


Figure 6. Without noise, the region of attraction for $\vec{V}_1$ ($\vec{V}_0$) is the entire region above (below) V_2=V_1. Any state trajectory starting from an initial state in a region of attraction of an equilibrium state will reach it eventually.



(a)



(b)

Figure 7. For noise amplitudes higher than the SNM (I_n>I_{snm}) only one equilibrium state exists (above V_2=V_1): (a) state trajectory; (b) time-domain node voltage evolution.

4. TRANSIENT BEHAVIOR OF SRAM UNDER NOISE

The effect of the transient noise on the stability of the SRAM cell is as follows. Suppose the cell is in the equilibrium state $\vec{V}_0 = (V_{dd}, 0)$ before the transient noise occurs:

$$\vec{V}(t \le 0) = \vec{V}_0 \qquad (12)$$

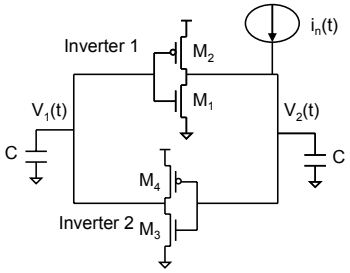


Figure 8. Schematics of the SRAM cell with a transient current noise being injected.

During $0 < t < pw$, the noise current and driving currents of the two inverters force the state vector of the cell to move away from $\vec{V}_0$. If by the time the transient noise vanishes ($t = pw$), the state vector is still within the attraction region of $\vec{V}_0$ (i.e., $V_2 < V_1$), the state trajectory will eventually terminate on $\vec{V}_0$, causing no error. Otherwise, the final state would be $\vec{V}_1$ or $\vec{V}_m$ (unstable), and a state flip will occur. Therefore, the criterion for dynamic stability is whether the state trajectory can cross the boundary of attraction regions *by the time* the transient noise vanishes.

First, we analyze the dynamic stability of the SRAM cell for $I_n \le I_{snm}$, i.e. the noise amplitude is smaller than the SNM. The appearance of noise at $t = 0$ relocates the equilibrium state originally at $\vec{V}_0$ to a higher position, Figure 5. From the symmetry of the system and the monotonicity of the drive curves, we can see that this new equilibrium state is always below the line $V_2 = V_1$. During $0 < t < pw$, the driving currents of the inverters and the noise current move the state vector in the direction of this new equilibrium state. When the transient noise vanishes, the state vector is guaranteed to be within the attraction region of $\vec{V}_0$, i.e. below the line $V_2 = V_1$, regardless of how large the noise pulse width pw is, and thus no state flip will occur. This is consistent with the conclusion of the SNM analysis. Second, we consider the case when $I_n > I_{snm}$. For $0 < t < pw$, according to the discussion in Section 3, there exists only one equilibrium state, which is above the line $V_2 = V_1$. Figure 7(a) shows that under the attraction of this equilibrium state, the state trajectory moves toward the line $V_2 = V_1$ and may cross the line if the noise pulse width pw exceeds the *critical pulse width*, T_{crit}, which is the time the trajectory arrives at the line. T_{crit} is a function of the pulse amplitude I_n. Figure 7(b) shows the corresponding SPICE-generated plot of the time evolution of the state vector. Thus, the criterion of dynamic stability becomes whether the pulse width pw is greater than T_{crit}. To determine T_{crit}, we only need to analyze the transient behavior before the line $V_2 = V_1$ is reached. Once the line is crossed, a flip will occur. During this time, the noise current is being drained through NMOS transistor M1, and the output voltage of inverter 2 switches low through the NMOS transistor too, Figure 8. The majority of the driving currents of both inverters are carried by the NMOS transistors. For inverter 2, since $V_2 < V_1$, M3 is either in cutoff mode or saturation mode, making the driving current primarily controlled by V_2 through

transconductance. It is useful to divide the operation of the SRAM cell, Figure 7, into two regions: *weak coupling* and *strong feedback mode*. The boundary between the two regions is at the saturation voltage of M1, that is when $V_2 = V_{dsat}$ of M1 with $V_{gs} = V_{dd}$.

The state trajectory starts in the weak coupling mode, in which M1 is in the linear region and M3 is mostly in cut-off if M3's threshold voltage V_{thn} does not differ from V_{dsat} too much. The state trajectory enters the strong feedback mode when M1 becomes saturated. In the weak coupling mode, V_1 nearly remains at V_{dd} and V_2 increases from 0 to V_{dsat}. Figure 7(b) shows that the increase of V_2 gets slower with time in the weak-coupling mode, an asymptotic behavior that is common for RC circuits. This feature will be used later for developing the model. In the strong feedback mode ($V_2 > V_{dsat}$), the driving currents of both inverters are primarily controlled by transconductance, i.e., the change of V_1 is primarily determined by V_2 and vice versa, which results in positive feedback. The positive feedback makes both nodal voltages change in a nearly exponential fashion, as shown in Figure 7(b).

The above regions of operation can be modeled using the linear gate model of (6). Specifically, in the weak coupling region, Figure 9(a), inverter 1 behaves like a resistor from V_2 to the ground, with resistance being the linear-region resistance of M1, R_n. Inverter 2 is approximated as not switching. This approximation is only true if V_1 nearly remains at V_{dd} in this mode. We found that for a cell designed in the 100nm technology V_1 can differ from V_{dd} by at most 0.05V for $V_2 \le V_{dsat}$, indicating that the approximation is reasonable. From the above analysis, the time behavior of the SRAM cell in the weak coupling mode is governed by the following equations:

$$dV_2/dt = -V_2/(R_nC) + I_n/C \tag{13}$$

$$V_1 = V_{dd} \tag{14}$$

with initial state being $(V_{dd}, 0)$, and final state being (V_{dd}, V_{dsat}), which is the point where the state trajectory exits the weak coupling mode. From (13), V_2 increases asymptotically toward the steady state value I_nR_n until the working mode is switched.

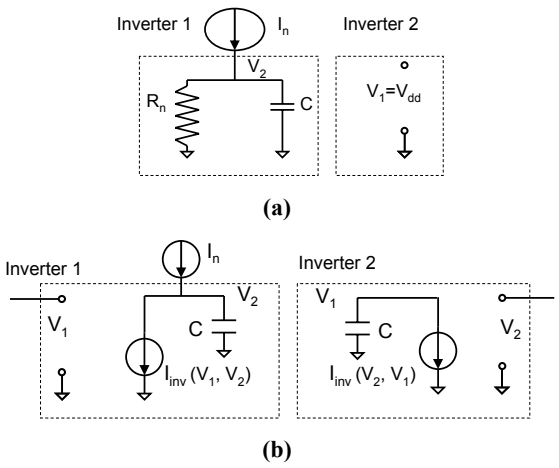


Figure 9. Using piece-wise linear gate model, SRAM operation is divided into regions of (a) weak coupling and (b) strong feedback.

Solving (13) with the initial and final conditions, we find that the time until the trajectory of the state vector leaves the weak coupling mode, T_{wc}, is given by:

$$T_{wc} = -R_nC\ln[1 - V_{dsat}/(I_nR_n)] \tag{15}$$

It is clear from (15) that to reach the strong feedback mode, the noise amplitude I_n must be larger than V_{dsat}/R_n. Otherwise, positive feedback will not occur and no state-flip will result. Therefore, the SNM I_{snm} can be approximated by V_{dsat}/R_n.

In the strong feedback mode, Figure 9(b), both inverters are modeled as voltage controlled current sources between the inverters' output nodes to the ground. By applying the linear gate model (6) in the saturation mode for both inverters, we obtain two cross-coupled linear equations:

$$\begin{bmatrix} dV_1(t)/dt \\ dV_2(t)/dt \end{bmatrix} = -\begin{bmatrix} 0 & g_{mn}/C \\ g_{mn}/C & 0 \end{bmatrix}\begin{bmatrix} V_1(t) \\ V_2(t) \end{bmatrix} + \begin{bmatrix} g_{mn}V_{thn}/C \\ (g_{mn}V_{thn} + I_n)/C \end{bmatrix} \tag{16}$$

with initial conditions

$$V_1(t = T_{wc}) = V_{dd}, \quad V_2(t = T_{wc}) = V_{dsat} \tag{17}$$

Equation (16) describes a linear system. The linear system theory [16] can then be used to solve for the state trajectory analytically. Specifically, linear transformation is first performed on equation (16) so that the matrix is diagonalized. In this way, we obtain two equations governing the dynamics of two characteristic functions which are not cross-coupled. The two transformed equations are similar to equation (13) and can be solved easily. By performing a reverse-transformation on the characteristic functions, we can obtain a closed-form solution for the state vector. Since our goal is to find the critical pulse width T_{crit}, we impose the final condition: $V_2 = V_1$. This gives:

$$\begin{aligned} T_{crit}(I_n) = {} & -R_nC\ln[1 - V_{dsat}/(I_nR_n)] - \\ & C/g_{mn}\ln[1 - g_{mn}(V_{dd} - V_{thn})/I_n] \end{aligned} \tag{18}$$

In (18), the first term is T_{wc}, the time spent in the weak-coupling mode, and the second term is the time spent in the strong feedback mode before the attraction boundary is reached. It can be proven that if the first term in (18) is finite, i.e., $I_n > V_{dsat}R_n$, and $V_{dsat} > V_{thn}$, the second term is finite too. In other words, if the noise amplitude exceeds the SNM, given sufficiently long (but still finite) duration for the noise pulse, state-flip can occur. The expression (18) for the critical pulse width gives the minimum pulse duration that a noise pulse must have in order to cause a state-flip.

It is instructive to check the asymptotic behavior of the critical pulse width when $I_n \to \infty$. In this case, we have a simple bound on the critical pulse width:

$$T_{crit}(I_n) = CV_{dd}/I_n \tag{19}$$

This expression states that when the noise amplitude is sufficiently high, the SRAM cell will undergo a state-flip once the total charge deposited by the transient noise is equal to the charge stored on the node, CV_{dd}. However, if the noise amplitude is comparable to those of the driving currents, the critical pulse width given by (18) is greater than that given by (19), as confirmed by the experiments in the next section. This is because for smaller noise amplitude, it will take longer time to build

enough charge on the victim node for a state-flip. During this time, however, M1 can drain more current from the transient noise, requiring more charge to be deposited by the noise source for a state-flip to occur.

5. EXPERIMENTAL RESULTS

In order to verify the validity of the developed theory, and specifically of the critical pulse width T_{crit} (Eq. 18), we designed an SRAM cell using the PTM 100nm technology [19], with V_{dd}=1.2V, $(W/L)_{nmos}$= 2, $(W/L)_{pmos}$=3, and $C = 4.8fC$. We performed SPICE simulation on the SRAM cell as follows. The SRAM cell was put in standby mode. A transient noise having a square-pulse shape was injected into the node where the nodal voltage is zero. The critical pulse width for certain noise amplitude was taken as the smallest pulse width that will result in a state-flip. The SNM was measured as the minimum noise amplitude that can possibly result in a state-flip, with no limitation on pulse width. Figure 10 shows the comparison between the SPICE simulation result and the result calculated by (18). For the convenience of viewing, Figure 10 plots critical pulse width in the horizontal direction. This gives the maximum tolerable noise amplitude given a pulse width. Figure 10 also shows the SNM and the asymptotic behavior of the critical pulse width in (19) as the simple bounds. It can be seen that excellent match is achieved between the analytical solution and the SPICE simulation result. The error is only significant when the noise amplitude is very close to the SNM. This is expected since in this region, the system performance is very sensitive to the circuit parameters. The dependency of T_{crit} on V_{dd} is shown in Figure 11 for three transient noise amplitudes. When adjusting V_{dd}, there is a tradeoff between the amounts of stored charge and feedback delay. Larger V_{dd} leads to shorter feedback delay time, making it easier for a transient noise to be latched. However, larger V_{dd} also results in more charge stored, making it more difficult for a transient noise to be latched. Figure 11 shows that this later effect dominates. It can be seen that as V_{dd} increases, the critical pulse width also increases, indicating that the SRAM cell is more resilient to transient noises for higher supply voltages.

In order to show the validity of the analytical solution for different circuit sizing, we fix the PMOS size at $(W/L)_{pmos}$=3 and vary the NMOS transistor size. Adjusting the transistor size faces similar tradeoff as adjusting V_{dd}: larger transistor size leads to shorter feedback delay time, making it easier for a transient noise to be latched. However, larger transistor size also results in more charges stored and more noise current drained, making it more difficult for a transient noise to be latched. Figure 12 shows that this later effects dominates.

An important application of the critical pulse width is to evaluate the impact of an SEU on the stability of an SRAM cell. An SEU induced by cosmic rays typically injects an "exponential" current pulse into an internal node of the cell [9]:

$$I(q,t) = \frac{2q}{\sqrt{\pi}T_s}\sqrt{\frac{t}{T_s}}e^{-\frac{t}{T_s}} \qquad (20)$$

where q is the collected charge, or total amount of charge deposited by the current pulse, and T_s is the technology-dependent charge-collection time constant It can bee seen from (20) that the pulse duration is technology-dependent, and the pulse amplitude is determined by q. The *critical charge* is defined as the minimum amount of collected charge that is able to flip a cell. To compute the critical charge of an SRAM cell, a mapping that needs to be enabled is between the noise profile in (20) and the square-pulse noise model on which the theory developed here relies. The mapping criteria we used are: (a) matching the total charge between the square pulse and the "exponential" pulse, and (b) injecting the square pulse and its "exponential" equivalent on an SRAM cell should generate similar effect, i.e. similar state trajectories. It was found through extensive experimentation that setting the pulse width to a value of $3T_s$ produces a reasonable match. The magnitude of the current pulse can then be easily found as $q/(3T_s)$, which can be shown to correspond to 69% of the maximum amplitude of the "exponential" pulse. Figure 13 shows a good match between the state trajectories induced by such equivalent pulses for two different collected charges. The procedure of computing the critical charge is as follows. First, we set the critical pulse width to be $3T_s$, and use (18) to compute the corresponding noise amplitude I_n, which is the maximum tolerable noise amplitude for this noise duration. Then, the critical charge is obtained as $3T_sI_n$. To verify the result, SPICE simulation was performed to iteratively search for the critical charge by injecting the "exponential" current pulse into the node where the nodal voltage is zero. The comparison in Figure 14 shows that good match is achieved and the average estimation error is 11%. Figure 14 indicates that higher V_{dd} makes the cell more soft-error resilient, since for higher V_{dd}, more charge is needed to flip a cell.

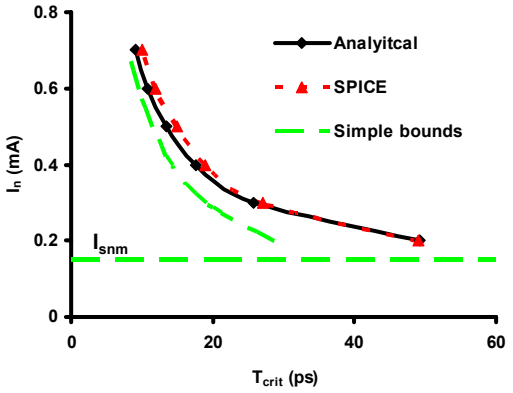


Figure 10. Noise amplitude vs. pulse width: region below is safe. Analytical prediction of dynamic margins matches a simulation well. A substantial improvement over the bounds based on the SNM and a simple bound is also verified.

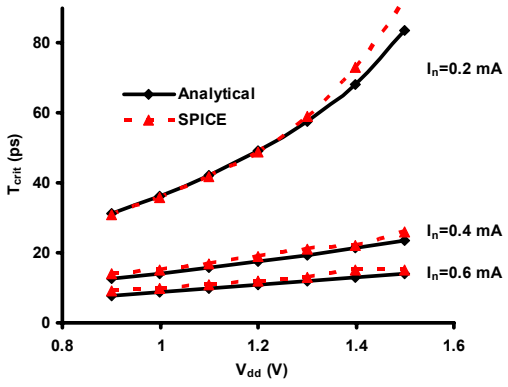


Figure 11. The developed analytical model allows SRAM design space exploration, e.g., studying critical pulse T_{crit} at different V_{dd} and noise amplitudes.

6. CONCLUSIONS

In this paper, for the first time, we present an analytical study of transient error susceptibility of SRAM cells. We model the transient current noise as a square-pulse. By using piece-wise linear modeling of the nonlinear feedback behavior of an SRAM cell, we obtain analytical closed-form solution of critical pulse width that can flip the state of the SRAM cell. Experiments show that excellent match is achieved between the analytical prediction and the SPICE simulation results.

7. ACKNOWLEDGMENTS

This research was supported in part by GSRC, SRC, NSF, SUN, Intel as well as the Office of Naval Research through the Electric Ship Research and Development Consortium.

8. REFERENCES

[1] E. Seevinck, F. J. List and J. Lohstroh, "Static-noise margin analysis of MOS SRAM cells," *JSSC*, 1987, pp. 748-754.

[2] J. Lohstroh, E. Seevinck and J. D. Groot, "Worst-case static noise margin criteria for logic circuits and their mathematical equivalence," *JSSC*, 1983, pp. 803-807.

[3] H. H. Chen and J. S. Neely, "Interconnect and circuit modeling techniques for full-chip power supply noise analysis," *IEEE Trans. Components, Packaging, and Manufacturing Technology*-Part B, 1998, pp. 209-215.

[4] R. Gharpurey and R. G. Meyer, "Modeling and analysis of substrate coupling in IC's," *IEEE Custom-Integrated Circuit Conf.*, 1995, pp. 125–128.

[5] H.-R. Cha and O.-K. Kwon, "An analytical model of simultaneous switching noise in CMOS systems," *IEEE Trans. Advanced Packaging*, 2000, pp. 62-68.

[6] P. E. Dodd and L. W. Massengill, "Basic mechanisms and modeling of single-event upset in digital microelectronics," *IEEE Trans. Nucl. Sci*., 2003, pp. 583–602.

[7] D. C. Pham, *et. al*., "Overview of the architecture, circuit design, and physical implementation of a first-generation cell processor," *JSSC*, 2006, pp. 179-196.

[8] J. S. Fu, C. L. Axness and H. T. Weaver, "Two-dimensional simulation of single event induced bipolar current in CMOS structures," *IEEE Trans. Nucl. Sci*., 1984, pp. 1155–1159.

[9] P. Hazucha and C. Svensson, "Impact of CMOS technology scaling on the atmospheric neutron soft error rate," *IEEE Trans. Nucl. Sci.*, 2000, pp. 2586–2594.

[10] K. Mayaram, J.-H. Chern and P. Yang, "Algorithms for transient three-dimensional mixed-level circuit and device simulation," *IEEE Trans. Computer-Aided Design*, 1993, pp. 1726–1733.

[11] K. Mohanram, "Closed-form simulation and robustness models for SEU-tolerant design," *VLSI Test Symp*., 2005, pp. 327-333.

[12] A. Dharchoudhury, *et al*., "Fast timing simulation of transient faults in digital circuits," *ICCAD*, 1994, pp. 719-726.

[13] M. Omana, *et al*., "A model for transient fault propagation in combinatorial logic," *Intl. On-line Testing Symp*., 2003, pp. 111-115.

[14] R. Heald and P. Wang, "Variability in sub-100nm SRAM designs," *ICCAD*, 2004, pp. 347-352.

[15] M. Horowitz, "Timing models for MOS circuits," Ph. D. Dissertation, Stanford University, 1984.

[16] Z. Vukic, *et. al.*, *Nonlinear Control Systems*, Marcel Dekker Inc., 2003.

[17] N. Hedenstierna and K. O. Jeppson, " CMOS circuit speed and buffer optimization, " *IEEE Trans. Computer-Aided Design*, 1987, pp. 270-281.

[18] T. Sakurai and A. R. Newton, "Alpha-power law MOSFET model and its applications to CMOS inverter," *JSSC*, 1990, pp. 584-594.

[19] PTM, http://www.eas.asu.edu/~ptm/.

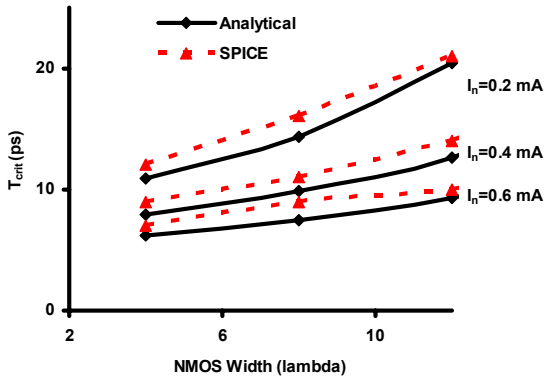


Figure 12. Dependency of T_{crit} on the NMOS size for different transient noise amplitudes. PMOS size is $(W/L)_{pmos}$=3.

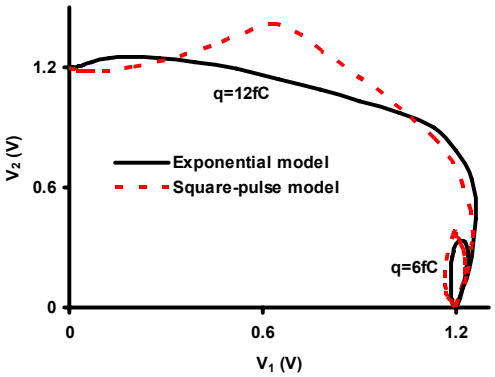


Figure 13. A square pulse model approximates here exponential current source by matching total charge. State-trajectories for pulses with two total charge values are shown (q=6 fC and q=12 fC). Trajectories are obtained via SPICE.

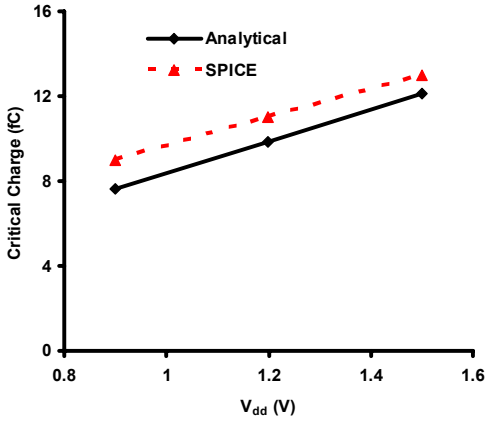


Figure 14. Critical charge that will cause a state-flip by a single event upset.