Improving Integrated Circuit Performance Through the Application of Hardness-by-Design Methodology

Ronald C. Lacoe, *Member, IEEE*

***Abstract*—Increased space system performance is enabled by access to high-performance, low-power radiation-hardened microelectronic components. While high performance can be achieved using commercial CMOS foundries, it is necessary to mitigate radiation effects. This paper describes approaches to fabricating radiation-hardened components at commercial CMOS foundries by the application of novel design techniques at the transistor level, the cell level, and at the system level. This approach is referred to as hardness-by-design. In addition, trends in the intrinsic radiation hardness of commercial CMOS processes will be discussed.**

***Index Terms*—Hardness-by-design (HBD), multibit upsets, single-event latchup, single-event transient, single-event upset, total-ionizing dose radiation.**

I. Introduction

DESIGNERS of space systems typically want to achieve the highest performance systems possible consistent with budgetary and schedule constraints. Satellite systems use microelectronics for command and control functions, for signal acquisition and processing functions, and for data storage. The more signal processing throughput available, the more onboard signal processing can be performed, often resulting in increased overall system performance and a reduction in off-board ground signal processing. Signal processing throughput is a function of the number of gates and the gate switching frequency. In today's world, most signal processing is done by circuits using CMOS technology, where gates are constructed from complementary field effect transistors. Using this technology, advanced processors, memories, and application-specific integrated circuits (ASICs) are built. CMOS technology has been characterized by continued scaling, where characteristic feature parameters are shrinking, resulting in increased performance at lower power. This is the exact trend that satellite system designers desire in microelectronic components. Clearly, the ability to leverage the most aggressively scaled CMOS technology would greatly benefit space systems.

Technology scaling can be defined as the process of reducing the sizes of both active and passive devices in order to improve packing density and circuit speed. In 1974, Denard *et al.* presented the first systematic study of the impacts of technology scaling on circuit performance [1]. If the dimensions and voltages were scaled down by the same factor then: 1) the operating frequency would increase; 2) more transistors could be packed in the same sized integrated circuit (IC); and 3) the active power per gate would decrease. This is "constant-field scaling," sometimes referred to as "classic scaling." In terms of the scaling parameter K, as technology is scaled to the next generation, the surface dimension decreases by 1/K, the gate density increases by K^2, the gate delay time decreases by 1/K, and the throughput by K^3, while the power density is maintained constant [2]. Typically, from one generation to the next the surface dimension is reduced by $\sqrt{2}$ ($\mathrm{K} = \sqrt{2}$), resulting in a throughput increase of $2\sqrt{2}$.

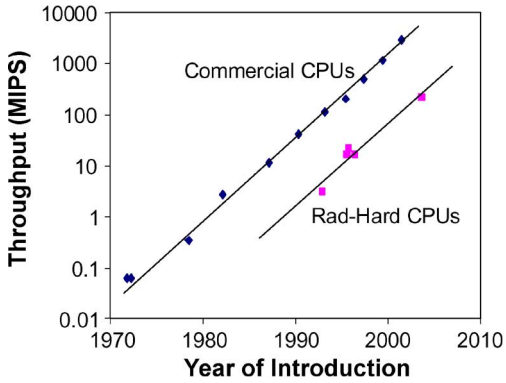

Fig. 1. Performance of commercial and rad-hard microprocessors as a function of time [3].

To use microelectronics in space, they need to be hardened against the natural radiation in space. This includes total-ionizing dose (TID) effects associated with interactions between ICs and trapped particles in the magnetic Van Allen belts and single event effects (SEEs) resulting from the interaction of highly energetic particles (including protons, galactic cosmic rays and other heavy particles). Until recently, the primary method of achieving radiation hardness in space-qualified electronics has involved the use of specialized radiation-hardened manufacturing processes developed specifically for space applications. These dedicated "rad-hard" foundries have successfully supplied hardened components for many space systems. However, the inherent added complexity of these specialty processes combined with their low-volume demand has led to an inevitable gap in scaling when the rad-hard foundry processes are compared with processes used at contemporary commercial foundries. As described above, the lag in scaling translates to a significant lag in performance. This is illustrated in Fig. 1 where microprocessor throughput (in MIPS) as a function of year-of-introduction is graphed for microprocessors fabricated at rad-hard foundries and commercial foundries [3]. As can be observed, there is a lag of approximately eight

Manuscript received October 31, 2007; revised April 14, 2008. Current version published September 19, 2008.

The author is with The Aerospace Corporation, Los Angeles, CA, USA (e-mail: Ronald.C.Lacoe@aero.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNS.2008.2000480

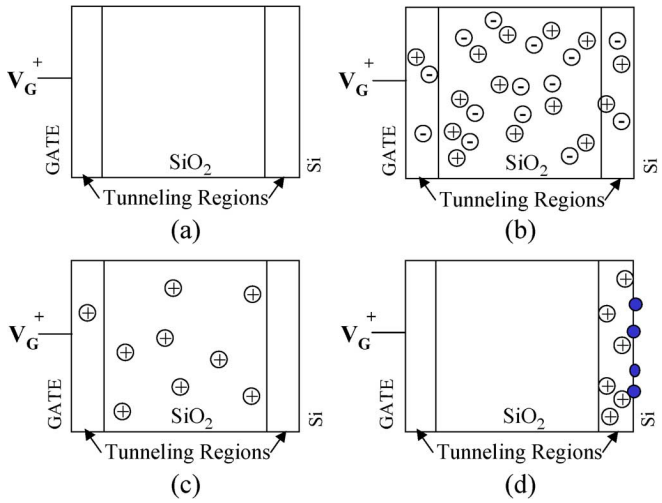

Fig. 2. Charge distribution in a gate oxide at three times after exposure to a pulse of irradiation ($t = 0$) for a thick gate oxide, as described in the text. (a) $t = 0^{-}$, (b) $t = 0^{+}$, (c) $t = 0^{++}$, and (d) $t \gg 0^{++}$.

years, or three CMOS generations, between the two. Clearly, if an approach to producing rad-hard components that has less of a technology penalty can be developed, the impact would be to allow space systems to be designed with higher performance. Hardness-by-design (HBD) is an approach that has been developed over the last decade that directly leverages the commercial microelectronics infrastructure and offers the potential of added performance with respect to the "rad-hard foundries." Hardness-by-design mitigates radiation effects using innovative design and layout techniques at the transistor level, the component level and the system level to assure performance and radiation-hardness requirements are met. The fabrication of HBD components is primarily at commercial microelectronics foundries using standard commercial processes and process flow, although HBD techniques are also used at the rad-hard foundries. In this paper, HBD techniques for assuring TID and SEE hardness, the penalties associated with these approaches, and strategies for employing HBD methodology will be discussed.

II. Toward Increased TID Hardness in Commercial CMOS Processes

The effect of TID radiation on CMOS devices primarily involves charging in the oxides and the resultant effects of this charging. This section will discuss the effects of total ionizing dose radiation on: 1) gate-oxides, which can result in the degradation of a number of transistor performance parameters; 2) transistor edges, which can result in increased intradevice edge leakage; and 3) isolation oxides, which can result in the loss of interdevice isolation.

A. Gate-Oxide Effects

The effects of TID irradiation on the gate oxide of a MOS transistor biased positively at the gate electrode can be described as a four-step process, as shown in Fig. 2 (see [4, ch. 3 and references within], [5], [6, and references within], [7], and [8]). In the first step, the ionizing radiation creates electron-hole (e-h) pairs. When an energetic particle (proton, electron, or heavy ion) impacts a solid, the particle loses energy at nearly a constant rate as it passes through the solid, as long as it is not near the end of its range. The mechanism by which electrons and protons lose energy is primarily by inelastic Coulomb scattering in which the incident particle ejects an outer shell electron from an atom [4]. The incident particle repeats this scattering process as it continues through the solid, producing a line of e-h pairs. The rate at which a particle loses energy normalized by the density of the material in which the energy is deposited is referred to as the particle linear energy transfer (LET) rate. The elements most sensitive to total dose effects from ionizing radiation in a CMOS device are the gate and isolation oxides, which are most often fabricated with SiO_2.

Immediately after the creation of a line of e-h pairs, a fraction of the pairs recombine. The temporal window during which recombination can occur is very short and limited by the time for an electron (which has high mobility in SiO_2) to transit and be removed from the gate oxide in response to the gate electric field (typically less than 0.1 ps). Even with no voltage applied to the gate, recombination will be completed in at most a few picoseconds due to the built-in gate electric field. During this window where recombination can occur, the amount of recombination depends on the density of e-h pairs as well as the applied electric field. For high LET particles, the density of e-h pairs is high and the e-h pairs interact, the Coulomb interaction between an isolated e-h pair is effectively screened out, and the fraction of e-h pairs that recombine can be high. For lower LET particles, the density of e-h pairs is less, the interaction between different e-h pairs is negligible, and the Coulomb interaction between the isolated electron and hole of a pair dominates the recombination process. The effect of an electric field on the recombination processes is to separate the electrons and holes, which will result in less recombination. Furthermore, both electrons and holes that are within a characteristic tunneling length of SiO_2, 4–5 nm, tunnel out of the gate oxide rapidly after a radiation pulse.

At this point, the only charges left in the gate oxide are holes. The second step is the drift of the remaining holes to the oxide/silicon interface under the applied electric field from the positively biased gate electrode. This process has been studied in depth, and it has been determined that the transport by holes is primarily by hopping between localized sites in the gate oxide [4], [10], [11]. This process is both thermally and field-activated [12]. The hole transport is highly dispersive in time, occurring over many decades in time after a radiation pulse, consistent with a wide dispersion in transit times of the holes through the oxide [4]. The transport of these holes to the interface is believed to occur via polaron hopping where the holes become self trapped by their own deformation potential [4], [5].

In the third step, a fraction of the holes are trapped at the interface. This "oxide-trapped charge" is positively charged, and can be neutralized over time from either electron tunneling from the silicon or by thermal emission of an electron from the oxide valence band [12]. Because the activation energy for this process is relatively low, significant recovery can occur for radiation exposure over a long time period such as for a space mission [13].

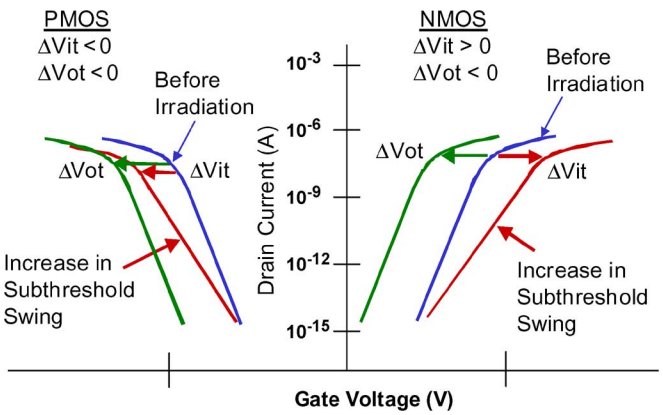

Fig. 3. Threshold voltage shifts and subthreshold swing changes for NMOS and PMOS transistors are shown relative to the preirradiation curves.

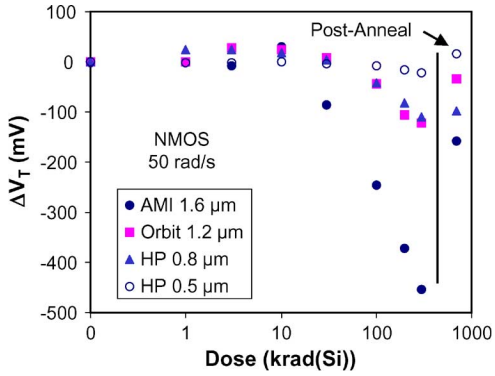

Fig. 4. Change in threshold voltage as a function of total dose for minimum geometry NMOS transistors for four different commercial CMOS processes [14].

In the fourth step, the formation of "interstate traps" can occur. A simplistic view is that interface states are associated with dangling bonds between the silicon and the SiO_2. Interface states exist within the silicon band gap at the interface with SiO_2. For NMOS transistors, the interface states act as negative charges in the gate-oxide of a NMOS transistor, or positive charges in the gate-oxide of a PMOS transistor. These steps are described pictorially in Fig. 2. It should again be noted that this description of total dose effects in CMOS devices is simplistic in nature, and does not address many of the subtleties of these processes. For more detailed presentations on the issues discussed in this section, see, for example, [4]–[8].

The introduction of these new charge sources, can affect device performance. The effect of introducing charge in the gate oxide and/or at the gate-oxide/silicon interface is to shift the CMOS transistor threshold voltage ΔV_T. The shift in threshold voltage is determined by integrating the weighted additional charge density (ρ) over the oxide thickness $(t_{\rm ox})$

$$\Delta V_{\rm ot,it} = \frac{-1}{C_{\rm ox} t_{\rm ox}} \int_0^{t_{\rm ox}} x\rho(x)dx \qquad (1)$$

where $x = 0$ is the gate polysilicon/gate oxide interface and $C_{\rm ox}$ is the gate capacitance. The radiation-induced trapped-hole charge (ot) is always positive. Hence, $\Delta V_{\rm ot}$ for both NMOS and PMOS transistor types are always negative. For NMOS transistors, interface-state charge (it) is negative, while for PMOS transistors, interface-state charge is positive. As a result, $\Delta V_{\rm it}$ is positive for NMOS transistors, and negative for PMOS transistors. The net threshold voltage shift ΔV_T will be the sum of $\Delta V_{\rm ot}$ and $\Delta V_{\rm it}$. The effect of oxide- and interface-trapped charge on CMOS device characteristics is shown in Fig. 3.

The trapped-hole spatial charge distribution depends on many factors, including the polarity and magnitude of the applied gate voltage, the gate-oxide processing technology, and the amount of hole neutralization that has occurred. As shown in (1), the closer the holes are to the gate/silicon interface, the larger the associated threshold voltage shift will be. It follows that a positive gate bias, that pushes the holes toward the gate/silicon interface, will have a larger $\Delta V_{\rm ot}$ than a negative gate bias, which pushes the same amount of hole-charge toward the poly-gate/gate-oxide interface Hence, the worst case bias condition to maximize $\Delta V_{\rm ot}$ is to bias the gate positively, and to ground the source, drain, and body contacts. This represents an actual bias condition an NMOS transistor typically sees in CMOS logic gates, and hence, these bias conditions are used for the worst case testing of NMOS transistors. However, in CMOS digital logic that set of bias conditions is never applied to a PMOS transistor. In an inverter, for example, when the gate of the PMOS transistor is biased high, the drain is also biased high. In addition, the body contact for a PMOS device is always biased positively. Hence, a more realistic bias during radiation testing for PMOS transistors that maximizes both trapped-oxide and interface-state charge is to have the gate, source and body biased high, while the drain is grounded.

There has been a trend in commercial CMOS processes toward increased total-dose hardness over the last ten years [7], [8]. We will first examine the effect of total-dose radiation on gate-oxide threshold voltage shifts in commercial CMOS processes as the transistor critical dimension continues to shrink to values below 100 nm. Fig. 4 shows the change in the measured total threshold voltage shift ΔV_T as a function of total dose for minimum geometry NMOS transistors processed at four different CMOS foundries [14]. These processes are an AMI 1.6-μm process, an Orbit 1.2-μm process, an HP 0.8-μm process, and an HP 0.5-μm process. As can be observed in Fig. 4, ΔV_T increases with increasing total dose and exceeds 100 mV at 300 krad for all but the HP 0.5-μm process. In fact, for most applications, these processes would become unusable when radiation-induced threshold voltage shifts approach the range of 50–100 mV in value. Hence, all processes except for the HP 0.5-μm process would become unusable after only low exposure to radiation. The negative values of ΔV_T for the NMOS transistors indicates that V_T shifts are dominated by trapped-oxide charge (holes). In addition, Fig. 4 indicates that for a given total dose, the older the technology node the larger ΔV_T. This can be explained in terms of the dependence of ΔV_T on oxide thickness. For a CMOS transistor,

$$\Delta V_T \propto \frac{\Delta Q_T}{C_{\rm ox}} \propto t_{\rm ox}^2 \qquad (2)$$

where ΔQ_T is the total oxide trapped charge $(\Delta Q_T = \Delta Q_{\rm ot} + \Delta Q_{\rm it})$ and is proportional to $t_{\rm ox}$, and $C_{\rm ox}$ is inversely

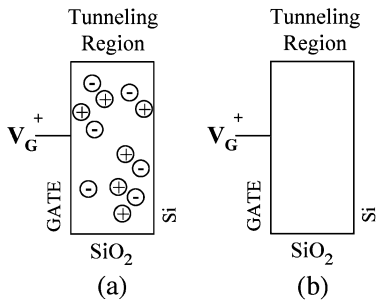

Fig. 5. Charge distribution in a gate oxide after exposure to irradiation for a thin gate oxide biased positively at the gate electrode (a) immediately after irradiation and (b) a short time later.

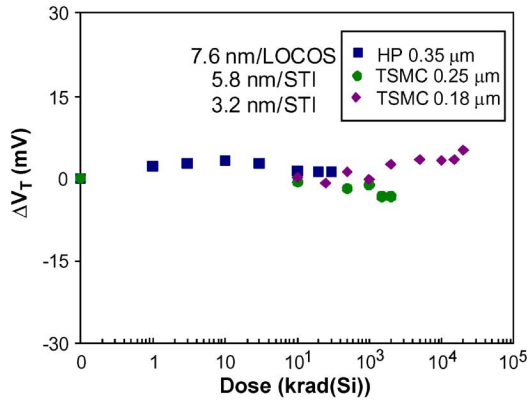

Fig. 6. Measured threshold voltage shifts as a function of total dose for NMOS transistors from three different commercial CMOS processes [3], [8], [15], [16].

proportional to t_{ox}. As technology scales, the vertical dimension scales linearly with the scaling parameter, indicating the oxide thickness scales quasi-linearly with the channel length. The oxide thickness for the AMI 1.6-μm, the Orbit 1.2-μm, the HP 0.8-μm, and the HP 0.5-μm processes are 32, 23, 17, and 9.4 nm, respectively. When the threshold voltages are normalized by the square of the oxide thickness for a given total dose, the values are very tightly grouped, indicating that the decrease in the magnitude of the threshold voltage shift with increased technology scaling is dominated by the quadratic dependence on oxide thickness [15]. Other possible factors, such as oxide quality variances, appear to play only a minor role in the processes discussed here.

The above analysis might suggest that as technology scaling continues indefinitely, threshold voltage shifts would continue to shrink to lower and lower values as the square of the oxide thickness, but that is not the case! The gate oxide thicknesses of the four processes described above are large compared to the characteristic tunneling length of $\sim$5 nm for holes within SiO_2. As shown in Fig. 5, when the oxide thickness is comparable to or smaller than the characteristic tunneling length, the radiation-induced hole charge tunnels out of the oxide and the probability that the hole can be trapped as oxide-trapped charge or contribute to the formation of interface states becomes very small. The result is that ΔV_T is expected to be very small, even after exposure to high total dose values.

Fig. 6 presents data on the measured threshold voltage shift as a function of total dose for three different commercial CMOS processes [3], [8], [15], [16]. As seen in Fig. 6, the ΔV_T does not exceed $\sim$5 mV up to 30 Mrad, consistent with no charge trapping in the oxide in advanced CMOS technologies with oxide thickness less than the intrinsic carrier tunneling length in SiO_2. Further evidence of the lack of radiation-induced charge trapping in thin gate oxides can be seen in the behavior of the transconductance and the subthreshold swing as a function of radiation dose, both of which depend on the density of interface states. Recent results for thin gate-oxide technologies indicate neither of these parameters for thin gate-oxide technologies shows changes after exposure to megarad levels of radiation [7], [14], [17]. It is clear from these data that negligible levels of charge trapping, both in the form of oxide and interface trapped charge, occur in advanced CMOS processes with oxide thickness less than the characteristic tunneling length for carriers in SiO_2.

While the decrease in gate-oxide thickness has resulted in negligible radiation-induced charge trapping, it also results in increased gate-substrate tunneling current. This increased tunneling current results in increased FET off-state current and an associated increase in quiescent power [18]. Since this additional current depends exponentially on the gate dielectric physical thickness, this is becoming more of an issue as technology scales below 100 nm. The semiconductor industry is addressing this issue by researching potential new materials to replace SiO_2 as the gate dielectric material. The key property they are looking for is materials with dielectric constant (K) values larger than that for SiO_2. These new potential gate dielectric materials are often referred to as high-K dielectrics. Because the dielectric constant is larger than for SiO_2, the thickness can be increased to maintain the same capacitance while decreasing the tunneling current. Among the materials under consideration are Hf-O/SiO_2/Si and Al_2O_3/SiO_xN_y/Si gate stacks [19]–[21]. Recent results on the effect of TID on capacitors and transistors fabricated with novel high-K gate dielectrics indicate that while for thick oxides they can have high trapping efficiency, for test structures with dielectric thickness less than $\sim$5 nm, the amount of charge trapping is very small, resulting in very small ΔV_T shifts [22]–[24]. Since the required high-K dielectric thickness in advanced CMOS technologies is likely to be less than 5 nm, the results look encouraging with respect to using these newer technologies for space applications.

The interaction of ionizing radiation with CMOS devices has been reviewed and the potential for trapping charge in the form of oxide trapped charge and interface states is a serious concern. It was shown that for commercial processes with thick oxides, significant charge trapping occurs, while for commercial processes with oxides near or below the characteristic tunneling length for carriers in SiO_2, the effective radiation-induced charge trapping approaches zero. This transition into a regime where gate-oxide threshold voltage shifts are no longer an issue is the primary reason that HBD has become a realistic option over the last few years. As you will see later in this article, there are design techniques that can mitigate all other radiation effects. There is, however, no easy way to design around gate-oxide threshold voltage shifts. Since this appears to no

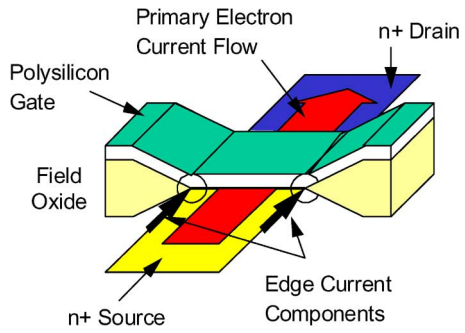

Fig. 7. Off-axis view of edge transistor current in parallel with the main transistor current.

longer be an issue for advanced CMOS technologies, HBD now becomes a realistic option.

B. Radiation-Induced Edge Leakage

Transistors are electrically isolated from one another by a thick field oxide (FOX). The edges for a standard transistor layout are defined by where the gate oxide interfaces with the FOX. Fig. 7 shows an off-axis view of a standard MOS transistor. The combination of the gate poly, the relatively thick oxide in the transition region and the p+ diffusion under-layer form a parasitic "edge" transistor, which can trap holes after exposure to TID radiation. Since the quality of the oxide near the FOX/gate oxide interface is likely to be poorer than that of the gate oxide, it may be more efficient in trapping charge. Radiation-induced trapped holes in these edge transistors can result in a negative shift in the threshold voltage large enough so that the edge transistors invert becoming conductive resulting in a source-drain current in the transistor off state ($V_G = 0\,V$). The end result is that the off-state current can increase due to radiation-induced edge effects. The consequences of increased transistor off-state current in circuits can be signal corruption and reduced margins, and for high transistor count circuits the total supply current can rapidly exceed component specifications. In some cases this can lead to functional failure, as the transistors behave as if they are always on. Edge leakage is not a problem for PMOS transistors, since for PMOS transistors the effect of oxide-trapped charge and interface trapped charge is to shift an already negative threshold voltage even further in the negative direction.

Until recently, edge leakage in a commercial CMOS process could exceed 1 μA for total dose levels as low as 10 krad(Si). Clearly, these technologies have little potential utility for space applications. As commercial CMOS technology has advanced, the inherent radiation-hardness has greatly increased.

The I_D–V_G characteristics as a function of total dose for a 0.35-μm minimum geometry NMOS transistor fabricated at Chartered Semiconductor are shown in Fig. 8 [7], [25]. This process uses LOCOS isolation and has a gate-oxide thickness of 7.6 nm. For exposure up to 50 krad, the I_D–V_G characteristics are unchanged. At 70 krad, there is an increase in the off-state current to $\sim 10^{-11}$ A. At 100 krad the off-state current has increased to a value slightly larger than 1 nA. At 300 krad, the off-state current has increased to within a factor of $\sim 5\times$ the drive current at $V_G = 3.3\,V$. The effect of a 100 °C/168 hr anneal (curve PA in Fig. 8) is to restore the I_D–V_G characteristics to those of the preirradiated device, indicating that the edge leakage is associated primarily with trapped-oxide charge in the parasitic edge transistors. The postanneal characteristics also show a small increase in the subthreshold swing, consistent with some small degree of interface state creation. The insert in Fig. 8 allows extrapolation of the radiation data to determine the total dose exposure level at which the NMOS transistor off-state current exceeds 1 nA (an arbitrary failure criteria). For this 0.35-μm process, the total-dose hardness level is ~92 krad under high-dose rate, worst case bias conditions, high enough to meet the requirements of many space programs.

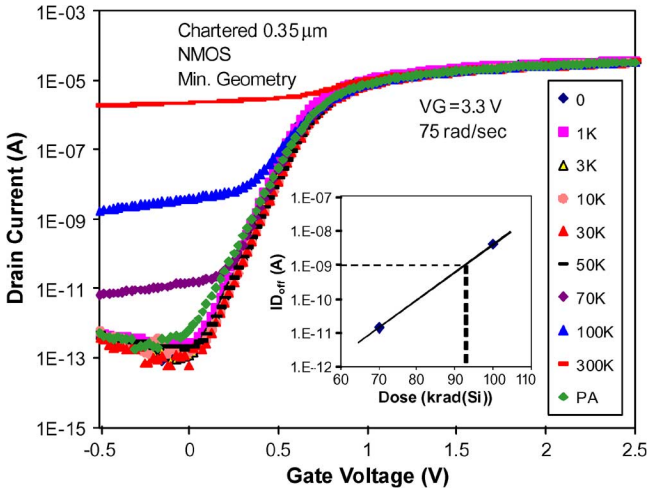

Fig. 8. I_D–V_G curves as a function of total dose for a 0.35-μm minimum geometry NMOS transistor. Insert is graph of I_D (0 V) versus dose [7], [25].

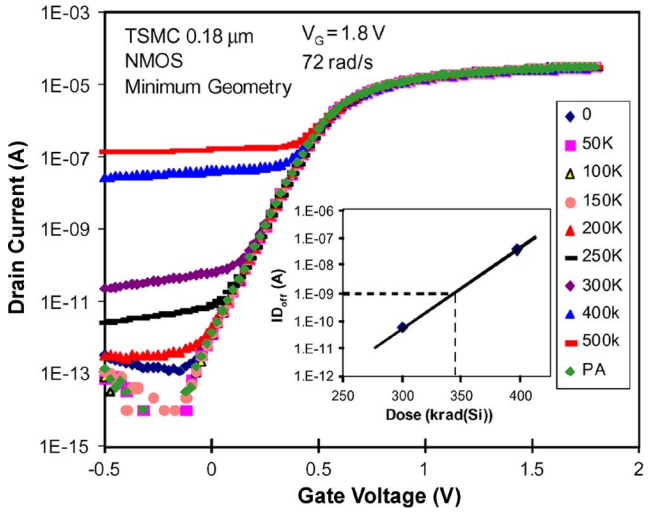

Fig. 9. I_D–V_G curves as a function of total dose for a 0.18-μm minimum geometry NMOS transistor. Insert is graph of I_D (0 V) versus dose [3], [7].

The I_D–V_G characteristics as a function of total dose for a 0.18-μm minimum geometry NMOS transistor fabricated at TSMC are shown in Fig. 9 [3], [7]. This process uses shallow-trench isolation (STI) technology and has a gate-oxide thickness of 3.2 nm. The behavior is similar to that of the 0.35-μm process, except for this 0.18-μm process, the total-dose hardness level is ~345 krad under high-dose rate, worst case bias

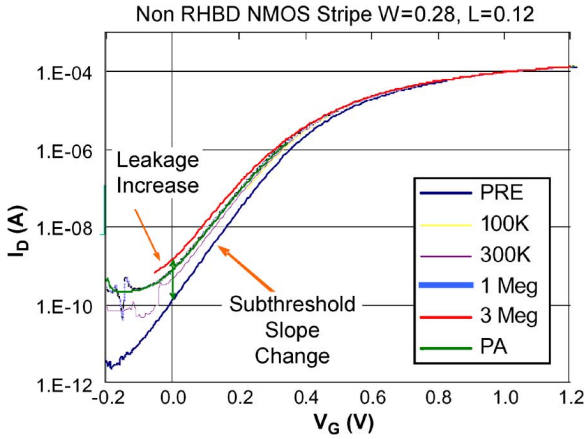

Fig. 10. I_D–V_G curves as a function of total dose for a 130-nm minimum geometry N MOS transistor [26].

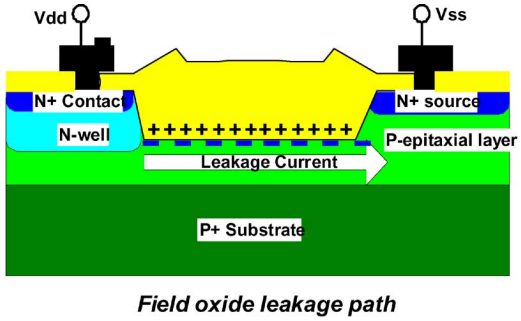

Fig. 11. Radiation-induced hole trapping in thick isolation field oxides can drive the parasitic field oxide transistor into inversion, resulting in leakage between adjacent devices and a lack of device isolation.

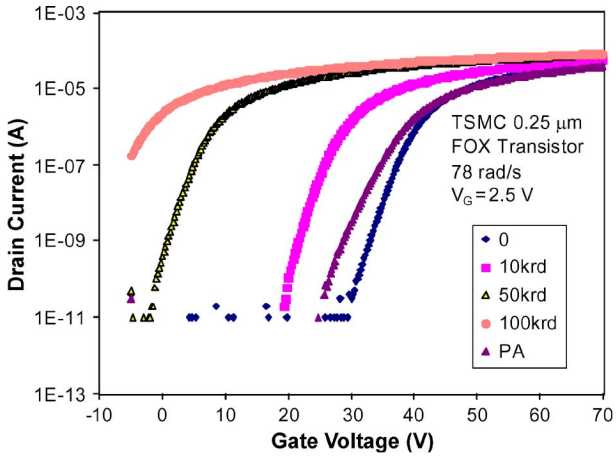

Fig. 12. I_D–V_G curves as a function of total dose for FOX transistor. The transistor was exposed to 500 krad before the postanneal process [3].

conditions. This level of intrinsic TID hardness would meet the TID requirements of many space programs.

Recently, the radiation hardness of advanced commercial CMOS processes with 130-nm nominal gate lengths was characterized with respect to the effects of TID. The I_D–V_G characteristics for minimum geometry NMOS transistors fabricated at a 130-nm commercial foundry are shown in Fig. 10, and are different than those observed for less advanced technology nodes. One significant difference is that the effect of total dose radiation on the zero-volt leakage current is much less for the advanced technology nodes. The zero-volt leakage current increased by a factor of $\sim 8\times$ at 1 Mrad for the 130-nm process, while for the 180-nm process it increased by $\sim 10^6$. While some of the difference can be explained by the fact that the advanced processes start out with increased off-state leakage current, the differences by at least 10^3 are still present and represent a different behavior than previously observed. Another difference can be seen in the behavior of the subthreshold swing, where the advanced processes show a slight increase with increasing radiation dose, while the less advanced processes showed no change. While more study is required to fully understand these new results, the concept of viewing the transistor as consisting of an intrinsic device and two parasitic edge devices may no longer be applicable. For these narrow devices, the intrinsic transistors and the edges may become inseparable, and charge at the edges can affect the behavior of the intrinsic device. This explanation is consistent with recent results where a width dependence in the radiation-induced V_T shifts is observed for both NMOS and PMOS transistors [27]. The authors explain this width dependence as a manifestation of the "narrow channel effect" when radiation-induced charge is introduced at the isolation edges.

C. Loss of Interdevice Isolation

As discussed previously, the isolation oxide is much thicker than the gate oxide. If exposed to an electric field during irradiation, a considerable number of holes can become trapped at the isolation oxide/silicon layer. Whether there is an electric field over this parasitic field-oxide transistor or not depends on whether a conducting line (metal or poly) happens to be routed over the FOX. Assuming a positive bias is present during exposure to radiation, the effect of radiation on the parasitic field-oxide transistor will be to create negative V_T shifts that drive the transistor into inversion, and can conduct even when no voltage is present on the metal line. In this case, as shown in Fig. 11, there is a leakage path from the n-well contact through the n-well along the interface of the isolation oxide and the low resistance epitaxial layer to the source of the NMOS transistor. This interdevice leakage results in a lack of device isolation. There is another leakage path between the n+ source/drain regions of adjacent NMOS transistors. As with edge leakage, interdevice leakage can result in signal corruption, reduced margins, and additional supply current and the associated increase in rail voltage drop.

To investigate this effect, special FOX transistors were fabricated with a polysilicon gate over STI field oxide between adjacent n-wells, which were contacted as the source/drain of the field-oxide transistors for a 0.25-μm commercial CMOS process. The I_D–V_G characteristics as a function of total dose for field oxide transistors are shown in Fig. 12 [3]. Prior to irradiation, the V_T of the FOX transistor was ~42 V. Exposure to radiation shifts the V_T negatively. At 100 krad, the V_T has nearly crossed zero $(V_T = 0.5\,V)$. The FOX transistor was further irradiated to 500 krad in 100 krad steps. These results are not shown here for simplicity. The data shows that the threshold voltage shifts have nearly saturated at 400 krad. Following irradiation,

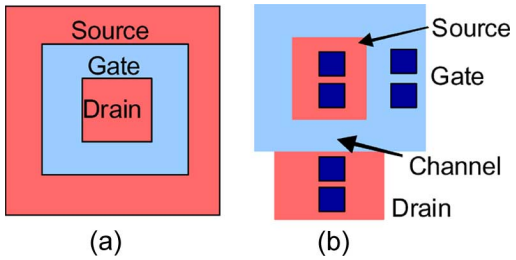

Fig. 13. (a) Edgeless and (b) an enclosed-source transistor.

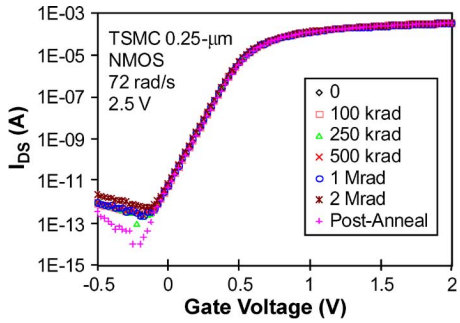

Fig. 14. I_D–V_G curves as a function of total-dose for an edgeless transistor [3].

the FOX transistor was annealed at 100 °C/168 hr (curve PA in Fig. 12) and the threshold rebounded to 41.5 V, nearly the preirradiation value. After the post anneal, an increase in the subthreshold swing is observed, consistent with the formation of interface states, which are partially responsible for the V_T rebound. There are certainly many variables that could potentially affect STI hardness, such as the geometry, STI profile and the interface from gate oxide to STI. It should be noted, however, that for the low dose rates in space, the formation of interface traps may prevent the inversion of the FOX for any radiation dose.

D. Hardness-by-Design Techniques for Mitigating Total-Dose Effects

Hardness-by-design is a new approach that is being developed to enhance the radiation hardness of microelectronic components without using special radiation hardening processing techniques. Through a combination of the application of specific design techniques and the leveraging of the intrinsic radiation hardness of leading edge commercial CMOS processes, it is now possible to fabricate radiation-hardened components for many military and space applications using standard CMOS process flows. This section will present HBD approaches to mitigate total-dose effects.

Novel transistor design topologies can be used to eliminate radiation-induced edge leakage. One such transistor topology is shown Fig. 13(a). This transistor layout is often referred to as an edgeless transistor, and has no active diffusion edges overlapped by polysilicon that separates the source and the drain. That is to say, this transistor has no edges, and hence, there is no edge leakage. The efficacy of the edgeless transistor in mitigating radiation-induced edge leakage is illustrated in Fig. 14 [3]. Here, the I_D–V_G curves as a function of total dose for a 0.25-μm edgeless NMOS transistor shows that there is no significant increase in the off-state current up to a total dose of 2 Mrad.

While edgeless transistors are highly effective in mitigating NMOS transistor edge leakage, this does not come without costs. Perhaps the most important cost is in transistor area. The minimum transistor W/L ratio for a standard-edged transistor is $\sim 4\lambda/2\lambda = 2$, assuming generic design rules where the minimum source/drain size is $4\lambda \times 4\lambda$ (associated with the minimum size to open up a contact hole) and the minimum channel length is 2λ. An edgeless transistor can be viewed as four separate trapezoidal transistors, each with an inner width of 4λ and an outer width of 8λ [7]. Using a highly simplified model it can be shown that the effective W/L ratio for an edgeless transistor is $\sim 24\lambda/2\lambda = 12$. The footprint of the minimum standard-edged transistor is $80\lambda^2$, while the area for the edgeless transistor is $256\lambda^2$, resulting in an area increase by a factor of $256\lambda^2/80\lambda^2 = 3.2$. In actuality, the area penalty is much less than this factor, since most digital logic applications do not use minimum geometry NMOS transistors, but rather use transistors with larger W/L ratios to increase transistor drive currents and increase switching speed.

The model used above is oversimplified. In reality, a rectangular edgeless transistor is not characterized by a single channel length. The effective channel length and the corresponding electric field for carriers traversing the device near the corners is variable and is different from that of carriers traversing the device away from corners. Furthermore, good design practices avoid sharp corners where the electric fields can become large and result in undesirable reliability consequences. All these factors make the modeling of an edgeless transistor a difficult challenge. Another factor to consider in the application of edgeless transistors to mitigate edge leakage is that it requires the development of a custom cell library and corresponding spice models.

When compared to standard-edged transistors, edgeless transistors have increased gate and source/drain capacitances [3], [28], [29]. Furthermore, edgeless transistors are not symmetrical devices. The circuit designer has the choice of contacting the source and the drain to the inner and outer ring of the gate, respectively, or vice versa. Choosing the inner ring as the drain minimizes the drain area, and hence the drain to substrate capacitance. This approach is preferred for maximizing performance. It should be noted, however, that there is a trade-off between maximizing performance and maximizing reliability that is associated with the choice of the location of the source/drain contact [3], [30]. Furthermore, there is an asymmetry in the output characteristic of edgeless transistors that is associated with the nonsymmetrical geometry of the device. The output conductance ($\partial I_{DS}/\partial V_{DS}$ at constant V_{GS}) for an edgeless transistor with the drain on the inside (g_{di}) is greater than that for the same transistor with the drain on the outside $(g_{\rm do})$. The fact that $g_{\rm do}$ is lower than $g_{\rm di}$ can be explained as follows: the distance between the pinch-off point and the drain, due to the conservation of space charge region for the same bias, will be smaller when the drain is outside. For this case, an increase in V_{DS} potentials will result in less of an increase in drain current, resulting in a lower $g_{\rm do}$.

A second class of closed geometry transistors that has been shown to be effective in eliminating edge leakage effects is the

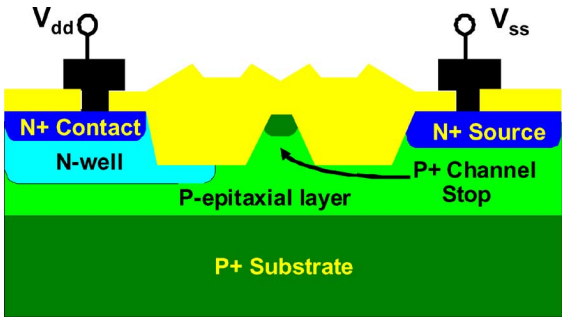

Fig. 15. Cross-section of a CMOS process with a p+ channel stop designed into the field-oxide isolation to mitigate interdevice isolation.

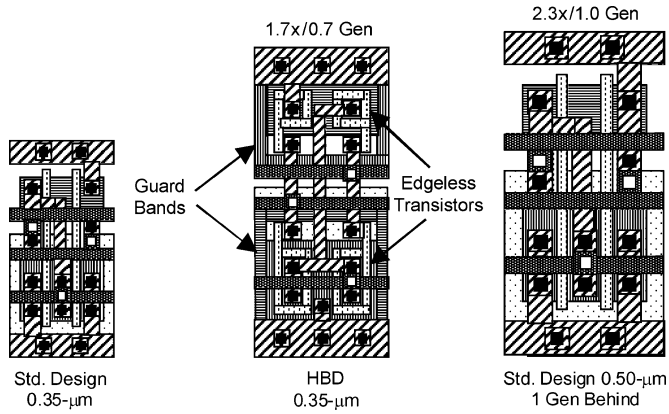

Fig. 16. Area comparison for two-input NAND gates designed for a 0.35-μm process using standard-edged and edgeless transistors and for a one-generation behind 0.50-μm process using standard-edged transistors [3].

enclosed-source/enclosed-drain transistor, shown in Fig. 13(b) and [29].[1] In this case, rather than being radial, the channel is more nearly transverse as in a standard-edged transistor. The remainder of this paragraph closely follows [29]. For this design, the polysilicon overlaps the active diffusion around the source/drain to eliminate the radiation-induced source-to-drain leakage path. While this technique does not eliminate the transistor edges, it inserts an insulator between the enclosed contact and the parasitic edge transistor and removes the low resistance leakage path between the source and the drain. This insulator is polysilicon over gate oxide, which does not invert with exposure to radiation. For these transistors, the W/L ratio can be maintained closer to that found for standard transistors. There is, however, a much larger gate-to-source/gate-to-drain capacitance due to the excess polysilicon that encloses the source/drain. The excess polysilicon can also have significant resistance associated with it, especially in large devices. The designer has the option of enclosing either the source or the drain, and sometimes for compactness the designer will alternate between enclosed sources and enclosed drains. When the drain is enclosed, the parasitic gate-to-drain capacitance of the excess polysilicon appears as a Miller capacitance (excess parasitic capacitance multiplied by the gain of the transistor) when referred to the input. The combination of Miller capacitance and excess polysilicon resistance in large transistors can introduce a significant RC delay into a circuit. For this reason, the usual choice is to enclose the source. A further consideration for advanced CMOS technologies is the gate leakage associated with enclosed poly layer separated from the silicon by a thin gate oxide. This additional tunneling current can become a power issue for certain applications. Enclosed source/drain transistors are probably best suited for small width devices, while for larger width devices edgeless transistors may offer less performance penalties. Furthermore, radiation-induced anomalies in the subthreshold regime of the I–V characteristics can occur in wide enclosed-source transistors and are a manifestation of the existence of edges in this transistor topology [31].

A design solution to mitigate radiation-induced interdevice leakage is to surround each transistor with a p+ diffusion ring, as shown in Fig. 15. This can be performed as part of a standard CMOS process flow and does not require the insertion of additional masks or processing steps. The ring performs an important dual function [3], [29], [32]. As a channel stop, the p+ diffusion prevents the inversion of the field oxide at that location by adjusting the local threshold voltage to a very high value. This maintains the integrity of the isolation between adjacent NMOS transistors and eliminates the n-well to n+-source leakage path. When using a p+ channel stop, the polysilicon gate cannot be allowed to cross the p+ ring, which can block the p+ implant and create a gap in the channel stop. This gap would provide a potential leakage path. Because of this restriction, a local polysilicon interconnect cannot be used in signal routing when channel stops are employed. Furthermore, there will be an area penalty for using p+ channel stops. Since the p+ diffusion ring must surround a transistor, the area penalty will depend on the design of the transistor (standard-edged, enclosed-source or edgeless transistors) that is being surrounded, as well as the number of transistors enclosed by a single ring.

To evaluate the effect of the application of total-dose HBD techniques on integrated circuit area, power and speed, the effects of HBD techniques were evaluated for a two-input NAND (2NAND) gate [3]. Two-input NAND gates designed with standard-edged and edgeless transistors targeting a 0.35-μm process and standard-edged transistors targeting a one-generation behind 0.5-μm process are shown in Fig. 16 [3]. The minimum size of the HBD cell was fixed by the minimum area required for an edgeless NMOS transistor. The PMOS transistor was scaled to keep the W/L ratio close to twice the NMOS ratio. Channel stops were used around both NMOS and PMOS devices. The area penalty for the HBD layout was found to be 1.7×, which is less than one generation. Similar results were found for a two-input NOR gate [29].

The effects of the application of HBD techniques on other performance characteristics were also evaluated in this study [3]. The performance of each technology was estimated assuming a simple long-channel approximation and assuming that the output was loaded by a capacitance equivalent to a fanout of two inputs, plus a fixed additional interconnect capacitance. The derived propagation delay of the HBD cell (0.09 ns) is slightly faster than that of the same cell built with standard-edged transistors (0.10 ns) because of the offsetting effects of increased drive current and increased capacitances

[1]Original concept by D. Mavis

from the edgeless transistors. The propagation delay of the HBD cell is more than 50% faster than that for the cell designed for a 0.5-μm technology (0.14 ns) because of the increased drive current associated with a minimum W/L ratio for an edgeless transistor of at least 12. The increased drive current will make HBD design faster than the 0.5-μm standard-edged design if the interconnect capacitance is large, but slower if the interconnect capacitance is small compared to the output and gate capacitances. Because of the larger capacitances associated with the edgeless layout, the power dissipation of the HBD cell is higher than for the standard-edged cell in the same technology, but is still less than the power dissipation for the cell designed using one-generation-behind technology. This assessment of the speed and power performance of each technology was made using a number of simplifying assumptions. A more careful extraction of the capacitances associated with the transistor nodes and interconnects and an assessment of the impact of the floating node capacitance on circuit performance is required for an accurate circuit timing assessment. A number of second order effects associated with the edgeless transistors have also not been considered, but may have a significant impact on the performance of the HBD technology. These include the asymmetrical behavior of the edgeless devices when source and drain are interchanged, the impact of the corners of the gates on the effective W/L ratio, and the impact of Ohmic losses on the distribution of current around the periphery of the edgeless transistors. Analysis of cell speed does not consider the fact that when designing a circuit using HBD techniques, the circuit size will grow, which will increase the number and length of speed-limiting global interconnects.

The analysis above is for relatively mature CMOS technologies. Recently, a program to investigate the application of HBD to a commercial 130-nm technology developed a basic cell library consisting of 47 cells [26]. These cells were designed using standard-edged transistors (soft), dogbone transistors, and annular edgeless transistors (hardest). In addition, the same cell library was designed for a 180-nm commercial process to be used as a reference one-generation behind technology. The area penalties were determined from the design layout, while the delay and power penalties were evaluated through simulation for each cell. The averages for these data were calculated and are shown in Fig. 17. These results indicate the hardest cells have an average area penalty of 1.1 generations, while the average delay penalty is 0.3 generations and the average power penalty is one generation. The cell libraries targeting the 130-nm process were fabricated and the delay and power penalties were measured and validated the simulation results. These results are consistent with earlier penalty evaluations and suggest that from a total-dose perspective, the HBD penalties are approximately one generation or less. It should be pointed out that if a given parameter is deemed critical (e.g., power), the cell library can be optimized for that parameter at the expense of increasing the penalties in the other parameters. Finally, it should also be noted here that dogbone transistors have not been discussed in this paper because it has been found that they are ineffective at suppressing radiation-induced edge leakage.

Using edgeless or enclosed transistors requires the development of a custom cell library and the development of relevant

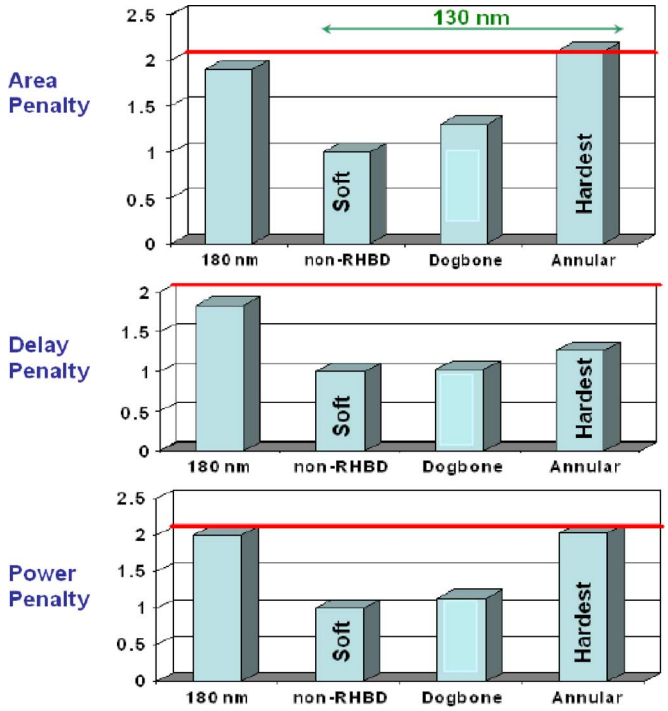

Fig. 17. Effects of the application of HBD techniques to mitigate total-dose effects for a cell library for a commercial 130-nm CMOS process averaged over all 47 cells [26].

SPICE models for these nonstandard transistors. It is critical to note that all of these design approaches can be implemented as part of a standard process flow, and hence, can be used at commercial foundries without violating design rules.

III. Single-Event Effects

A. Single-Event Upsets—Basic Mechanisms

A single-event upset occurs in a digital circuit when a particle strike causes data to change states in a storage element such as a flip-flop, latch, or memory bit [33]. The energetic particles that are the cause of SEUs in integrated circuits can arise from many sources, including galactic cosmic rays, solar protons, trapped protons in the earth's radiation belts, or trace radioactive materials in the package or board. Direct ionization is the primary charge deposition mechanism for upsets caused by heavy ions, which are rather loosely defined as an ion with atomic number $z \geq 2$. Lighter particles, such as protons and neutrons, usually do not produce enough charge by direct ionization to cause upsets in memory circuits, although this can change as technologies scale and the critical charge to create an upset decreases. These particles can, however, produce upsets due to indirect mechanisms. When a high-energy proton or neutron enters a semiconductor, it may undergo an inelastic collision with a target nucleus. This can generate a nuclear reaction resulting in the emission of alpha particles or gamma particles and the recoil of a daughter nucleus, or a spallation reaction in which the target nucleus is broken into two fragments, each of which can recoil. Any of these reaction products can now deposit charge along their paths by direct ionization. Because these particles are much heavier than the original protons and

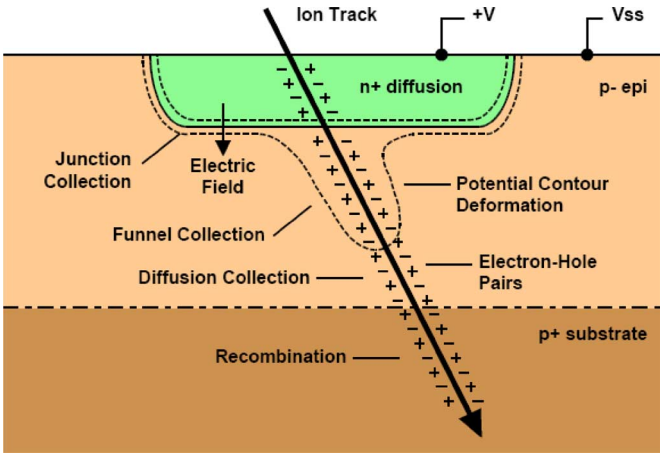

Fig. 18. When an energetic particle strikes a p-n junction, a line of electron-hole pairs are created, extending the depletion region as a funnel. This increases the charge collected at the junction [34].

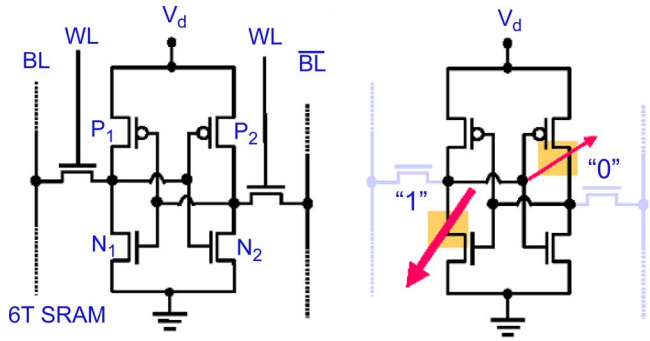

Fig. 19. Schematic diagram of cross-coupled inverters in a six-transistor CMOS SRAM. A particle strike at one storage node can cause a change in the node's stored voltage state, which can propagate to the other node, upsetting the cell [44].

neutrons, they can deposit greater linear charge density as they travel through the semiconductor, which is more likely to cause upsets than the ionization created by the original particle.

Fig. 18 shows a conceptual cross-section of an ion strike through a reverse-biased p-n junction, typical of the p-n junctions that are widespread in bulk CMOS circuits and tend to be the most sensitive sites for upset. For a particle with an LET of 100 MeV-cm^2/mg, the charge created along a silicon track per micron of track length is $\sim$1 pC. As shown in Fig. 18, a "funnel" extension of the depletion region under the junction is caused by the projection of the electrostatic potential of the heavily doped surface electrode along the "wire" of charge [35]–[37]. In response, current flows creating an effective short across the p-n junction for the period of time needed to dissipate the generated charge, typically a few nanoseconds. The initial "prompt" current flow is due to the drift component associated with the original depletion region and the funnel depletion region. The later contribution to the current flow is associated with diffusion of the charge toward the critical node. The time scale for this diffusion current can be as long as nanoseconds for processes with lightly doped substrates [38]–[40, and references within]. The impact of this current on circuit operation depends strongly on the location and angle of the particle trajectory through the circuit, and the circuit response of an inadvertent current pulse at that particular location. It should be noted that the concept of a funnel of a specific length is a misnomer but can be used as a quasi-physical "fudge factor" in error rate calculations [41]–[43].

The SRAM serves as a useful example of an SEU-sensitive component, both because of its extensive use of floating storage nodes and because of its ubiquitous presence in space systems. The heart of the device is composed of cross-coupled inverters that store the logic state and act to provide stability to the logic state under noise or other extraneous signals. Fig. 19 shows the four MOSFETs comprising the two cross-coupled inverters that form the latch that stores a single bit of information in an SRAM, along with the two access transistors at each edge of the cell. When the access transistors are deselected by driving the Word Line (WL) low, charge representing the rail voltage V_{DD} is stored at one of the floating common-drain nodes in the cell, and the other side is discharged to ground. Reversing the voltages on these two floating nodes by a write operation causes the memory state to switch from 1 to 0 or vice versa. A particle strike at one of these sensitive nodes can generate an effective wire of charge between the floating node and the V_{DD} or ground electrode in the well or substrate. If enough charge is generated along the ion track, the resulting current can discharge the charged side of the cell. It should be noted that a consideration for charge collection is whether the junction is located inside a well or in the substrate, because the well-substrate junction provides a potential barrier that prevents charge deposited deep within the substrate from diffusing back to the struck drain junction [40]. The voltage transient is essentially similar to a write pulse, and can cause the wrong memory state to be locked into the memory cell. Explained from a different point of view, the particle strike on the NMOS transistor causes a transient that tries to change the state of that inverter. The PMOS pull-up transistor tries to respond quickly enough with enough drive current to restore the original state. If this action is quick, no bit flip occurs. If not, the bit flip will stabilize through the cross-coupled inverter and an upset will have occurred. The cell feedback time is simply the time required for the distributed node voltage to feed back through the cross-coupled inverters and latch the struck device in its disturbed state. This time can be thought of as the RC delay of the inverter pair and is related to the SRAM write time. The smaller the RC delay, the faster the cell can respond to voltage transients and the more susceptible the SRAM cell will be to SEU. The higher the particle LET, the more charge collected at the drain of the sensitive node. If the particle LET is less than the LET threshold for that node, a voltage glitch occurs, but will not propagate to the opposite node. If the LET threshold is exceeded, both inverters switch, resulting in a reversal of the stored information in the bit. This results in a single-event upset or SEU. The minimum charge associated with the integrated single-event current at a critical node that can cause an upset is referred to as the critical charge (Q_{crit}).

B. Single-Event Upsets—Mitigation Approaches

A number of SRAM cell designs have been proposed to reduce the susceptibility of the cell to SEUs by the use of charge

dissipation techniques [45]. The first approach effectively increases the cell LET threshold by increasing the width of the transistors in the cell, thereby increasing their drive current capacity and their conductance. In the event of a particle strike on a critical node, the pull-up transistor will now be able to supply additional drive current to maintain the original latched logic state. Large high drive transistors also have increased node capacitance, which reduces the voltage excursion caused by the SEU injected charge ($\Delta V = \Delta Q_{\text{inj}}/C$). The application of this SEU mitigation approach results in an increase in cell area and power dissipation which is roughly proportional to the ratio of the transistor widths. There is no significant speed penalty using this approach. A technique that has been widely used by the dedicated rad-hard foundries involves the insertion of resistors between the cross-coupled inverters in the SRAM cell [40]. These resistors increase the effective RC delay impeding the propagation of the erroneous signal triggered by an ion strike, allowing the pull-up PMOS device enough time to restore the collapsed node voltage before a switch in the cell's memory state, thereby improving its immunity to SEUs. The feedback resistors can increase the delay of the cell to normal write operation. An alternative approach involves the use of excess capacitance at each of the floating nodes to increase the critical charge needed to switch cell states. While feasible, this approach is not commonly employed because it consumes excess area and slows down the read and write speeds, and because well-characterized capacitors are also not always supported by CMOS foundries.

An alternate approach to mitigating SEU effects is to reproduce the information spatially on the silicon. If each block of information is separated so that the probability of upset of two different blocks is negligibly small, then the information can be correctly reconstructed even if one block of information experiences an upset. By requiring quasi-simultaneous multiple node hits to create an upset that cannot be reconstructed, the effective SEE cross-section of the device is greatly reduced.

Single-event upset hardening by redundant circuit design approaches is based on three fundamental concepts: 1) information storage redundancy maintains a source of uncorrupted data after an SEU; 2) feedback from the noncorrupted data storage location can cause the corrupted data to recover after a particle strike, and/or maintain the correct output; and 3) the "intelligence" needed in the feedback to cause recovery of the proper location can be derived from the fact that the current induced by a particle hit flows from n-type diffusion to p-type diffusion.

The most straightforward implementation of this SEU hardening concept is triple modular redundancy (TMR) [46]. A schematic of a typical TMR cell is shown in Fig. 20. In this circuit, a single unhardened logic latch is replaced by three unhardened logic latches. Each of these latches is connected to the same clock and the same data line. The output of the three latches goes to majority voting logic. If a single energetic particle strikes one and only one of the latches and creates an upset in that latch, the other two latches will retain the correct logic state, and the majority voting logic will output the correct state. The area and power penalties associated with this approach are approximately 3×–4×. This is a brute force approach to mitigating SEU effects. There are some potential problems with this technique. If the state of the latch is not refreshed frequently in comparison to the expected upset rate, errors can accumulate such that the cells making up the TMR latch contain double errors.

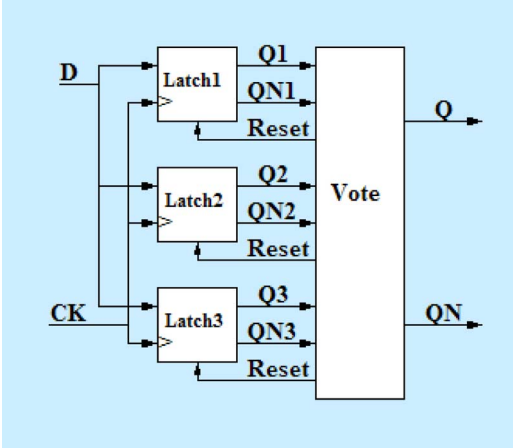

Fig. 20. Triple modular redundancy applied to latched logic.

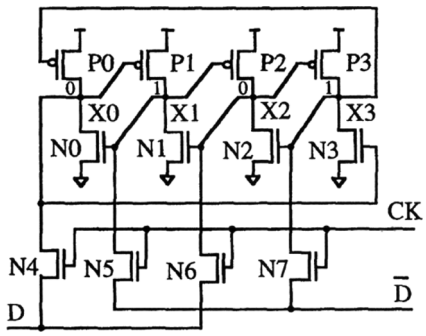

Fig. 21. Schematic diagram of the dual interlocked storage cell [47].

Triple modular redundancy is inefficient for SRAMs because of the excessive associated area and power penalties, but is very effective for individual latches. The voter circuit can be implemented using only combinational logic (i.e., no memory elements). Thus the voter can be relatively immune to SEUs. However, the voter may still be susceptible to single-event transients (SETs), as discussed below, and thus may add an additional element of susceptibility. Special care must be taken to design the voter circuit to operate reliably in the space environment.

A more elegant, area-efficient storage cell that uses a four-node redundant structure to mitigate SEUs is the dual interlocked storage cell (DICE) [47]. This cell is suitable for replacing latches and flip-flops distributed within logic blocks in CMOS in order to make them tolerant to upsets. It may also be used to implement SEU-hardened SRAMs. The schematic diagram of the DICE storage element is shown in Fig. 21 [47]. It employs two conventional cross-coupled (horizontal) inverter latch structures $N_0 - P_1$ and $N_2 - P_3$ connected by bidirectional feedback (vertical) inverters $N_1 - P_2$ and $N_3 - P_0$. The four nodes, $X_0, \ldots, X_3$, store the data as two pairs of complementary values (1010 or 0101) that are simultaneously accessed using transmission gates for the write or read operation. It uses dual node feedback control in order to achieve immunity to upsets. In order to upset a DICE cell, two nodes must be simultaneously upset: X_0 and X_2, or X_1 and X_3. For older technologies,

TABLE I
HAMMING CODE REQUIREMENTS FOR SINGLE-ERROR CORRECTION DOUBLE-ERROR DETECTION FOR DATA BLOCKS OF VARYING LENGTHS [49]

Data Bits	Check Bits	Total Bits
4	4	8
8	5	13
16	6	22
32	7	38

the probability of this occurring from a single particle can be made very low through proper layout (by physically separating the sensitive node pairs), and the area penalty as compared to a 6T SRAM cell is $\sim 2\times$. This area penalty can be larger for DICE flip-flops implemented in more advanced CMOS processes which may require proportionally greater critical node separation. Additionally, the DICE storage element requires additional drive current to write the cell because of the additional transistors used in the design.

The SEU mitigation techniques described in the previous two subsections represent attempts to prevent an SRAM/register cell from upsetting. In this section, upset mitigation techniques that allow for individual bit upsets to occur, but correct for these upsets by the application of error detection and correction (EDAC) techniques are discussed. As its name suggests, EDAC is a methodology in which extra bits are added to a block of bits to detect when one or more bits have been corrupted, and to correct the corrupted bit, if possible. For example, a single parity bit can be added to a bit string of arbitrary length. The parity bit is adjusted to ensure that the arithmetic sum of all bits in the string is even. Then if an upset occurs anywhere in the string, simple addition of the bit string provides an indication (an odd parity) that an error has occurred, but determination of the location of the specific bit error is not possible. Furthermore, the parity bit itself may upset, creating an additional source of errors. The single-parity-bit approach also cannot detect an even number of upsets in the same bit string. TMR can be viewed as a very inefficient form of EDAC because it requires two bits of overhead for every bit in the circuit, but unlike parity checks, for TMR the errors are corrected in "real" time.

An ingenious code known as the Hamming code has been devised to provide both detection of an upset bit and its location in the string [49]. The Hamming EDAC technique adds one parity check bit to each of multiple arrangements of the bits within a data string in sequences that enable reconstruction of the erroneous bit in case of an upset, including protection of the check bits themselves. The most common application of Hamming EDAC is for single error correction and double error detection (SECDED). SECDED means if there is a single error in a word of any length it can be detected and corrected, but if there are two errors in a word, it can only be detected but not corrected. The Hamming technique requires overhead, in the form of additional bits, as shown in Table I for SECDED. Note that the overhead becomes more efficient as the bit string length increases. For example, a 16-bit string can be corrected with six additional bits using this approach (38% overhead), while a 32-bit string requires seven additional bits (22% overhead). Furthermore, additional check bits can be incorporated to correct multiple bit errors if desired. A more thorough discussion on EDAC, including Reed–Solomon (RS) and Bose, Chauduri and Hocquenghem (BCH) codes can be found elsewhere [49].

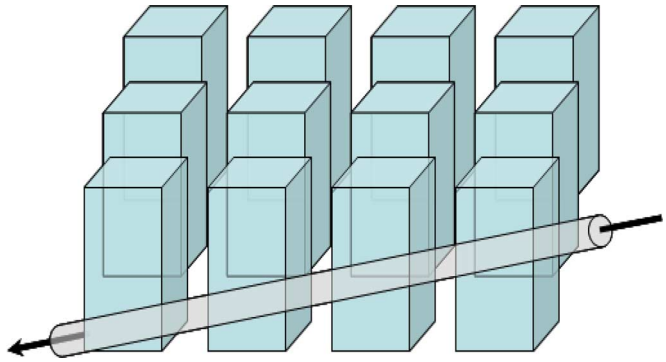

Fig. 22. Schematic diagram of a single energetic ion strike at off-normal incidence. The particle intersects multiple sensitive volumes along the charge path, potentially resulting in MBUs/MNUs.

C. Single-Event Upsets—Multibit Upsets

Multiple-bit upsets occur when a single energetic particle strike results in the upset of more than a single bit [50]. For older CMOS technologies, where the distance between adjacent devices is relatively large, when an energetic particle strikes perpendicular to an IC, the induced ionization charge path will likely only intersect one device and its associated sensitive volume. But radiation in space is omnidirectional. As the angle of incidence of an energetic particle with respect to normal increases, it becomes more likely that the induced ionization path can intersect two sensitive volumes. This situation is illustrated in Fig. 22. If these sensitive volumes are associated with separate stored logic elements (bits), it can result in the upset of two adjacent bits. The number of bits upset will depend on the angle of incidence, the particle LET and the dimensions of the sensitive volume and the distance between sensitive volumes. As the angle of incidence approaches parallel to the surface of the integrated circuit, the number of bit upsets associated with a single particle strike can continue to increase. As the sensitive volumes become closer together and/or the vertical dimension of the sensitive volume increases, the minimum angle of incidence with respect to normal at which an MBU can occur decreases. Hence, the associated effective solid angle over which MBUs can result increases, and the associated bit error rate increases. Clearly, as technology scales and the dimensions between devices decreases, the expected susceptibility to MBUs will increase [51]–[53].

The above discussion assumed that an MBU occurred from a single energetic particle strike that intersected adjacent bits at different points along the charge track. The collected charge for each sensitive volume came from the ionization charge from different locations along the ion track. Another regime for MBUs exists where two adjacent bits are upset from energetic particle strike ionization charge from the same segment of the ionization path. That is to say that the charge cloud at a given point along an ionization track can be shared among adjacent bits, resulting in an upset in both bits. This charge is primarily the diffusion component of the charge generation. In this case, MBUs can

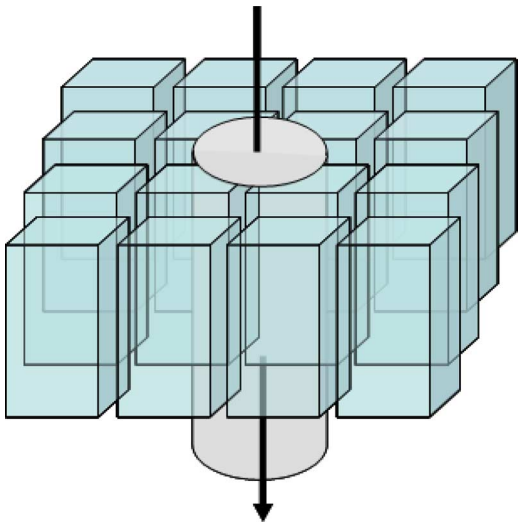

Fig. 23. Schematic diagram of a single energetic ion at normal incidence that intersects multiple sensitive volumes at any given point along the charge path, potentially resulting in MBUs/MNUs.

occur from particle strikes normal to the integrated circuit surface. This situation is illustrated in Fig. 23.

A short discussion on terminology. When TMR is applied, a single register or "bit" is replaced by three "bits." In this paper, since one TMR cell represents one bit of information, each of the three registers that make up that bit will be referred to as sensitive nodes. When DICE is applied, two specific transistors must be simultaneously hit for an upset. We will refer to each of these transistors as sensitive nodes. Hence, when two sensitive nodes are quasi-simultaneously hit resulting in a cell/bit upset (e.g., TMR cell, DICE cell), we will refer to that as a multinode upset (MNU). When EDAC is applied as a mitigation strategy, extra bits are added to a word to implement SECDED. Each one of those bits will be referred to as a bit. Hence, when two bits are quasi-simultaneously upset by a single energetic particle strike, such as two bits in an EDAC word, we will refer to that as an MBU.

The impact of MBUs/MNUs on HBD mitigation strategies will now be discussed. When charge dissipation techniques are used to mitigate SEUs, the bit LET threshold is increased, and the likelihood of an MBU is decreased simply by the fact that the critical charge necessary to upset a bit is increased. When spatial redundancy approaches are employed to mitigate SEUs, the impact of an MNU must be carefully considered since the essence of these techniques is that the information stored in a cell is spatially redundant. If two of the three redundant sensitive nodes are upset in a TMR cell from a single energetic particle strike, the mitigating effect of TMR is lost and an SEU error will be propagated. Similarly, for a DICE cell, if a single energetic particle strike results in the disturbance of two critical nodes (e.g., $X_0 - X_2$, $X_1 - X_3$), an upset can result. The design solution, in theory, is to keep critical nodes from a single cell sufficiently separated so that the probability of an MNU is kept small. Clearly, as technology scales and the sensitive volumes become closer together, this becomes more important and presents a more difficult challenge. This can result in cell area and speed penalties. One approach to minimizing these penalties is to interleave adjacent cells. This has been demonstrated on a 130-nm commercial process for a DICE cell where four DICE cells were interleaved and resulted in a decrease in the SEU error rate. This approach greatly complicates the interconnect routing, and as a result it is difficult to extend to higher number of cells. Another consideration is in how the cells are implemented. A TMR cell can be implemented as a single custom cell in a cell library that consists of three registers and a voting circuit. In this case, the designer has control over register placement, but it becomes increasingly difficult to keep sensitive nodes spatially separated without incurring significant area penalties. It is also possible to implement TMR by using individual register cells and individual logic cells. In this case, the automated layout tool may place the registers close together or far apart, and the designer loses control of a critical layout parameter.

When SECDED EDAC is applied to mitigate SEUs, typically for memories, it is essential that at most one error occurs in an EDAC word. Hence, it is essential that the individual bits that make up one word are physically separated from each other. For example, for a $n \times 39$ bit memory array (32 bit word + seven check bits), the n first bits are all grouped together, the n second bits are all grouped together, etc. Hence, if an energetic particle strike results in MBUs, they will all be in the same bit for different words, and are correctable by SECDED. Depending on the bit error rate, errors will begin to accumulate within an EDAC memory. To avoid double bit errors in the same word, it is necessary to periodically scrub the memory, where every EDAC word is checked and corrected, if necessary. The time between scrubbing depends on the intrinsic bit SEU sensitivity (LET threshold, saturation cross-section) and the environment. The more sensitive the memory is to upset and/or the more demanding the environment, the higher the required scrub rate. During scrubbing, parts of the memory are not available, so high scrubbing rates can result in reduced system availability. Furthermore, power usage can become an issue for high scrub rates. All these factors must be considered in designing an EDAC protected memory for a given system.

D. Single-Event Upsets—Advanced Technology Challenges

To fully understand charge sharing in advanced CMOS processes, it is necessary to understand the role of bipolar amplification in the charge collection process. For CMOS technologies, it has been shown that parasitic bipolar action affects the collected charge [54], [55] after an energetic particle strike, and with decreasing gate length, the bipolar gain increases. As shown in Fig. 24, the lateral parasitic bipolar transistor is formed by the drain, well, and source regions [56]. The drain acts as the collector, the well as the base, and the source as the emitter.

It is important to distinguish between the charge resulting directly from an energetic particle strike and the charge due to parasitic bipolar action [57]. The charge deposited directly from an energetic particle is subject to charge sharing with other nodes in the proximity. The charge collected due to bipolar amplification is associated with a single device. However, charge that reaches an adjacent device can result in bipolar amplification if the voltage perturbations on the adjacent device are sufficient to turn on the parasitic bipolar amplification. Recent

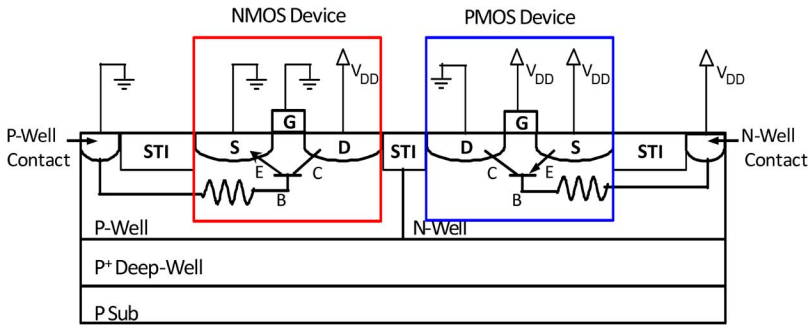

Fig. 24. CMOS cross-section, showing parasitic elements. The NMOS device has a lateral parasitic npn bipolar transistor, while the PMOS device has a lateral parasitic pnp bipolar transistor [57].

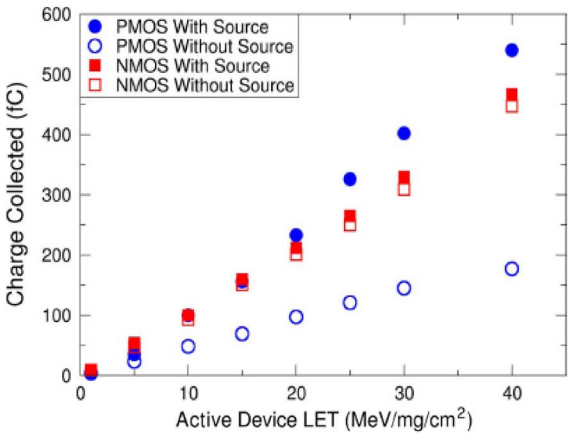

Fig. 25. Charge collected at the drain of a transistor for an energetic particle strike on the same transistor as a function of LET [57]. The difference between the "with source" and "without source" components represents the parasitic bipolar contribution.

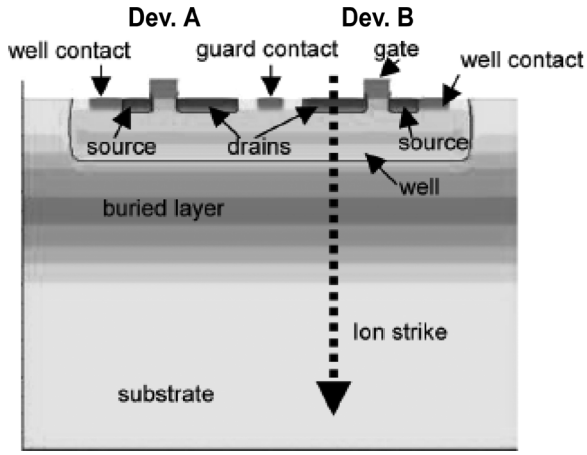

Fig. 26. Two-dimensional cut plane view of a test structure used to evaluate the efficacy of guard bands in reducing charge sharing between adjacent transistors in a common well [59].

modeling/simulations has shown a bipolar parasitic effect in a common well region, increasing the probability of multiple node charge collection [58].

Technology computer-aided design (TCAD) simulations were performed to investigate the degree of bipolar amplification in a twin-well 130-nm commercial CMOS process [57]. The charge collected at the drain of a transistor for an energetic particle strike on the same transistor as a function of LET was simulated and is shown in Fig. 25 [57]. The simulations were performed using devices with and without the source implant. The difference between these two represents the bipolar contribution. The simulations for the PMOS devices show high parasitic bipolar amplification when compared to that for the NMOS devices. It is explained that the PMOS devices show high parasitic bipolar amplification compared to the NMOS device due to a voltage perturbation in the n-well during the charge collection process [57]. Substrate/p-well voltage variations caused by an energetic particle strike are much smaller than those in the n-well since the connection of the p-well to the p-substrate provides effective control of the well potential. The effect of bipolar amplification is that the total charge collected can be greater than the total charge deposited by the incident energetic particle, resulting in increased vulnerability to SEUs and MBUs/MNUs.

The effectiveness of HBD techniques to mitigate charge sharing between adjacent transistors has recently been investigated [59]. TCAD simulations were performed to estimate charge collected by two adjacent nodes due to a single energetic particle strike, and to compare the effectiveness of different HBD isolation techniques. A 2-D cut plane of the PMOS test design for a 130-nm commercial CMOS process used in this simulation is shown in Fig. 26. This structure is used to determine the effectiveness of an HBD guard contact in reducing charge sharing effects, and hence, reducing MBUs/MNUs in this technology. For NMOS transistors, three different isolation techniques were examined: two NMOS transistors isolated by a guard contact, two NMOS transistors in a p-well isolated by an n-well, and two NMOS transistors in separate p-wells isolated by an n-well between and a buried third well below. In the simulations, an energetic particle strikes the drain of device B, and the resultant drain current and associated integrated charge is evaluated for device A. Simulations were performed for particles with LET of 2.5, 20, and 50 MeV-cm^2/mg [59]. The results are represented as the charge collected at device A expressed as a percentage of charge collected at device B for a direct strike on the device B drain. The results for PMOS devices with and without a guard band, are shown in Fig. 27 for a particle strike with an LET of 50 MeV-cm^2/mg. For simulations of a structure without a guard band, the charge

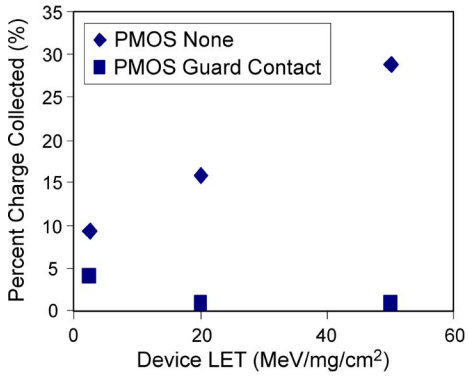

Fig. 27. Charge collected on device A as a percentage of the charge collected on device B for a direct hit for simulations on PMOS test structures [59].

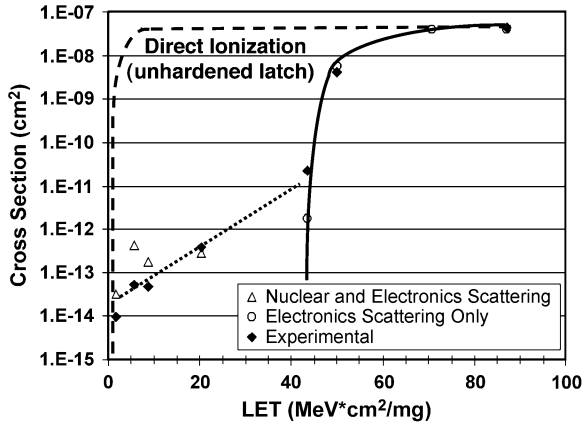

Fig. 28. Comparison of cross-section versus LET curves resulting from simulations and experiments for a rad-hard memory [62]. The dashed line represents a typical cross-section curve for an unhardened memory.

shared on the adjacent node is 28.8% of that of the charge collected at the drain of device B for a hit on device B. The effect of inserting a guard contact is to reduce the device A charge by a factor greater than 15 for the higher LETs. The guard band acts to remove charge, more effectively restoring the well potential and turning off bipolar amplification in the secondary device. It was shown for the device with the guard contact that by 1 ns after a particle strike, the potential in the well has nearly fully recovered [59]. Furthermore, the guard contact reduces the charge collected at the drain of device B as compared to without a guard contact.

The impact of a guard contact for NMOS devices is relatively modest as compared to the effectiveness of a guard contact on PMOS devices. This can be understood in terms of the relative role of bipolar amplification in PMOS and NMOS charge collection from an energetic particle strike. In the NMOS device, the well does not collapse to give bipolar amplification. As a result, the guard contact, which acts to reduce bipolar amplification, is less effective. Two other n-well isolation approaches also did not provide significant improvement [59]. An n-well full isolation approach that implements a triple well design, however, was shown to be very effective in mitigating well sharing effects in NMOS devices. The role of bipolar amplification in MBU/MNU is an active area of research. The need to isolate critical nodes in separate wells may emerge as a necessary mitigation technique for some applications. As technology scales, it is likely new mitigation techniques related to these effects will need to be developed.

Up to this point, we have discussed the case where upsets are caused by the direct ionization charge associated with an energetic particle strike. For light particles (e.g., $Z < 1$), such as protons and neutrons, there is usually not enough ionization charge to lead to a node or cell upset. Nucleons, however, can interact with matter through nuclear reactions and recoils to produce secondary ions that can now produce direct ionization and result in upsets. Furthermore, while heavy ions must cross several media before reaching a sensitive region in the cell, secondary ions induced by nucleons can be produced directly anywhere in the device. Although the probability for nuclear reactions is small relative to direct ionization from the primary ion, the reaction products may have LETs exceeding that of the primary particle [60]–[62]. If the charge generated by these secondary particles in a sensitive volume exceeds the critical charge, the cells may upset, resulting in an MBU.

With some exceptions, analysis of SEU error rates for heavy ion interactions with microelectronic components have been assumed dependent on a single parameter, the particle LET. Underlying this assumption is that the LET represents an average value of stopping power for a given ion type at a given energy. This analysis does not take into account the statistical variations of the energy deposition characteristics. Recent results [62] suggest that this assumption may not be valid, and that nuclear reactions and recoils from relatively light heavy ions can be important in determining the upset cross-section at lower LET values. In that work, a Monte Carlo simulation that includes the possibility of nuclear reactions (elastic and inelastic) and screened Coulomb scattering is run for 10^8 particle strikes for a 0.25-μm commercial CMOS process. An unambiguous sensitive volume is defined and all incident particles were assumed normal to the surface. These simulations can be performed for a set of primary ions of differing LET values. For a circuit with a given $Q_{\rm crit}$, the effective cross-section for each LET value can be obtained, and a cross-section as a function of LET can be constructed. Similarly, if experimental cross-section as a function of LET is known for a given circuit, cross-section versus LET graphs can be generated as a function of $Q_{\rm crit}$ and compared with the experimental data. This was done in simulation for a memory that was hardened by increasing the LET threshold (defined as 10% of the saturation cross-section) significantly above the unhardened equivalent memory. The result of this analysis is shown in Fig. 28, where the best fit was obtained for a $Q_{\rm crit} = 1.21\ pC$. As expected, the cross-section drops rapidly as the beam LET decreases below the SEU threshold of the cell. Instead of going to zero just below this LET value, as expected when nuclear reactions are not included, there is a low cross-section tail in the experimental data that extends to LET values down to ~ 2 MeV-cm^2/mg. The simulation results that include nuclear reactions fit this low LET data fairly well.

Other recent results indicate that the effects of nuclear reactions on the SEU error rate become even more pronounced when

a layer of tungsten is placed above the silicon emulating tungsten plugs that are used in modern CMOS processes to electrically connect the silicon, polysilicon, and metal layers [62], [63]. In this work, it was found the low LET tail in the cross-section versus LET curves becomes even more pronounced and is sensitive to the location of the tungsten plug. The effect of the nuclear reaction tail can increase the error rates by up to two orders of magnitude [64].

The reason this is important for HBD is that with the exception of the application of charge dissipation techniques, which are usually not employed due to penalties, SEU mitigation is achieved through redundancy. The redundancy approach does not harden by increasing the Q_{crit} of a single sensitive node, as is often the case for hardening by process approaches (e.g., feedback resistors), but rather, is achieved by having the information stored at more than one critical node (TMR, DICE, EDAC). The Q_{crit} for advanced commercial latched logic cells upset by a single sensitive node strike tend to be very low, as does the associated LET threshold, as shown schematically by the dashed line in Fig. 28. As discussed above, the rad-hard cell error rate can be increased by up to two orders of magnitude from nuclear reactions, severely limiting the actual achieved mitigation. For HBD cells (without design modifications to change the intrinsic threshold LET of isolated nodes) that are designed so that two sensitive nodes cannot be simultaneously upset by a single particle strike up to high LET values, such as a DICE cell or TMR cell with wide critical node spacing's or an EDAC memory where bits within the same word are widely separated, the LET cross-section for each node is dominated by direct ionization effects down to very low LET values. Although there can be small nuclear effects in this LET regime, for these HBD cells the upset cross-section remains unchanged. That is to say the nuclear contribution does not produce secondaries with high enough LET to simultaneously upset two critical nodes and the nuclear contribution is negligible down to low LET, resulting in an error rate that is unaffected by the nuclear contribution. Hence, nuclear reactions do not appear to be an issue when mitigation is achieved through HBD techniques that utilize information redundancy to prevent simultaneous critical node upset up to high LETs. However, if the HBD cells are not designed to prevent simultaneous upset of two critical nodes at moderate LET values or above, such as for DICE cells with small critical node spacing and or cells that have critical nodes in the same well, secondaries from nuclear interactions can increase the below threshold cross-section and result in an increased error rate, as illustrated in Fig. 28.

E. Singe-Event Transients

Single-event upsets are the consequence of an energetic particle strike directly on a critical node of a latched or sequential logic element, resulting in a change in the logic state of that element. Single-event transients address errors that can result from an energetic particle strike on nonlatched elements, such as combinatorial logic, clock line and global control lines. In this section, a description of the basic mechanism which leads to SETs in digital circuits is discussed. Following that is a discussion on SET mitigation approaches using HBD techniques, and a discussion of recent research into characterizing and understanding the distribution of pulse widths that result from an energetic particle strike found in advanced commercial CMOS technologies.

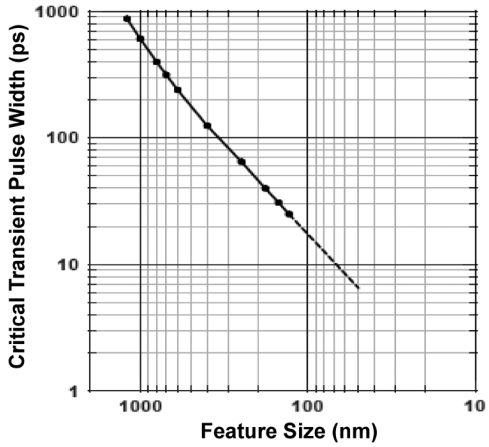

Fig. 29. Critical transient width versus feature size for unattenuated propagation [65].

A single-event transient is the result of a voltage disturbance on a signal line caused by a particle strike. The current produced by a particle strike can cause a capacitive element, such as the output load of a combinatorial logic element, to charge or discharge, causing a temporary change in the voltage on that signal line. Whether this voltage transient will propagate any significant distance through the combinatorial logic depends on both the width of the transient voltage spike and the speed (capacitance) of the CMOS transistors. The critical width is the minimum width of the transient pulse required for the transient to propagate through an infinitely long chain of inverters. If the transient width is narrower than this critical width, the transistors do not have the speed to respond to the transient and the transient will be attenuated and die out after passing through only a few gates. The critical transient width versus feature size for unattenuated propagation is shown in Fig. 29 [65, and references within]. As technology advances and feature size shrinks, the critical transient width narrows. Hence, as CMOS is scaled and gate delays become shorter, narrower and narrower transients can propagate through the circuit. Because of this, as CMOS technology advances, SETs are expected to become more important in contributing to the overall error rate. It was pointed out that while a transient pulse greater than or equal to the critical transient width would propagate infinitely, a pulse of half this width would terminate at the first gate [66].

To understand how a transient in nonlatching logic can result in a circuit error, consider the case of a series of combinatorial logic elements terminating at a sequential logic element. If an energetic particle strikes the combinatorial logic and if the induced transient is of sufficient width to propagate unattenuated through the remaining combinatorial logic element to the input of the register, there is a potential that this transient will be interpreted as a signal and the state of the register could be put in error. If this occurs or not will depend on whether the transient signal is coincident with the clock edge that latches data into the register. As shown in Fig. 30, the transient will be

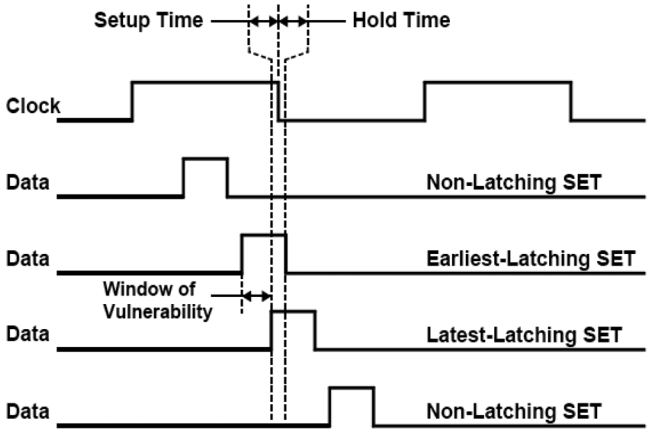

Fig. 30. When the transient on a data line occurs during the setup and hold times for a latch, it can produce an SET error [68].

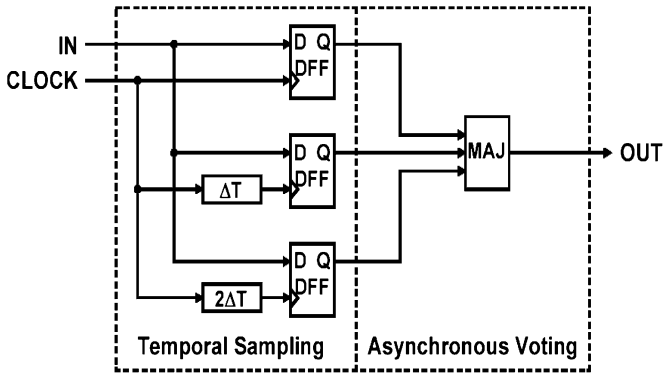

Fig. 31. SET-hardened latch design that incorporates temporal offsets in the clock inputs to three flip-flops to ensure that an SET on the data input does not trigger more than one error in the inputs to the voter [65].

incorrectly interpreted as a valid signal if it arrives during the time period extending from a setup time before the clock edge to a hold time after the clock edge. Furthermore, the fanout at each logic gate can cause these errors to propagate into many independent branches of the logic, potentially inducing more than one error from a single transient. Since the probability of a transient being captured depends on the number of falling clock edges arriving at the latch per unit time, it is expected that the SET upset rate will depend linearly on clock frequency. More specifically, error cross-section for an SET will be given by the total sensitive area within the combinatorial logic multiplied by the window of vulnerability and divided by the clock period. The window of vulnerability is not the pulse-width, but rather the transient pulse width minus the setup and hold time (if the pulse width is equal to the setup and hold time, the probability of those being coincident in time is zero). Not only does each combinatorial gate in a circuit potentially contribute to the SET error rate (because transients are no longer attenuated), but the probability of storing any given transient pulse into a latch, resulting in an error, will also increase.

The total circuit upset error rate is the combination of SEU and SET rates. Until recently, SEU errors for static-latch cells dominated the overall soft error rate in the logic portions of microcircuits. For technologies below 0.25-μm, SET errors have become dominant in circuits that use SEU-hardened static-latch cells [67]–[69]. Recent experiments for 0.25- and 0.18-μm technologies demonstrated the linear dependence of the SET cross-section as a function of clock frequency [67], [69], [70]. In addition, for a 0.18-μm technology the transient pulse width has been shown to depend linearly on particle LET, and exceed 1 ns for heavy ions with LET values above ~ 55 MeV-cm^2/mg [68], [69], [71]. In these experiments, SET saturated cross-section values exceeded a value of 1×10^{-8} cm^2/bit at $\sim$150 MHz, a typical value for the SEU saturated cross-section in these technologies, and would be the dominant upset mechanism in circuits using hardened static latches.

Single event transients can be mitigated using HBD techniques. One technique involves slowing down circuit speed, increasing the critical minimum pulse width for propagating a transient unattenuated through a chain of combinatorial logic. This can be done by capacitively loading the logic chain [45]. Alternatively, increasing the drive of the transistors in the combinatorial logic cell will effectively decrease the transient width and voltage change, reducing the chance of propagating an SET.

The primary approach to mitigating SETs involves the incorporation of temporal delays and/or redundancy. Fig. 31 shows a triple voting scheme that uses temporal filtering to prevent an SET from corrupting the voter output [65]. In this circuit, the data input is connected to three edge-triggered D-flip-flops in parallel. The clock signal goes directly to the top flip-flop, is delayed by Δt at the middle flip-flop, and is delayed by $2\Delta t$ at the bottom flip-flop. If a transient is induced on the data input, it will arrive at all three flip-flops simultaneously. The clock signals, however, will not arrive at the same time, and at most, only one clock signal will arrive coincident with the transient on the data line. This assumes the transient pulse width is less than Δt. When the asynchronous voting occurs, two inputs will contain the correct signal, while at worst only a single input will contain an erroneous signal, and thus the output of the voter will be the correct logic signal. Alternately, the delay can be imposed on the input signal instead of the clock signal. For these cells the clock/data delays are generated internally in the temporal latch itself. For either case, a transient on the clock pulse can result in an error in the cell output. Since CLKB and CLKC phases are derived from the CLKA phase, a transient on CLKA will produce a transient on CLKB and CLKC. This may cause incorrect data to be latched into multiple branches of the latch producing an error at the output of majority gate. This circuit incorporates both temporal and spatial redundancy TMR into a single cell. It should be noted, however, that if there is a transient on the input (potential SET) and a strike on one of the flip-flops (potential SEU) during one clock cycle, an error may result.

A more elegant application of temporal sampling uses a multiplexer (MUX) circuit to achieve temporal and spatial redundancy without actually physically replicating the circuitry [65]. The following discussion closely follows Mavis [65]. An alternative approach to implementing a level sensitive transparent latch is to use a two-input MUX with its output fed back to one of its inputs, the data fed into the input, and the select line controlled by the clock signal. Extending this concept, a temporal latch that mitigates SEUs and SETs can be constructed and is

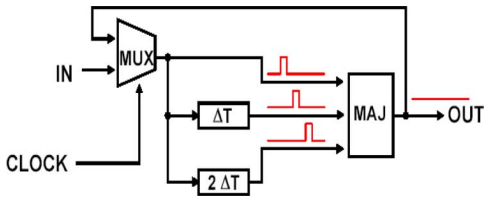

Fig. 32. Minimal temporal sampling latch replicating itself in time to suppress both SEUs and SETs [65].

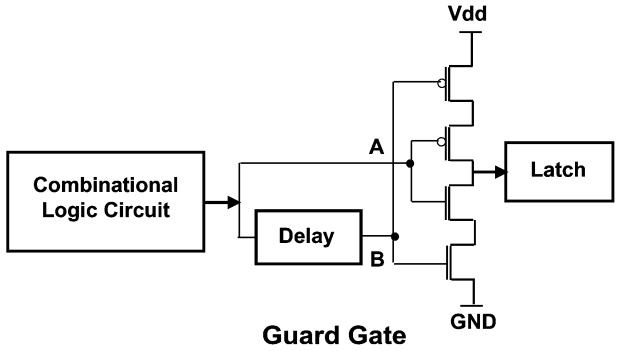

Fig. 33. Alternative circuit approach used to block the transmission of transients on the input from reaching a latch [70].

shown in Fig. 32. This circuit differs from the level sensitive transparent latch in that instead of the output of the MUX being fed back directly to the MUX input, the MUX output is triplicated and delayed as in the temporal latch shown in Fig. 32. By doing this, the equivalent of triple spatial redundancy is achieved without physically replicating the circuitry. As with the earlier described form of temporal latch, for the latch shown in Fig. 32 a transient will be "blocked" if the transient pulse width is less than or equal to ΔT. If the transient pulse width is greater than ΔT, the circuitry will be ineffective in "blocking" the transient and an error will occur. This particular latch is also immune to transients occurring on the input clock node. Any clock transient momentarily switches the selected MUX input producing a possible transient at the MUX output. The data input is momentarily selected, but will be rejected by the voting circuitry. Thus, unlike the temporal latches described in Fig. 31 and the related latch where the delays are incorporated into the data lines, this latch does not require SET hardened clock nodes. Finally, since the temporal latch is level sensitive, two of these cells in series, with clock signals inverted from one another, are required to make an edge-triggered latch.

The insertion of two extra sampling times employed in the temporal latches described above increases the latch setup time by $2\Delta t$. This increased setup time clearly results in a lower maximum clock operating frequency which is given by

$$1/f_{\text{eff}} = 1/f_o + 2\Delta t \tag{3}$$

where f_o is the prehardened maximum latch frequency, and f_{eff} is the effective maximum latch frequency after hardening. It can easily be shown that the factor by which the frequency is reduced is given by $f_{\text{eff}}/f_o = 1/(1+2\Delta T f_o)$. Application of the temporal latch to block transients as long as 1 ns can impede operation at high clock frequencies. For example, if the maximum clock frequency was 500 MHz without temporal latches, with temporal latches it is necessary to add $2\Delta T$ (2 ns) to the clock period $(2\,\text{ns} + 2\,\text{ns}) = 4$ ns, or a maximum operating frequency of 250 MHz.

Another approach that uses temporal delay and a guard gate to block transients is illustrated in Fig. 33 [70]. The guard gate, also referred to as a Muller C-element, is a buffer circuit with two inputs and one output. When both inputs A and B are the same, the guard gate acts as an inverter. When the two inputs are different, the output floats in the high-impedance state, and the output voltage will maintain its value until degraded by leakage current. The circuit approach to SET transient suppression can be understood as follows. When no transient has occurred, the output of the combinatorial logic will be the correct value, the guard gate will act as an inverter, and the correct signal reaches the latch. When a transient is produced in the combinatorial logic circuit and propagates to the combinatorial out, it is applied directly to one of the guard gate inputs, while the other input receives the same transient but delayed in time. The result is the output of the guard gate does not change over this time period. As with the temporal latch, this approach is only successful when the delay time is longer than the pulse width. Also similar to the temporal latch is a penalty in the maximum operating frequency. It has recently been proposed to apply guard gates to improve the SEU performance of DICE cell designs [71].

A very active area of investigation over the last three years has been to quantify and understand SET cross-sections and pulse widths as a function of LET and technology node in advanced commercial CMOS technologies. Since the window of vulnerability depends on the transient pulse width, it is important to understand the distribution of pulse widths for a given technology, combinatorial circuit, LET, frequency, and voltage. One approach to characterizing SET pulse widths is to use the temporal latch (Fig. 32). In these experiments the test structures consisted of shift registers designed from temporal latches [67]–[69]. As discussed above, the temporal latch rejects transients of width less than ΔT, and passes transients of widths greater than ΔT. By incrementally increasing the value of ΔT in the temporal latch until the error rate reaches zero, the maximum pulse width can be determined. The SET cross-section as a function of LET value for various values of temporal latch delay for a 0.18-μm commercial process is shown in Fig. 34. Each point for a given LET and delay value represents the cross-section that includes all transients with pulse widths greater than the specified delay value. As the delay is reduced, more transients are included and the SET cross-section increases. In addition, by taking the data from Fig. 34 [68], [72], it is possible to generate a graph of the SET cross-section as a function of pulse width for various LET values.

Another variable that can affect the SET cross-section is the value of the supply voltage. The SET cross-section as a function of pulse width at different supply voltages is shown in Fig. 35 for a 0.18-μm technology [68]. The solid lines represent a fit to an inverse log-normal cumulative distribution. These results indicate for a given technology node, and at a fixed value of pulse width, as the supply voltage is lowered the SET cross-section

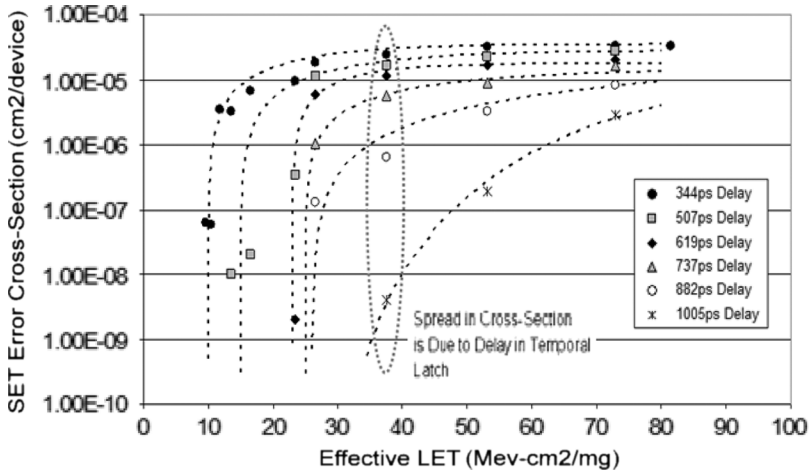

Fig. 34. Heavy ion SET cross-section as a function of LET for various temporal latch delay values [68], [72].

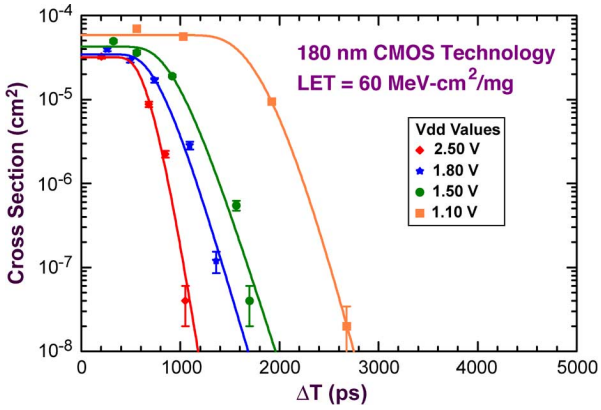

Fig. 35. Heavy ion SET cross-section as a function of temporal filter delay/pulsewidth and at various supply voltages [68].

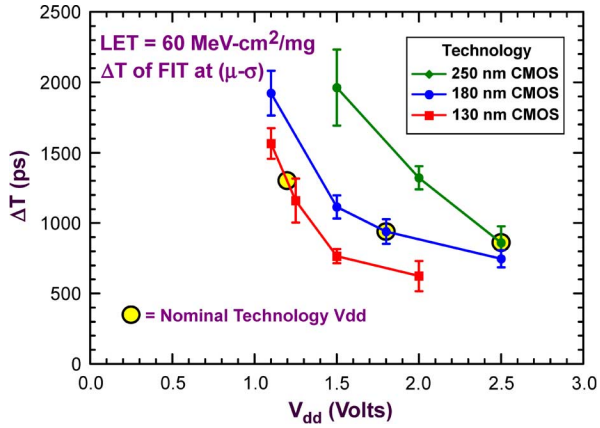

Fig. 36. Temporal latch delay necessary to reduce the SET errors to 30% as a function of supply voltage for three different technology nodes. The circled value represents the delay at the nominal supply voltage value for a given technology [68].

increases. In addition, as the supply voltage is reduced, the maximum pulse width generated from a heavy ion strike increases. The implications of this can be calculated. Using the distribution fitting parameters, it is possible to calculate the delay necessary at a fixed LET value to reduce the SET errors to 50%, 30%, 5%, and 0.1% (0σ, 1σ, 2σ, 3σ). Similarly, results were obtained for 0.25- and 0.13-μm CMOS technologies, and using the same approach, the delay necessary to reduce SET errors a given amount can be calculated. These results are plotted in Fig. 36 for a 70% (1σ) reduction in SET errors as a function of supply voltage and technology node [68]. The circled values are for the nominal supply voltage at each technology node. The results indicate that the transient widths are increasing with technology scaling, presenting greater challenges to the designer of radiation-hardened microelectronic components for space. There is not yet a complete understanding on the basic mechanism(s) that determine pulse widths at a given technology node.

F. Single-Event Latchup

An energetic particle strike can cause an additional problem in CMOS circuits known as single event latchup. Single-event latchup can occur in bulk CMOS circuits as a direct result of the numerous parasitic pnpn circuit paths between VDD and ground

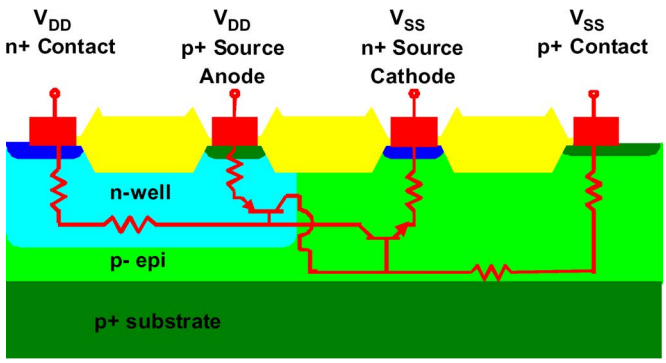

Fig. 37. Cross-section of a CMOS circuit showing the substrate, well, and diffused regions that form a parasitic pnpn SCR [73].

normally occurring in a CMOS circuit. These parasitic transistors are usually comprised of the source of a PMOS transistor, the n-well, the p-substrate, and the source of an NMOS transistor, as shown in Fig. 37. This pnpn structure is the configuration for a classic semiconductor-controlled rectifier (SCR), as shown in Fig. 38, that can be stable in either its off or on (conducting) state. Normally the n-well is maintained at the same

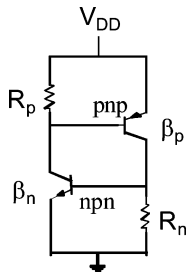

Fig. 38. Lumped equivalent circuit of the parasitic pnpn SCR. Injection of current into the base of either of the transistors associated with an energetic particle strike can trigger the SCR into a high current state.

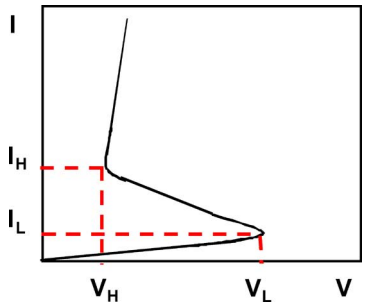

Fig. 39. Current–voltage characteristics of an SCR. If the injected current increases the voltage to above the latch voltage V_L, the SCR goes into a high conducting state. If the supply voltage is greater than the holding voltage V_H, the high current state will be maintained unless power cycling occurs.

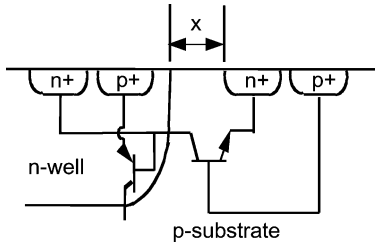

Fig. 40. The n-well to diffusion spacing X between can be increased to increase the latchup threshold LET [3], [75], [76].

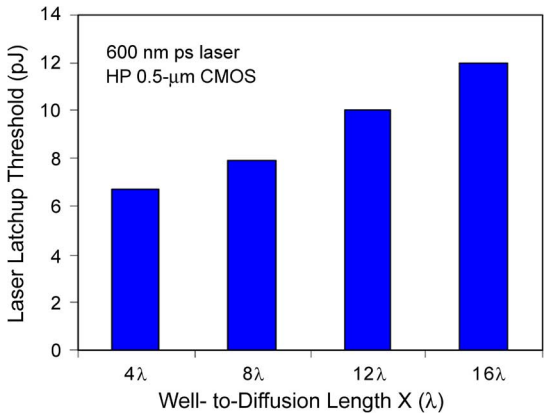

Fig. 41. Latchup threshold measured using an SEU laser simulation technique as a function of the well-to-diffusion spacing (X) in units of the design scaling constant λ [3], [76], [77].

potential as the PMOS source, and the p-substrate is held at the NMOS source potential, which prevents the SCR from latching. However, after a particle strike, the generated photocurrents can cause voltage drops in the substrate or well that may forward bias one of the parasitic bipolar transistors in the pnpn circuit, triggering the device into a highly conducting state. If the injected current increases the voltage above the latch voltage V_L, the SCR goes into a highly conducting state, as illustrated in Fig. 39. If the supply voltage is greater than the holding voltage V_H, the high current state will be maintained unless power cycling occurs. In some cases, sufficient currents can be generated in the latched state to damage sensitive metal lines, which are not designed to carry the high currents that can result from a latchup condition. In other cases, although electrically the part appears to function normally, imaging indicates the presence of latent damage [74]. Although supply voltages have decreased over the last two decades, the supply voltage is expected to remain constant at $\sim$1 V over the next five years. In addition, it can be shown that a necessary condition for latchup to be sustained is that the product of the parasitic transistor betas be greater than one $(\beta_{\rm p}\beta_{\rm n} > 1)$. The minimum LET of an energetic particle that triggers latchup is referred to as the latchup LET.

Single-event latchup is typically avoided in CMOS circuits by reducing the resistance of the well or substrate, thereby increasing the injection current required to trigger latchup. This can be accomplished by the proper attention to well and substrate contacts in the design or by choosing a process that incorporates an epitaxial structure over a highly conductive substrate. Latchup can also be prevented by utilizing an insulating substrate, such as SOI, which prevents current flow between adjacent transistors.

Commercial foundries are concerned about latchup in their commercial products triggered by voltage transients (not heavy ion strikes). Thus, design rules are typically in place to prevent adjacent well and active area edges from approaching too closely, thereby avoiding lateral high-gain parasitic bipolar devices in the CMOS substrate.

While CMOS processes that are susceptible to SEL are normally avoided for space applications, several HBD techniques can be applied to improve a CMOS circuit's SEL immunity. Design rules that extend the distance between well edges and active regions, as shown in Fig. 40, can reduce the gain of the lateral parasitic bipolar transistors that create the latching SCR, and therefore increase the latchup LET threshold [3], [75], [76]. The latchup threshold measured using an SEE laser simulation technique as a function of the well-to-diffusion spacing (X), in units of the design scaling constant λ is shown in Fig. 41 [3], [75]–[77]. The minimum design rule for X for this 0.5-μm process was 4λ. By quadrupling X, the latchup threshold increases by approximately a factor of two. It should be noted that quadrupling X could result in minimal area impact for lower density logic where diffusions approach well edges in only a small percentage of the circuit layout.

The application of guardbands around active regions is also effective in mitigating latchup. Guardbands are p+ or n+ diffusion pockets placed around the well or substrate (i.e., n+ guardbands are placed around the n-well, and p+ guardbands

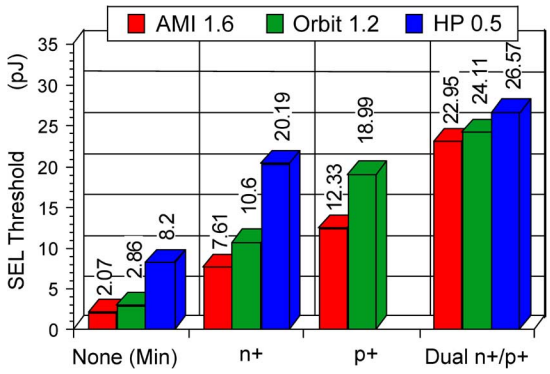

Fig. 42. Laser latchup threshold for test structures with no guardbands, a single n+ guardband, a single p+ guardband, and dual n+/p+ guardbands for three different CMOS technologies [3], [76].

are placed around the p-substrate). The insertion of a guardband increases the separation between the n-well edge and the diffusion region, decreasing the lateral transistor bipolar gain. The heavily doped guardbands also act as a carrier sink, removing carriers that otherwise would be available to initiate the latchup process. Guardbands also decrease the resistance in parallel with the anode and cathode gates of the SCR, especially if the guardbands are silicided to reduce their resistivity. In general, guardbands control the potential in the latch path and prevent latchup from being initiated. Application of this technique does not require process changes or additional masks. Laser latchup thresholds for test structures with no guardbands, a single n+ guardband, a single p+ guardband and dual n+/p+ guardbands for three different CMOS technologies are shown in Fig. 42 [3], [76]. The three CMOS processes are the AMI 1.6-μm, the Orbit 1.2-μm and the HP 0.5-μm process, discussed earlier. As can be observed, the addition of guardbands increases the latchup threshold significantly. The increase in latchup threshold between a design with no guardbands and one with dual guardbands is a factor of $3\times$–$10\times$, depending on the technology.

The sensitivity to latchup for advanced CMOS technology nodes in a nonradiation environment has been reduced through innovative process engineering. Technology downscaling has resulted in reduced nwell depth (and hence, a shorter base for the pnp), but at the same time, higher channel stopper implants (and hence a higher Gummel number) [78]. The result of this is a reduction in vertical bipolar gain as a function of technology node (e.g., the vertical gain for a 90-nm technology is 20% that for a 0.25-μm technology) [78]. Furthermore, the insertion of a deep p+ implant buried layer also has the effect of reducing latchup susceptibility. Because of these innovations, today's available advanced processes tend to be resistant to latchup up to high LET values.

IV. Conclusion

This paper discussed the system drivers which has led the space community to adopt an HBD approach targeting commercial CMOS foundries for the fabrication of radiation-hardened microelectronic components. In this approach, radiation effects are mitigated by the application of novel design techniques. The effects of total dose radiation on CMOS technology was presented, as well as recent trends toward increased total dose hardness in advanced commercial technologies. An important transition occurred when the gate oxides became thinner than the carrier tunneling length, resulting in a complete lack of radiation-induced charge trapping in the gate-stack of advanced CMOS technologies. Hardness-by-design techniques to further mitigate total dose effects were described, as well as the associated penalties in area, power and speed. A lengthy presentation on SEEs was presented, including sections on SEUs, SETs, and SELs. This included discussions on basic mechanisms and HBD mitigation approaches. A detailed section on MBUs was also presented, which included discussions on charge sharing and nuclear effects. Because of the advancements in HBD over the last decade, many space programs are targeting commercial foundries to meet their microelectronic needs. There is no indication that this will change in the near future.

Acknowledgment

The author gratefully acknowledges the contributions of Jon Osborn, Donald Mayer and Steven C. Moss for their work on The Aerospace Corporation's program on hardness-by-design. The author also gratefully acknowledges Dr. Michael Fritze, Lew Cohn, and the DARPA Hardness-by-Design Program for supporting this work.

References

[1] R. H. Dennard, F. H. Gaensslen, H. N. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, "Design of ion-implanted MOSFETs with very small physical dimensions," *IEEE. J. Sol. State Circuits*, vol. SC-9, pp. 256–268, 1974.

[2] P. K. Ko, "Approaches to scaling," in *Advances in CMOS Device Physics*, N. G. Einspruch and G. Gildenblat, Eds. Dordrecth, The Netherlands: Academic, 1989, ch. 1.

[3] R. C. Lacoe, J. V. Osborn, R. Koga, S. Brown, and D. C. Mayer, "Application of hardness-by-design methodology to radiation-tolerant ASIC technologies," *IEEE Trans. Nucl. Sci.*, vol. 47, no. 6, pt. 3, pp. 2334–2341, Dec. 2000.

[4] F. B. McLean, H. E. Boesch, Jr., and T. R. Oldham, "Electron-hole generation, transport and trapping in SiO2," in *Ionization in Radiation Effects in MOS Devices and Circuits*, T. P. Ma and P. V. Dressendorfer, Eds. New York: Wiley, 1989.

[5] J. R. Schwank, "Total dose effects in MOS devices," in *2002 IEEE Nuclear and Radiation Effects Conf. Short Course*, Phoenix, AZ, Jul. 2002.

[6] J. L. Leray, "Total dose effects: Modeling for present and future," in *1999 IEEE Nuclear and Radiation Effects Conf. Short Course*, Norfolk, VA, 1999.

[7] R. C. Lacoe, "CMOS scaling, design principals and hardening-by-design methodologies," in *2003 IEEE Nuclear and Radiation Effects Conf. Short Course*, 2003.

[8] R. C. Lacoe, "Improving system performance through the application of hardness-by-design methodology," in *2007 RADECS Short Course*, 2003.

[9] R. C. Hughes, "High electric field properties of SiO2," *Sol. State Elec.*, vol. 21, pp. 251–1978.

[10] F. B. McLean, G. A. Ausmandl, Jr., H. E. Boesch, Jr., and J. M. McGarrity, "Application of stochastic hopping transport to hole conduction in amorphous SiO_2," *J. Appl. Phys.*, vol. 47, no. 4, pp. 1529–11532, Apr. 1976.

[11] F. B. McLean and G. A. Ausmandl, Jr., "Simple approximate solutions to continuous-time random-walk transport," *Phys. Rev. B*, vol. 15, no. 2, pp. 1052–1061, Jan. 1977.

[12] P. J. McWhorter, S. L. Miller, and W. M. Miller, "Modeling the anneal of radiation-induced trapped holes in a varying thermal environment," *IEEE Trans. Nucl. Sci.*, vol. 37, no. 6, pp. 1682–1689, Dec. 1990.

[13] D. M. Fleetwood, S. L. Miller, R. A. Reber, Jr., P. J. Whorter, P. S. Winokur, M. R. Shaneyfelt, and J. R. Schwank, "New insights into radiation-induced oxide-trap charge through thermally stimulated current measurements and analysis," *IEEE. Trans. Nucl. Sci.*, vol. 39, no. 6, pp. 2192–2196, 1992.

[14] J. V. Osborn, R. C. Lacoe, D. C. Mayer, and G. Yabiku, "Total dose hardness of three commercial CMOS microelectronic foundries," *IEEE Trans. Nucl. Sci.*, vol. 45, no. 3, pp. 1458–1463, Jun. 1998.

[15] R. C. Lacoe, J. V. Osborn, D. C. Mayer, S. Brown, and D. R. Hunt, "Total dose radiation tolerance of a commercial 0.35- μm CMOS process," in *1998 IEEE Radiation Effects Data Workshop Record*, 1998, pp. 104–110.

[16] R. C. Lacoe, "Performance impact of radiation-hardness-by-design," presented at the 2001 Hardness-by-Design Workshop, Albuquerque, NM, Jul. 2001.

[17] G. Anelli, M. Campbell, C. Dachs, F. Faccio, A. Giraldo, E. Heijne, P. Jarron, A. Marchioro, E. Noah, A. Paccagnelle, P. M. Signe, W. Snoeys, and K. Vleugels, "Total dose behavior of submicron and deep submicron CMOS technologies," in *Proc. 3rd Workshop on Elec. for LHC Experiments*, Sep. 1997, pp. 139–143.

[18] D. J. Frank, "CMOS design near the limit of scaling," *IBM J. Res. Develop.*, vol. 46, pp. 213–222, 2002.

[19] G. D. Wilk, R. M. Wallace, and J. M. Anthony, "Hafnium and Zirconium slicates for advanced gate dielectrics," *J. Appl. Phys.*, vol. 87, pp. 484–492, 2000.

[20] G. D. Wilk, R. M. Wallace, and J. M. Anthony, "High-k gate dielectrics: Current status and materials properties considerations," *J. Appl. Phys.*, vol. 89, pp. 5243–5275, 2001.

[21] B. H. Lee, L. Kang, W. J. Qi, R. Nieh, Y. Jeon, K. Onishi, and J. C. Lee, "Ultrathin hafnium oxide with low leakage and excellent reliability for alternative gate dielectric application," in *IEEE IEDM Tech. Dig.*, 1999, pp. 133–135.

[22] J. A. Felix, D. M. Fleetwood, R. D. Schrimpf, J. G. Hong, G. Lucovsky, J. R. Schwank, and M. R. Shaneyfelt, "Total-dose radiation response of hafnium-silicate capacitors," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 6, pp. 3191–3196, Dec. 2002.

[23] S. Dixit, X. J. Zhou, R. D. Schrimpf, D. M. Fleetwood, S. T. Pantelides, R. Choi, G. Bersurker, and L. C. Feldman, "Radiation induced charge trapping in ultrathin HfO_2-based MOSFETs," *IEEE Trans. Nucl. Sci.*, vol. 54, no. 6, pt. 1, pp. 1883–1890, Dec. 2007.

[24] J. A. Felix, M. R. Shaneyfelt, D. M. Fleetwood, T. L. Meisenheimer, J. R. Schwank, R. D. Schrimpf, P. E. Dodd, E. P. Gusev, and C. D'Emic, "Radiation-induced charge trapping in thin Al2O3-SiOxNy-Si(100) gate dielectric stacks," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 4, pt. 1, pp. 1910–1919, Aug. 2006.

[25] R. C. Lacoe, J. V. Osborn, D. C. Mayer, S. Brown, and D. R. Hunt, "Total dose radiation tolerance of a chartered 0.35-μm CMOS process," in *1999 IEEE Radiation Effects Data Workshop Record*, 1999, pp. 82–86.

[26] W. Snapp, "New results from the DARPA hardness-by-design program—Boeing," presented at the Microelectronic Reliability Qualification Workshop, Manhattan Beach, CA, 2005, unpublished.

[27] F. Faccio and G. Cervelli, "Radiation-induced edge effects in deep submicron CMOS transistors," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 6, pt. 1, pp. 2413–2420, Dec. 2005.

[28] D. R. Alexander, "Design issues for radiation tolerant microcircuits for space," in *1996 IEEE Nuclear and Radiation Effects Conf. Short Course*, 1996.

[29] N. Nowlin, J. Bailey, B. Turfler, and D. Alexander, "A total-dose hardening-by-design approach for high-speed mixed-signal CMOS integrated circuits," *Int. J. High Speed Elec. Syst.*, vol. 14, pp. 367–378, 2004.

[30] D. C. Mayer and R. C. Lacoe, "Enhanced reliability in commercial CMOS processes by the application of hardness-by-design techniques," presented at the 2002 Micro, Elec. Rel. Qual. Workshop, Manhattan Beach, CA, Dec. 2002 [Online]. Available: http://www.aero.org/conferences/mrqw/2002papers/B_Mayer.pdf

[31] R. N. Nowlin, S. R. McEndree, A. L. Wilson, and D. R. Alexander, "A new total-dose-induced parasitic effect in enclosed-geometry transistors," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 6, pt. 1, pp. 2495–2502, Dec. 2005.

[32] R. C. Lacoe, "Performance impact of radiation-hardness-by-design," presented at the 2001 Hardness-by-Design Workshop, Albuquerque, NM, Jul. 2001.

[33] T. C. May and M. H. Woods, "Alpha-particle induced soft errors in dynamic memories," *IEEE Trans. Elec. Dev.*, vol. ED-26, no. 1, pp. 2–9, Jan. 1979.

[34] D. Mavis, "Single event transient phenomena—Challenges and solutions," in *Proc. Microelectronic Reliability Qualification Workshop*, 2002 [Online]. Available: http://www.aero.org/conferences/mrqw/2002-papers/A_Mavis.pdf

[35] C. M. Hsieh, P. D. Murley, and R. R. O'Brien, "Dynamics of charge collection from alpha-particle tracks in integrated circuits," in *Proc. IEEE Int. Reliability Physics Symp.*, 1981, pp. 38–42.

[36] F. B. McLean and T. R. Oldham, "Charge funneling in n and p-type Si substrates," *IEEE Trans. Nucl. Sci.*, vol. 38, pp. 2018–2023, 1982.

[37] L. D. Edmonds, "A simple estimate of funneling-assisted charge collection," *IEEE Trans. Nucl. Sci.*, vol. 38, no. 2, pt. 1–2, pp. 828–833, Apr. 1991.

[38] C. M. Hsieh, P. C. Murley, and R. R. O'Brien, "Collection of charge from alpha-particle tracks in silicon devices," *IEEE Trans. Elec. Dev.*, vol. 30, no. 6, pp. 686–693, Jun. 1983.

[39] P. E. Dodd, F. W. Sexton, and P. S. Winokur, "Three-dimensional simulation of charge collection and multiple-bit upset in Si devices," *IEEE Trans. Nucl. Sci.*, vol. 41, no. 6, pt. 1, pp. 2005–2017, Dec. 1994.

[40] P. E. Dodd, "Basic-mechanisms for single-event effects," in *1999 IEEE Nuclear and Radiation Effects Conf. Short Course*, 1999.

[41] L. D. Edmonds, A theoretical analysis of steady-state photocurrents in simple silicon diodes Jet Propulsion Lab., Pasadena, CA, JPL Pub. 95-10, Mar. 1995.

[42] L. D. Edmonds, "Charge collection from ion tracks in simple EPI diodes," *IEEE Trans. Nucl. Sci.*, vol. 44, no. 3, pt. 3, pp. 1448–1463, Jun. 1997.

[43] L. D. Edmonds, "Electric currents through ion tracks in silicon devices," *IEEE Trans. Nucl. Sci.*, vol. 45, no. 6, pt. 3, pp. 3153–3164, Dec. 1998.

[44] R. Baumann, "Soft errors in commercial semiconductor technology: Who should be worried and what can we do about it?," presented at the 2002 SEE Symp., Manhattan Beach, CA, Apr. 2002, unpublished.

[45] M. P. Baze, J. C. Killens, R. A. Paup, and W. P. Snapp, "SEU hardening techniques for retargetable, scalable, sub-micron digital circuits and libraries," presented at the 2002 SEE Symp., Manhattan Beach, CA, Apr. 2002.

[46] W. Pederson, *Error-Correcting Codecs*, 2nd ed. Cambridge, MA: MIT Press, 1980.

[47] T. Calin, M. Nicolaidis, and R. Velazco, "Upset hardened memory design for submicron CMOS technology," *IEEE. Trans. Nucl. Sci.*, vol. 43, no. 6, pt. 1, pp. 2874–2878, Dec. 1996.

[48] T. Calin, R. Velazco, M. Nicolaidis, S. C. Moss, S. D. LaLumondiere, V. T. Tran, V. T. Tran, R. Koga, and K. Clark, "Topology-related upset mechanisms in design-hardened memory cells," in *Proc. Radiation and Its Effects on Components and Systems, 1997*, 1998, pp. 484–488.

[49] W. F. Heidergott, "System level mitigation strategies," in *1999 IEEE Nuclear and Radiation Effects Conf. Short Course*, 1999.

[50] Y. Song, K. N. Vu, J. S. Cable, A. A. Witteles, W. A. Kolasinski, R. Koga, J. H. Elder, J. V. Osborn, R. C. Martin, and N. M. Ghoniem, "Experimental and analytical investigation of single event multiple bit upsets in polysilicon load 64 k NMOS SRAMs," *IEEE Trans. Nucl. Sci.*, vol. 35, no. 6, pp. 1673–1677, Dec. 1988.

[51] Y. Tosaka, H. Ehara, M. Igeta, T. Uemura, H. Oka, N. Matsuoka, and K. Hatanaka, "Compreshensive study of soft errors in advanced CMOS circuits with 90/130 nm technology," in *Proc. IEDM*, 2004, pp. 941–944.

[52] H. Quinn, P. Graham, J. Krone, M. Caffrey, and S. Rezgui, "Radiation-induced Multi-bit-upsets in SRAM-based FPGAs," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 6, pt. 1, pp. 2455–2461, Dec. 2005.

[53] F. Wrobel, J. M. Palau, M. C. Calvert, O. Bersilon, and H. Duarte, "Simulations of nucleon-induced nuclear reactions in a simplified SRAM structure: Scaling effects on SEU and MBU cross-sections," *IEEE Trans. Nucl. Sci.*, vol. 48, no. 6, pt. 1, pp. 1946–1952, Dec. 2001.

[54] J. S. Fu, C. L. Axness, and H. T. Weaver, "Memory SEU simulations using 2-D transport calculations," *IEEE Elec. Dev. Lett.*, vol. 6, no. 8, pp. 422–424, Aug. 1985.

[55] R. L. Woodruff and P. J. Rudeck, "Three-dimensional numerical simulation of SEU of an SRAM cell," *IEEE Trans. Nucl. Sci.*, vol. 40, no. 6, pt. 1–2, pp. 1795–1803, Dec. 1993.

[56] P. E. Dodd, F. W. Sexton, G. L. Hash, M. R. Shaneyfelt, B. L. Draper, A. J. Farino, and R. S. Flores, "Impact of technology trends on SEU in CMOS SRAMs," *IEEE Trans. Nucl. Sci.*, vol. 43, no. 6, pt. 1, pp. 2797–2894, Dec. 1996.

[57] O. A. Amusan, A. F. Witulski, L. Massengill, B. L. Bhuva, P. R. Fleming, M. L. Alles, A. L. Sternber, J. D. Black, and R. D. Schrimpf, "Charge collection and charge sharing in a 130 nm CMOS technology," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 6, pt. 1, pp. 3253–3258, Dec. 2006.

[58] B. D. Olson, D. R. Bell, K. M. Warren, L. W. Massengill, N. Haddad, and R. Richards, "Impact of scaling on SEU immunity for a fixed SRAM topology," presented at the Hardened Electronics Radiation Technology (HEART) 2005, Tampa, FL, 2005, unpublished.

[59] J. D. Black, A. L. Sternber, M. L. Alles, A. F. Witulski, B. L. Bhuva, L. W. Massengill, J. M. Benedetto, M. P. Baze, L. Wert, and M. G. Hubert, "HBD layout isolation techniques for multiple node charge collection mitigation," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 6, pt. 1, pp. 2536–2541, Dec. 2005.

[60] M. Huhtinen and F. Faccio, Computational method to estimate single event upset rates in an accelerator environment CERN [Online]. Available: http://www.cern.ch/Atlas/GROUPS/FRONTEND/WWW/seu.pdf

[61] A. D. Tipton, J. A. Pellish, R. A. Reed, R. D. Schrimpf, R. A. Weller, M. H. Mendenhall, B. Sierawski, A. K. Sutton, R. M. Diestelhorst, G. Espinel, J. D. Cresller, P. W. Marshall, and G. Vizkelethy, "Multiple-bit upset in 130 nm CMOS technology," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 6, pp. 3259–3264, Dec. 2006.

[62] K. M. Warren, R. A. Weller, M. H. Mendenhall, R. A. Reed, D. R. Ball, C. L. Howe, B. D. Olson, M. L. Alles, L. W. Massengill, R. D. Schrimpf, N. F. Haddad, S. E. Doyle, D. McMorrow, J. S. Melinger, and W. T. Lotshaw, "The contribution of nuclear reactions to heavy ion single event upset cross-section measurements in a high-density SEU hardened SRAM," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 6, pp. 2125–2131, Dec. 2005.

[63] A. S. Kobayashi, D. R. Ball, K. M. Warren, M. H. Mendenhall, R. D. Schrimpf, and R. A. Weller, "The effect of metallization layers on single event susceptibility," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 6, pp. 2189–2193, Dec. 2005.

[64] C. L. Howe, R. A. Weller, R. A. Reed, R. D. Schrimpf, L. W. Massengill, K. M. Warren, D. R. Ball, M. H. Mendenhall, K. A. LaBel, and J. W. Howard, "Role of heavy-ion nuclear reactions in determining on-orbit single event rates," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 6, pp. 2182–2188, Dec. 2005.

[65] D. G. Mavis and P. H. Eaton, "Soft error rate mitigation techniques for modern microcircuits," in *Proc. 2002 Int. Rel. Phys. Symp.*, 2002, pp. 216–225.

[66] M. P. Baze and S. P. Buchner, "Attenuation of singel event induced pulses in CMOS combinatorial logic," *IEEE. Trans. Nucl. Sci.*, vol. 44, no. 6, pt. 1, pp. 2217–2223, Dec. 1997.

[67] J. Benedetto, P. Eaton, K. Avery, D. Mavis, M. Gadlage, T. Turflinger, P. E. Dodd, and G. Vizkelethyd, "Heavy-ion induced digital single-event transients in deep submicron processes," *IEEE Trans. Nucl. Sci*, vol. 51, no. 6, pp. 3480–3485, Dec. 2004.

[68] D. G. Mavis and P. H. Eaton, "Soft error rate mitigation techniques for modern microcircuits," in *Proc. 2002 Int. Rel. Phys. Symp.*, 2007.

[69] P. Eaton, J. Benedetto, D. Mavis, K. Avery, M. Sibley, M. Gadlage, and T. Turflinger, "Single event transient pulsewidth measurements using a variable temporal latch technique," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 6, pp. 3365–3368, Dec. 2004.

[70] M. P. Baze, J. Wert, J. W. Clement, M. G. Hubert, A. Witulski, O. A. Amusan, L. Massengill, and D. McMurrow, "Propagating SET characterization technique for digital CMOS libraries," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 6, pp. 3472–3478, Dec. 2006.

[71] A. Balasubramanian, B. L. Bhuva, J. D. Black, and L. W. Massengill, "RHBD techniques for mitigating effects of single-event hits using guard-gates," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 6, pp. 2531–2535, Dec. 2005.

[72] J. M. Benedetto, P. H. Eaton, D. G. Mavis, M. Gadlage, and T. Turflinger, "Variation of digital SET pulse widths and the implications for single event hardening of advanced CMOS processes," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 6, pp. 2114–2119, Dec. 2005.

[73] D. R. Alexander, "Hardened by design cell library issues," presented at the 2001 Hardness-by-Design Workshop, Albuquerque, NM, Jul. 2001.

[74] H. N. Becker, T. F. Miyahira, and A. H. Johnston, "Latent damage in CMOS devices from single-event latchup," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 6, pp. 2923–2929, Dec. 2002.

[75] S. C. Moss, S. D. LaLumondiere, J. R. Scarpulla, K. P. MacWilliams, W. R. Crain, and R. Koga, "Correlation of picosecond laser-induced latchup and energetic particle-induced latchup in CMOS test structures," *IEEE Trans. Nucl. Sci*, vol. 42, no. 6, pp. 1948–1956, Dec. 1995.

[76] J. V. Osborn, D. C. Mayer, R. C. Lacoe, S. C. Moss, and S. D. LaLumondiere, "Single event latchup characteristics of three commercial CMOS processes," in *Proc. 7th NASA Symp. VLSI Design*, Albuquerque, NM, 1998, pp. 4.2.1–4.2.14.

[77] S. D. LaLumondiere, R. Koga, J. V. Osborn, D. C. Mayer, R. C. Lacoe, and S. C. Moss, "Wavelength dependence of transient laser-induced latchup in epitaxial CMOS test structures," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 6, pp. 2662–2666, Dec. 2002.

[78] G. Boselli, V. Reddy, and C. Duvvury, "Impact of scaling on the high current behavior of RF CMOS technology," in *Proc. 2003 Int. Rel. Phys. Symp.*, 2003, pp. 229–234.