Soft Error Trends and New Physical Model for Ionizing Dose Effects in Double Gate Z-RAM Cell

Nauman Z. Butt, Student Member, IEEE, P. D. Yoder, Senior Member, IEEE, and Muhammad Ashraful Alam, Fellow, IEEE

Abstract—We model the soft error trends and total ionizing dose (TID) effects in floating body capacitorless DRAM (Z-RAM) cell. We find that soft error rates (SER) in Z-RAM scales strongly with cell body thickness. We propose a new physical model for TID effects in thin (< 10 nm) oxides. This model is based on injection of hot carriers into the oxide which are generated by particle strikes and is relevant for thin oxides in any MOS device. We utilize hydrodynamic approach and Monte Carlo method to implement our model. The primary radiation interaction is simulated by using Geant4—a toolkit for the simulation of particle interaction and transport through matter. The preliminary results of our model match well with radiation-induced leakage current (RILC) vs. dose data.

Index Terms—Geant4, Monte Carlo, soft errors, total ionizing dose, Z-RAM.

I. INTRODUCTION

THE scaling of basic memory cell is essential for getting bigger memory sizes on a limited area available in system-on-chip (SoC) designs. The embedded memory takes more than 70% of the chip area in many of today’s processors and this trend is expected to reach to 90% by year 2011 [1], [2]. Although SRAM is the fastest memory, it needs 4–6 transistors for one memory cell and is therefore more suitable for limited size memories where performance is more crucial than size, e.g., L1 caches for microprocessors. DRAM is not as fast but requires smaller cell area which makes it an attractive option for embedded memories of larger size, e.g., L2 caches. The conventional 1T-1C DRAM has however significant challenges for scaling to smaller fabrication processes [3], [4]. The storage capacitor in 1T-1C cell requires complex geometries which often result in poor yield and added cost. In recent years, capacitor-less floating body DRAM (Z-RAM) has gained a huge interest in embedded memories because of its smaller area, simpler design and easier process integration with logic [1]. A Z-RAM cell consists of only one transistor made of SOI technology which makes it ~5x denser than SRAM and ~2–3x denser than DRAM [2]. Because of its advantages, Z-RAM has a potential of widespread use in commercial embedded memory applications in near future. While the current Z-RAM products are made of single gate partially or fully depleted SOI MOSFETs, the double gate (DG) MOSFETs with ultrathin body is a promising candidate for scaling to future Z-RAM generations. We focus our attention to study the radiation performance of DG Z-RAM cell with dimensions relevant for sub 45 nm technologies.

Radiation effects are amongst the biggest concerns for SoC architectures [9]. With continuous scaling, the critical charge ($Q_{\text{crit}}$) to upset a bit gets smaller which makes single event upsets (SEU) more probable. The collected charge ($Q_{\text{col}}$) by a single particle strike can often become larger than $Q_{\text{crit}}$, hence resulting in high soft error rates (SER). On the other hand, the sensitive volume (SV) for charge collection also becomes smaller with scaling which helps in mitigating SER. A design optimization between the cell dimensions and $Q_{\text{crit}}$ is therefore necessary to ensure a better SEU immunity. We study this optimization for Z-RAM cell by using simple two dimensional (2D) modeling. We show that SER in Z-RAM can be greatly reduced by scaling body thickness of the cell.

Apart from SEU, hard errors pertaining to total ionizing dose (TID) effects remain very important even for devices having ultra-thin gate oxides [10]. In thicker oxides, radiation induced hole trapping used to be a major concern which resulted in threshold voltage shift ($\Delta V_T$). For ultra-thin oxides (<10 nm), the hole trapping can be negligible because the electron tunneling from both electrodes neutralizes the trapped hole charge [10]. There are however some newer concerns for ultra-thin oxides which need careful attention. These include radiation induced leakage current (RILC), radiation induced soft breakdown (RSB) and single event gate rupture (SEGR) [11]–[13]. In addition, the radiation-induced stuck bits remain a major challenge for the error detection and correction schemes employed in memory arrays [14]. There are reports of latent damage which appears in the degradation of time dependent dielectric breakdown (TDBB) characteristics after oxides are exposed to heavy ion irradiation [15]. The MOSFET electrical characteristics have also been reported to show degradation after irradiated by heavy ions [16]. In spite of the large number of experimental findings, the basic physics of TID damage in ultrathin oxides is yet not clearly understood. The physical models which were successfully developed and calibrated for thicker oxides may not be applicable at all for ultrathin oxides.

In this paper, we propose a new physical model which can explain the basic mechanisms of the TID damage in ultrathin oxides. This model is based on the injection of radiation-induced hot carriers from body into the oxide. We show that the oxide damage is proportional to the amount of carriers injected into oxide at a given total dose. The model is first implemented by...
the hydrodynamic approach which can be applied for a limited range of total dose. The Monte Carlo model is then developed to include a wide range of total dose. The primary radiation interaction is simulated using Geant4—a toolkit for simulating particle interaction and transport through matter [17]. The simulation results agree well with experimental data.

This paper is divided in 7 sections. Section I is the introduction. In Section II, we discuss Z-RAM cell operation. Section III is about SEU performance of Z-RAM. In Section IV, we switch from soft errors to TID modeling. Sections V and VI discuss the hydrodynamic and Monte Carlo methods for implementing TID model respectively. Finally, we provide conclusions in Section VII. The list of various topics covered in this paper is summarized in Fig. 1.

II. Z-RAM OPERATION

The WRITE, READ and HOLD operations of a DG Z-RAM cell are illustrated in Fig. 2 [8]. A Z-RAM cell stores a bit by making use of floating body charge. The state “1” is written by applying high drain bias which creates electron-hole pairs near the drain by impact ionization process. The electrons are rapidly swept out of the cell by the electric field while a portion of holes remains trapped in the body, resulting in a net positive charge. The state “0” is written by forward biasing the source to body junction which pushes holes out from body through the source. This results in a shortage of holes in the body. The excess or shortage of holes in the body modulates body potential (V_B) which consequently shifts threshold voltage (V_T) from its equilibrium value.

In the HOLD state, holes are trapped in the body by creating large energy barriers at the source and drain junctions through a suitable applied bias. The state written previously is hence retained for a refresh cycle. Finally, the state is read by sensing the drain current which has distinct levels for each of the two states.

Fig. 3 illustrates the noise margin and retention time characteristics of a Z-RAM cell. The HOLD state of cell is shown for both “0” and “1”. The density of holes in the cell body is consequently either lower or higher than its value at equilibrium. The loss mechanisms, which mainly involve Shockley-Read-Hall (SRH) recombination and the S/D junction leakage, bring holes density back towards its equilibrium value. The rate of the loss mechanisms determine the retention time (~1 sec for the cell shown). One should note that the Q_{coll} from a particle strike can only decrease the noise margin if the state “0” is stored in the cell. For state “1”, Q_{coll} will just add to the excess holes already stored in the body. The cell is hence susceptible to SEU for state “0”. The magnitude of the charge stored in cell in state “0” (Q_{stored}) is equal to the number of holes in body at equilibrium. This has an important role for SEU performance of Z-RAM as discussed in Section III.

III. SOFT ERROR PERFORMANCE

A. Simulation Approach

We use a simple two-dimensional (2-D) model for the soft error analysis of Z-RAM cell. This model is implemented in the device simulator Medici. The generated charge has a Gaussian distribution in both time and the distance perpendicular to the axis of strike, where the characteristic time and distance are
50 fsec and 25 nm respectively. The charge density is obtained by considering classical bulk energy bands and the drift-diffusion equation is solved for the carrier transport. We simulate the transient response of the cell in the **HOLD** state after the data is written. The particle strike is introduced at an initial time $\sim 1\mu s$ during the **HOLD** state and the effects on $Q_{\text{stored}}$ and retention time are observed.

### B. Threshold LET

In a Z-RAM cell, the charge is stored inside the body. The whole body can therefore be considered as the sensitive volume (SV). Unlike DRAM, a Z-RAM cell does not utilize any additional capacitor structures to store a bit. Although this provides a big advantage in terms of density, it can have a negative effect on SER since $Q_{\text{stored}}$ in the body is relatively small for scaled Z-RAM cells. For example, $Q_{\text{stored}} \sim 0.007$ fC for a cell with $L_{\text{ch}}/W = 50\,\text{nm}/100\,\text{nm}$ and $t_{\text{body}} = 10\,\text{nm}$ as shown in Fig. 3. For the same cell dimensions, Fig. 4 shows the simulation result when a particle with increasing LET strikes through the source/body junction of a cell which is in **HOLD** state. With increasing LET, the $Q_{\text{cell}}$ increases which reduces the noise margin. When $Q_{\text{cell}}$ becomes closer to $Q_{\text{stored}}$, the two states essentially become indistinguishable. The $Q_{\text{crit}}$ of a Z-RAM cell is therefore $\sim Q_{\text{stored}}$ (0.007 fC for this cell). Fig. 4 shows that the threshold LET ($\text{LET}_{\text{th}}$) is $\sim 1.5$ fC/um for this cell. Considering a typical range of alpha particle LET (6–10 fC/um), this is quite a small value of $\text{LET}_{\text{th}}$ and implies that every alpha particle hit will flip the bit.

The relative significance of $Q_{\text{stored}}$ to the $\text{LET}_{\text{th}}$ is obvious from the above discussion. For device optimization for soft errors, the challenge is to increase $Q_{\text{stored}}$ without increasing the SV or the $Q_{\text{cell}}$. We will show that scaling $t_{\text{body}}$ in Z-RAM cell can be very effective in fulfilling this design challenge in Section III-D.

### C. Angular Dependence

The angular dependence of $\text{LET}_{\text{th}}$ is described in Fig. 5. Two curves are shown for cells having $t_{\text{body}} = 10\,\text{nm}$ and $L_{\text{ch}}/W = 75\,\text{nm}/100\,\text{nm}$ and $50\,\text{nm}/100\,\text{nm}$. The $\text{LET}_{\text{th}}$ for both cells decreases with increasing angle and is lowest for angle close to 90 degree. This is consistent with the classical RPP model and is due to the fact that the strikes which are more along the horizontal axis can deposit more charge because of longer strike lengths [19]. An important distinction for modeling the angular strikes in case of very thin SV is that the radial track structure can be greater than the cell dimensions. Since SV thickness is equal to $t_{\text{body}}$ for Z-RAM, which is $\sim 10\,\text{nm}$ or smaller for scaled cells, the exact radial shape of the track may therefore play an important role for precise error rate predictions.

### D. Body Thickness Effects

Let us consider two cells X and Y with equal aerial dimensions but with $t_{\text{body},X} = 2 \times t_{\text{body},Y}$. For a normal incident strike, we know that $Q_{\text{cell}}X = 2 \times Q_{\text{cell}}Y$ by simply considering the relative volumes of cells. It would also appear from the same consideration that $Q_{\text{stored}}X = 2 \times Q_{\text{stored}}Y$ implying that $Q_{\text{crit}}$ is the same for the two cells. We show in the following that $Q_{\text{stored}}$ (and hence $Q_{\text{crit}}$) in cell Y can actually be greater than that in cell X. Fig. 6 shows $Q_{\text{stored}}$ as...
as a function of $t_{\text{body}}$ for a cell with $L_{\text{ch}} = 50$ nm. We note that $Q_{\text{stored}}$ first show a sharp increase with decreasing $t_{\text{body}}$ and then settles below $t_{\text{body}} = 3$ nm. The reason of this behavior is that decreasing $t_{\text{body}}$ is very effective against the short channel effect known as drain induced barrier lowering (DIBL) [8]. This is illustrated in the inset B of Fig. 6. The profiles of valence band edge ($E_v$) are plotted for various $t_{\text{body}}$ with $L_{\text{ch}} = 50$ nm along the longitudinal cut through body as shown by dotted line in the inset A. When $t_{\text{body}}$ is decreased, $E_v$ moves up and gets closer to the Fermi level ($E_{\text{FP}}$) due to reduced DIBL. The $Q_{\text{stored}}$ is given by Fermi-Dirac distribution, $Q_{\text{stored}} = N_{v} W \int_{0}^{L_{v}} \int_{0}^{W} \frac{3}{2} [(E_v(x,z) - E_{\text{FP}})/k T] dx dz$, where $E_{\text{FP}} = E_{\text{FP}}(x,z)$ is a Fermi-Dirac integral of order $1/2$, $dx$ and $dz$ are incremental lengths along the channel and thickness direction respectively, $W$ is the width and the $N_{v}$ is a constant. When $t_{\text{body}}$ is reduced from 10 nm to 5 nm, the cell volume is although halved but $Q_{\text{stored}}$ still increases because of its strong dependence on $\Delta E = (E_v - E_{\text{FP}})$. For $t_{\text{body}} \leq 3$ nm, where $\Delta E$ becomes mostly positive, $Q_{\text{stored}}$ dependence on $\Delta E$ does not remain as strong and is cancelled by the effect of smaller $t_{\text{body}}$. The $Q_{\text{stored}}$ is therefore settled around one value in this $t_{\text{body}}$ range.

The increase of $Q_{\text{stored}}$ with decreasing $t_{\text{body}}$, has an important consequence on SER of Z-RAM. Considering again cells $X$ and $Y$ having $t_{\text{body}}$, of 10 nm and 5 nm respectively, then $Q_{\text{cal}}(X) = 1/2 \times Q_{\text{cal}}(X)$ but $Q_{\text{cal}}(Y) > Q_{\text{cal}}(X)$. We thus get two advantages of reducing $t_{\text{body}}$: (i) smaller SV which results in lesser $Q_{\text{cal}}$ and (ii) higher $Q_{\text{cal}}$ due to lesser DIBL. Fig. 7 shows the LET$_{10}$ as a function of $t_{\text{body}}$ for cells having various $L_{\text{ch}}$. The LET$_{10}$ increases considerably while going from $t_{\text{body}} = 10$ nm to $t_{\text{body}} = 2$ nm. In order to get an estimate of SER, we assume an alpha particle fluence of 0.01 hits/cm$^2$-hr having energies between 2–5 MeV which corresponds to LET between 6–10 fC/µm [20]. Three regions indicating different SER ranges are identified based on LET$_{10}$ of the cells. In the lower region where LET$_{10}$ is below 6 fC/µm, every strike will cause an upset. The (failure in 10$^9$ device hours) FIT taken as the product of fluence and the cell area, is 750/Mb, 500/Mb, and 300/Mb for $L_{\text{ch}} = 75$ nm, 50 nm, and 30 nm respectively. In the middle region, LET$_{10}$ is within the range of alpha particle LET spectrum. Not every hit can cause SEU in this region and the corresponding FITs are therefore lower than the maximum values for each cell dimension. Finally in the top region, LET$_{10}$ lies above 10 fC/µm and FIT is minimal. The FIT $\sim 1000$/Mb was reported by an accelerated alpha particle test on a 90 nm Z-RAM technology [21]. The reported FIT is within the high FIT region for $L_{\text{ch}} \sim 90$ nm and is expected for $t_{\text{body}} > 10$ nm as shown in Fig. 7. Although there is no data available for $t_{\text{body}} < 10$ nm, we predict a significant reduction in FIT from our simulation results. We should also note that as we go lower in $L_{\text{ch}}$, the room for improvement in LET$_{10}$ by $t_{\text{body}}$ scaling becomes narrower (compare $L_{\text{ch}} = 75$ and $L_{\text{ch}} = 30$ nm in Fig. 7). The latter has much more DIBL as compared to the former and requires extremely small $t_{\text{body}}$ for improving LET$_{10}$.

E. Discussion

The SER in DG Z-RAM are comparable to the typical SER found in the embedded DRAM which ranges from 100–1000 FITS/Mb [21]. The main challenge for a low SER design in ultra dense Z-RAM is to increase $Q_{\text{stored}}$ in a cell without adding volume. The ultrathin $t_{\text{body}}$ which is effective against short channel effects is a good way of reducing SER. The use of multi gate structures like FINFETs and tri-gate MOSFETs which have inherently better short channel control are also expected to provide a lower SER.

We assume an ideal limit where we ignore any manufacturability problems and the real world defects. Modern techniques for SOI thinning are routinely used in research laboratories to fabricate sub 5 nm thick wafer with the uniformity of $\sim 3$ Å [28]. With the continuing scaling of channel length, the thicknesses studied in this paper are likely to become a necessary part of the Z-RAM cell design in order to control the short channel effects [8], [28]. The manufacturing defects such as thickness and random dopant fluctuations etc. can be considered as second order effects. In addition, the random dopant effects in the Z-RAM cells that are discussed here should be minimal because the cells have intrinsic doping in the body. We therefore expect that the qualitative nature of the soft error trends studied in the ideal limit should not be altered significantly by the manufacturing defects.

Finally, we emphasize that while the relative simplicity of our 2-D model may not be highly accurate in quantitative predictions; the trends discussed are expected to be robust and would provide useful guidance for low SER Z-RAM design.

IV. Hard Errors

Let us now switch our focus from soft errors to hard errors. The hard errors pertaining to TID in the ultra-thin oxides have been of growing interest in recent years. For thicker oxides, the traditional model nicely explains most of the TID effects. Our model focuses on the TID damage in ultra-thin ($\leq 10$ nm) oxides. A radiation strike on a transistor creates a large number of highly energetic e-h pairs in the body which relax back to the thermal energy in a time $\sim 0.1–1$ psec after their generation. During thermalization, however, many of these high energy
carriers can be injected into the oxide either by crossing over 
the energy barrier at the silicon/oxide interface or by tunneling 
through the barrier. This hot carrier injection from the body into 
the oxide can break Si-O bonds and create oxide defects by 
this process. The number of high energy carriers injected into 
the oxide increase with the total dose. At a given dose, the amount 
of oxide damage is proportional to the number of injected carriers. 

The injection of hot carriers into the oxide has been shown 
to be responsible for stress induced oxide defects in thin oxides 
which lead to stress induced leakage current and oxide break-
down [22], [23]. It has been shown that hot holes and hot elec-
trons both can create damage when injected into the oxide. We 
propose in this paper that hot carriers generated by radiation can 
be responsible for TID damage in ultra-thin oxides.

For thick oxides the TID damage is usually imagined to be 
uniformly generated by the total dose which is gradually intro-
duced over a period of time. The contribution from individual 
ionizing particles is often not specifically studied. In contrast 
to this, we model the individual contribution of the particles to 
the damage. The net damage is then estimated from the cumula-
tive contribution of all particles. Since many of the TID effects 
on thin oxides including RSB and SEGR are sensitive to local-
ized damage done by a very few number of particles [11]–[16], 
our model is particularly relevant for these effects. Moreover 
for a given radiation dose, the dependence on particle energy 
and LET can be studied in our model.

We develop two approaches for the model implementation. 
The first approach is based on the hydrodynamic (HD) or energy 
balance method for the radiation deposited carriers and energy. 
This approach is relatively simple but efficient. In the second ap-
proach, we develop a Monte Carlo (MC) method for the relax-
ation of generated carriers back to the thermal energy. In both 
approaches, the oxide damage is estimated by calculating the 
number of injected carriers from the body into the oxide.

V. HYDRODYNAMIC MODEL

The basics of HD model are illustrated in a flowchart of energy 
shown in Fig. 8. A radiation dose deposits energy to electrons and 
holes by the process of ionization. During thermalization, the 
carriers lose their energy in two major ways: (i) one part of energy is 
lost by different scattering mechanisms predominantly with op-
tical phonons and (ii) the other part of it is lost because of the car-
rier injection into the oxide. The optical phonons then further lose 
their energy to acoustic phonons. The part of the carriers which 
is injected into the oxide results in oxide damage.

A. Rate Equations

We solve the following energy balance rate equations for the 
electrons (similar equations apply to holes) [24]:

\[ \frac{dn}{dt} = G_{inj} - J_{ec} - R_{nth} \]  
(1)

\[ \frac{dW}{dt} = G_{inj} - J_{w,ec} - E_{op} \]  
(2)

In above equations, \( n \) is the total number of generated carriers 
(electrons or holes), \( G_{inj} \) and \( G_{inj,ec} \) are the injection rates for 
carriers and energy respectively, \( J_{ec} \) and \( J_{w,ec} \) are the carrier 
and energy flux respectively which is flowing because of the 
charges leaving to oxide from high energy tail, \( R_{nth} \) is the re-
combination rate for the carriers, \( W \) is the total energy, \( E_{op} \) is 
the energy lost due to carrier-phonon scattering.

Similarly, we can write energy balance equations for hot phonons

\[ \frac{dN_{op}}{dt} = (N_{op} + 1)R_{em} - (N_{op}) \times R_{ab} \]  
(3)

\[ \frac{dW_{op}}{dt} = [(N_{op} + 1)R_{em} - (N_{op})R_{ab}]\omega \]  
(4)

where \( N_{op} \) is the number of optical phonon, \( R_{em} \) and \( R_{ab} \) are 
the emission and absorption phonon scattering rates and \( \tau_{op} \) is 
the time constant for acoustic to optical phonon conversion.

B. Simulation Approach

We calculate the total energy deposited and electron-hole 
pairs generated in the cell from the total dose. An electron-hole 
pair creation energy of 3.6 eV is used to estimate the generated 
number of carriers. The total energy and carriers are generated 
in the cell body over a time period of 1 psec. In HD model, we 
need to assume a shape of energy distribution for the generated 
carriers. We adopt a very simplified approach by assuming that 
both electrons and holes maintain Fermi-Dirac distributions 
which are broadened by the deposited energy. The consequence 
of this assumption will be discussed later. The (1)–(4) are then 
solved in a transient simulation.

Fig. 9 shows the energy band diagram along the vertical cut 
through the cross-section of a cell. The initial carrier distribu-
tions in the body are also shown as insets. The initial distribu-
tions are very broad in energy due to the applied dose and have 
long tails reaching above the oxide barrier energy. The distribu-
tions relax to lower energies as a function of time because of the 
phonon scattering and oxide injection. We solve the Schrodinger 
equation in the vertical direction in the transient simulation to 
get the carrier flux through oxide. The simulation completes 
when distributions cool down to the pre-strike thermal energy.

C. Results

We integrate the flux through oxide \( J_{ec} \) due to carrier in-
jection in order to estimate the oxide trap density (\( N_{ox} \)). The 
relation between \( N_{ox} \) and \( J_{ec} \) can be determined by considering a 
simple rate equation for \( N_{ox} \) generation

\[ \frac{dN_{ox}}{dt} = \alpha J_{ec}(N_0 - N_{ox}) \]  
(5)

where \( N_0 \) is the initial density of Si-O bonds and \( \alpha \) is a constant 
which depends on the capture cross-sections.
For the case when $N_0 \gg N_{\text{ox}}$, (5) can be written as

$$dN_{\text{ox}}/dt = \alpha J_{\text{ox}}(N_0).$$  \hspace{1cm} (6)

The integration of above equation gives the relation between $N_{\text{ox}}$ and $N_{\text{inj}}$

$$N_{\text{ox}} = \alpha N_0 \int_0^t J_{\text{ox}} \, dt = \alpha N_0 N_{\text{inj}}.$$ \hspace{1cm} (7)

Equation (7) shows that $N_{\text{ox}}$ is directly proportional to $N_{\text{inj}}$ with the assumption of $N_0 \gg N_{\text{ox}}$. Finally, the RILC is given by

$$J_{\text{RILC}} = \lambda N_{\text{ox}} (f_1 - f_2)$$ \hspace{1cm} (8)

where $f_1$ and $f_2$ are the occupation factors for the anode and cathode respectively and $\lambda$ is a constant which depends on the transmission probability from the trap to the anode or cathode assuming the trap position is at the oxide center. From (7) and (8), it should be clear that RILC is proportional to both $N_{\text{ox}}$ and $N_{\text{inj}}$.

The $N_{\text{ox}}$ vs. dose plot is shown in Fig. 10. This is compared with the RILC vs. dose data from [12]. Fig. 10 shows that the slope of $N_{\text{ox}}$ vs. dose matches well with RILC vs. dose data up to 10 Mrad (Si).

### D. Discussion

Despite the simplicity and efficiency of HD model, its application is rather limited. The main shortcoming is in the assumption that the generated carriers maintain equilibrium-like energy distribution shapes before the thermalization is completed. In the real situation, the distributions of the generated carriers are likely to have significantly non-equilibrium distributions before relaxing back to the thermal energy. Even with such a simplification, the fact that the results match the experimental data for a range of dose (0.1–10 Mrad) is surprising. Perhaps the reason is that for total dose below a certain limit, the integrated density of the generated carriers is relatively small and the Fermi-Dirac distributions broadened by the deposited energy incidentally provide right trends in the results. Regardless of its limitations the HD model provides a convenient way of understanding the theoretical framework of our TID model. The limitations of HD model necessitate the use of Monte Carlo method which is discussed next.

### VI. MONTE CARLO MODEL

In the MC model, we start with the simulation of primary radiation interaction in the body of Z-RAM cell. This is done in Geant4—a high energy based simulation toolkit for the passage of particles through matter [18]. We use Geant4 to get the initial distribution of carriers that are generated inside the cell body by a radiation dose. After getting the initial carrier distribution, we perform full-band Monte Carlo simulations with required scattering mechanisms which solve the Boltzmann equation to get carrier distributions as a function of time. We then solve the transport equation along the vertical direction to calculate the transient flux of the carriers through the oxide.

#### A. Primary Radiation Interaction—Geant4 Simulation

The primary radiation strike in silicon produces secondary (delta) electrons with a wide spectrum of energies ranging form a few eV up to a few keV through the ionization of silicon atoms. The high energy delta electrons then lose their energy by creating more secondary electrons through further ionizations. This is simulated by using the low energy electromagnetic package provided by Geant4. The low energy package is based on cross sections data for ionization and atomic relaxation for energies down to a few eV. The cross sections are extracted from a set of publicly distributed data libraries [25].

Fig. 11(a) shows a two-dimensional view of primary and secondary particle trajectories from a Geant4 simulation which includes 50,000 events of 8 MeV electron strikes through a Z-RAM cell. The primary particles are normally incident from top of the cell. The lines that branch out of the vertical lines are tracks of secondary electrons. The circles indicate a delta incremental step to the track. There are several levels of branches in a track which indicate the hierarchical generation of secondary electrons. Fig. 11(b) shows the initial energy of secondary electrons as a function of generation time measured from the start of the primary electron track. The secondary electrons have a broad spectrum of initial energy with most of them in 1–10 eV range while the time span of the initial generation goes up to a few fsec.
We let the Geant4 based simulation run until all of the high energy secondary electrons relax down to $\sim 6$ eV. As a result of this, we get a distribution of the generated electrons in the range of 0.1–6 eV. This distribution is then used as an initial distribution for the Monte Carlo simulation.

An important point to note is that although the individual radiation strikes can be far apart in time depending on the dose rate, relaxation is simulated cumulatively for the sum of all carriers generated by the total dose. The distributions at any time in our simulations are hence the cumulative distributions of the generated carriers.

B. Monte Carlo Simulation

The low energy package of Geant4 toolkit does not include all aspects of the carrier transport when the energy is very low [25]. In order to properly account for the carrier relaxation in the lower energy range (0.1–6 eV), we use a full-band ensemble Monte Carlo simulator, described in [26]. The Monte Carlo simulation included phonon scattering, ionized impurity scattering and impact ionization. We use an approach called Transition matrix approach (TMA) to solve the time dependant but space independent Boltzmann equation: 

$$
\partial f/\partial t = S_{\text{op}} f = \sum_{p'f'} f(p') S_{\text{op}} f(p') - f(p) S_{\text{op}} f(p),
$$

where $S_{\text{op}}$ is the scattering operator, $p$ and $f$ are the current and the outgoing momentum states respectively [27]. The TMA is a table based Monte Carlo approach which is used to reduce a large computational time and the statistical noise of a direct Monte Carlo simulation. In TMA the scattering probabilities are pre-computed and tabulated. The transition matrix (T) is a square matrix in which each element $T_{ij}$ represents the probability of carrier transition from an input energy $E_j$ to output energy $E_i$ in a short time step $\Delta t$.

We pre-computed the elements $T_{ij}$ by running the full band ensemble Monte Carlo simulation in which 50,000 electrons were launched at random locations on the isoenergy surfaces for a given energy $E_j$ and the subsequent relaxation is observed for $\Delta t = 10$ fs. $T_{ij} = N_j/N_i$, where $N_j$ and $N_i$ are the number of carriers in the output and input energy bins respectively. Once the transition matrix is constructed the $f(t)$ is simply calculated by iterating the matrix multiplication: 

$$
f(t+n\Delta t) = T^n f(t),
$$

where $n$ is an integer and $f(t+n\Delta t)$ is the distribution function after time $n\Delta t$. In this way the time consuming full-band ensemble Monte Carlo simulation is amortized over many simulations and since a large number of particles are used to define $T_{ij}$, statistical noise is virtually eliminated. Fig. 12 shows the $f(t)$ vs. energy curves for various times. It should be noticed that the distribution tail is still populated for times $> 100$ fsec. While solving for $f(t)$, we also calculated the carrier flux through oxide by solving the Schrodinger equation along the vertical (thickness) direction of the cell. The $N_{\text{ox}}$ vs. dose results from the MC model are plotted in Fig. 13. The RILC vs. dose results from two separate experiments are also plotted [11, 12]. The slope of MC model curve ($\sim 1$) is close to the slopes of RILC vs. dose data for a wide range of total dose. The MC model result however do not show any decrease in the slope for total dose $> 30$ Mrad(Si) which is observed in data from [12]. The decrease in slope can be due to the cumulative effect of $N_{\text{ox}}$. By looking at (5) in the Section V-C, we note that if $N_{\text{ox}}$ becomes closer to $N_0$, the $N_{\text{ox}}$ vs. $N_{\text{inj}}$ relation would not be linear as the drop in slope can be expected. Since we assume $N_0 \gg N_{\text{ox}}$ in our calculations, our results implicitly neglect any possible cumulative effect due to $N_{\text{ox}}$.

Fig. 11. (a) Cumulative track structure of 50,000 events of 8 MeV electrons normally incident on a Z-RAM cell. The secondary electron tracks can be seen as lines branching out of the primary (vertical) tracks. The dots indicate a delta step to the tracks. The hierarchical structure of secondary electron generations can be observed (b) The initial energy of secondary (delta) electrons is plotted as a function of the time of generation. The time is measured as the interval between the beginning of the primary (8 MeV) track and the generation of the secondary electron.

Fig. 12. Energy distributions of generated electrons for various times from the MC model. The initial distributions have a large number of high energy carriers. The scattering mechanisms relax the high energy carriers to lower energies with the passage of time. The high energy carriers contribute to the flux through oxide which can create oxide damage.

Fig. 13. Oxide trap density vs. dose from MC model is plotted and compared with RILC vs. total dose data from two separate experiments. The slopes of the simulation curve matches well with the experiments for a wide range of dose. The change of slope observed at high dose in one of the experiments may be due to the cumulative effects of the oxide traps.
VII. CONCLUSION
The better scaling prospects of Z-RAM and its easier process integration with logic leverage its potential use for widespread commercial embedded memory applications in near future. Radiation effects in Z-RAM are important not only for terrestrial microelectronics but also for space applications because of their increasing interest in commercial parts. We have modeled the soft error and the hard (TID) errors for the capacitor-less floating body Z-RAM cell. We study the effect of cell scaling on SER performance. The body thickness scaling of cell is found to be very effective in mitigating SER. We propose a new physical model for TID damage in ultra-thin (< 10 nm) gate oxides. This model is based on the mechanism of charge injection into oxide due to the hot carriers generated in the cell body by radiation strikes. Although we apply this model to Z-RAM, it is relevant for thin oxides in any kind of MOS device. We develop hydrodynamic (HD) and Monte Carlo (MC) methods to implement the theoretical model. The HD model is based on the energy balance solution for carriers and phonons generated from a radiation dose. Even though we find a good match with experimental data in some dose range, this model has limitations due to the assumption of oversimplified carrier distributions. In MC model, we use full-band ensemble Monte Carlo with appropriate scattering mechanisms to simulate carrier distributions for energy range (≤ 6 eV). For carriers generation and the subsequent relaxation in higher energy range (from 6 eV to keVs) we utilize Geant4 based high energy physics simulations. We obtain a good match with experimental data by our model for a wide range of total dose.

ACKNOWLEDGMENT
The authors would like to thank Geant4 collaboration team for their helpful discussions. They would also like to thank the Geant4 Team at Stanford Linear Accelerator for teaching a free Geant4 tutorial course.

REFERENCES