Identification of Physical Mechanisms of Radiation Effects in 3D NAND Solid-State Memories using Monte Carlo Analysis

By

Matthew L. Breeding

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Interdisciplinary Materials Science & Engineering

August 13th, 2021

Nashville, Tennessee

Approved:

Robert A. Reed, Ph.D.

Ronald D. Schrimpf, Ph.D.

Michael L. Alles, Ph.D.

Sokrates T. Pantelides, Ph.D.

D. Greg Walker, Ph.D.

Acknowledgments

First, I would like to thank my dissertation committee, Vanderbilt University, and the interdisciplinary materials science program for the chance to present this research and for an excellent experience in graduate school. Thank you to my adviser, Prof. Robert Reed, for your patience, encouragement, and mentorship over the years, and for allowing a good deal of independence as my research progressed.

Thank you to Prof. Ron Schrimpf for giving me the initial opportunity to pursue research in radiation effects at Vanderbilt, for your insightful and thought-provoking questions, and for the best lectures in device physics I have seen. Thank you as well to Prof. Mike Alles for the tirelessness with which you champion the research efforts of the graduate students, and for your commitment to student development and growth.

I am grateful to the Defense Threat Reduction Agency for funding the research presented here, and for Sandia National Laboratories for providing an excellent experience as an intern and the chance to work on interesting projects with great people. Thank you especially to Jeff and Dolores Black for all that they've done to help develop my skills at Sandia and as a researcher.

Thank you to Prof. David Thompson at Sam Houston State University for providing my first opportunity to conduct scientific research. It is difficult for me to overemphasize the impact of Professor Joel Walker in physics at SHSU, and it is safe to say I would not be where I am today if not for his extensive guidance, instruction, and mentorship.

I am grateful for the abundant opportunities he provided to learn and apply beyond the standard coursework, for the opportunities he enabled to present my research at conferences, and for introducing me to vim. His influence and support continued well beyond Sam Houston,

helping me through the challenges and uncertainties of starting a new research program at Vanderbilt. Thank you for all that you have done, Professor Walker.

A special thank you as well goes to Prof. Kevin Warren, whose patience in hindsight was remarkable, and who spent numerous hours teaching the details and constructively challenging my understanding of Monte Carlo techniques with MRED. Thank you to Ted Wilcox at NASA for an engaging collaboration, and Prof. Brian Sierawski for putting us in contact. Thank you to all of the teachers at Providence Classical School, especially Mr. Keefe and Donna Grimley who laid a strong foundation for a lifelong love of learning.

Thank you to all of the students in the radiation effects research group at Vanderbilt; you have been excellent people to work with, and I especially appreciate your patience with my tendency to stretch the bounds of (and at times trample) acceptable ACCRE usage over the years. Thank you to Ellis Richards for your engaging conversations and insights, to Drew Tonigan for your inquisitive spirit and helpfulness, and to Daniel Vibbert for the unique perspectives you're always sure to bring to the table.

Thank you to my mother and father-in-law Acenith and Brian, who have been supportive and encouraging to me, Liz, and Jo while we have been finishing up at Vanderbilt.

A very grateful thank you to my parents, Steve and Barbara, for all of the sacrifices you have made over the years, for your commitment to my education and development, and the importance of hard work and curiosity that you continuously instilled from an early age.

Finally, thank you to my beautiful wife Elizabeth and daughter Josephine -- your encouragement, support, patience, and faith has made this possible.

Table of Contents

Acknowledgments ..... **iii**

List of Figures ..... **vii**

Chapter 1 Introduction and background **1**

I. Relevant radiation effects physics and terminology **3**

A. Direct Ionization 4

B. Indirect Ionization 8

C. Photon-matter interactions and dose enhancement 10

II. Monte Carlo radiation transport to predict device characteristics using MRED **15**

III. General structures and features common to 2.5D and 3D integrated circuits **23**

IV. Charge trap and floating gate 3D NAND memories **29**

A. General 3D NAND structures 29

B. Floating gate (FG) vs charge trap (CT) 3D NAND 34

C. Literature review of radiation effects in 3D NAND 40

Chapter 2 Simulation results of radiation effects in relevant 3DIC structures **50**

I. TSVs in neutron environments **51**

II. Dose-enhancement effects from vertically stacked metallization layers **59**

Chapter 3 Direct ionization from low-energy protons in CT 3D NAND **70**

I. Experimental details **71**

II. SRIM simulations **74**

III. Proton testing results 76

IV. Initial MRED simulation structure for 3D NAND 81

V. Resolving differences with experimental data 87

VI. Updated and higher-fidelity simulations of the 3D NAND structure 91

Chapter 4 Heavy-ion-induced upsets in 3D NAND 105

I. Experimental details for SEU in 72-layer CT 3D NAND 106

II. Characterizing heavy-ion SEU in multi-die 3D NAND packages 109

Chapter 5 Total dose effects and mechanisms in 3D NAND 131

I. Monte Carlo simulation details for dose enhancement in CT and FG 3D NAND 132

II. Contributions to TID-induced ΔV_{th} based on physical memory configurations 134

Chapter 6 Conclusion 139

Appendix A 3D Dose Enhancement Effect – Extra Data for monolayers, alternative spacings 142

References 149

List of Figures

Figure 1.1. Cartoon of ion traversing a silicon cube, generating electron-hole pairs along the track 6

Figure 1.2. Cartoon representation of a heavy-ion striking the sensitive junction of a transistor and generating a transient current in the inverter [3]. 7

Figure 1.3: The Bragg peak for various ions in silicon [4]. 8

Figure 1.4 Possible neutron reaction classified according to the paradigm of elastic vs. inelastic (top), or absorption vs. scatter (bottom) [5]. 9

Figure 1.5. 10-keV peaked X-ray spectrum, used in TID testing [9]. 10

Figure 1.6. 57-keV peaked Bremsstrahlung X-ray spectrum, used in TID & dose enhancement testing for [10]; the spectrum was produced using an X-ray converter with a Pelletron at Sandia. 11

Figure 1.7. Dominant mechanisms of photon-solid interactions as a function of photon energy and atomic number of the target material [10]. 12

Figure 1.8. Basic representation of the dose enhancement mechanism for X-rays and gamma-rays [12]. 14

Figure 1.9. Block diagram of the components of the MRED program [15]. 16

Figure 1.10. The default configuration and axes used in the MRED simulation world. Device geometries are conventionally constructed from negative to 17

positive in the Z-direction, with a randomization sphere entirely enclosing the device.

Figure 1.11. Track structure of neutron interactions in MRED. The primary neutron undergoes inelastic scatters which produces secondary particles. These recoiling secondaries deposit energy in the silicon substrate with direct and indirect ionization mechanisms, both of which are trackable in MRED. 18

Figure 1.12. More Moore vs. More-than-Moore paradigm diagram. Efforts began over a decade ago to shift the focus to novel integration architectures as opposed to fabricating smaller transistors [25]. 24

Figure 1.13. Flowchart of various fabrication methods for 3DICs [28]. 25

Figure 1.14. Overview of the 3-D technology landscape, as defined by the 3-D insertion point in the electronic system wiring hierarchy" (from [29]). 26

Figure 1.15. Three examples of common methods for achieving 3D integration between two chips. From left to right: face-to-face (F2F), face-to-back (F2B) and back-to-back (B2B) [29]. 26

Figure 1.16. Circuit-level diagram showing the p-BiCS NAND configuration. 31

Figure 1.17. SEM images showing the evolution of the SK Hynix 3D NAND stack from initial configurations with 32 layers to the subject of the present study at 72 active layers [38]. 31

Figure 1.18. Close-up SEM image of the interface between the upper and lower region in the NAND stack [38]. 32

Figure 1.19. SEM image of a Hynix 32-layer NAND. The vertically etched slits are clearly visible as dark regions between the word-line contacts [39]. 33

Figure 1.20. Representation of the different operating windows for NAND memory cells. Increased memory density is achieved in multiple-bit cell operating modes at the cost of decreased reliability due to the more narrow Vth windows [41]. 34

Figure 1.21. Planar (2D) charge trap transistor (left) and floating gate transistor (right) [43]. 35

Figure 1.22. SEM images of a FG 3D NAND cell. Note the lack of metal layers; FG processes use polysilicon for electrical connectivity throughout the memory array [44]. 36

Figure 1.23. Fabrication process for FG-3D NAND. This is a gate-first method, as opposed to CT NAND. Alternating layers of oxide polysilicon are deposited on a substrate, followed by high aspect ratio etches for the string channels. The control gate layers are recessed and an inter-poly dielectric is deposited, followed by the floating gate material, a tunnel oxide, and then the polysilicon channel which runs the entire length of the string [44]. 37

Figure 1.24. Gate-last process used to fabricate CT NAND [48] 38

Figure 1.28. Observation of increased bit error rates as a function of cell operating conditions [46]. 42

Figure 1.29. SEU for CT 3D NAND in heavy-ion environments and as a function of operating mode [39]. 43

Figure 1.30. The proton-induced SEU cross section for the 3D CT NAND operated at SLC and MLC mode [39]. 44

Figure 1.31. The fluence dependence of SEU in CT 3D NAND. Increasing fluence 44

results in close to an order of magnitude variation in the SEU susceptibility of the chip [39].

Figure 1.32. Results presented for 32-layer FG 3D NAND irradiated with a 10-keV X-ray source [52]. 45

Figure 1.33. Proposed mechanism for the increase in TID-induced ΔVth observed in the lower layers of a FG 3D NAND [52]. 46

Figure 1.34. Predicted ΔV_{th} variability in FG cells as a function of feature size [53]. 47

Figure 1.35. SEU cross section as a function of LET for 3D FG NAND in SLC and MLC mode [54]. 48

Figure 2.1. Simulation structure for TSV in MRED. 52

Figure 2.2. Example of integrated cross section plot generated from the energy deposition histograms for sensitive regions below a TSV. The vertical lines denote energy thresholds, i.e., the relative probability of depositing a certain amount energy (or generating a certain amount of charge) greater than or equal to the threshold value. 54

Figure 2.3. Critical charges for SEUs in SRAM technology as a function of feature size [61]. 55

Figure 2.4. Material response to 30 MeV neutrons; secondaries produced in the metal layers of the given species tend to not produce energetic recoils as frequently as those generated in silicon at 30 MeV. 56

Figure 2.5. 300 MeV neutron response. At higher thresholds, tungsten and ruthenium result in significantly greater charge generation within the sensitive volumes than observed in other metals. 56

Figure 2.6. Integrated cross section ratio for all four metal configurations with Qcrit = 0.5 fC. A large material-dependent response is seen for sensitive volumes the closer they are to the TSV structure. This response becomes less pronounces at higher Qcrit. 57

Figure 2.7. 300 MeV neutrons result in a much less pronounced materials-dependent effect. 58

Figure 2.8. Schematic of nuclear reaction occurring within a via layer in the BEOL of a circuit [3]. 59

Figure 2.9. Cartoon representation of the simulation structure for examining dose enhancement in chips with multiple vertical levels of metallization. 61

Figure 2.10. Dose enhancement at low and medium X-ray energies for 100 nm of metal (spaced every 1 μm). 62

Figure 2.11. Dose enhancement at low and medium X-ray energies for 100 nm of metal (spaced every 5 μm). 63

Figure 2.12. Visualization of electron tracks within the simulated structure for incident X-ray-energy gamma primary particles. There is no discernible directionality to the spreading of the electrons (qualitatively; quantitatively there are directional dependencies) leading to a dose-enhancement effect on both sides of the metal layers relative to the beam path. 65

Figure 2.13. Dose enhancement at low and medium X-ray energies for 25 nm of metal (spaced every 1 μm). 65

Figure 2.14. Dose enhancement at low and medium X-ray energies for 25 nm of metal (spaced every 5 μm). 66

Figure 2.15. Dose enhancement at the two characteristic γ-ray energies for 250 nm of metal (spaced every 1 μm). 68

Figure 2.16. Dose enhancement at the two characteristic γ–ray energies for 250 nm of metal (spaced every 5 μm). 69

Figure 3.1. Decapsulated eight-die H25QFT8F4A9R-BDF 3D NAND package mounted on test board for heavy ion (see Chapter 4) and proton experiments. 72

Figure 3.2. Reverse view of the test board mounted in the Van de Graff accelerator for proton testing (top), with a closeup of the aluminum masking plate on the front of the PCB (bottom). 73

Figure 3.3. Close-up SEM of the 72-layer memory array. 75

Figure 3.4. Simulation of proton LET as a function of depth in a monolithic silicon block for the proton energies used experimentally. Protons of different energies enter from the left at a depth of 0 μμ and slow as they travel deeper through the material. The energy loss from this slowing is plotted as the instantaneous LET on the ordinate. 75

Figure 3.5. SEU cross-section vs. initial proton energy. 77

Figure 3.6. Histograms of errors by layer as proton energy increases from 500 keV (left) to 1000 keV (right). 78

Figure 3.7. Expected number of cells with single and multiple particle strikes for given fluences, overlaid with heavy ion and proton experimental data. 79

Figure 3.8. SEU cross section as a function of proton energy. Pre-irradiation errors 81

which do not re-appear in the post-irradiation data are not considered.

Figure 3.9. Initial simulation structure in MRED. 83

Figure 3.10. Depth profile of simulated proton-induced upsets in 3D NAND memory array using the simplified configuration as shown in Figure 3.9. 86

Figure 3.11. Simulation data (in red) shown together with the experimental proton cross section. 86

Figure 3.12. Peripheral-under-cell (PUC) architecture is used for Hynix 3D NAND devices [49].

Figure 3.13. Two different peripheral circuitry configurations used in 3D NAND devices [63]. 91

Figure 3.14. Upset distribution as a function of depth for selected energies. 92

Figure 3.15. Render of the advanced 3D NAND simulation structure in MRED, including the contact fanouts and TSVs. 94

Figure 3.16. Render of the 3D NAND cell geometry viewed in the bit line direction. The WL material is shown in green (tungsten and silicon used), with the pink showing N-filled air gaps. The circle regions correspond to the gate-all-around transistors which make up the cells of the memory array. 95

Figure 3.17. MRED 3D NAND model, with focus on the transistor structure. 95

Figure 3.18. Detailed MRED simulation structure as viewed in the WL direction. The BEOL is structure based on the p-BiCS configuration, with monolithic copper placed at level M3 and above in the simulation. 96

Figure 3.19. Simulated SEU probability distribution per layer for 800 MeV protons. 97

Figure 3.20. Depth distribution of proton-induced upsets from the high fidelity MRED model. 98
Figure 3.21. Single event upset cross section range based on range of Qcrit values corresponding to SLC operation as a function of initial proton energy. 99
Figure 3.22. Depth distribution of proton-induced upsets from the high fidelity MRED model for proton incidence of 15◦. 99
Figure 3.23. Depth distribution of proton-induced upsets from the high fidelity MRED model for proton incidence of 45◦. 100
Figure 3.24. Depth distribution of proton-induced upsets from the high fidelity MRED model for proton incidence of 70◦. 100
Figure 3.25. Single event upset cross section range based on range of Qcrit values corresponding to SLC operation as a function of initial proton energy for a range of incidence angles. 102
Figure 3.26. Single event upset cross section range based on range of Qcrit values corresponding to MLC operation as a function of initial proton energy for a range of incidence angles. 102
Figure 3.27. Single event upset cross section range based on range of Qcrit values corresponding to TLC operation as a function of initial proton energy for a range of incidence angles. 103
Figure 3.28. Trapped proton spectrum used for event-rate simulations. 103
Figure 3.29. SEU error rate for the 512 Gb/die / 8 die 72-layer Hynix 3D NAND as a function of critical charge. 104
Figure 4.1. Simulation results (left) using the simplified 3D NAND model and 108

experimental data (right) for SEU cross section as a function of LET.

Figure 4.2. Heavy ion tracks in the 3D NAND memory stack. 109

Figure 4.3. Multi-die simulation structure. 110

Figure 4.4. Example of multi-die 3D NAND package [64]. Each die for stacks up to and in some cases over 1 mm contain individual stacks of 3D NAND memory. 111

Figure 4.5. SEU cross section, as predicted by MRED, as a function of die in a multi-die package and depth within each die for 3D NAND operating in SLC mode. 113

Figure 4.6. SEU cross section, as predicted by MRED, as a function of die in a multi-die package and depth within each die for 3D NAND operating in MLC mode. 113

Figure 4.7. SEU cross section, as predicted by MRED, as a function of die in a multi-die package and depth within each die for 3D NAND operating in TLC mode. 114

Figure 4.8. Cumulative cross sections, as predicted by MRED, over all NAND arrays in 8-die package for cells operated in SLC mode (CT region sensitive). 116

Figure 4.9. Cumulative cross sections, as predicted by MRED, over all NAND arrays in 8-die package for cells operated in MLC mode (CT region sensitive) 117

Figure 4.10. Cumulative cross sections, as predicted by MRED, over all NAND arrays in 8-die package for cells operated in TLC mode (CT region sensitive). 118

Figure 4.11. Cumulative cross sections, as predicted by MRED, over all NAND arrays in 8-die package for cells operated in SLC mode (CT + tunnel oxide 119

sensitive).

Figure 4.12. Cumulative cross sections, as predicted by MRED, over all NAND arrays in 8-die package for cells operated in MLC mode (CT + tunnel oxide sensitive). 120

Figure 4.13. Cumulative cross sections, as predicted by MRED, over all NAND arrays in 8-die package for cells operated in TLC mode (CT + tunnel oxide sensitive). 121

Figure 4.14. Cumulative cross sections, as predicted by MRED, over all CT NAND arrays in 8-die package for cells operated in all three modes. Energy deposition collected for both the charge trap regions for each cell. 124

Figure 4.15. Cumulative cross sections, as predicted by MRED, over all CT NAND arrays in 8-die package for cells operated in all three modes. Energy deposition collected for both the charge trap and tunnel oxide regions for each cell. 125

Figure 4.16. Cumulative cross sections, as predicted by MRED, over all FG NAND arrays in 8-die package for cells operated in SLC mode 126

Figure 4.17. Cumulative cross sections, as predicted by MRED, over all FG NAND arrays in 8-die package for cells operated in MLC mode. 127

Figure 4.18. Cumulative cross sections, as predicted by MRED, over all FG NAND arrays in 8-die package for cells operated in TLC mode. 128

Figure 4.19. Cumulative cross sections, as predicted by MRED, over all FG NAND arrays in 8-die package for cells operated in all three modes. Energy deposition collected for both the charge trap and tunnel oxide regions for 129

each cell.

Figure 4.20. The cross section for cells positioned closer to the tungsten word lines is noticeably larger at higher LET than for those positioned farther away. 130

Figure 5.1. Simulated total dose deposited per layer normalized to the dose in the top layer for each specified run. These data correspond to FG NAND (with polysilicon WLs). 135

Figure 5.2 Simulated total dose deposited per layer normalized to the dose in the top layer (closest to the BEOL) for CT NAND (tungsten metal WLs). 136

Figure 5.3. The distribution of simulated DEF as a function of depth and testing environment with normal incidence photons (through the BEOL). 137

Figure 5.4. The distribution of simulated DEF as a function of depth and testing environment with photons incident through the backside. 138

Chapter 1

Introduction and background

This dissertation presents an analysis of the unique physical configurations used in 3D NAND memories and how the presence of 3D structures is responsible for the modulation of local radiation environments with respect to single-event effects (SEE) and total ionizing dose (TID) effects. This physical approach requires a functional counterpart dealing with architectural and circuit-level considerations for a full characterization of these highly complex chips when exposed to radiation effects, but such functional analysis is beyond the scope of this work. Here, the emphasis is on understanding how the many physical configuration options available in 3D NAND technologies contribute to and modulate the radiation response of the devices.

The uniquely 3D structure of these memory chips gives rise to multiple physical mechanisms for modulating local radiation environments and partially determining device performance. These structures include U-shaped bit lines, multiple vertical stacks of high-Z metals, N-filled voids between word lines, channel irregularities including non-uniform radii as a function of depth and misalignments between columns in multi-etch processing, multiple chips included within the same package up to total depths greater than 1 mm, word line contact fan-outs adjacent to rather than above or below active devices, and the gate-all-around transistor construction.

Each of these features found in 3D NAND (and in 3DICs in general) has the potential to significantly alter the distribution and manifestation of energy deposition in and near sensitive

device regions both with respect to TID and SEE. Focused on the physical rather than functional, simulation and modeling tools with the capabilities of arbitrarily complex device construction and physics-based radiation transport provide a powerful tool in understanding and unpacking these mechanisms.

The first chapter addresses background information regarding the physics and mechanisms of radiation effects in microelectronic circuits with a focus on relevance to the later discussion of 3D NAND memories, including stopping power and direct ionization, indirect ionization, and photon-matter interactions.

The Monte Carlo radiation transport tool MRED is then introduced, which is used extensively throughout the work both in constructing models based on experimental data to account for radiation effects in the presence of 3D integrated circuit (3DIC) components, both at the general and specific level. Validated models of the physical device are extended in later chapters to account for overall chip response in multi-die packages in heavy-ion environments and to provide a distinction between total-dose response of two of the most common commercially available 3D NAND chips.

The evolving landscape of 3DICs in general with its unique terminologies and technologies is also briefly introduced, followed by a much more detailed specific look at 3D NAND memories, including configurations, fabrication, and architectures. The current state of the art research in understanding radiation effects in 3D NAND is then presented from the available literature for both SEE and TID.

Following these introductions, the next chapter presents a series of simulation results using MRED which explain how the characterization of energy deposition from heavy-ion, proton, and neutron environments is partially determined by the presence of general features

common to 3DICs (included in most 3D NAND devices) such as through-silicon vias (TSVs) and multiple vertically stacked metallization layers.

After the review of device-agnostic general 3DIC structures, experimental and simulated data are presented in detail for a 72-layer 3D NAND chip from SK Hynix in which single-event upsets (SEU) were observed from directly ionizing protons at energies ranging from 500 keV to 1.2 MeV. Iterative improvements to the MRED model are described which follow a trend of increasing complexity in order to better account for the physical mechanisms involved in actual devices.

The fourth chapter addresses heavy-ion SEU in the 72-layer memory, both at the experimental and simulation level. The modeling efforts extend the analysis to similar technologies from other manufacturers which use a floating-gate process, and multi-bit upsets (MBU) are discussed within the context of the different cell operating modes of the device. The validated simulation model is then extended to provide insight into the cross-die characterization of heavy-ion SEU at both high and low kinetic energies, with implications for future testing and parts qualification.

Chapter 5 concludes the technical section of the dissertation with an analysis of the dose enhancement effect compared between the two dominant types of 3D NAND (floating gate and charge trap) which take into account the presence or absence of multiple vertically stacked metallization layers.

I. Relevant radiation effects physics and terminology

This section will introduce the basic mechanisms requisite to understanding the subsequent treatment of energy deposition in specific 3D NAND memory circuits. Specifically,

direct and indirect ionization mechanisms and the associated proton and heavy-ion environments are presented, along with photon-matter interactions at X-ray and gamma energy ranges and the dose-enhancement.

Particles traveling in matter undergo energy loss through a variety of physical mechanisms which are dependent on the particle species and energy, as well as the properties of the material through which it travels. Particles are generally classified as either ionizing or non-ionizing, which refer to the primary mechanism of energy deposition for the initial particle of interest.

A. Direct Ionization

Charged particles passing through materials lose energy in a process known as stopping. Equivalently, energy is deposited in a target material when an interacting particle travels through it. The energy loss per unit path length dE/dx describes how much energy is lost by the particle (or deposited in the material) and is described by the Bethe-Bloch equation [1]:

$$-\frac{dE}{dx} = 2\pi N_a r_e^2 m_e c^2 \rho \frac{Z z^2}{A\beta^2}\left[ln\left(\frac{2m_e\gamma^2 v^2 W_{max}}{I^2}\right) - 2\beta^2 - \delta - 2\frac{C}{Z}\right] \qquad (1.1)$$

where N_a is Avogadro's number, r_e and m_e are the classical electron radius and mass, c is the speed of light, r is the density of the target material, Z and A are the atomic number and mass of the target material, z is the charge of the incident particle, v is the velocity of the particle, b is the standard v/c, g is relativistic term $\sqrt{1 - {v^2}/{c^2}}$, I is the mean excitation potential, and W_{max} is the maximum energy transfer possible in a single collision, and C, d are the shell and density correction terms, respectively.

Stopping power, when used in the context of this dissertation and in the field of radiation effects in microelectronics in general, refers to the density-normalized energy loss per unit path length in a given material. This is commonly called Linear Energy Transfer, or LET:

$$LET = -\frac{1}{\rho}\frac{dE}{dx} \tag{1.2}$$

Figure 1.1 shows a cartoon example of an ion traversing a block of silicon material. Electron-hole pair production is shown as a column along the ion track. In an actual microelectronic circuit, other materials will be present which change the stopping characteristics of the ion. For the case of direct ionization, the electrons of a target material exert Coulomb forces on the ion which slows it down by absorbing energy and resulting in the excitation of the electrons. Electron-hole pairs created around the path of a traversing ion are then subject to the mechanisms of recombination, drift, and diffusion.

If a group of electron-hole pairs generated from an ion strike is near the junction of an active device as shown in Figure 1.2, the electric field can sweep away carriers resulting in a transient current which can upset the operation of the circuit, or in the case of digital circuits such as 3D NAND, cause a stored bit to flip states resulting in a single-event upset (SEU). In silicon, the conversion between energy deposited and charge generated is based on an energy of approximately 3.6 eV required to excite electron-hole pairs resulting a conversion factor of 22.5 keV/fC [2].

The equation for LET provides a way of estimating the energy deposited over small path lengths. Because LET is a differential quantity, the total energy deposited within a region of interest (1-dimensionally, between x_f and x_o) is given by the integral:

$$E_{dep} = \rho \int_{x_0}^{x_f} LET(x)dx \tag{1.3}$$

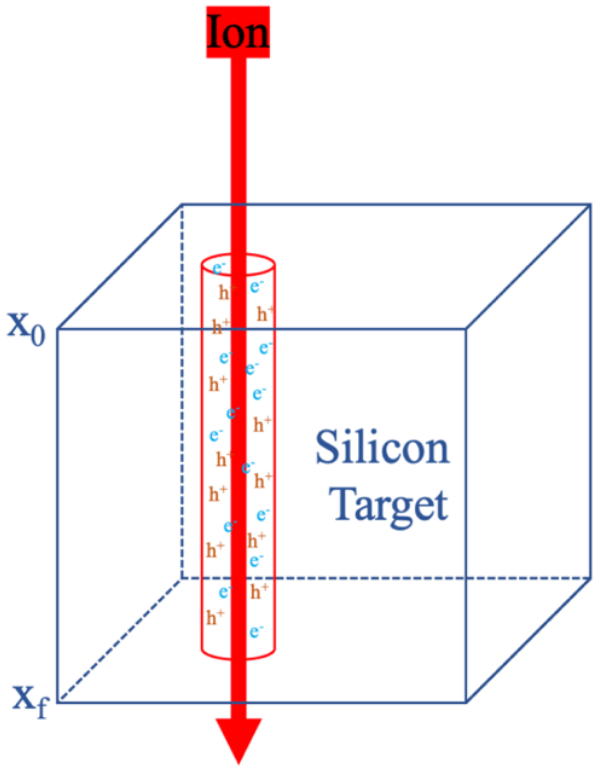


Figure 1.1. Cartoon of ion traversing a silicon cube, generating electron-hole pairs along the track

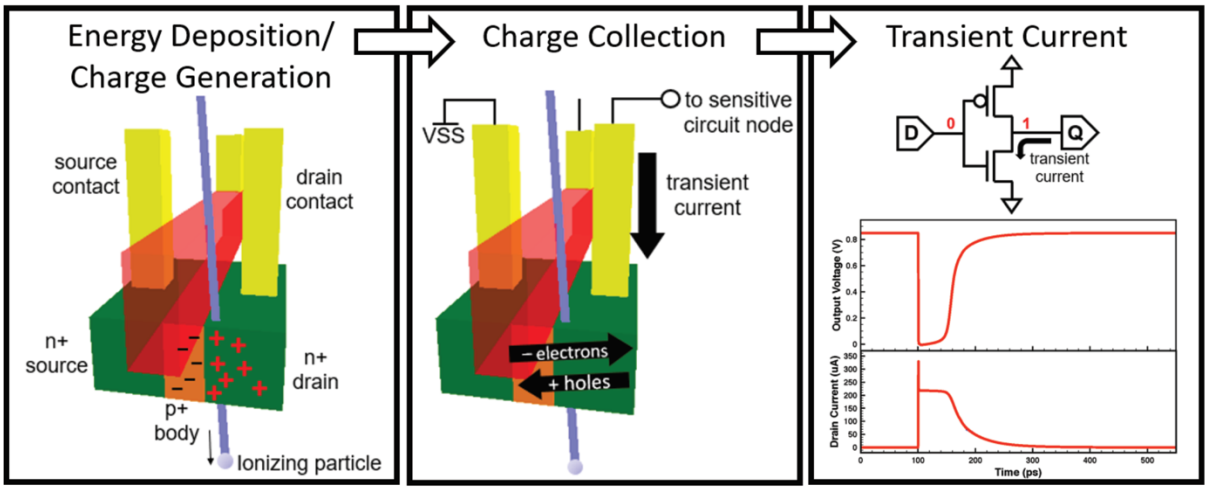


Figure 1.2. Cartoon representation of a heavy-ion striking the sensitive junction of a transistor and generating a transient current in the inverter [3].

LET is variable dependent on the ion's kinetic energy, with a maximum stopping power occurring at what is called the Bragg peak as shown in Figure 1.3 for several heavy ions. Estimating the amount of energy deposited within a material thus requires small enough steps such that the LET is approximately constant. This is achieved in Monte Carlo sampling as introduced later. If the path length of a step in the track is much smaller than the range of the ion, the energy deposition can be approximated as:

$$E_{dep} = \rho \; LET \;\; (x_f - x_0) \tag{1.4}$$

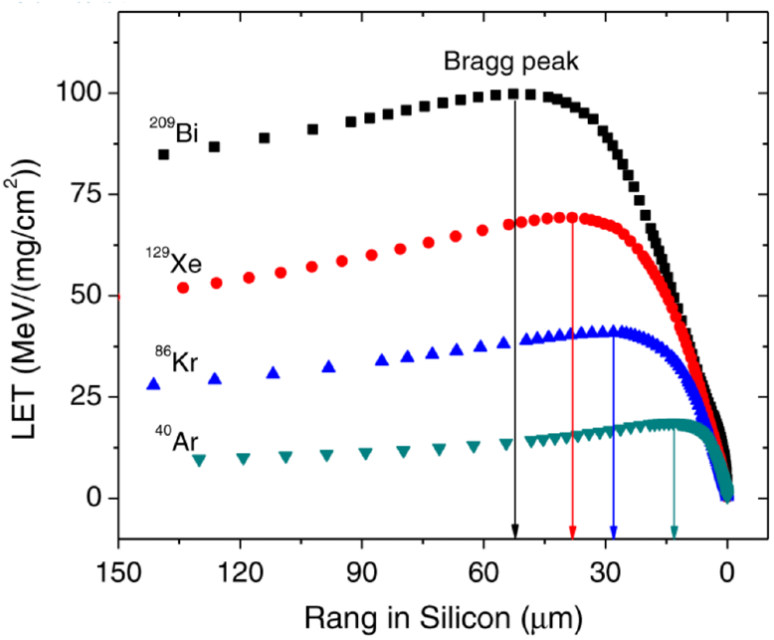


Figure 1.3: The Bragg peak for various ions in silicon [4].

B. Indirect Ionization

Indirect ionization refers to charge generation that is the result of some secondary particle interaction with a material system, as opposed to a primary interaction as, for example, in the case of direct ionization from protons or heavy ions. Neutron environments specifically are a good example of indirect ionization; neutrons have no charge and thus do not interact via Coulomb forces while traversing a material system. While ionizing particles can and do interact with materials via non-ionizing energy loss, the concepts are best introduced in the context of the neutron which does not participate in Coulomb forces. An outline of the different ways of classifying neutron interactions with matter is given in Figure 1.4.

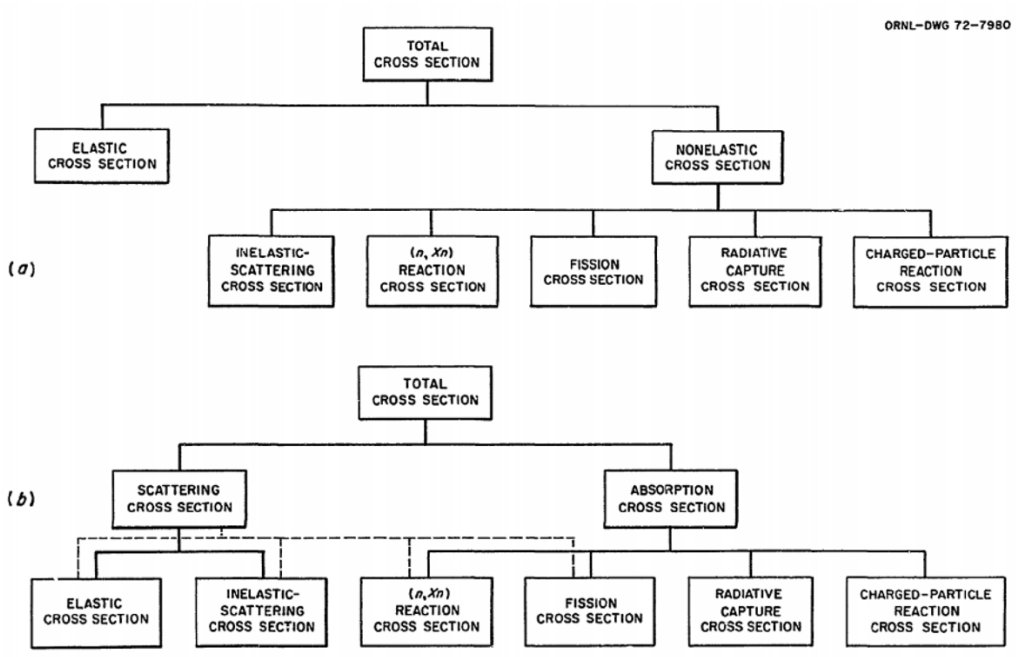


Figure 1.4 Possible neutron reaction classified according to the paradigm of elastic vs. inelastic (top), or absorption vs. scatter (bottom) [5].

Recoils for neutron interactions with the nuclei of a target material can be classified as either elastic or inelastic. Elastic collisions are those in which the species of the colliding particles are preserved after the interaction, i.e., an incoming neutron hitting a silicon nucleus produces a recoiling neutron and a recoiling silicon nucleus. Inelastic scattering is when an incoming neutron enters the nucleus, resulting in an excited energy state which then undergoes de-excitation via cascade and evaporation. Inelastic scatters tend to dominate for $E_{neutron} > 10$ MeV, at which energies the nucleus is much more likely to interact with individual nucleons rather than the nucleus as a whole. Elastic scattering, to contrast, "sees" the entire nucleus as a single object rather than individual nucleons [1]. Inelastic scatters are primarily responsible for the generation of energetic secondaries for neutrons with $E_n > 10$ MeV [6].

C. Photon-matter interactions and dose enhancement

This section will introduce the concept of the dose-enhancement effect. Simulations using MRED have been validated in the study of dose enhancement [7][8] and will be extended later in this dissertation. Photon interactions are of specific importance for TID testing of microelectronics. Most TID testing is done at relatively low dose rates in either a gamma environment (commonly ^{60}Co with a mean characteristic photon energy of ~1.25 MeV) or with lower energy X-ray sources such as an ARACOR. An example spectrum of X-ray energies used in ARACOR testing (and in the simulations described below) is shown in Figure 1.5, while a slightly higher peak is given in the Bremsstrahlung spectrum of Figure 1.6.

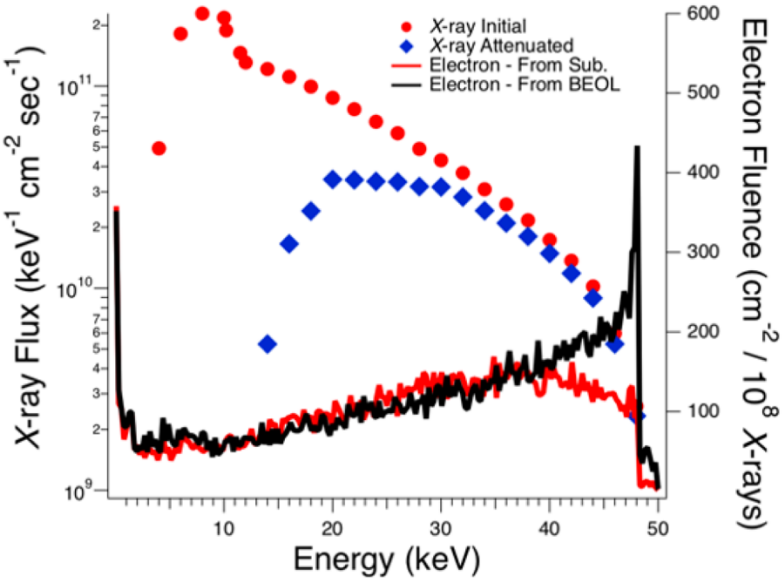


Figure 1.5. 10-keV peaked X-ray spectrum, used in TID testing [9].

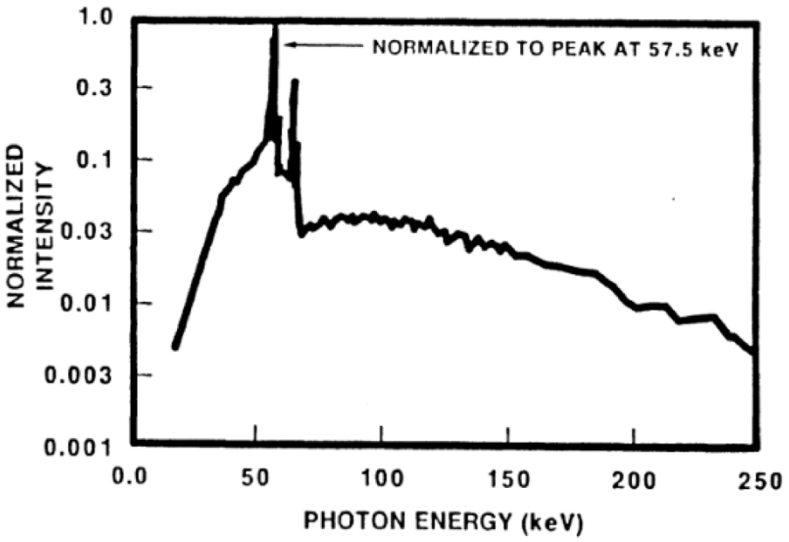


Figure 1.6. 57-keV peaked Bremsstrahlung X-ray spectrum, used in TID & dose enhancement testing for [10]; the spectrum was produced using an X-ray converter with a Pelletron at Sandia.

Dose enhancement is a materials-dependent effect that occurs when two (or more) adjacent layers of materials with different atomic numbers (Z) are irradiated by photons resulting in a non-equilibrium condition in the instant after irradiation for the electrons produced [10] [11][12]. Dose enhancement is often reported using the dose-enhancement factor (DEF) which is defined as the ratio of excess to equilibrium dose in a given system. DEFs have been established with respect to material configurations and test environments. There are three mechanisms which govern photon-material interactions that are sufficiently energetic to produce ionization. These include: the photoelectric effect, Compton scattering, and pair production. Which mechanism applies in a given system depends on the energy of the photon and the Z of the target material [13]. While the probability for all three mechanisms is non-zero at most relevant energy ranges, three distinct regions of the parameter space emerge where one mechanism dominates in terms of carrier production. Figure 1.7 shows the regions as a function of photon energy and Z for which each of these mechanisms is dominant.

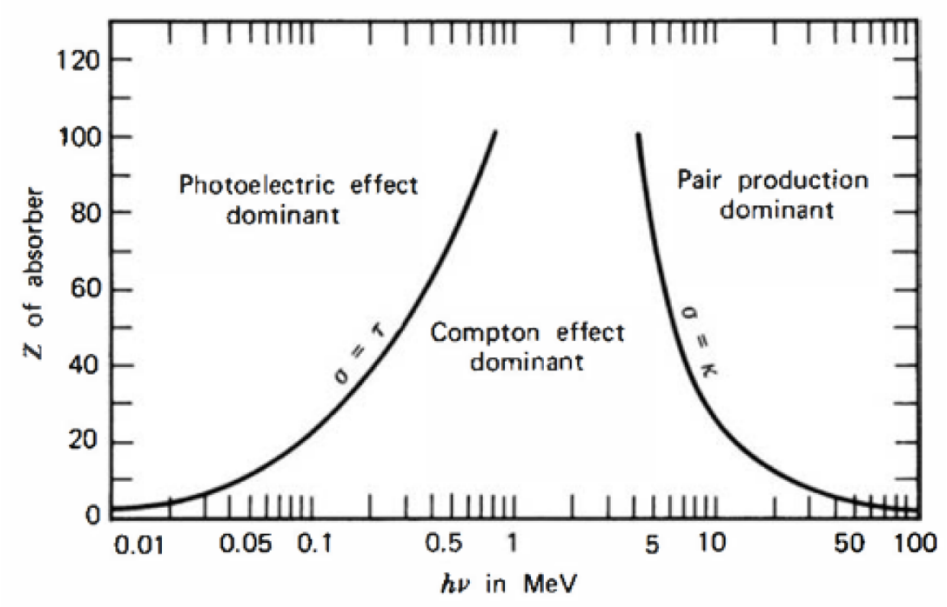


Figure 1.7. Dominant mechanisms of photon-solid interactions as a function of photon energy and atomic number of the target material [10].

Because Compton scattering is directional and photoelectrons are emitted isotropically, there is a directional dependence only in the case of higher-energy photons (into gamma-ray energies) as shown in Figure 1.8. The range of the excess deposited dose is defined by the maximum range of the most energetic electrons generated in either layer. High-Z materials produce more electrons at a given incident photon energy than are produced in lower-Z materials, and so the resulting distribution of charge carriers in two adjacent material layers is discontinuous at the instant following irradiation. Such a discontinuity is considered the non-equilibrium initial condition of the device post-irradiation; the time-evolution of the charge carriers therefore entropically tends towards establishing equilibrium through diffusion. In other words, the concentration gradient of charge carriers in a high-Z/low-Z material stack results in a net flow of carriers from high-Z to low-Z. The total dose per unit fluence in each layer is

therefore enhanced in the case of the low-Z material and reduced/depleted in the case of the high-Z material.

It is important to note that these are relative terms; enhancement and reduction are understood to change the expected dose within a specific layer, but the expected dose must be explicitly and reasonably established. For example, it would not make sense to compare deposited dose in an oxide layer using only the oxide itself as a baseline for expected dose if, for unalterable design reasons, there must be a metal layer nearby. The question is rather which metal species results in a larger enhancement effect for a given environment and fluence relative to the other viable metals. Because there is very little relative enhancement for gamma-rays, gamma-irradiated devices have been experimentally used to establish the equilibrium dose in a device for a dose enhancement factor in X-ray environments. New 3D integrated designs present a multitude of such cases requiring detailed analysis.

Monte Carlo methods are uniquely suited to study dose enhancement especially in more complex structures such as 3DICs; photon transport and scattering as well as electron generation, transport, scattering, and energy deposition are all highly stochastic processes which frustrate efforts aimed at closed-form solutions [14].

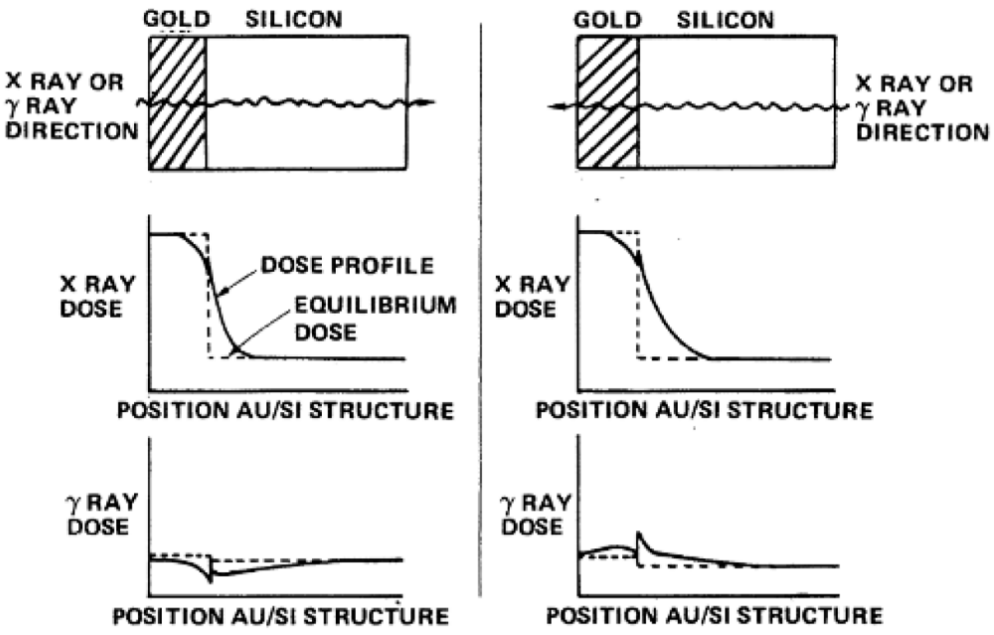


Figure 1.8. Basic representation of the dose enhancement mechanism for X-rays and gamma-rays [12].

II. Monte Carlo radiation transport to predict device characteristics using MRED

Monte Carlo Radiative Energy Deposition (MRED) [15][16] is the primary research tool used in the efforts of this dissertation. This section will introduce MRED and its capabilities, as well as important related concepts such as cross section, energy threshold. MRED is an application developed at Vanderbilt University based on the Geant4 toolkit [17] which provides the capability to conduct highly customizable simulations specific to radiation transport and energy deposition in semiconductor devices, and has been validated and used to reliably study a wide variety of radiation effects, including, among many others, dose enhancement [7][8], 3D integration [18][19], nuclear event contributions to Single-Event Effects (SEE) [20][21], Multiple-Bit Upset (MBU) [22], and Single-Event Upset (SEU) rate predictions for digital circuits [23].

A block diagram of the architecture of MRED is given in Figure 1.9. Geometric structures are defined in MRED most often using a rectangular parallel-piped (RPP) -based geometry. Regions may be defined using common materials found in semiconductor systems or by implementing custom materials with unique densities and compositions. Structures in MRED are encapsulated within a randomization sphere, defined with a radius according to the half-maximum chord length of the physical structure in the simulation as shown in Figure 1.10. Particles are generated on this randomization sphere according to one of many possible spatial and angular distributions. The most commonly implemented random spatial sampling methods are the directional flux, which approximates a unidirectional broad-beam environment, and isotropic flux which corresponds to the GCR environment at geostationary orbit and interplanetary altitudes in which particles are uniformly incident as a function of direction and angle.

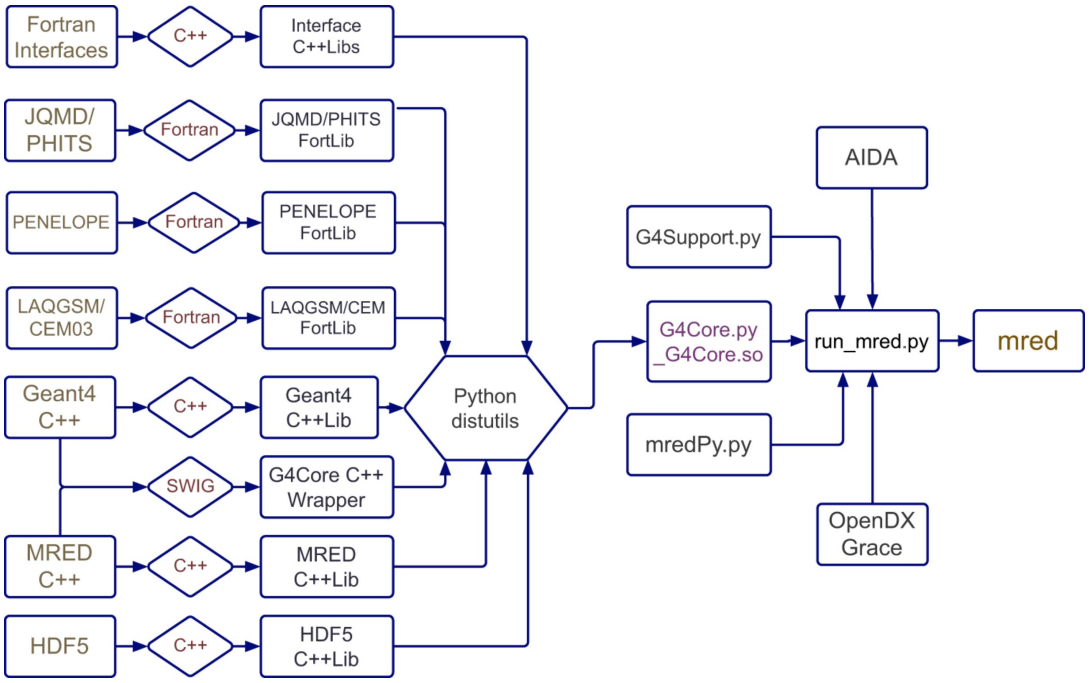


Figure 1.9. Block diagram of the components of the MRED program [15].

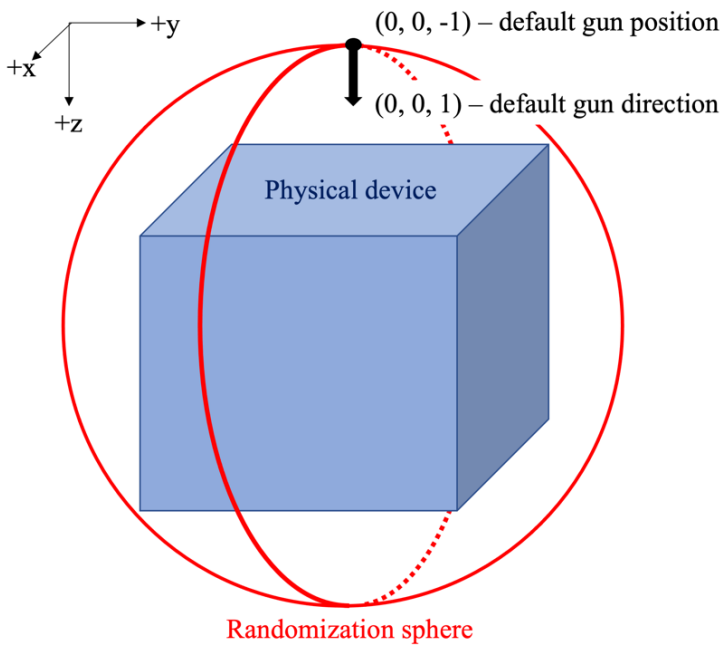


Figure 1.10. The default configuration and axes used in the MRED simulation world. Device geometries are conventionally constructed from negative to positive in the Z-direction, with a randomization sphere entirely enclosing the device.

MRED is, as the name suggests, a Monte Carlo analysis tool. Monte Carlo is a method of whereby the convergence of some parameter(s) approaches an accurate and physically meaningful value based on the random sampling of relevant processes according to known probability distributions. In general, the larger number of sample points, the better the convergence. Monte Carlo is used to solve problems which are unwieldy or intractable via analytical methods. The problem of particle transport, energy deposition, and accumulated event tracking in geometric structures which must at least come close to accurately representing the complexities of a microelectronic circuit are perhaps the perfect combination of conditions to justify a Monte Carlo approach. Figure 1.11 demonstrates the ability to track both ionizing and non-ionizing energy loss in MRED, as well as the ability to export and analyze detailed track structures and secondary production.

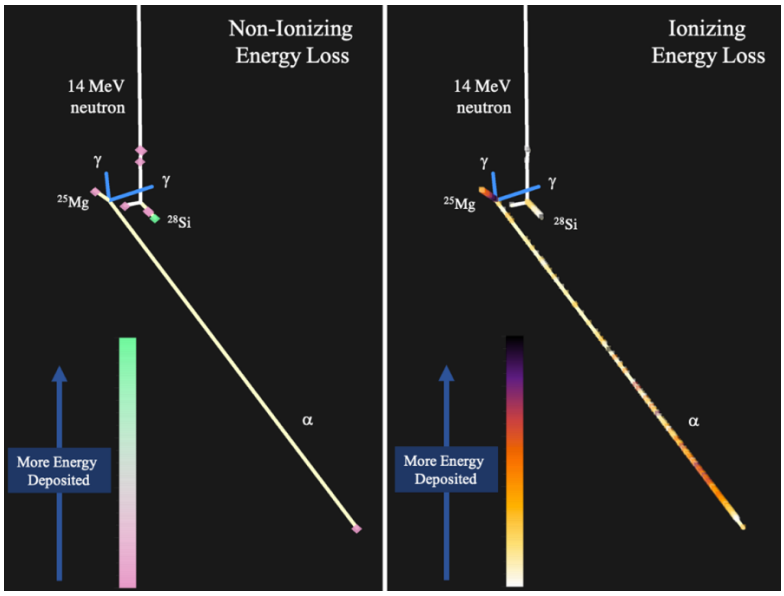


Figure 1.11. Track structure of neutron interactions in MRED. The primary neutron undergoes inelastic scatters which produces secondary particles. These recoiling secondaries deposit energy in the silicon substrate with direct and indirect ionization mechanisms, both of which are trackable in MRED.

The central method of presenting distributions within regions of interest in this dissertation is a plot of what is known as an SEU cross section, usually as a function of LET or particle energy. Cross sections have units of cm^2 and when normalized to the number of sensitive bits in, e.g., a memory chip such as 3D NAND, this is often reported as cm^2/bit. Despite having the same dimensionality as area, the cross section is better thought of as a relative probability of a certain event (here, SEU) occurring at or above a given energy threshold. Calculating cross sections using MRED is one of the common uses of the tool, and the mathematics is given here as a reduction of the generalized rate equation for Monte Carlo analysis under appropriate constraints. Weller et al. presented the following generalized rate equation for single event effects in [15]:

$$R_e(\xi,t) = \sum_z \int_{All E} dE \int d\Omega \left[\oiint dA(-\hat{n}(\vec{x}) \cdot \hat{u}) \int_{-\infty}^{t} dt' \Phi(z,E,\hat{u},\vec{x},t') P_e(z,E,\hat{u},\vec{x},t';\xi,t) \right]$$
$$+ \sum_z \int_{All E} dE \int d\Omega \iiint d^3x \int_{-\infty}^{t} dt' \left[\Gamma(z,E,\hat{u},\vec{x},t') P_e(z,E,\hat{u},\vec{x},t';\xi,t) \right] \quad (1.1)$$

where $P_e(z,E,\hat{u},\vec{x},t';\xi,t)$ is the probability per unit time that a particle z with initial energy E at $(\vec{x},t')$ traveling in the direction $\hat{u}$ will result in an effect during the interval dt at time t. ξ is a single configuration parameter or set of configuration parameters on which the overall event probability P_e depends (e.g., the critical charge Q_{crit}). The flux distribution of these external particles is given by $\Phi_e(z,E,\hat{u},\vec{x},t')$ such that the flux of all particles z, at all energies E traveling in all directions $\hat{u}$ as a function of position and time on the randomization sphere is defined. $H(x)$ is the Heaviside step function which in the surface integral constrains particles such that only those traveling inward from the surface $-\hat{n}(\vec{x} \cdot \hat{u})$ of some arbitrary external sphere contribute to the overall event rate. This is intuitively understood, as particles which at $t = 0$ in a simulated event begin with p_0 to move away from the structure of interest cannot contribute to the overall event rate of that structure. $\Gamma(z,E,\hat{u},\vec{x},t')$ is the local generation rate of a radiation source internal to the defined device, hence the volume integral is found in this term as opposed to the surface integral in the case of an external environment. For most MRED use cases, the internal environment of a semiconductor system is not radioactive and thus the second term can be safely ignored (this will be revisited later in the section on proxy environments).

Several assumptions can be further made to simplify the rate equation to the level of practical applications using MRED. These include:

- There is no internal radiation generation that contributes to the environment

$$\Gamma(z, E, \hat{u}, \vec{x}, t') = 0$$

- The flux is time independent $\Phi = \Phi(z, E, \hat{u}, \vec{x})$
- The flux is isotropic $\Phi = \Phi(z, E)$
- The effect occurs over a very short time period

$$\int_{-\infty}^{t} dt' \Phi P_e(z, E, \hat{u}, \vec{x}, t'; \xi, t) \rightarrow \Phi P_e(z, E, \hat{u}, \vec{x}; \xi)$$

- The effect parameter ξ is defined by the critical energy ($\xi \rightarrow E_c$). Any amount of energy deposited above this threshold will result in a single-event effect, while events which deposit a lower amount of energy do not produce an effect. This necessarily leads to the next item:
- The probability P_e is uniquely determined by the amount of energy deposited E_d within the region of interest

$$P_e(z, E, \hat{u}, \vec{x}; E_c) = \int p_d(z, E, \hat{u}, \vec{x}; E_d) P_e(E_d, E_c) dE_d$$

- In the case of a simple RPP model, $P_e(E_d, E_c)$ reduces to a step function $H(E_d - E_c)$ and the overall event probability is governed by the density function p_d.
- The continuous probability density function p_d is validly represented by the Monte Carlo engine through the event weight w_i and the energy deposited ε_i such that

$$p_d(z, E; E_c) = \frac{1}{N_z} \sum_{i=1}^{N_z} w_i \delta(E_c - \epsilon_i) \qquad (1.2)$$

The rate equation (1.1) taking these assumptions into account, is now given by:

$$R(E_c) = \sum_z \int_{AllE} dE\Phi(z,E) \int d\Omega \left[\oiint dA(-\hat{n}(\vec{x}) \cdot \hat{u}) \int p_d(z,E;E_d) H(E_d - E_c \, dE_d \right] \tag{1.3}$$

Integrating over the surface and all angles gives

$$R(E_c) = 4\pi^2\rho^2 \sum_z \int_{AllE} dE \; \Phi(z,E) \int p_d(z,E;E_d) \; H(E_d - E_c) dE_d \tag{1.4}$$

where ρ is the radius of the randomization sphere which encloses the world as previously introduced. The derivative of a Heaviside function is a Dirac delta function, and since E_c is the event parameter:

$$\frac{d}{dE_c} \int dE_d \; p_d(z,E;E_d) \; H(E_d - E_c) = p_d(z,E;E_c)$$

which, substituting the expression in (1.2) for p_d by applying the derivative to both sides, yields a differential rate equation:

$$\frac{dR(E_c)}{dE_c} = 4\pi^2\rho^2 \sum_z \int_{AllE} dE \; \Phi(z,E) \frac{1}{N_z} \sum_{i=1}^{N_z} w_i \delta(E_c - \epsilon_i) \tag{1.5}$$

To arrive back at a rate equation, the integral is taken from E_c to infinity, as all energies above the critical energy contribute to the rate.

$$\begin{aligned} R(E_c) &= \int_{E_c}^{\infty} dE \; \frac{dR(E)}{dE} \\ &= \int_{E_c}^{\infty} \left(4\pi^2\rho^2 \sum_z \int_{AllE} dE \; \Phi(z,E) \frac{1}{N_z} \sum_{i=1}^{N_z} w_i \delta(E_c - \epsilon_i) \right) dE \end{aligned} \tag{1.6}$$

Discretizing Eq. (1.6) is the final step, which accounts for the fact that Monte Carlo methods must necessarily sample on an event-by-event basis for the remaining continuous quantities in the rate equation (Φ, in this simple case), approaching the continuous (and hopefully physical) limit as the number of samples N_i becomes very large.

$$R(E_c) = \int_{E_c}^{\infty} \frac{4\pi^2\rho^2}{N_{tot}} \sum_z \sum_{i=1}^{N_{tot}} \Phi_{z,E_i}\ w_{z,E_i}(1 - H(E_c - \epsilon_{z,E_i})) \tag{1.7}$$

Variance reduction technique can be further implemented, which MRED accomplishes by defining the individual event weight, w_i, according to the product of the flux and the intrinsic weight along with a biasing factor β.

$$w_i = \beta\ \Phi_{z,E}\ w_{z,E} \tag{1.8}$$

In the case of a directional flux, the solid-angle integration is 1, not 4π, leading to

$$R(E_c) = \frac{\pi\rho^2}{N} \sum_i^N\ w_i \cdot (1 - H(E_c - \epsilon_i)) \tag{1.9}$$

Histograms which sort events into energy deposition bins in MRED are accumulated during run-time, and the rate according to Eq. (1.9) is calculated by summing the weights in all bins where $\epsilon_i > E_c$, starting from the maximum energy bin ϵ_{max} and proceeding to zero. MRED calculates the gun fluence unit which is defined as $(\pi\rho^2)^{-1}$ where ρ is the half-maximum chord length of the defined device which serves as the radius for the randomization sphere. Thus, normalizing sum of the weights as introduced and normalizing to the gun fluence unit and the total number of particles simulated in the run which generated those histograms yields an integrated cross section.

$$\sigma_i(E_{th}) = \sum_{\epsilon_{E_{th}}}^{\epsilon_{max}} w_i$$

The units which result are cm^2, which corresponds to the cross section for a device. More specifically, this is an integrated cross section for an event, which in practice is usually some single-event such as upset, latchup, functional interrupt etc. A differential (in energy) flux is calculated by normalizing the sum of the weights in each bin to the bin width ($\epsilon_{i+1} - \epsilon_i$) and then normalizing to the total number of events simulated. The integrated event cross section is best understood as the relative probability of an energy deposition at or above the given energy threshold E_{th} occurring.

Very often, the energy threshold is not known, or many energy thresholds must be considered for the same set of simulation configurations due to, for example, variations in sensitive regions of a device. It is typical and convenient to represent the integrated event cross section as a function of energy deposited, in which case the cross section for a range of possible critical energies can be read off. This means integrated from the highest energy bin to the lowest one, and can easily be accomplished in python by reversing an array with the normalized bin weights, executing a cumulative sum method, and the reversing the array again.

In this dissertation, the term cross section, when applied to MRED results and unless explicitly stated as otherwise, refers to the integrated cross section for an event at or above a threshold (this threshold is usually given in the discussions which follow) and is calculated according to these methods.

III. General structures and features common to 2.5D and 3D integrated circuits

Vertical integration of semiconductor devices is not an entirely new technology; patents for 3D integration methods go back to the 1980s [24], Nevertheless, the field is considered to be relatively new with rapid development occurring over the past decade. A large part of the motivation for this growth is driven by the desire in industry to achieve performance scaling in accordance with Moore's Law [25]; Figure 1.12 shows projected scaling trends (as originally presented in 2005 [25]) according to the "More Moore" paradigm, which refers to industry efforts to increase the number of transistors per unit area on a chip by reducing the size of the transistors. The "More Moore" paradigm (i.e., continued scaling by developing smaller transistors) is beginning to run into physical limitations; feature sizes less than a couple nanometers present issues with electron tunneling and are difficult to reliably fabricate [26][27]. The "More than Moore" paradigm is presented as an alternative focus of innovative energy, in which scientists and engineers develop clever ways of integrating existing transistor technology to achieve higher levels of performance year by year.

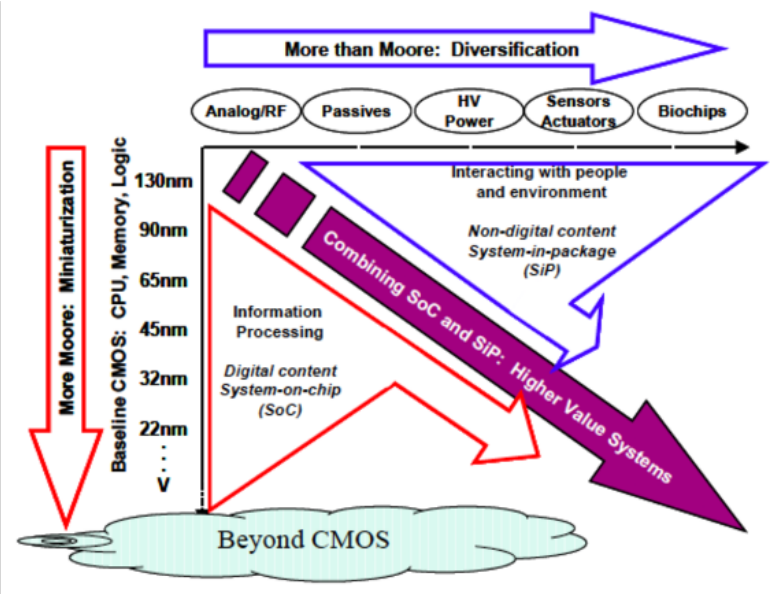


Figure 1.12. More Moore vs. More-than-Moore paradigm diagram. Efforts began over a decade ago to shift the focus to novel integration architectures as opposed to fabricating smaller transistors [25].

Fabrication methods are quite varied in the 3DIC world. Figure 1.13 demonstrates the logical classification of the most common methods for achieving vertical integration [28]. Combined with the detailed overview of the classification and methods of 3DICs given in [29] (see Figure 1.14 and Figure 1.15) it becomes clear that 3DIC is a rather broad term which is applied for devices integrated in three dimensions from the wafer level down to the transistor level. Monolithic 3DICs are thus considered the "most" 3D in that the multiple active layers vertically stacked are all fabricated within the same process; System-in-package (SiP) and 3D Wafer-Level Packaging (WLP) are examples of integration which rely on vertically connecting chips which are usually (but not necessarily) fabricated in a traditional planar process.

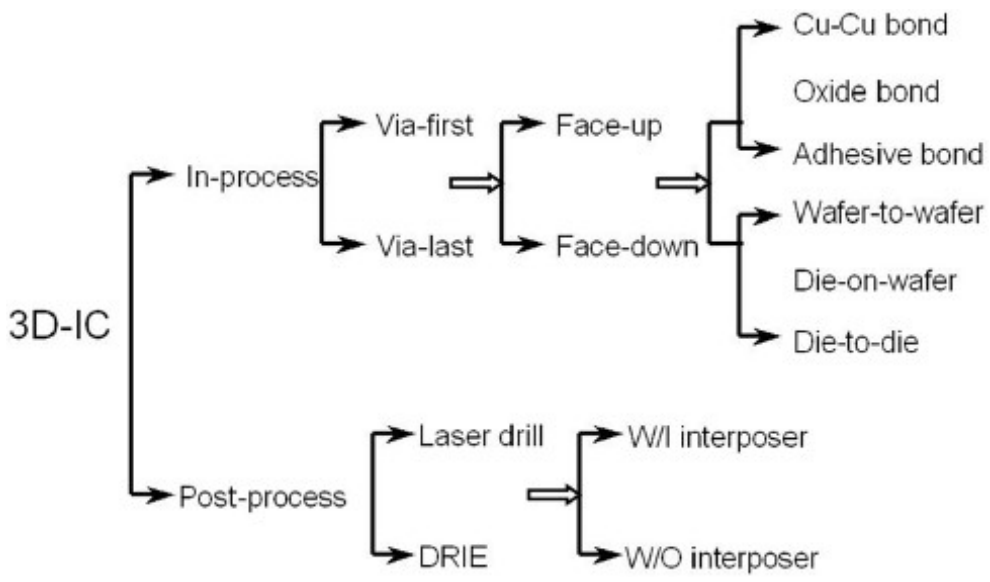


Figure 1.13. Flowchart of various fabrication methods for 3DICs [28].

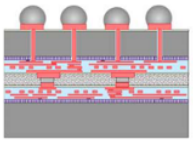
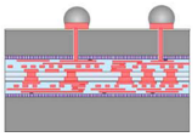
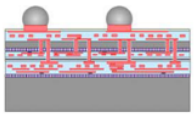
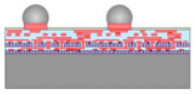
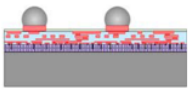
	3D-SIC	3D-SOC			3D-IC
3D-Wiring level	Global	Semi-global	Intermediate	Local	FEOL
Partitioning	Die	**blocks** of standard cells		**Standard cells**	**Transistors**
3D Technology	Die stacking Die-to-Wafer stacks Die-to-Si-interposer	Parallel FEOL wafer processing Wafer-to-Wafer bonding		Sequential FEOL processing Active layer bonding or deposition	
2-tier stack schematic				2nd FEOL after stacking 	Multi-tier FEOL 
Characteristic	Known Good Die 3D stacks or Si-interposer stacking	BEOL between 2 FEOL layers *Overlay 2nd tier defined by W2W alignment/bonding*		*Overlay 2nd tier defined by litho scanner alignment*	FEOL/FEOL stack
Contact Pitch *Relative density:*	40 ⇒ 20 ⇒ 10µm ⇒ *5µm* 1/16 ⇒ ¼ ⇒ 1 ⇒ 4	5 ⇒ 1µm 4 ⇒ 100	2 µm ⇒ 0.5 µm 50 ⇒ 400	200 ⇒ 100 nm 5000 ⇒ 10000	< 100 nm > 10000

Figure 1.14. Overview of the 3-D technology landscape, as defined by the 3-D insertion point in the electronic system wiring hierarchy" (from [29]).

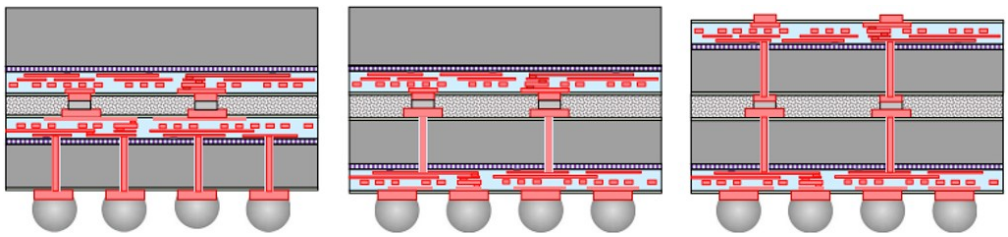


Figure 1.15. Three examples of common methods for achieving 3D integration between two chips. From left to right: face-to-face (F2F), face-to-back (F2B) and back-to-back (B2B) [29].

Some important advantages of 3DICs are decreased footprint, higher bandwidth, lower power consumption, faster operating frequencies, and support for heterogeneous integration [29][30]. Of course, there are trade-offs to consider as well: 3DICs are much more difficult to reliably fabricate, generally come at a higher cost, have considerable thermal and mechanical issues, and testing at intermediate fabrication steps is sometimes impossible. Within the different classifications of 3DICs there are likewise additional trade-offs to consider. Monolithic 3DICs

with gate-level and transistor-level designs for example have experimentally demonstrated advantages for overall package footprint, timing, and power consumption when compared to TSV-based designs, however they are generally more difficult and expensive to fabricate [30]. Some of the common applications for which there is a large demand for tighter vertical integration is in the field of stacked memories, CMOS+Sensor technologies, and high-density MEMS devices [31].

Putting all this nomenclature aside, there are certain features common to most 3DICs regardless of the specific implementation that may reasonably have a direct effect on the response of a device to various kinds of radiation. These are vertical-interconnect-access (via) lines, which are also known as Through-Silicon-Vias (TSVs) or Through-Package-Vias (TPVs). These are large metal columns used to achieve electrical connectivity between vertically layered chips or vertically stacked active devices within the same chip.

TSV Depth / Diameter [μm]	TSV Pitch [μm]	Material	Ref.
230 / 80	180	Cu	[34]
120 / 60	650	Cu	[34]
50 / 15	20 - 50	W	[35]
6 / 1.25	2 - 15	W	[18]
14 / 1.2	2.5 - 4	W	[33]
50 / 5 - 13	10	W	[36]

Table 1.1. Common TSV materials and dimensions

TSVs are present in a variety of configurations and materials depending on the technology. Tungsten and copper are the most common TSV materials, each with their own

advantages and disadvantages. Both materials usually require implementation using a via-middle or via-last integration. Tungsten has a Coefficient of Thermal Expansion (CTE) that is matched closely with that of silicon, but is generally more difficult to implement in 3D, especially for smaller TSVs [33] and introduces significantly more mechanical stress [37]. Copper has disadvantages as well, primarily the need for thermal pumping and failure due to electromigration when interfacing between TSVs and BEOL metals [33]. Table 1.1 shows the wide range of TSV geometries used in devices for both copper and tungsten. A glossary of common TSV- and 3DIC-related terms is given in [29] which is reproduced in part here to summarize this brief introduction to vertical integration technologies:

- 3D Bonding – an operation that joins two or more die or wafer surfaces together
- 3D Stacking – 3D bonding operation that realizes electrical interconnects between device levels
- 3D System-in-Package (SIP) – 3D integration using "traditional" packaging technologies
- 3D Wafer-Level Packaging (WLP) – 3D integration performed after wafer fabrication (e.g., flip-chop redistribution, redistribution interconnect)
- 3D System-on-Chip (SoC) – multiple-die SoC; 3D interconnects at the level of global on-chip interconnects
- 3D integrated circuit (IC) -- 3D approach using direct stacking of active devices. The 3-D stack is characterized by a stack of front-end devices, combined with a common back-end interconnect stack.

In summary, the two most common features of virtually all 3D integration methods are the vertically stacking of active layers and the presence TSVs or similar structures to achieve connectivity between these active layers. The focus of the proposed work in the following sections centers on explaining in detail how these features may contribute to the response of 3DICs to radiation environments, both in terms of total dose and single events.

IV. Charge trap and floating gate 3D NAND memories

Here, 3D NAND memories are discussed in detail. The general 3D NAND structure is introduced with distinctions made between charge-trap (CT) and floating-gate (FG) memories. Experimental results from proton, heavy-ion, and TID testing of both kinds (CT and FG) from the literature are summarized.

A. General 3D NAND structures

This section will introduce the specific memory device of interest and, use it as a representative guide of the technology, explore some common terminology and background information regarding the design and fabrication of 3D NAND starting at the materials level with the transistor and including a brief discussion of the circuit-level architecture. These foundations are important to contextualize the experimental data introduced below and will provide a framework for understanding the further discussion of those results.

3D NAND consists of vertical stacks of bit lines (BL) which are strings of transistors joined together from source to drain in a long line. A circuit-level schematic of the bit string is shown in Figure 1.16. The transistors store digital information in the form of trapped charge, either in a floating gate or in a charge trap material such as silicon nitride. The gate of each cell, which modulates the amount of stored charge during program operations, is controlled by the word line (WL) contacts.

The example device used here from SK Hynix (and throughout this dissertation) is a specific 3D NAND design architecture known as pipe-shaped bit cost scalable (P-BiCS), which means that each NAND bit string is configured in a U-shape which begins and terminates at the top of the device with the source and drain select transistors. The 72-layer 3D NAND device is shown in Figure 1.17. The SEM cross section in the figure shows the memory array partitioned into an upper and lower region, separated by two dummy layers. The bright regions correspond to the tungsten WLs, and the dark regions between adjacent word lines are air-gaps intentionally included as part of the fabrication process in order to reduce the coupling capacitance from nearby WLs. This feature is important later in the discussion of experimental and simulation results for low-energy protons.

A close-up image of the region around the dummy layers is given in Figure 1.18, which shows a slight misalignment of ~10 nm between the top and bottom half of the stack. There also appears to be a lack of uniformity in the cell structure for the dummy gates. The same report which contains the SEM images notes that these dummy layers are identical to the cells in the rest of the memory with the difference being that they are not electrically connected. No other structures are found at this interface [38].

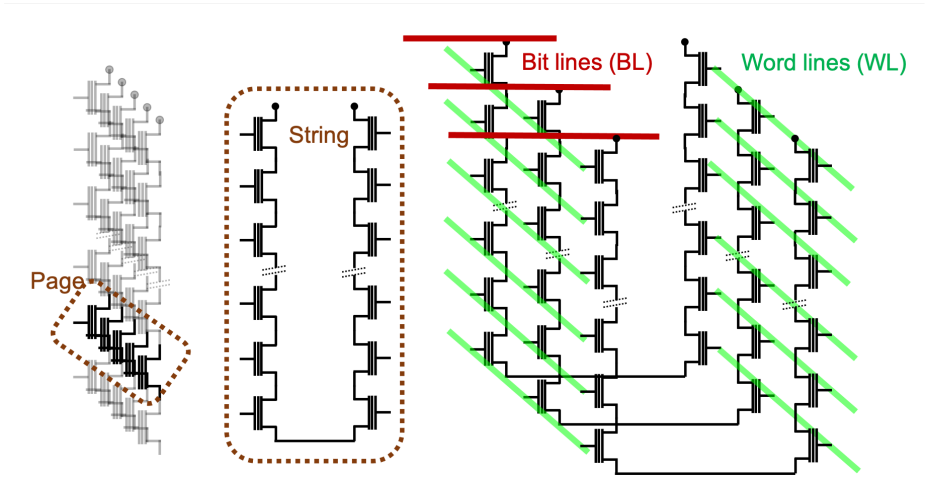


Figure 1.16. Circuit-level diagram showing the p-BiCS NAND configuration.

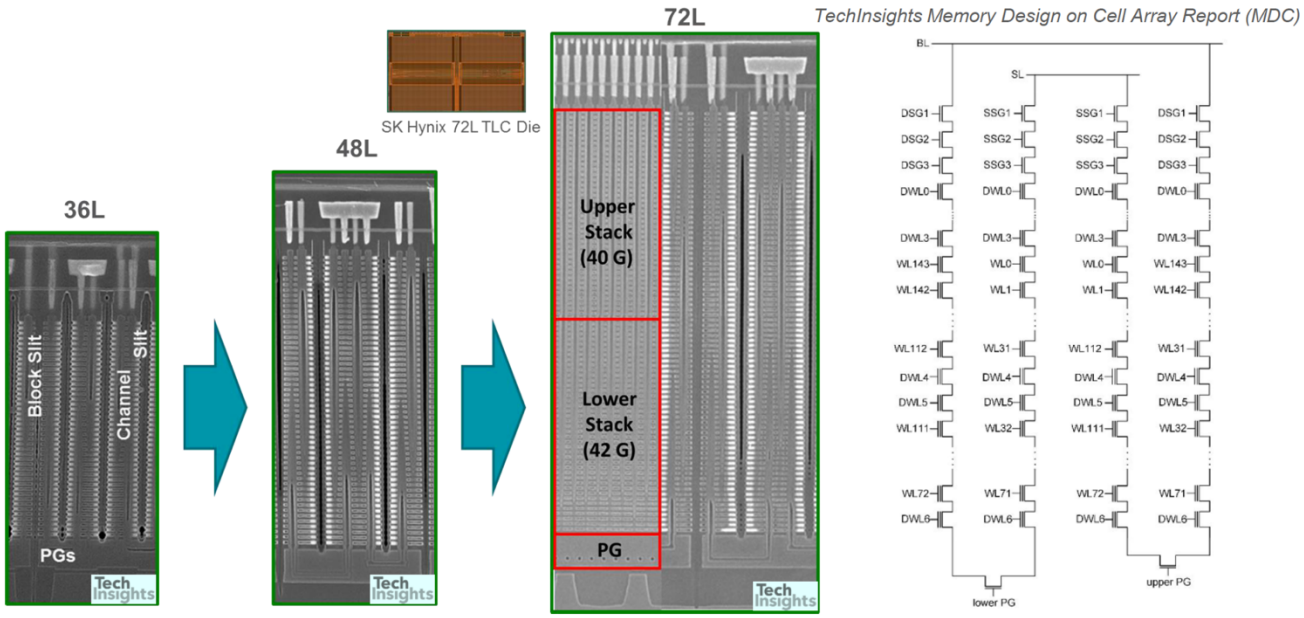


Figure 1.17. SEM images showing the evolution of the SK Hynix 3D NAND stack from initial configurations with 32 layers to the subject of the present study at 72 active layers [38].

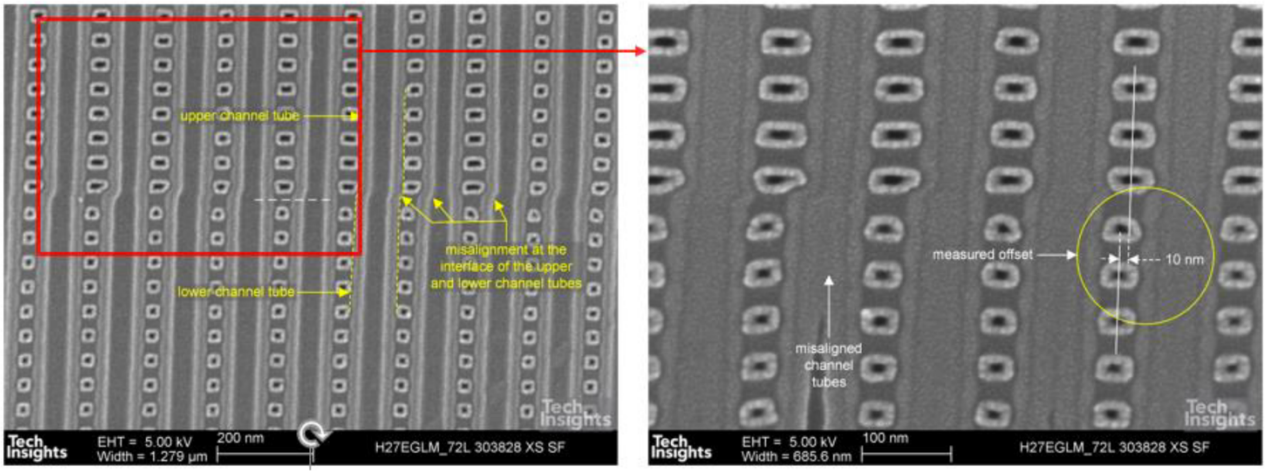


Figure 1.18. Close-up SEM image of the interface between the upper and lower region in the NAND stack [38].

Figure 1.19 shows a similar memory from the same manufacturer, where isolation voids are found between the WLs and shown much more clearly than in the figure above. These are included in NAND memories both at the planar and 3D level to reduce the capacitive coupling between adjacent word lines [39]. The voids in Hynix 3D NAND chips are filled with Nitrogen [39].

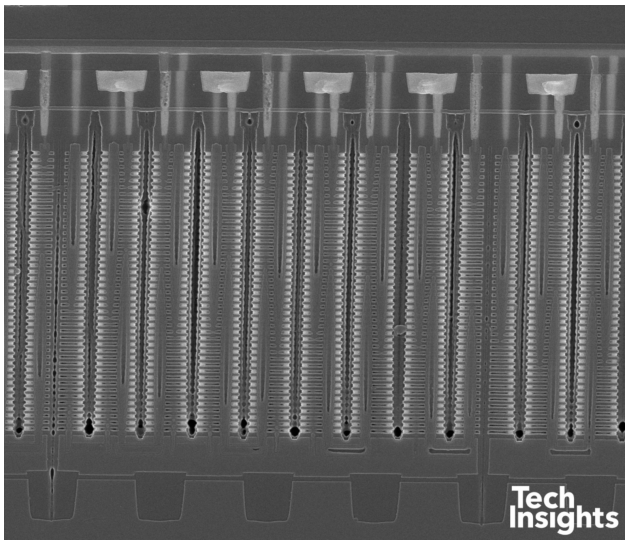


Figure 1.19. SEM image of a Hynix 32-layer NAND. The vertically etched slits are clearly visible as dark regions between the word-line contacts [39].

The gate-all-around charge trap transistor is the core device in the memory array and is found at the intersections of the WLs and BLs. The cylindrical structure of the transistor is found in both CT and FG 3D NAND, and operate by pushing and pulling electrons from the poly-silicon channel through the tunnel layer and into the trapping memory layer by applying strong electric fields. The trapped charge modifies the threshold voltage (V_{th}) characteristics of the transistor which is utilized for the read operation for a cell.

3D NAND memories may be operated in a single-level cell (SLC), or multi-level cell (MLC), or triple-level cell (TLC) configuration, which corresponds to the 1, 2, or 3 bits of digital information stored on that transistor. SLC operation is generally more reliable, as a broader V_{th} window defines the logical state of the node. Figure 1.20 provides an illustration of the narrowing of the threshold voltage window to achieve higher memory density.

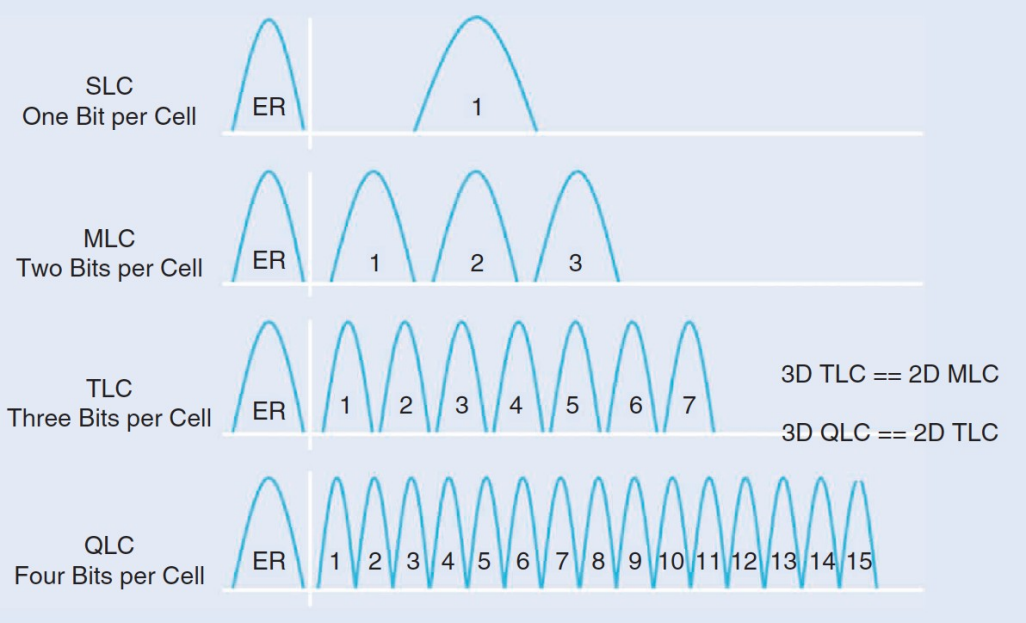


Figure 1.20. Representation of the different operating windows for NAND memory cells. Increased memory density is achieved in multiple-bit cell operating modes at the cost of decreased reliability due to the more narrow V_{th} windows [41].

B. Floating gate (FG) vs charge trap (CT) 3D NAND

Both 3D and 2D planar NAND technologies have used floating-gate (FG) and charge-trap (CT) transistors for the cells of the NAND string. Figure 1.21 shows a cartoon of the two kinds of transistors for a planar transistor structure. Floating gate 3D NAND uses highly doped polysilicon WLs in a gate-first fabrication method [43][44], while CT devices use metal (usually tungsten) in a gate-late fabrication process. In floating gate memories, the FG is totally isolated and charge spreading from WL layer to layer is not likely (this is not the case for CT NAND as discussed later). Electrons are stored in the floating gate for a programmed "0" state at the cell; erasing the cell thus involves the removal of electrons from the FG [45].

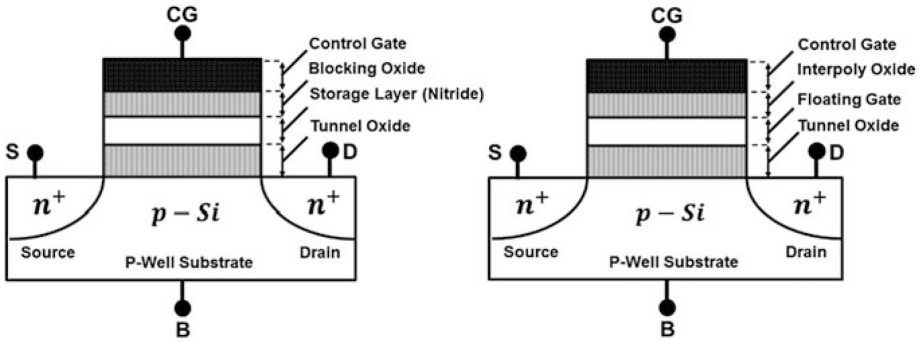


Figure 1.21. Planar (2D) charge trap transistor (left) and floating gate transistor (right) [43].

Figure 1.22 shows the stack-level (a) and close-up (b) view of a FG 3D NAND device. Compared with the figures above for CT NAND, the lack of metallization throughout the stack is obvious. Tungsten contact can still be observed in the BEOL as is standard in all 3D NAND memories.

The two different fabrication processes for FG and CT 3D NAND are shown in Figures 1.23 and 1.24, respectively. FG 3D NAND uses a gate-first approach in which alternating layers of polysilicon and oxide are deposited, followed by the etching steps for the formation of the floating gate as shown in the figure. CT 3D NAND uses a gate-last approach in which the tungsten word lines are deposited, often with air-gaps as discussed above, as one of the last steps in fabrication. Charge trap 3D NAND (CT) stores charge in an insulating layer as opposed to a conducting floating gate. This has advantages such as removing interference between cells, using a smaller cell area, and better read and write performance [47]. The material used to store the charge at each cell node in the memory is one with a high density of trap centers such as silicon nitride [49]. Due to the nitride layer consisting of a single continuous structure which spans the entire vertical length of the bit string, there is a significant reliability issue posed by the

possibility of charge spreading resulting in data retention degradation and poor performance with respect to cell state distributions.

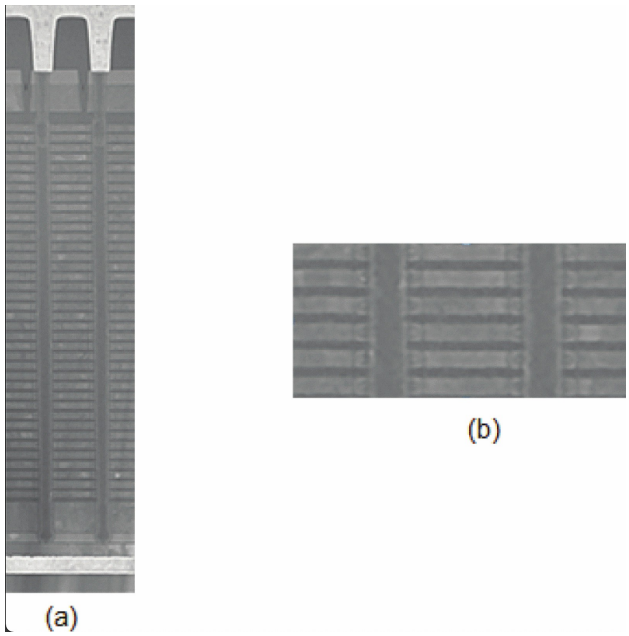


Figure 1.22. SEM images of a FG 3D NAND cell. Note the lack of metal layers; FG processes use polysilicon for electrical connectivity throughout the memory array [44].

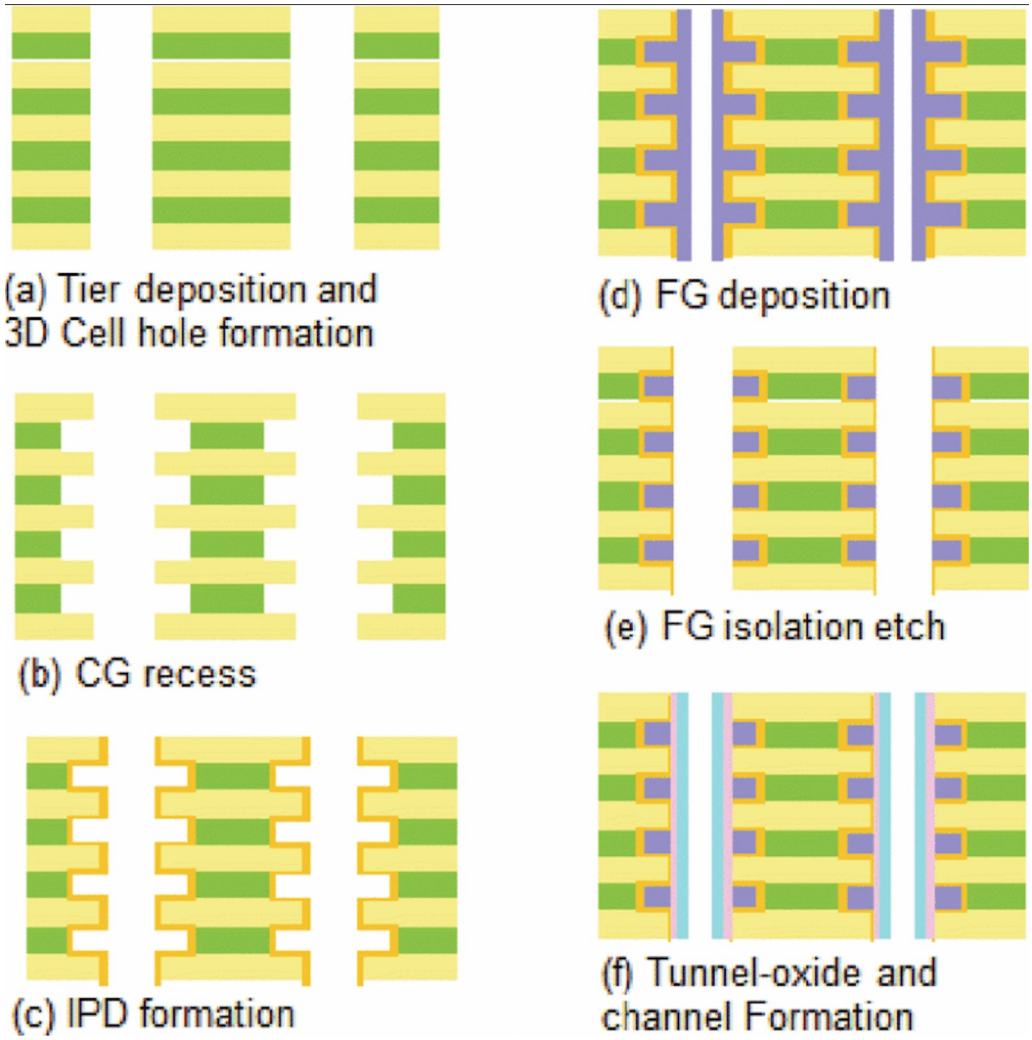


Figure 1.23. Fabrication process for FG-3D NAND. This is a gate-first method, as opposed to CT NAND. Alternating layers of oxide polysilicon are deposited on a substrate, followed by high aspect ratio etches for the string channels. The control gate layers are recessed and an inter-poly dielectric is deposited, followed by the floating gate material, a tunnel oxide, and then the polysilicon channel which runs the entire length of the string [44].

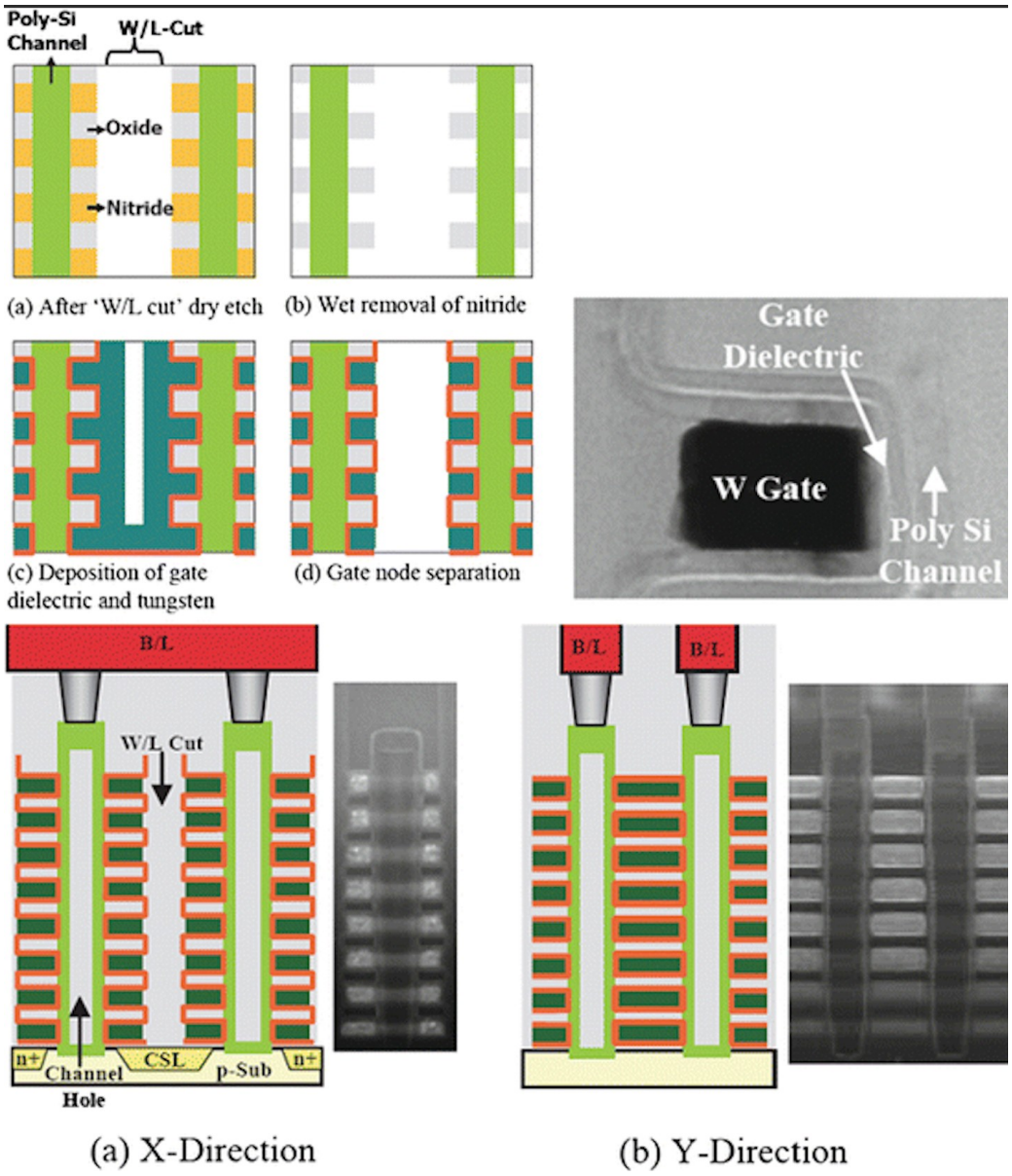


Figure 1.24. Gate-last process used to fabricate CT NAND [48]

A cross section view of the two different kinds of 3D NAND is shown in Figure 1.25. The center of the channel column in modern devices is filled with oxide, resulting in a macaroni-shaped ring of polysilicon (shown in yellow) which acts as the channel. The nitride layer in CT NAND is continuous along the channel, while FG regions are isolated by the oxide between each transistor layer and the inter-poly dielectric (IPD).

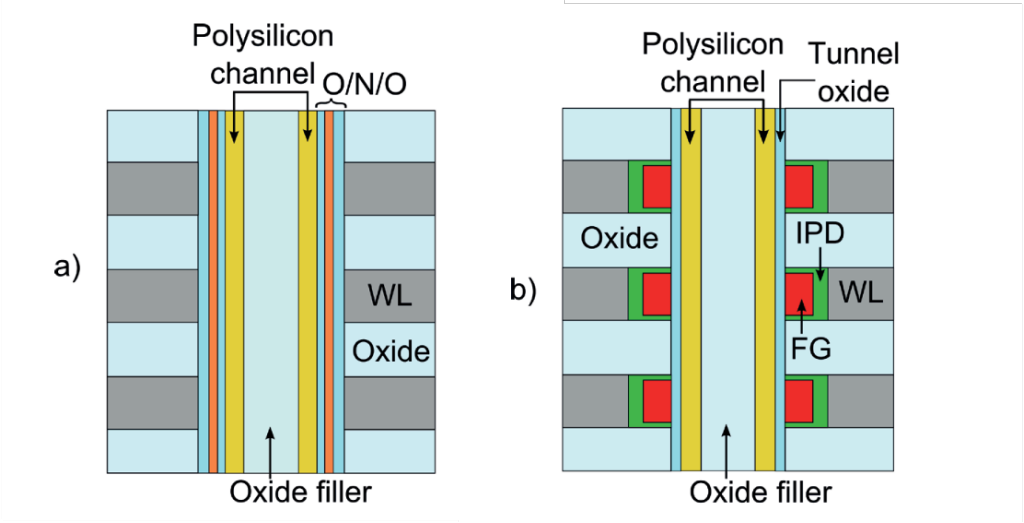


Figure 1.25. Cross sections of CT (a) vs FG (b) cells [48].

The way in which bit strings and pages are configured and connected varies between manufacturers. The SK Hynix device which is the main subject of this dissertation uses a piped-bit cost scalable (p-BiCS) configuration as shown in Figure 1.26. The channel forms a "U"-shape which both begins and terminates at the BEOL at the top of the memory stack. In the case of the 72-layer NAND, this means that there are actually 144 logical levels per BL.

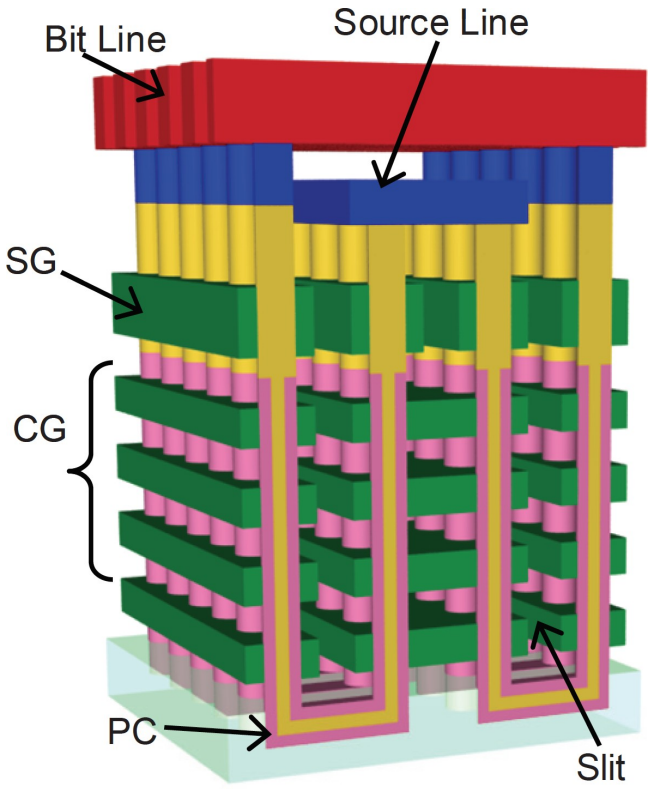


Figure 1.26. Schematic of p-BiCS memory, showing the bit lines (BL), source lines (SL), control gates (CG), select gates (SG), and pipe connection (PC) [50].

C. Literature review of radiation effects in 3D NAND

Gadlage et al. showed directional dependencies in the bit errors of both planar and 3D NAND devices in ^{60}Co environments, as well as a strong dependence on the operating mode of the NAND device (i.e., SLC, MLC, and TLC) [46]. These data are however not presented at the detailed level of effects as a function of depth in the NAND stack as in the subsequent studies.

Figure 1.27 shows the increase in bits with errors as a function of the angular configuration of the incident photons from ^{60}Co gamma cells. The three configurations used were directionally through the BEOL, through the backside, and a semi-isotropic configuration in which the DUT was placed in the center of a ring of gamma cells. All three operating modes

available on the chip (SLC, MLC, and TLC) were tested, with the increase in memory density in multi-level cell operation resulting in a much higher amount of data corruption as a function of total dose as shown in Figure 1.28.

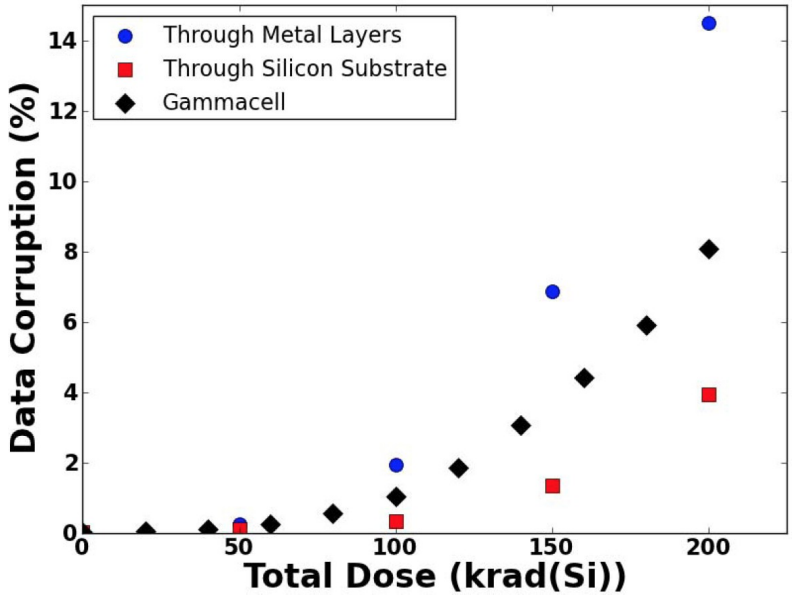


Figure 1.27. Angular dependence of TID-induced errors in ^{60}Co gamma testing [45].

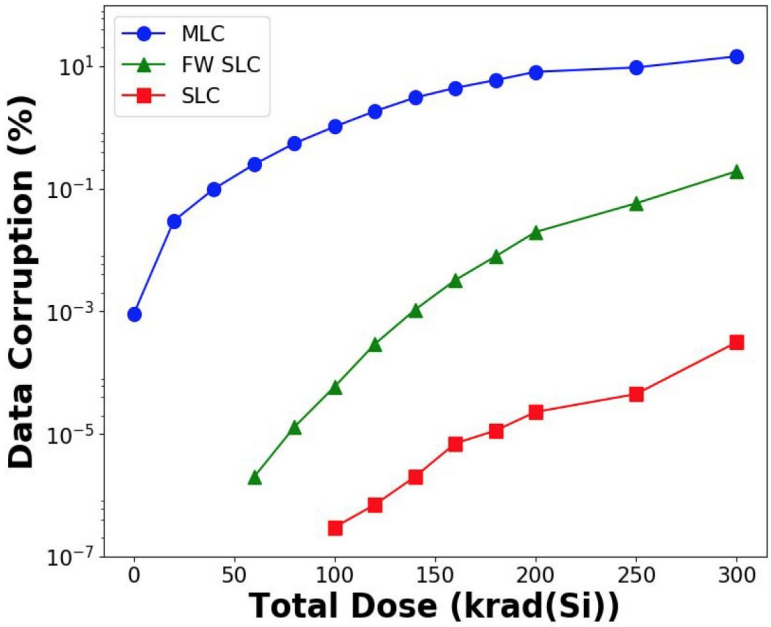


Figure 1.28. Observation of increased bit error rates as a function of cell operating conditions [46].

Chen et al. presented a detailed review of heavy-ion and proton-induced SEU and single-event functional interrupt (SEFI) in a SK Hynix CT NAND over a range of several LET and angular values [39]. Comparisons were made between the 3D NAND memory and a 2D planar NAND chip from Micron. While this is the first instance of a publication including both CT and FG NAND, only the CT device was 3D. A key finding in this work was that there is a data-dependence on the upset cross section in the case of 3D NAND, meaning some data patterns written to the array were much more likely to upset than others; this effect was not observed in the case of 2D NAND.

The heavy-ion induced SEU cross section as a function of LET from [39] is given in Figure 1.29. Similar to the increase in errors due to TID at higher density operation, MLC operation in this experiment resulted in a higher cross section. Figure 1.30 shows the upset cross section for proton-irradiated devices at SLC and MLC. As expected, the smaller threshold voltage window associated with MLC operation results in a much larger upset cross section.

Additionally, the cross section was observed to be higher at lower proton energies -- a result which is discussed and supported later as a part of this dissertation.

Figure 1.31 shows that there is a fluence-dependence for SEU cross section, regardless of LET. At higher fluences, upset cross sections tend to decrease. It is proposed that previous upsets caused at lower fluences reduce the cross section at higher fluences due to an accumulation of upset regions with the device. As new ions penetrate the stack, higher fluences mean that the ions are more likely to interact in a region already previously upset, and thus not contribute to the cross section.

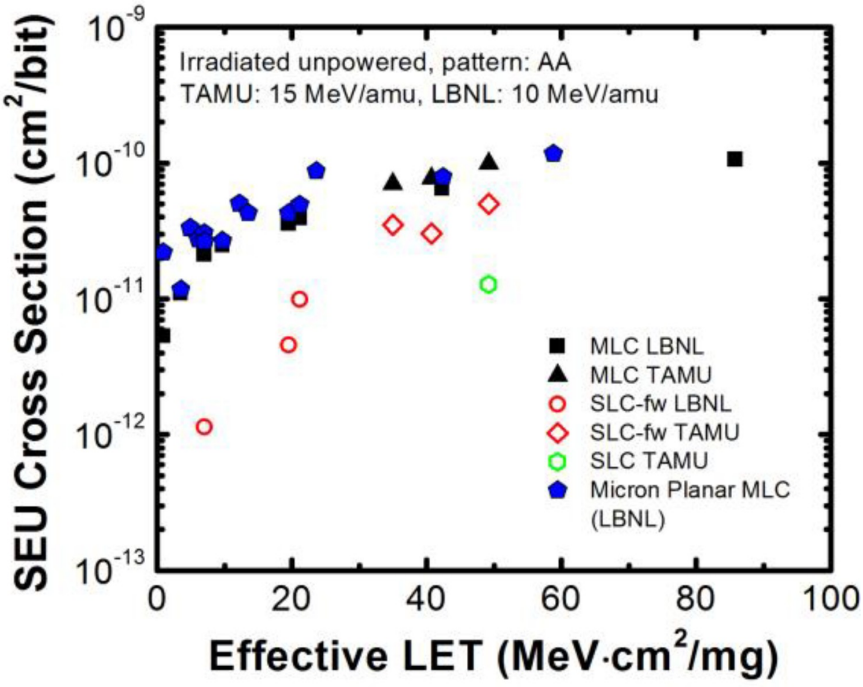


Figure 1.29. SEU for CT 3D NAND in heavy-ion environments and as a function of operating mode [39].

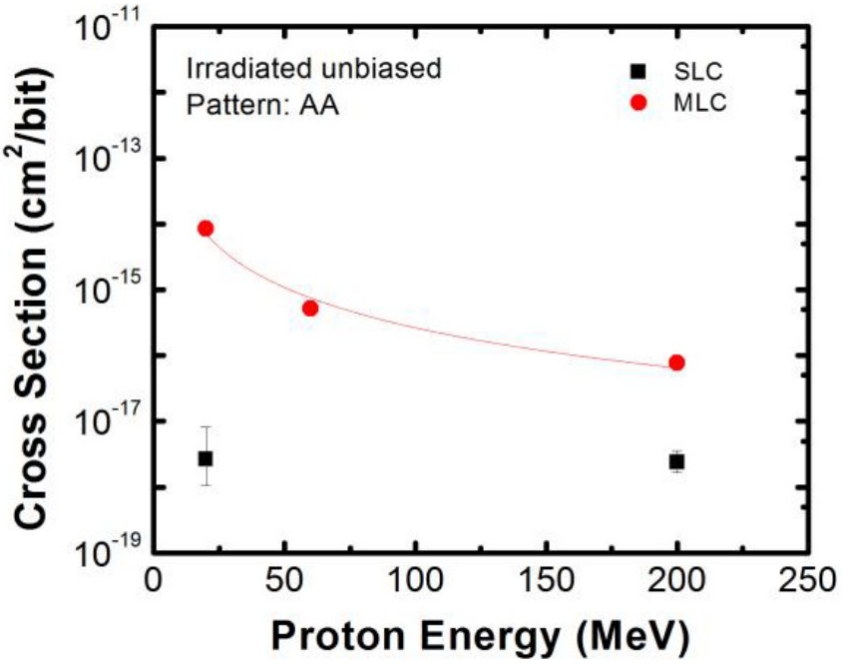


Figure 1.30. The proton-induced SEU cross section for the 3D CT NAND operated at SLC and MLC mode [39].

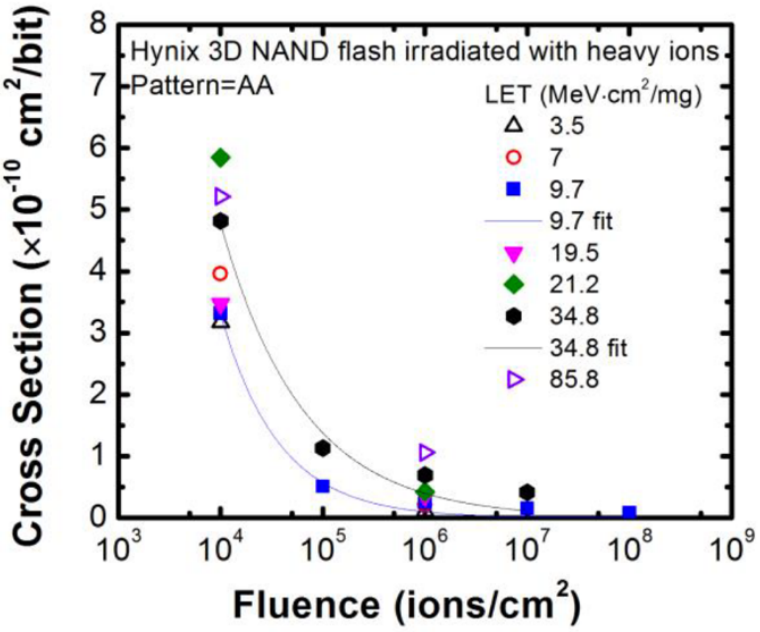


Figure 1.31. The fluence dependence of SEU in CT 3D NAND. Increasing fluence results in close to an order of magnitude variation in the SEU susceptibility of the chip [39].

Bagatin et al. present several studies of 3D NAND, most of which focus on shift in the V_{th} distribution for different cell programming levels in FG devices from Micron [51][52]. In [52], the depth distribution of ΔV_{th} for a 32-layer FG 3D NAND chip as a function of TID (X-ray) is given. The observed distribution in arbitrary units is given in Figure 1.32. The authors propose the narrowing of the channel as primarily responsible for the increased TID-induced ΔV_{th} in the bottom part of the stack as shown in Figure 1.33.

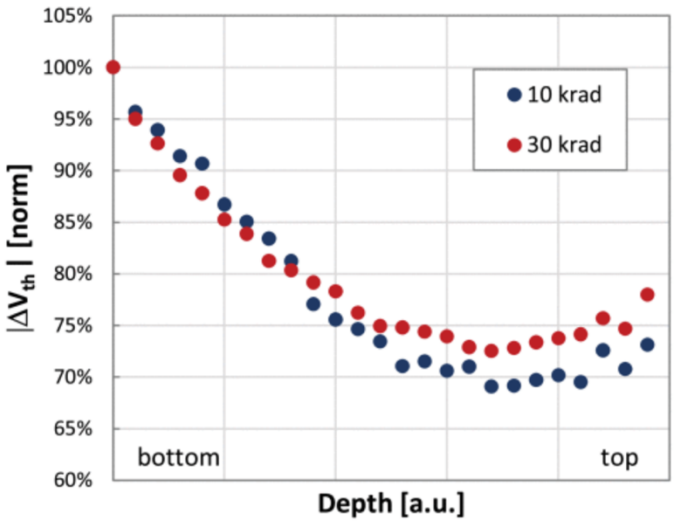


Figure 1.32. Results presented for 32-layer FG 3D NAND irradiated with a 10-keV X-ray source [52].

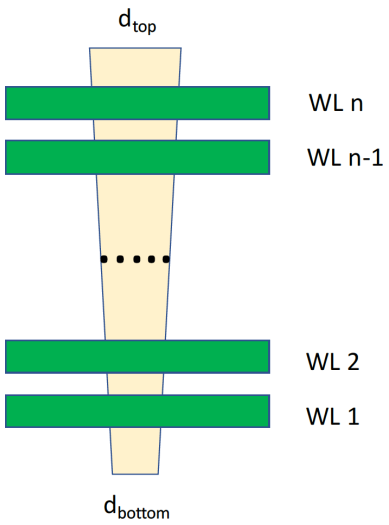


Figure 1.33. Proposed mechanism for the increase in TID-induced ΔVth observed in the lower layers of a FG 3D NAND [52].

The same authors presented a thorough study of the sample-to-sample variability in planar FG NAND technologies which showed significant ΔV_{th} within and between manufactured lots [53]. Figure 1.34 shows the projected variability in ΔV_{th} as a function of feature size in FG NAND technologies. It is noted that current 3D FG NAND is fabricated using a gate-all-around transistor with radii of less that 20 nm. The proposed dominant mechanisms for this variation were random discrete dopants and interface-trapped charge, this highlights the importance of cell-to-cell, sample-to-sample, and lot-to-lot variability within 3D NAND chips from the same manufacturer. While these results focus more on the functional rather than physical approach to understanding radiation effects in 3D NAND, the conclusions are important to keep in mind especially as cell-to-cell variations are a central part of the data presented in this thesis.

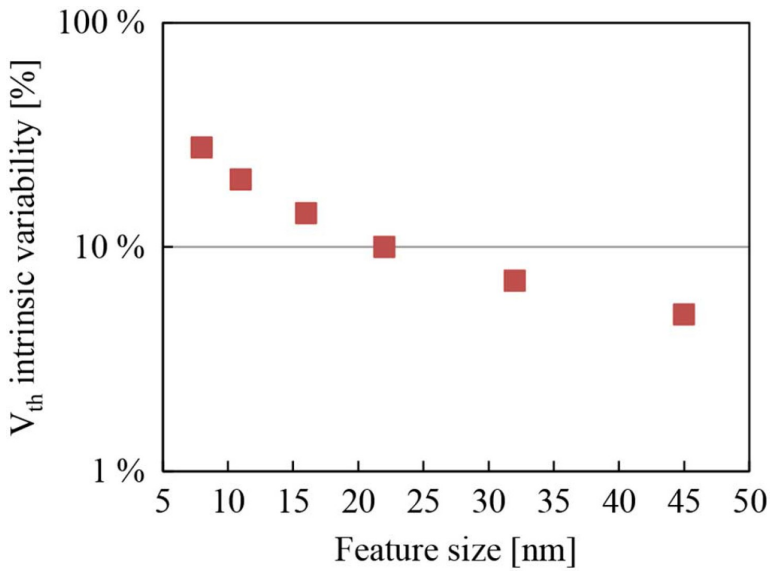


Figure 1.34. Predicted ΔVth variability in FG cells as a function of feature size [53].

Wilcox et al. presented TID and SEE data for a 3D FG NAND from Micron [54]. The SEU cross section for heavy-ion-induced SEU is shown in Figure 1.35 with the same trend significantly increased SEU cross sections observed for devices operated in MLC mode.

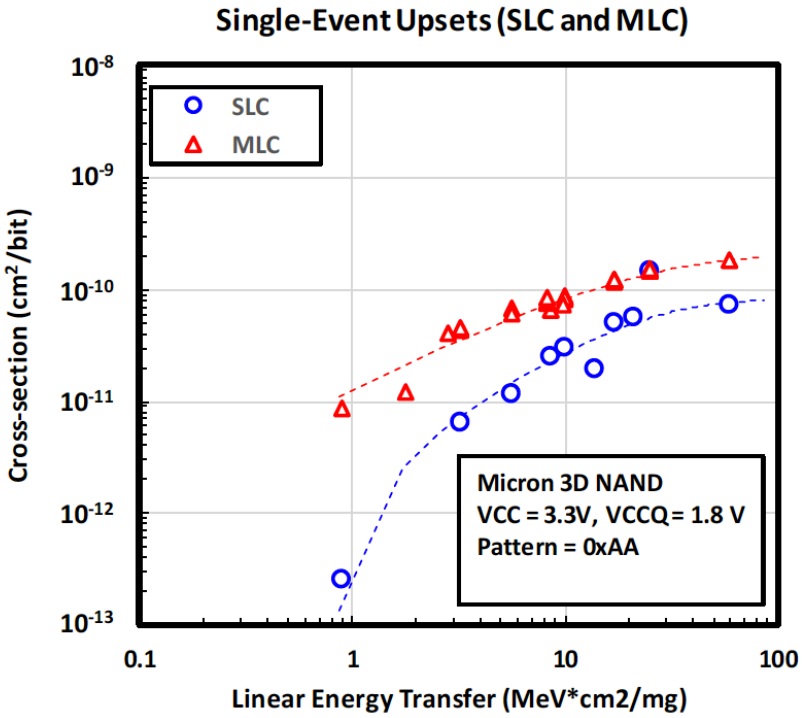


Figure 1.35. SEU cross section as a function of LET for 3D FG NAND in SLC and MLC mode [54].

Finally, Kumari et al. report a U-shaped profile in the distribution of bit-error rates (BER) due to shifted threshold voltages with a ^{60}Co gamma source as shown in Figure 1.36 [55]. These results are similar to those presented above by Bagatin et al. for X-ray irradiated devices. The number of errors in both cases increases with larger TID. There are no known data in the literature for TID effects in 3D CT NAND. In 2D NAND, FG transistors are much more commonly used than CT due to the increased control of the charge distribution for each cell. In most modern 3D NAND devices however, the CT SONOS configuration is ubiquitous, with the notable exception of Intel/micron which uses FG technology.

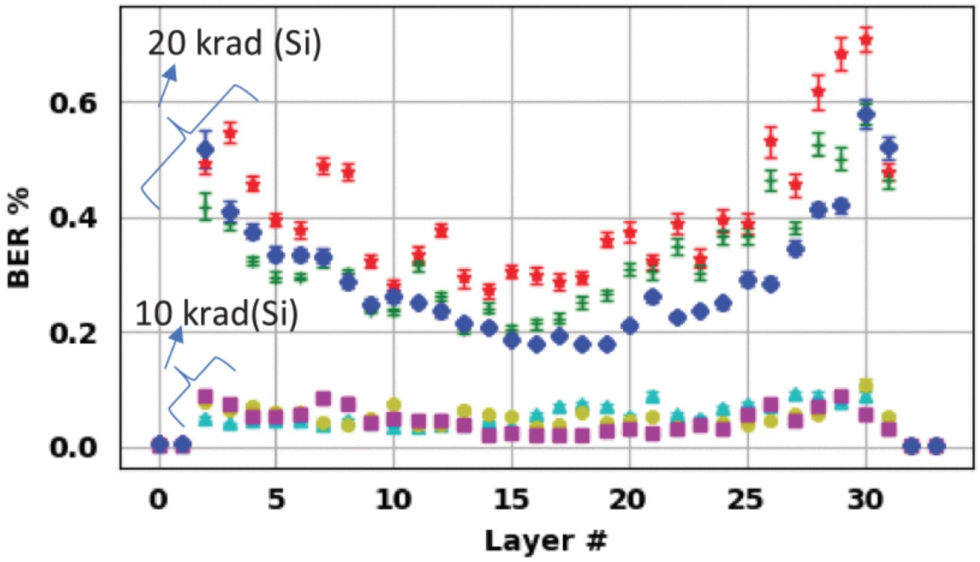


Figure 1.35. Distribution of bit error rates as a function of layer in the 3D FG NAND stack at two levels of TID under ^{60}Co irradiation [55].

The current state of radiation effects testing on 3D NAND devices is summarized in the table below:

Manufacturer/ Part number	NAND Type	Layers	Radiation testing
SK Hynix H27QDG822C8R-BCG	CT	32	^{60}Co γ [**44**] heavy ion, proton [**37**] neutron [**39**] (SLC, FW-SLC, MLC)
SK Hynix H27QDG822C8R-BCG	CT	32	Neutron (SLC, FW-SLC, MLC) [**39**]
Micron MT29F256G08CBCBBWP- 10	FG	32	^{60}Co γ (MLC) [**53**][**54**]
Micron (not disclosed)	FG	32	Heavy ion [**49**] X-ray [**50**]

Table 1.2: Summaryof radiation effects in 3D NAND from the literature.

Chapter 2

Simulation results of radiation effects in relevant 3DIC structures

This chapter presents general simulation and modeling results of vertical integration in ICs. While the focus of this dissertation is on explaining the mechanisms and specific device responses of various 3D NAND flash memories, the research presented here is intended to develop understanding about the general effects caused by structures commonly found in 3DICs and in most cases generally applies to 3D NAND flash devices.

A "3D structure" in this case broadly applies to any feature of a 2.5D or 3D IC without which vertical integration would be impossible. Specifically, the structures of interest are through-silicon-VIAs (TSVs) and metal layers configured beyond standard BEOL considerations. There is no direct comparison to a specific device or circuit here, as the motivation for the study of these structures is based on features commonly found in a wide variety of chips, some of which are introduced below. 3D NAND uses TSVs and (in some cases) many active metal layers vertically stacked; the discussion on interposers likewise provides a basis for understanding the modulation of radiation environment over relatively long vertical distances, which is then applied in greater detail during the discussion on multi-die packaged NAND.

Taking this circuit-agnostic approach means focusing strictly on the material species and geometry selected for use in 3D structures. The modulation of the radiation response of devices as a function of the presence of different material species is addressed.

In a sense, most physics-based simulation addresses the question of the impact of material selection and the effect of proximity to sensitive regions for those materials, albeit in a specific device of interest. For example, it is well understood that the presence of high-Z materials near sensitive circuit nodes results in the increased likelihood of a high charge generation single event in neutron, heavy-ion, and proton environments due to nuclear recoils [19][20][57][58][59][60]. The presence of high-Z materials can also result in an increased TID response due to the dose-enhancement effect [7][60].

Regardless of the material used in fabricating 3DICs with TSVs, it has been recognized for several years now that these structures pose potentially unique threats to device reliability in general and radiation reliability specifically; systematic studies are therefore required and have been called for [29].

I. TSVs in neutron environments[1]

In MRED, a 2 μm x 2 μm x 20 μm silicon wafer is defined as the substrate semiconductor material, with a 0.5 μm x 0.5 μm x 2.5 μm metal TSV structure embedded at the top (defined as the surface normal to incident neutron irradiation in the simulation world). Sensitive volumes were then constructed as regions in which energy deposition histograms are collected during the simulation. These were positioned directly below the TSV structure (again,

[1] Adapted from M. L. Breeding, R. A. Reed, K. M. Warren and M. L. Alles, "Exploration of the Impact of Physical Integration Schemes on Soft Errors in 3D ICs Using Monte Carlo Simulation," 2019 IEEE International Reliability Physics Symposium (IRPS), 2019, pp. 1-7.

relative to incident neutrons) at distances ranging from direct contact to up to 625 nm away in the substrate material. The dimensions of the sensitive volumes were set to 500 nm x 500 nm in the X-Y direction with varying levels of thicknesses from 10 nm to 250 nm. A cartoon of the general simulation structure is given in Figure 2.1. Neutrons were simulated as a directional flux (broad-beam) at energies of 14 MeV, 30 MeV, and 300 MeV.

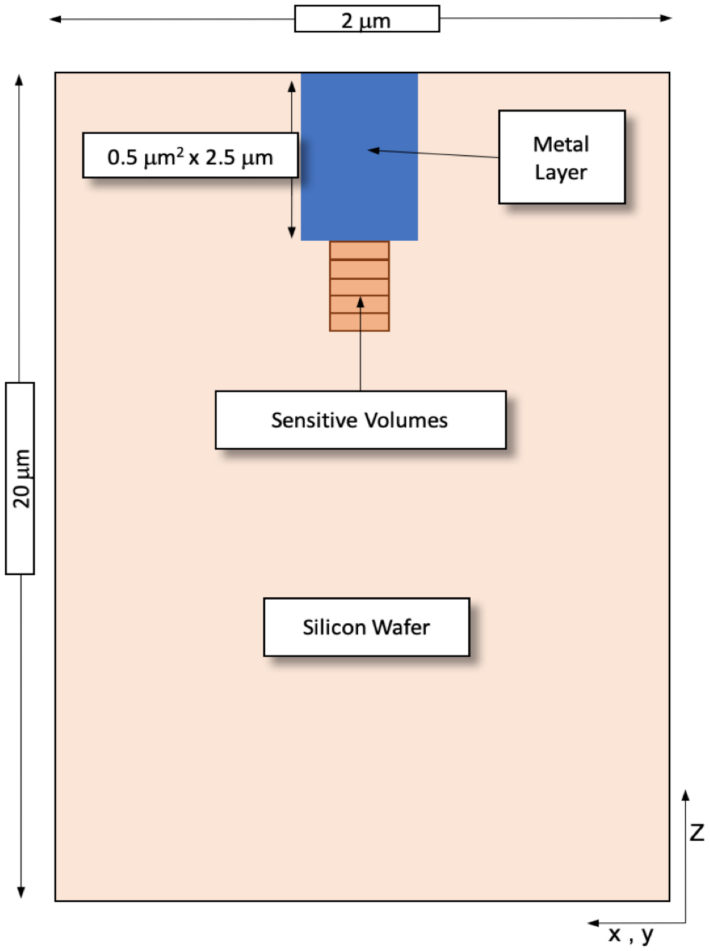


Figure 2.1. Simulation structure for TSV in MRED

Copper, cobalt, tungsten, and ruthenium are the metal species used as the TSV; additional simulations were performed with silicon as the TSV material to provide a reference point by

which differences in the energy deposition response in the sensitive volumes can be meaningfully compared.

The integrated cross section for a given threshold, $\sigma_{E,th}$, is given in units of cm^2 (or cm^2 bit^{-1}) and denotes the relative probability of depositing an energy E_{dep} greater than or equal to some threshold energy E_{th} required to produce an effect in the circuit (commonly an upset, but E_{th} can refer to other effects such as latchup, transients, burnout, etc.). The energy deposited within a region is directly proportional to the charge generated within the region, usually expressed by the conversion factor 22.5 keV/fC for Si.

Figure 2.2 shows the result of the integrated cross section for 30 MeV neutrons with Cu, W, and Si as the TSV materials. To compare the effects of these different materials on the energy deposited, a range of energy thresholds are considered (illustrated as vertical lines in Figure 2.2) where the intersecting points give the relevant cross section for a given system. This range is based on the minimum values of Q_{crit} as given in Figure 2.3 corresponding to relevant technology dimensions.

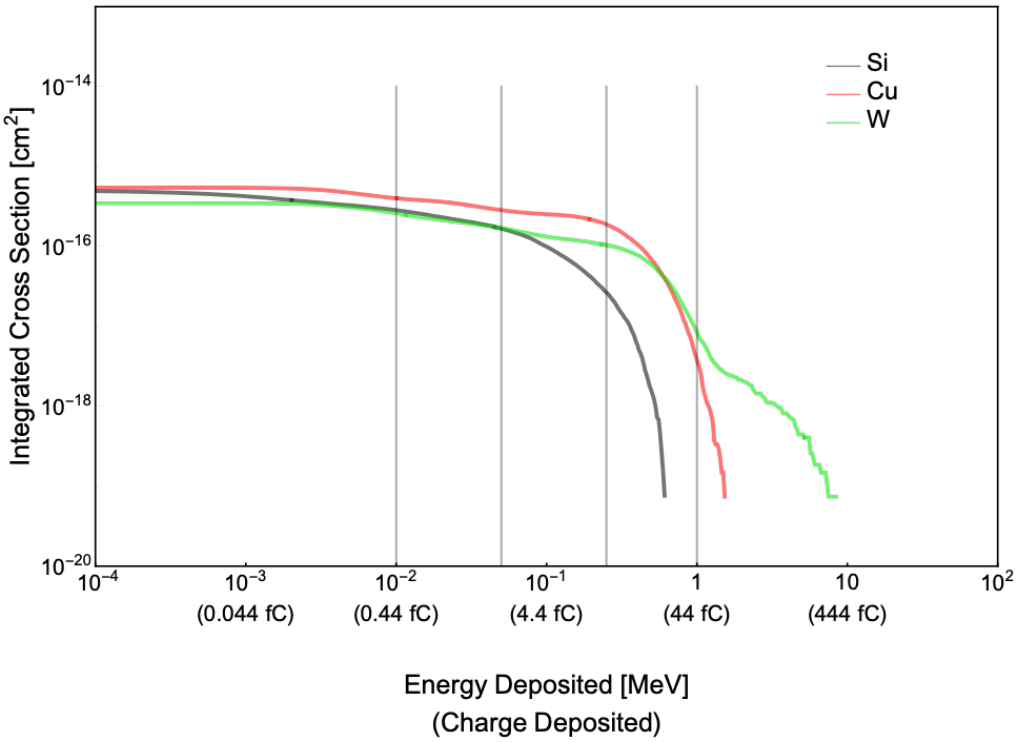


Figure 2.2. Example of integrated cross section plot generated from the energy deposition histograms for sensitive regions below a TSV. The vertical lines denote energy thresholds, i.e., the relative probability of depositing a certain amount energy (or generating a certain amount of charge) greater than or equal to the threshold value.

Sampled at lower and higher thresholds, Figures 2.4 and 2.5 highlight the differences in response between simulations with 30 MeV and 300 MeV neutron environments for sensitive volumes 100 nm x 100 nm x 10 nm placed directly below the TSV structures composed of the indicated metals. Simulations indicate that the metal layer effectively shields the sensitive region from energy deposition via neutron-produced secondaries regardless of metal species at sufficiently low thresholds. Low energy neutrons do not produce high LET secondary fragments from high-Z targets. Figure 2.4 demonstrates this effect, with all metal systems at lower cross sections than that of pure silicon for 0.44 fC. At higher thresholds and energies (Figure 2.5) this

shielding effect becomes a function of metal selection and thus of more importance for 3D ICs. High-Z materials such as tungsten are explored in detail as a function of energy in [19].

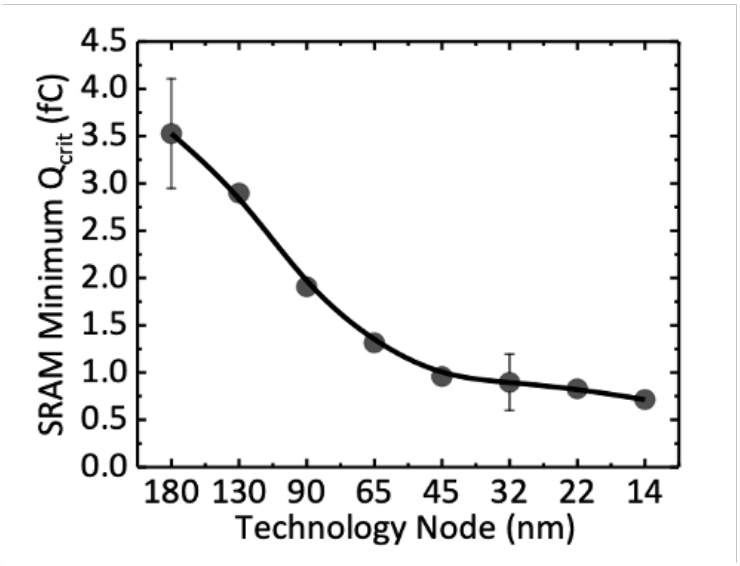


Figure 2.3. Critical charges for SEUs in SRAM technology as a function of feature size [61].

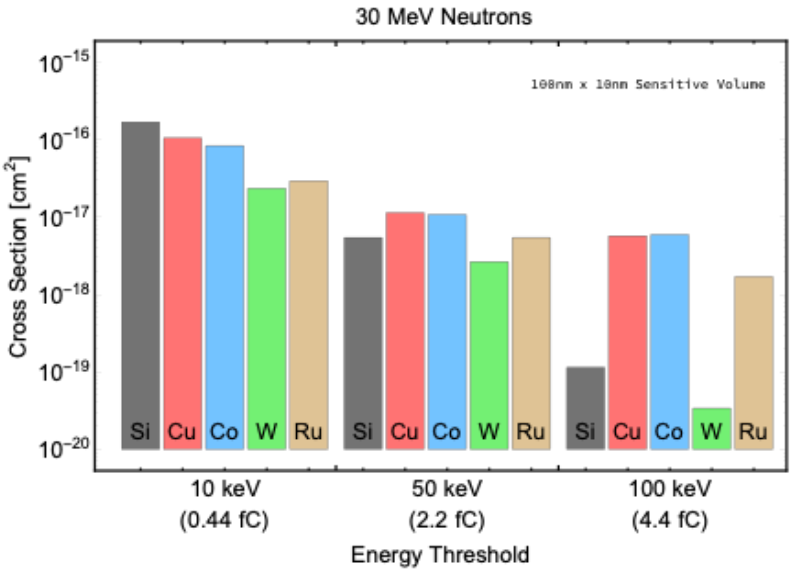


Figure 2.4. Material response to 30 MeV neutrons; secondaries produced in the metal layers of the given species tend to not produce energetic recoils as frequently as those generated in silicon at 30 MeV.

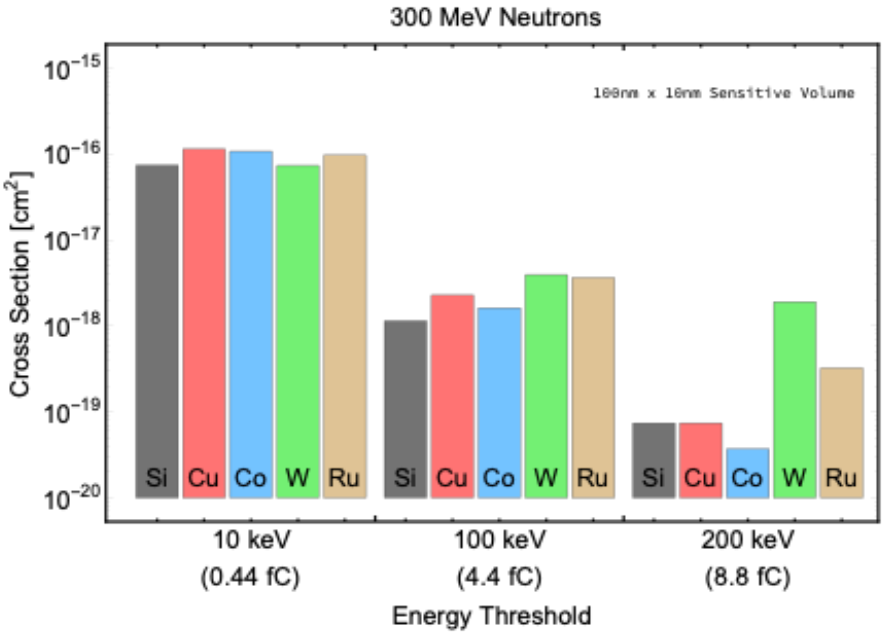


Figure 2.5. 300 MeV neutron response. At higher thresholds, tungsten and ruthenium result in significantly greater charge generation within the sensitive volumes than observed in other metals.

For 300 MeV neutrons, the presence and close proximity of some metals offer advantages preventing local soft errors, especially for Cu and Co at high thresholds. Whereas, the tungsten-modulated response is severely depressed in the 30 MeV neutron environment, a sharp rise in sensitivity for the 300 MeV results the highest thresholds. Ruthenium shows a response roughly equivalent to that of the more common copper and cobalt, with similarly amplified effects at higher thresholds for the 300 MeV results.

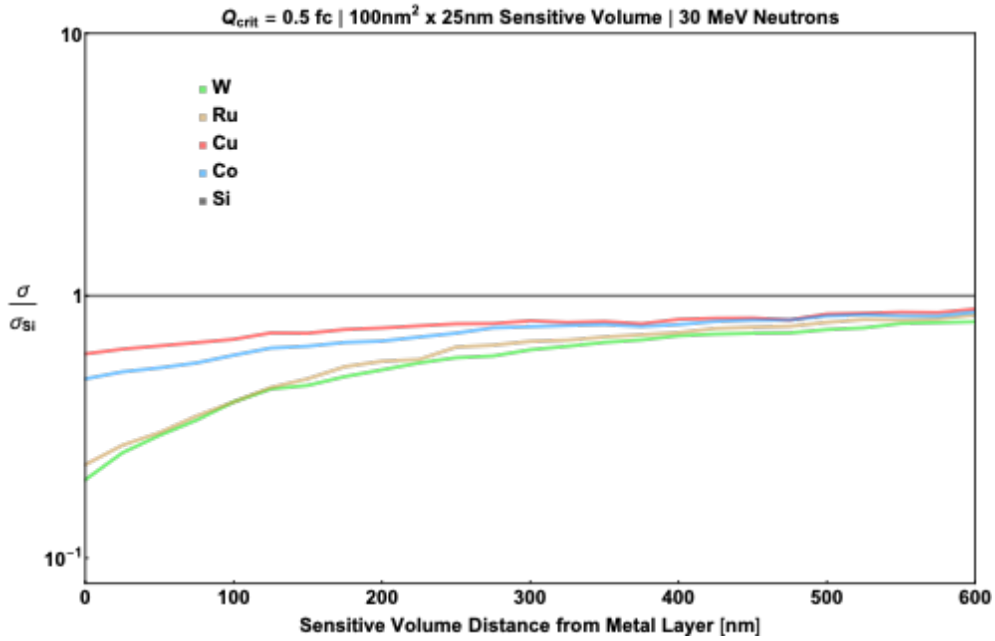


Figure 2.6. Integrated cross section ratio for all four metal configurations with Qcrit = 0.5 fC. A large material-dependent response is seen for sensitive volumes the closer they are to the TSV structure. This response becomes less pronounces at higher Q_{crit}.

As the differences in response between material systems are of primary importance, the ratio of cross sections within the full energy threshold range provides an intuitive way of comparing the response of material systems as the sensitive volume is moved farther away from

the metal layer. The cross-section ratio is the integrated cross section for simulations configured with a given metal is normalized to the configuration ran with silicon in the metal region, and is plotted in Figure 2.6 and 2.7 for 30 and 300 MeV neutrons; note that nearly an order of magnitude difference in soft error mitigation is suggested by simulations for sensitive volumes situated 50 nm or closer to bulk tungsten or ruthenium metal layers. The ratio for all four metal systems trends toward the baseline measurement (no metal present) as the sensitive region is moved further away. While differences still exist in the 30 MeV case out to half a micrometer, the shift is less than an order of magnitude. The same trends obtain for sensitive volumes of smaller dimensions. In the case of 300 MeV neutrons at a relatively sensitive threshold (1.0 fC), there does not appear to be significant variation as a function of material selection.

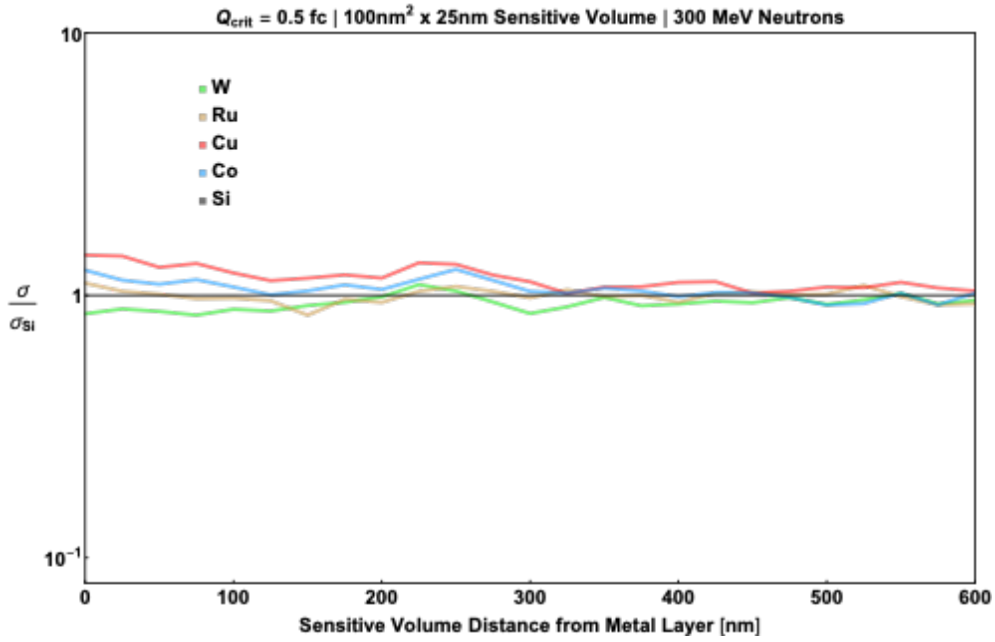


Figure 2.7. 300 MeV neutrons result in a much less pronounced materials-dependent effect.

While neutron effects are not included in the following analysis of 3D NAND structures, these results suggest further work to be done as TSVs are ubiquitous in all 3D NAND types,

whether CT or FG, often as contact vias for the WL staircase, and in some devices through sections of the memory stack itself between blocks to provide connectivity between the bit-select transistors and the control circuitry located underneath the stack. Additionally, secondary particles may be generated through nuclear reactions by neutrons in the BEOL as shown in Figure 2.8, or within the tungsten WL contacts in the memory array.

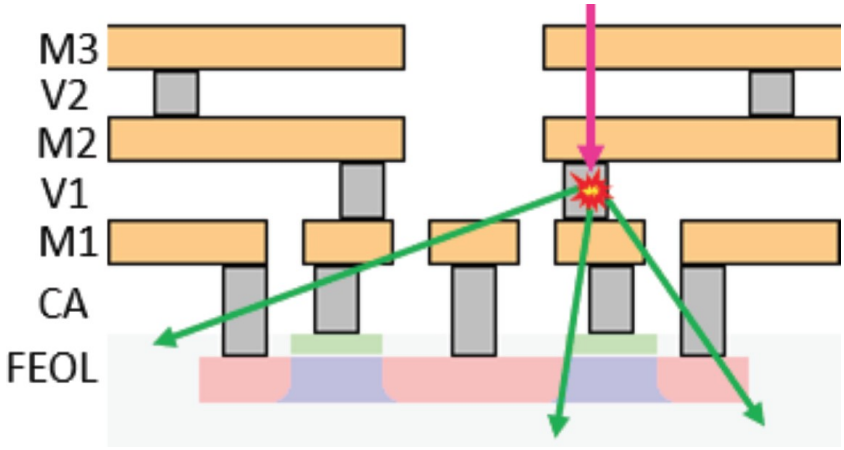


Figure 2.8. Schematic of nuclear reaction occurring within a via layer in the BEOL of a circuit [3].

II. Dose-enhancement effects from vertically stacked metallization layers

Several sets of simulations were performed using MRED to characterize the dose enhancement profile across various metal species commonly used in microelectronics as a function of overall depth at levels that correspond to 3D structures. These simulations are general in nature but apply specifically to 3D NAND both in the case of charge-trap (CT) NAND which uses metallization throughout the memory stack and in floating-gate NAND in which case multiple active die still contribute to a regenerative dose-enhancement effect due to the BEOL

and control circuitry under the memory array. The chief goal of this effort is to estimate the dose-enhancement effect profile as a function of material (including both substrate and metal species), photon energy, depth, and directionality. Developing this understanding is of specific concern for TID testing of 3D NAND flash memories, which has been reported in the literature with X-ray and ^{60}Co experimental configurations.

Figure 2.9 provides a cartoon representation of the simulation structure in MRED. Sheets of metal were placed at regular intervals throughout an RPP structure, with each substrate region between the metal comprised of 250 nm thick sensitive detector regions. The metal species included in simulations were copper, cobalt, aluminum, silver, ruthenium, tungsten, and gold. Additional simulations were performed with the substrate materials (Si, SiO_2) used in place of the metal. Recall that the dose-enhancement factor (DEF) is defined as the ratio of excess dose to equilibrium dose; structures of entirely substrate material composition thus provided the equilibrium dose. Photons were simulated at normal incidence, with an extra 2 μm above the top metal layer for equilibrium purposes. Simulations used the Fortran Penelope physics module in MRED which provides high fidelity electron tracking beyond the 250 eV limit in the standard MRED physics list. Monoenergetic X-rays and gamma-rays, as well as X-rays from the spectra presented in Figures 1.5 and 1.6 were selected. The thicknesses of the metals ranged from very small (25 nm) to rather large (250 nm) in order to cover a range of possible 3DIC-related effects (e.g., WL contacts within monolithic 3DICs to multi-die stacks in which contributions to the DEF from several BEOL regions might be expected).

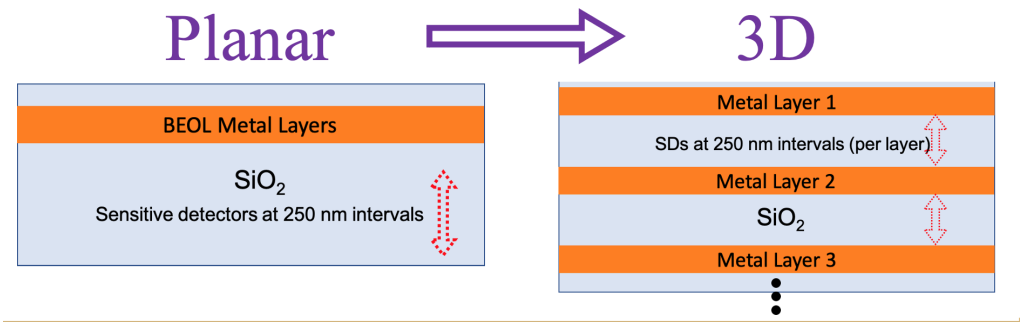


Figure 2.9. Cartoon representation of the simulation structure for examining dose enhancement in chips with multiple vertical levels of metallization.

Figure 2.10 shows the distribution of DEF as a function of depth and material for metals spaced every 1 μm for the two X-ray spectrum. Rather than falling off after the first metal layer, each additional layer of Z-mismatched material regenerates the excess dose relative to the equilibrium state. The effect is much more pronounced for higher-Z materials, and for the higher energy spectrum. In Figure 2.11, the spacing between the metals is extended to 5 μm, and the DEF distribution is seen to form a "U"-shaped profile between the metal regions. In the case of low-energy X-rays (10 keV), the DEF falls off to 1 (meaning no excess dose) beyond the 250 nm substrate regions directly in front of and behind the metal layers. This lack of DEF is not apparent in Figure 2.10 with smaller spacing.

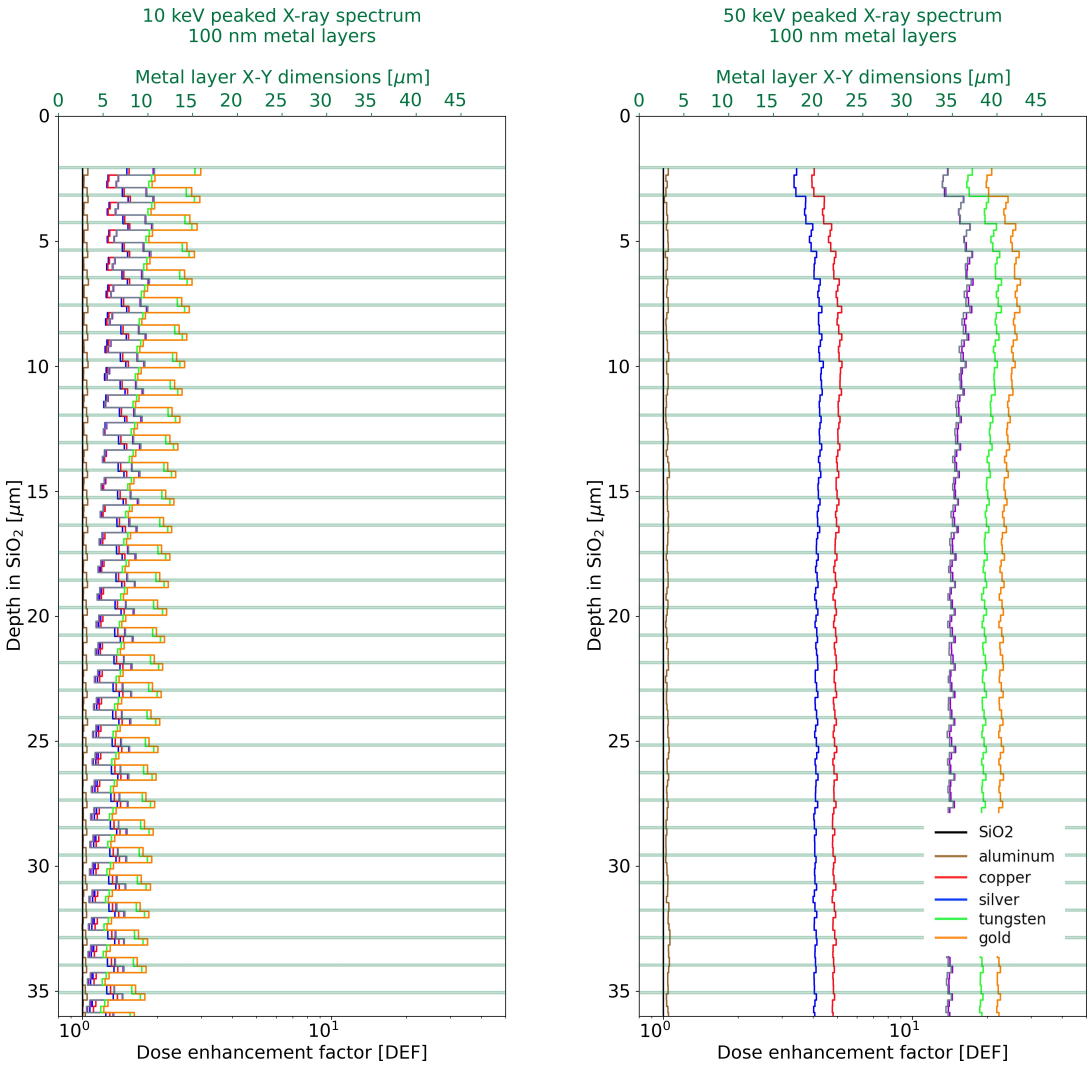


Figure 2.10. Dose enhancement at low and medium X-ray energies for 100 nm of metal (spaced every 1 μm).

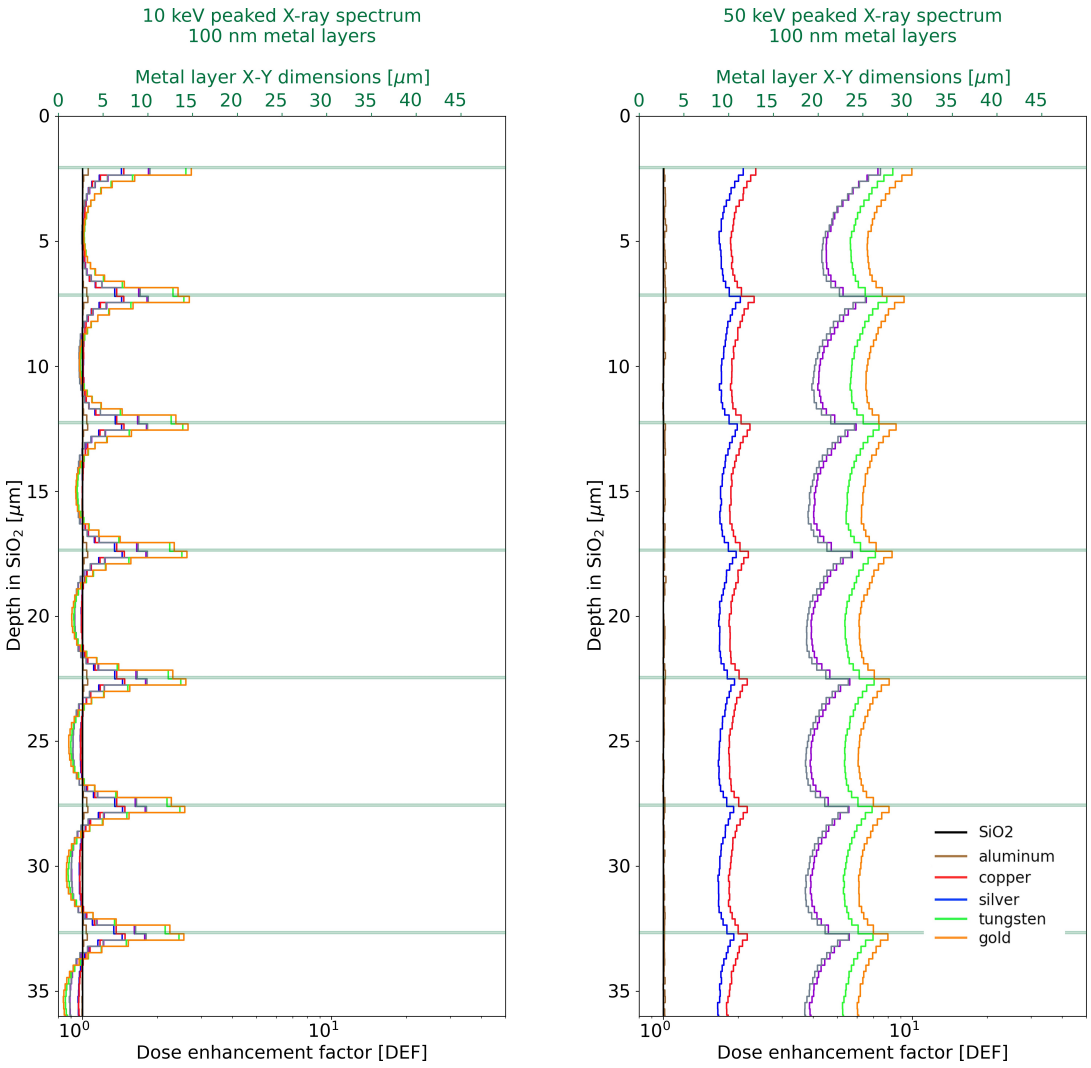


Figure 2.11. Dose enhancement at low and medium X-ray energies for 100 nm of metal (spaced every 5 μm).

Figure 2.12 shows individual tracks of electron generated within a metal layer taken from the same simulation configuration as the other figures for an increasing number, N, of incident photons. The smaller regions bound by white in the figure correspond to the metal, while several sensitive detectors make up the space between the layers, with the first ones closest to the metal region from which generated electrons are shown highlighted in green. Looking at the individual track structure, it becomes clear that the "U"-shape in the DEF profile is due to a significant

portion of electrons either scattering directly backwards relative to the incident photon, or eventually traveling in that direction after numerous scattering events.

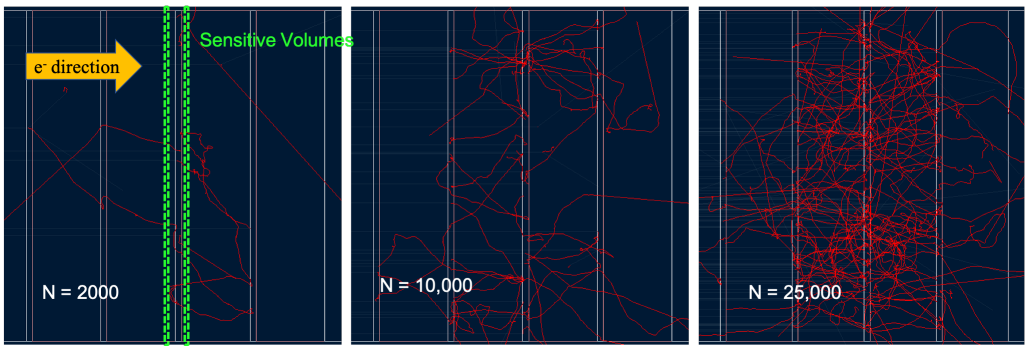


Figure 2.12. Visualization of electron tracks within the simulated structure for incident X-ray-energy gamma primary particles. There is no discernible directionality to the spreading of the electrons (qualitatively; quantitatively there are directional dependencies) leading to a dose-enhancement effect on both sides of the metal layers relative to the beam path.

It is also worth noting the rather strong dependence on incident photon energy in its ability to produce a dose-enhancement effect. The lower energy X-rays sampled from the 10-keV peaked spectrum have a much more sharply defined shape as well as a significantly smaller dose enhancement factor. The regions directly adjacent (within 250 nm according to the data as presented) to metallization are much more likely to experience a large DEF in the presence of high-Z metals such as tungsten. There is minimal effect due to lower-Z metals such as copper, with DEF falling to 1 for all metal species beyond a distance of around 1 μm from metal layers 100 nm thick and smaller in either the forward or backward direction (see Appendix 1 for profiles due to thicker mono-over layers).

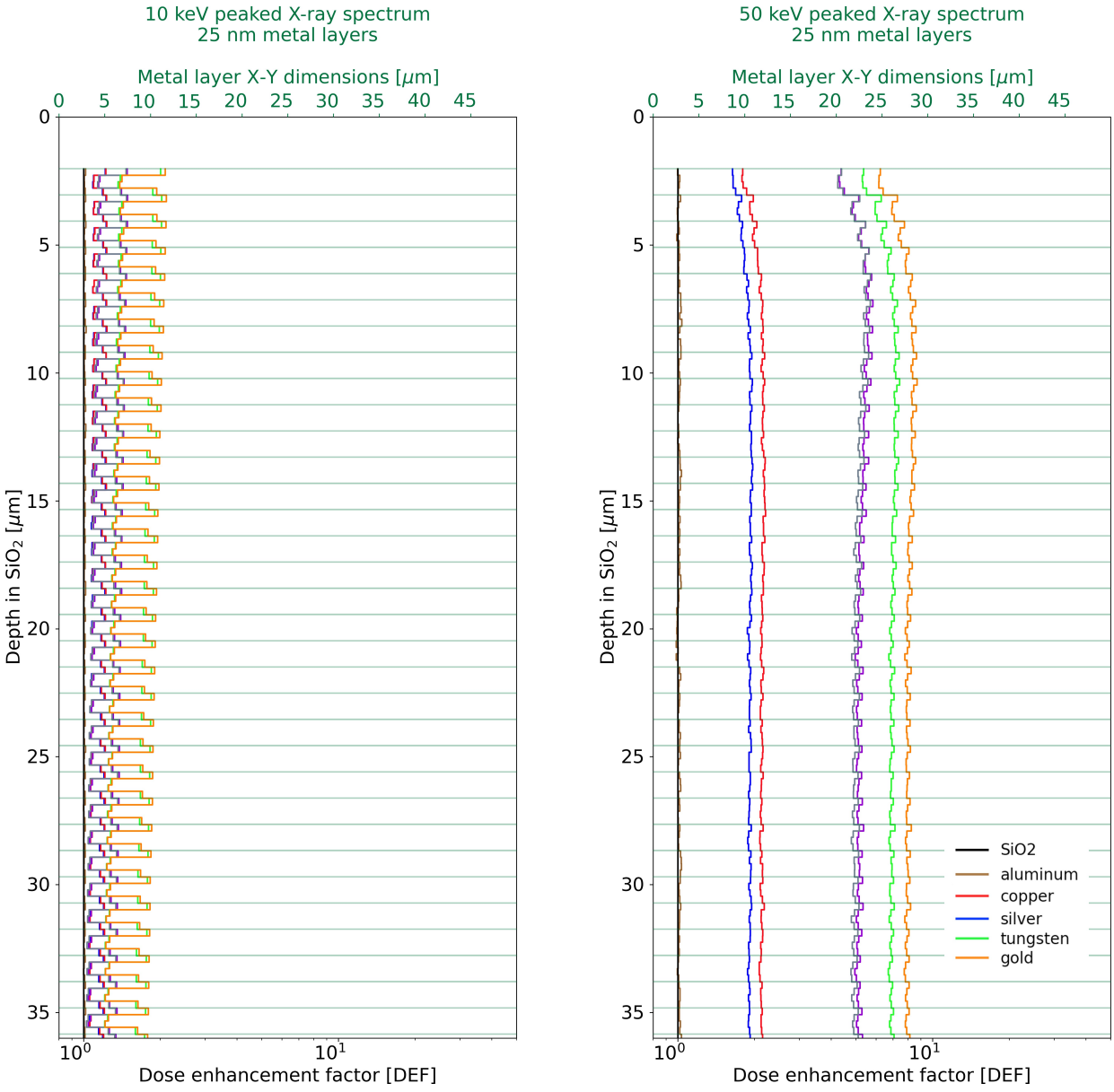


Figure 2.13. Dose enhancement at low and medium X-ray energies for 25 nm of metal (spaced every 1 μm)

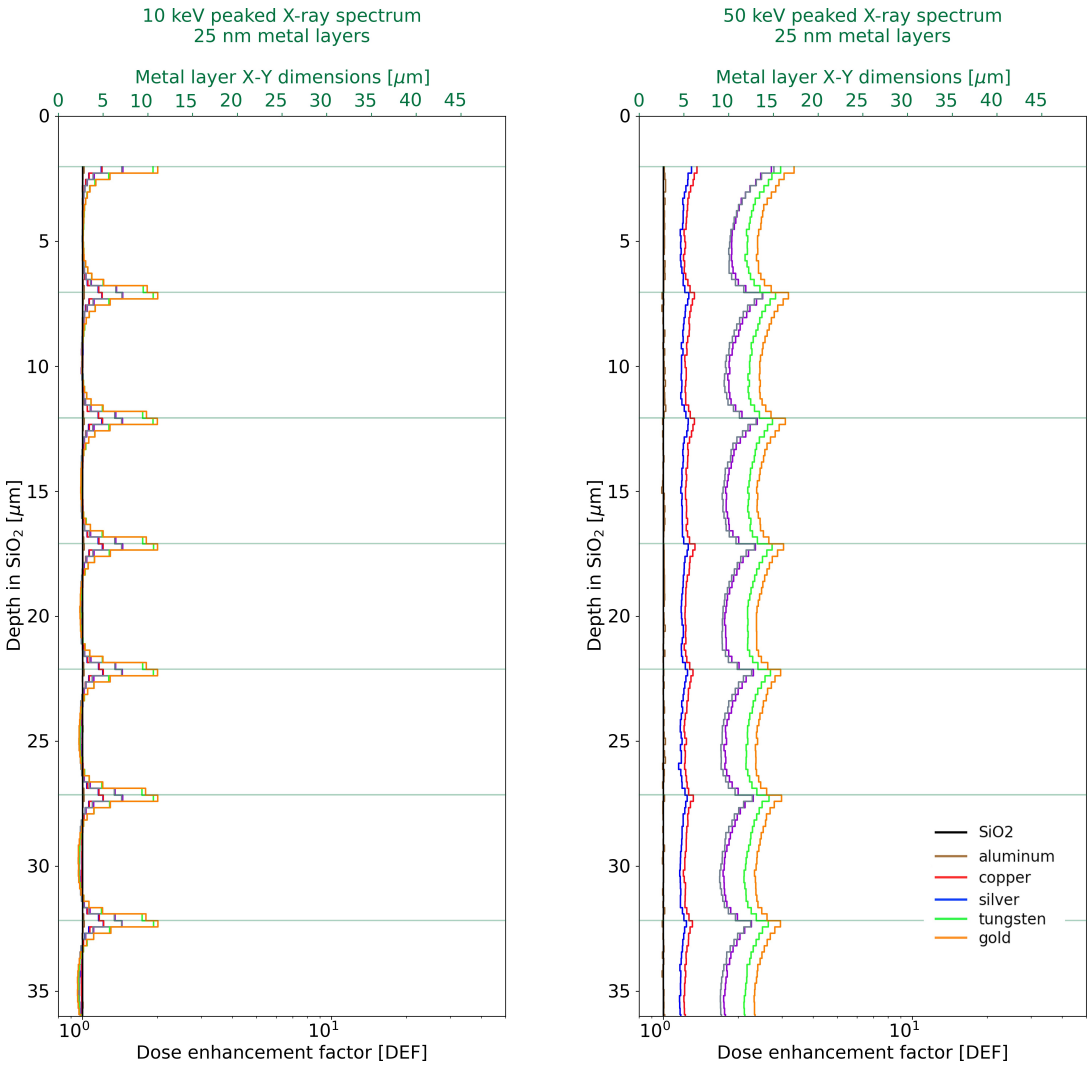


Figure 2.14. Dose enhancement at low and medium X-ray energies for 25 nm of metal (spaced every 5 μm)

Slightly higher X-ray energies however do result in a significant DEF value both for high- and low-Z metals even when spaced up to 5 μm apart. This starts to fall off with metals that are less thick (Figure 2.14), but only in the case of lower-Z metals such as copper, aluminum, and cobalt. Over twice the equilibrium dose can still be delivered to regions out to 5 μm from what is in practice a very small thickness of metal (25 nm). For example, this is of

particular note when considering the tungsten word line contacts used in 3D NAND structures, as the thicknesses involved in such devices are usually slightly more than 25 nm.

Finally, the results from ^{60}Co gammas are shown in Figures 2.15 and 2.16. Recall from Figure 1.8 that the dose-enhancement effect is much more pronounced in the with X-rays than γ-rays. Due to the long simulation times required for gammas in MRED, thicker metal layers were used and only copper and tungsten results are presented. There is still however a dose-enhancement effect observed in all cases, meaning that although considering this effect is much more important for TID testing in low- to medium-energy X-ray environments, there is potentially a significant contribution to the dose beyond equilibrium for devices tested in gamma environments, especially if the DUT has larger metallization layers. An effect due to < 50 nm thick tungsten word lines in a 3D NAND memory for example would be much less pronounced than that due to the BEOL metallization. DEF from the BEOL in gamma environments has been observed experimentally in 3D NAND before [45].

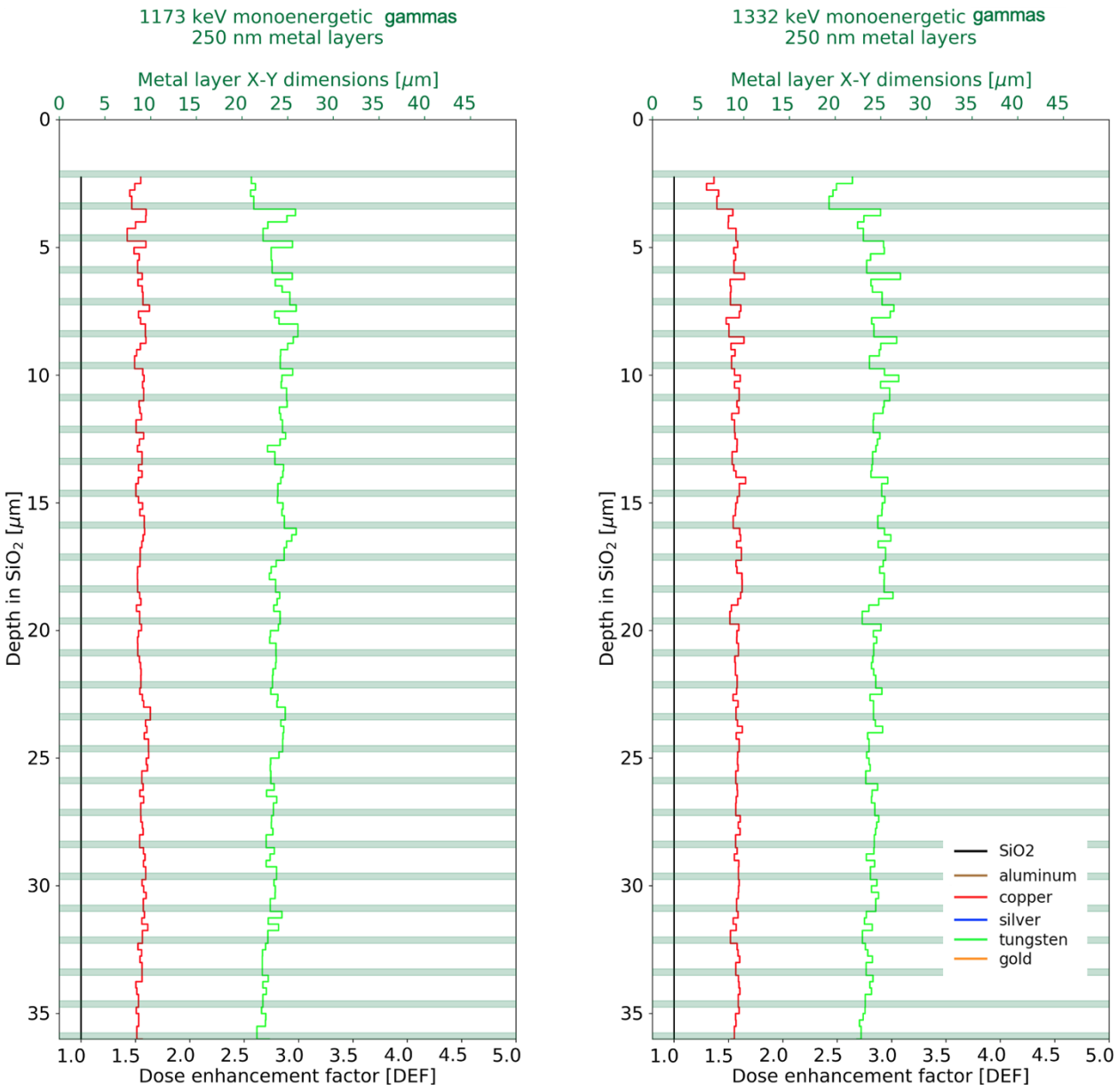


Figure 2.15. Dose enhancement at the two characteristic γ-ray energies for 250 nm of metal (spaced every 1 μm).

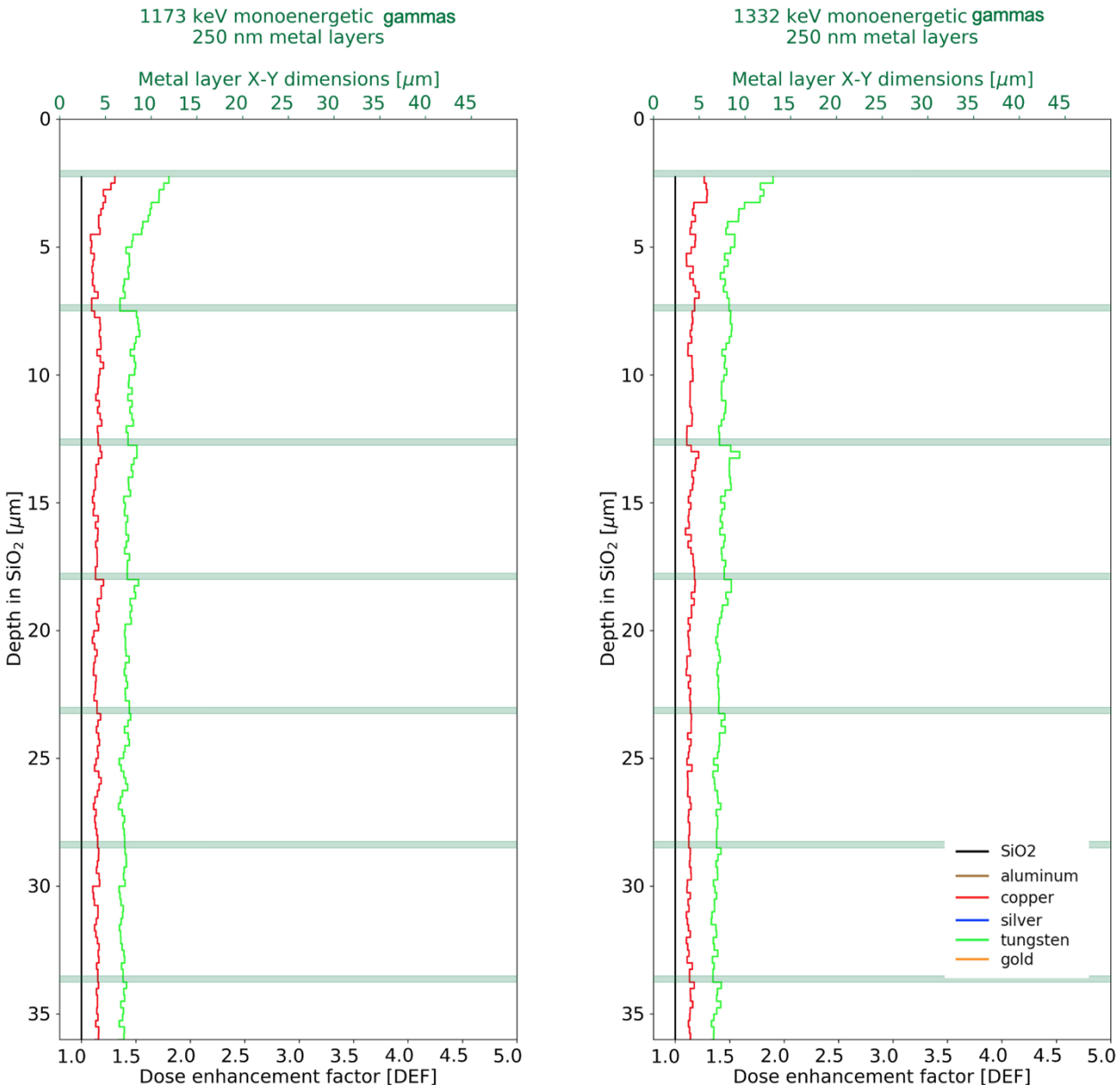


Figure 2.16. Dose enhancement at the two characteristic γ-ray energies for 250 nm of metal (spaced every 5 μm)

Chapter 3

Direct ionization from low-energy protons in CT 3D NAND[2]

This chapter presents experimental and modeling data for direct ionization from low-energy protons in a 72-layer charge trap (CT) 3D NAND chip from SK Hynix. Experimental data was taken by collaborators at NASA-Goddard, with results presented showing the SEU upset profile as a function of physical depth in the NAND array. Initial simulation efforts are presented, with unique characteristics in this experimentally-observed distribution motivating more detailed and higher-fidelity sets of simulations which provide mechanistic explanations for what was observed.

The work presented here has shown that 3D NAND is susceptible to SEU from protons with energies of at least 500 keV and that there is a large amount of variation in the number of upsets per WL level as a function of the incident proton energy which enables a direct observation of a traversing Bragg peak. Additionally, in 3D NAND chips which include high aspect ratio nitrogen-filled voids between adjacent WLs, simple methods of estimating radiation transport such as SRIM are insufficient for accurate predictions. A related finding is that the high aspect ratio etches result in a significant change in the SEU cross section between normal incident proton irradiation and isotropic or angular irradiation as is more commonly of interest for space environments. The very low-density voids extend the effective range of protons which

[2] Partially reprinted with permission from E. P. Wilcox et al., "Observation of Low-Energy Proton Direct Ionization in a 72-Layer 3-D NAND Flash Memory," *IEEE Trans. Nucl. Sci*, vol. 68, no. 5, pp. 835-841, May 2021.

boosts the overall cross section of the device (although not necessarily at the page level). Findings also suggest that the intrinsic variability in the electrical characteristics of the cells of 3D NAND arrays may still be large even after accounting for effects arising from materials and geometry which motivate future work to be done.

I. Experimental details

A 4 Tb, triple-level cell (TLC) 3D NAND flash device (part number Hynix H25QFT8F4A9R-BDF) which contains eight stacked die in a plastic 132-ball low-profile ball grid array (LBGA) was tested experimentally at NASA Goddard Space Flight Center (GSFC) using both heavy ion (presented in Chapter 4) and low energy (< 1 MeV) proton beams.

Each die within the package contains 512 Gb of TLC 3D NAND charge-trap flash memory. The core voltage (V_{CC}) is nominally 3.3 V, while the I/O (V_{QCC}) is 1.8 V. As discussed in the introduction, the memory density of 3D NAND depends on the cell operating condition; in the case of this device, the maximum memory capacities as introduced above correspond to TLC operation mode (3 bits per cell).

For the testing done in both proton and heavy-ion (discussed in the next chapter) environments, the device was operated exclusively in single-level cell mode (SLC) which results in a lower overall memory capacity but has the advantage of enhanced retention and increased device endurance, both of which are very likely to be desirable trade-offs for practical applications in space environments.

The decapsulated package is shown in Figure 3.1. Devices were prepared for testing by chemical/laser decapsulation to remove the plastic packaging material, which is necessary for ground-based heavy-ion testing using particles with lower range than the high energy particles

typically found in a natural space radiation environment. This is especially important for examining low-energy protons, as the range of protons at 1 MeV in silicon is just over 16 μm; with any significant amount of packaging material in addition to the BEOL layers, no protons would be able to make it far enough into the die to result in energy deposition events in the memory array.

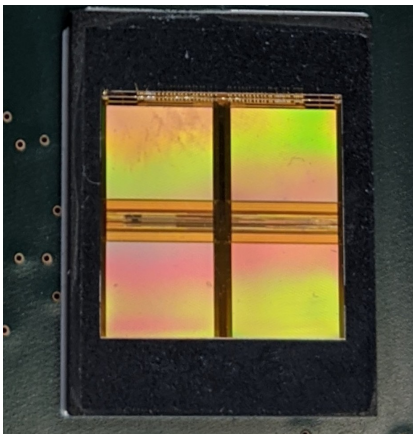


Figure 3.1. Decapsulated eight-die H25QFT8F4A9R-BDF 3D NAND package mounted on test board for heavy ion (see Chapter 4) and proton experiments.

The flash memory chip was powered by an external Keithley 2230-30-1 power supply for all erase, program, and read operations. The control signals and data were generated by a custom-built test board which included an ARM Cortex-M4 microcontroller running up to 240 MHz and providing effective test data rates of roughly 10 MB/s. The low-energy proton testing was performed at the NASA-GSFC Radiation Effects Facility (REF) using a 2 MeV Van de Graff accelerator. An aluminum masking plate was used to protect the test board from any possible electrical charging effects in the beamline as shown in Figure 3.2. The device was powered off during all proton irradiations and error correction codes (ECC) were not implemented. Fluences between 10^7 and 10^8 /cm^2 were achieved over a test period of 100 seconds, with the lower fluences used to evaluate the possibility of cross-section modulation as a

function of fluence; operational runs and the resulting data presented below were for higher fluences closer to $10^8/cm^2$. The total ionizing dose per run for proton irradiations ranged from 281 to 409 rad(Si), for a cumulative dose of 7.22 krad(Si) over all irradiation at all proton energies. Between runs, the device was verified for functionality, erased from the previous run, and re-programmed with pre-rad errors saved as a mask. A single decapsulated device not previously irradiated was tested with 45 blocks per pattern (around 255 MB per pattern) using SLC mode.

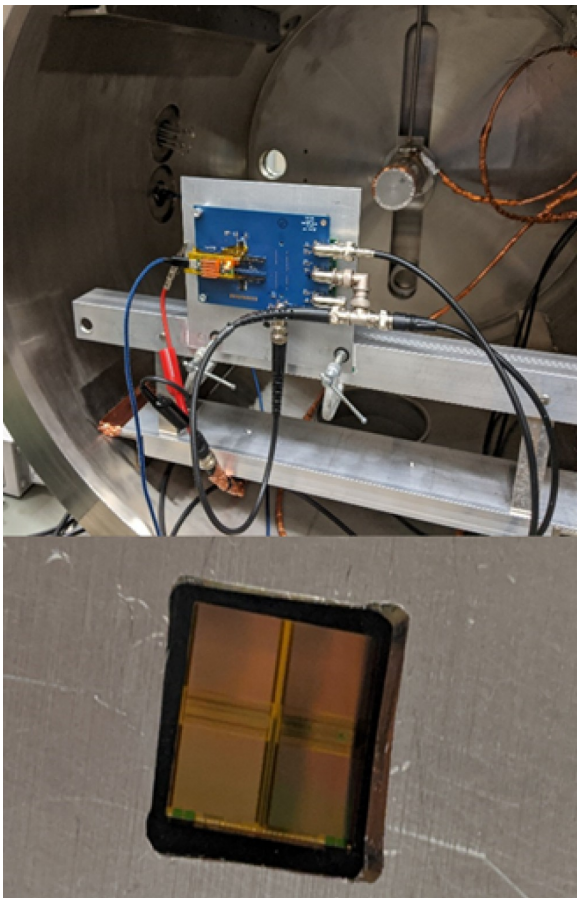


Figure 3.2. Reverse view of the test board mounted in the Van de Graff accelerator for proton testing (top), with a closeup of the aluminum masking plate on the front of the PCB (bottom).

The energy width of the beam was specified by the facility to with 2% at the DUT, with the energy level calibrated to a solid-state detector against a known source. The Faraday cup current used to compute flux is measured with an electrometer verified with a calibrated current source to within 10% agreement for each run.

II. SRIM simulations

Figure 3.3 shows a scanning electron microscope image of the 72-layer NAND structure used in this study. To achieve near stopping protons in the 3-10 μm layer of interest, appropriate proton energies must be chosen. Stopping Range of Ions in Matter (SRIM) calculations were performed to provide estimates of the range of protons in a simplified uniform silicon block (very roughly approximating the 3D NAND device), thus providing a general overview of the kinds of depth- dependence trends in energy deposition expected as the Bragg peak of the protons moves through the structure. Figure 3.4 presents the results of these simulations, which clearly show the Bragg peak shifting as a function of proton energy from a maximum deposition around 5 μm deep all the way past 20 μm. If direct ionization from protons irradiation is to be observed at all in these devices, energy deposition events from protons at these Bragg peaks are most likely to be the cause. Accounting for the 4-6 μm of BEOL materials (and recalling that the device is decapsulated), direct ionization from protons will be observed and furthermore the 3D structure of the NAND memory array will enable visualization of this traversal of the Bragg peak at different energies.

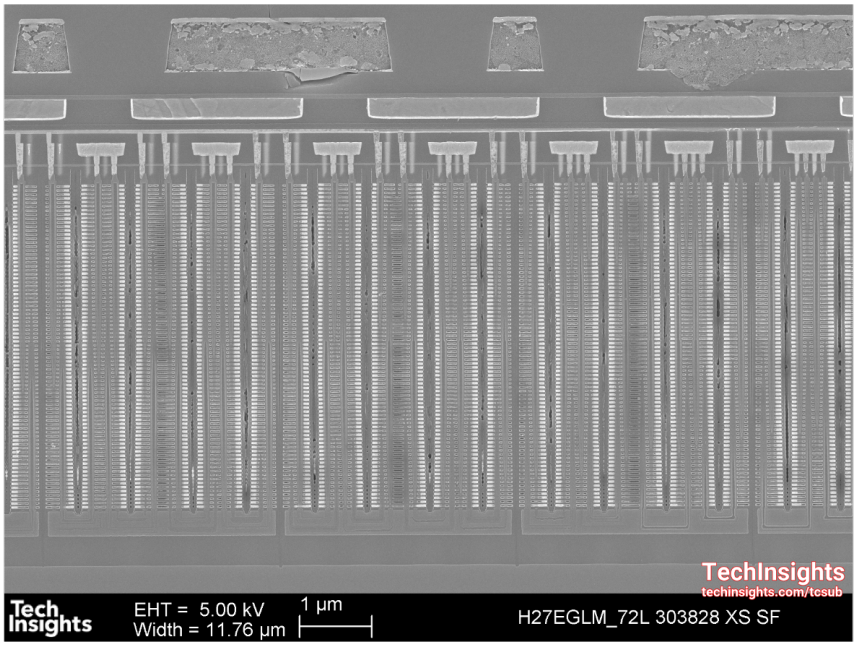


Figure 3.3. Close-up SEM of the 72-layer memory array.

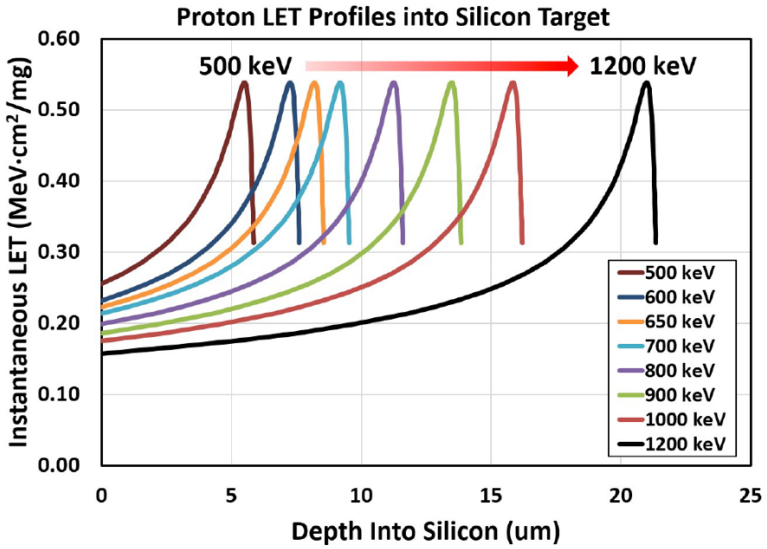


Figure 3.4. Simulation of proton LET as a function of depth in a monolithic silicon block for the proton energies used experimentally. Protons of different energies enter from the left at a depth of 0 µm and slow as they travel deeper through the material. The energy loss from this slowing is plotted as the instantaneous LET on the ordinate.

III. Proton testing results

The SEU cross-section from proton irradiation is shown in Figure 3.5, with a clear dependence on proton energy for all data patters tested. The characteristic peaking of direct ionization is present at an energy of approximately 700 keV; errors tail off rapidly as energy decreases because the beam is no longer able to penetrate the memory stack. At higher energies, the range is no longer a limiting factor, but the increased energy of each particle results in a lower LET, reducing total upset counts. Eventually as proton energy increases further, indirect ionization resulting from nuclear interactions should become the dominant source of proton-induced SEU [63].

Analysis of the proton SEU data within the three-dimensional volume confirms the predicted movement of the Bragg peak from the surface of the die to the base of the memory stack. Proton upsets were sorted by active physical device layer (0-71) then plotted as a series of histograms. Selected energies are shown in Figure 3.6 to illustrate this movement of the Bragg peak. The vertical axis of Figure 3.6 is the physical depth in the material stack up of the NAND memory array, while the horizontal axis corresponds to the number of errors counted in that layer. For the lowest energy tested (500 keV, left in Figure 3.6) only a small number of errors occur in the top-most memory layers. The distributions at 600 and 650 keV show a distribution of errors that fits remarkably well with the expected visualization suggested by SRIM over approximately the top 30 layers. At 1000 keV, relatively few errors occur throughout the entirety

of the memory stack, and the distribution is noticeably more uniform as a function of depth, which is attributed to the fact that the LET is more or less constant within the active device layers for these energies.

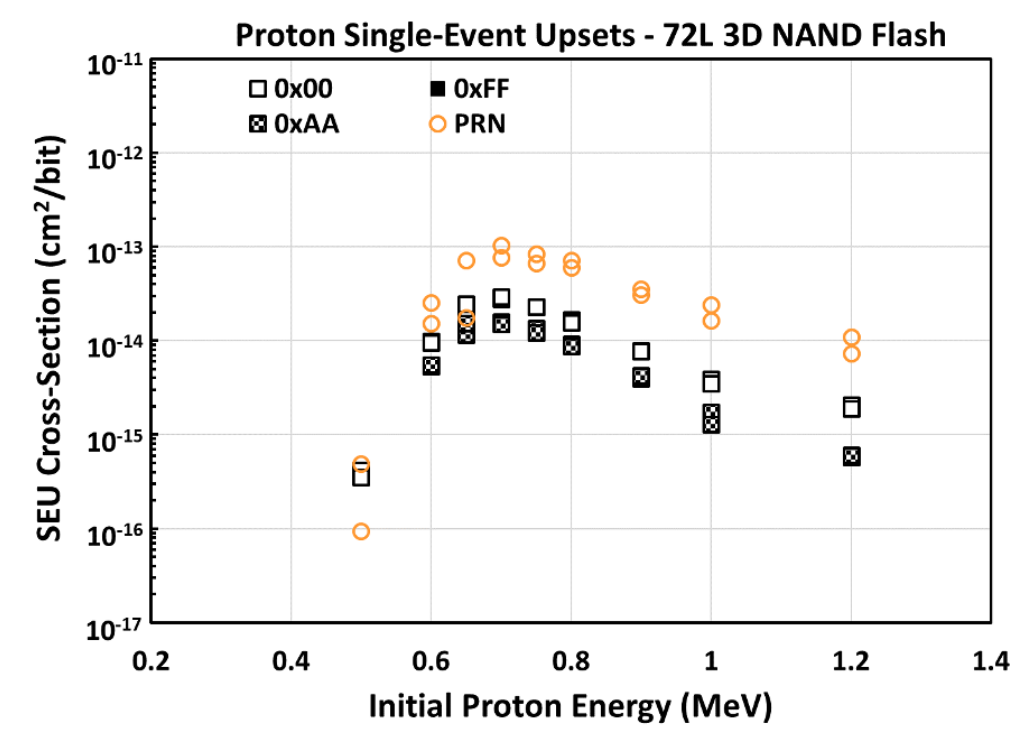


Figure 3.5. SEU cross-section vs. initial proton energy

Statistical evaluation indicates that the number of cells expected to receive multiple proton strikes at these fluences is insignificant and does not account for a significant portion of the proton-induced upsets. Figure 3.7 shows a mathematical expectation of the number of cells receiving a single hit versus two or more hits for a given fluence based on a bins and balls problem approach. The single-hit curve is calibrated to overlay the heavy ion data taken at low LET (again, heavy ion data is discussed in more detail in Chapter 4). The proton data have the expected 1 decade per decade slope of the single-hit line and are several orders of magnitude higher in occurrence than the estimated double-hit data.

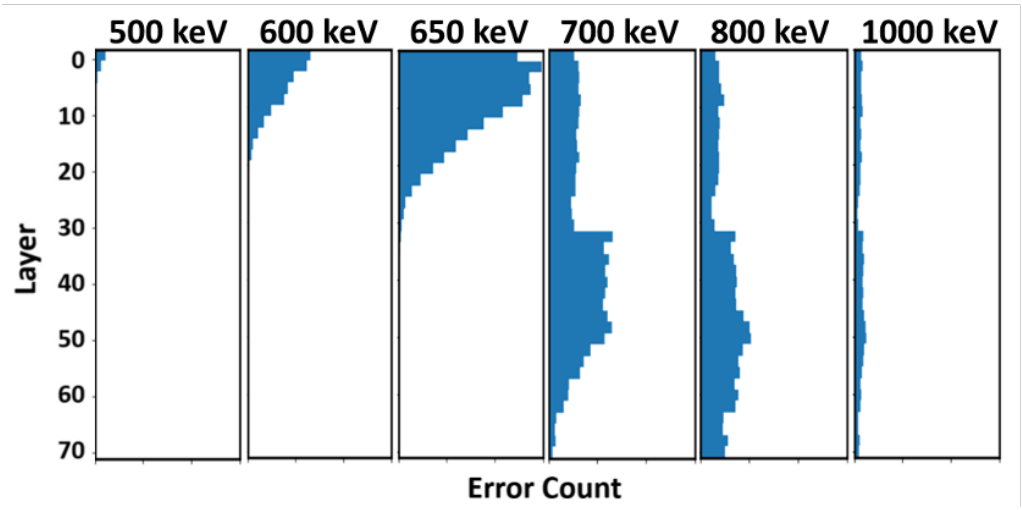


Figure 3.6. Histograms of errors by layer as proton energy increases from 500 keV (left) to 1000 keV (right).

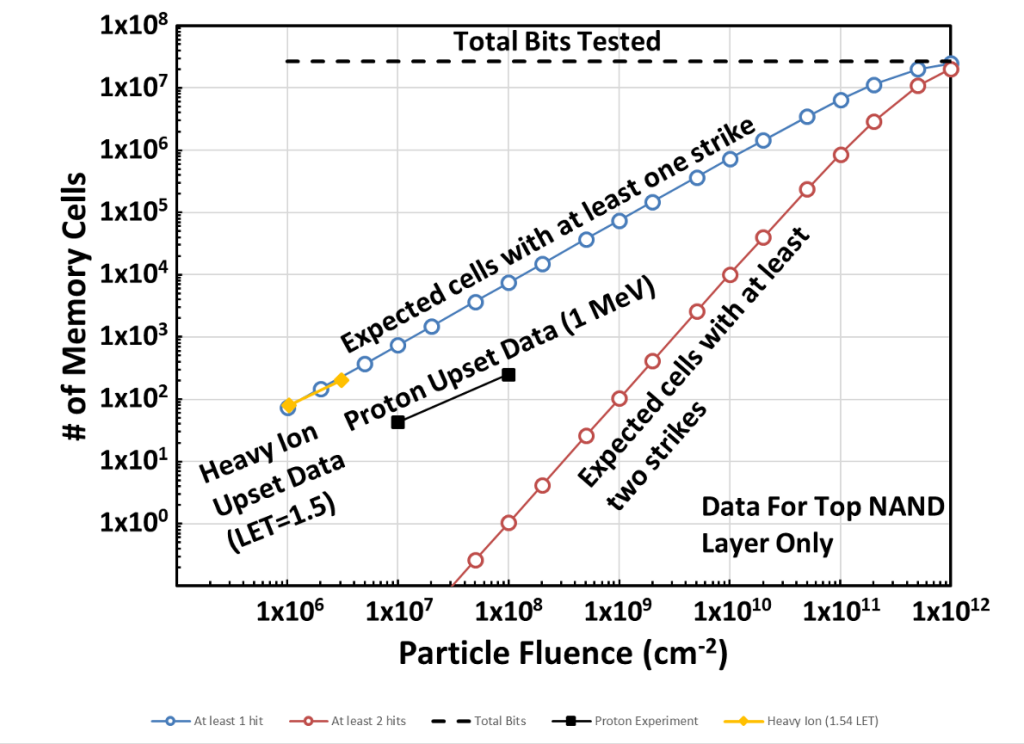


Figure 3.7. Expected number of cells with single and multiple particle strikes for given fluences, overlaid with heavy ion and proton experimental data.

Characterizing the pre- and post-irradiation behavior of the device leads to a couple different ways of determining an SEU cross section for a charge trap device. The previous cross sections reported were calculated using the total number of post-irradiation errors minus the total number of pre-irradiation errors within sets of 45 blocks for 4 different data patterns each (0x00, 0xFF, 0xAA, and PRN). However, a closer look at the data shows that there are some cases in which the raw number of pre-irradiation errors is greater than the number of post-irradiation errors, which would lead to a negative cross section.

Instead of a bulk subtraction to account for the intrinsic errors after the write operations which occur in most non-volatile memory devices, the explicit cell addressing can be mapped between the two sets of data (i.e., pre- and post-irradiation errors). The SEU cross sections using this method are reported in Figure 3.8. A consequence of this method of data analysis is that some of the intrinsic errors which appear in the pre-irradiation state are no longer observed post-irradiation. These are not included in the revised cross section calculation however, because although it is possible that proton-induced SEU 'fixes' some of the bad bits (i.e., moves the logical value in a previously bad cell to the correct one), during normal operation of the chip error correction codes mark and mask bad bits to ensure that data is not actually written to them.

The next section further explores the layer-by-layer error distribution, including the apparent discontinuity of data between 650 keV and 700 keV, a Monte Carlo simulation approach is also used as a part of this analysis.

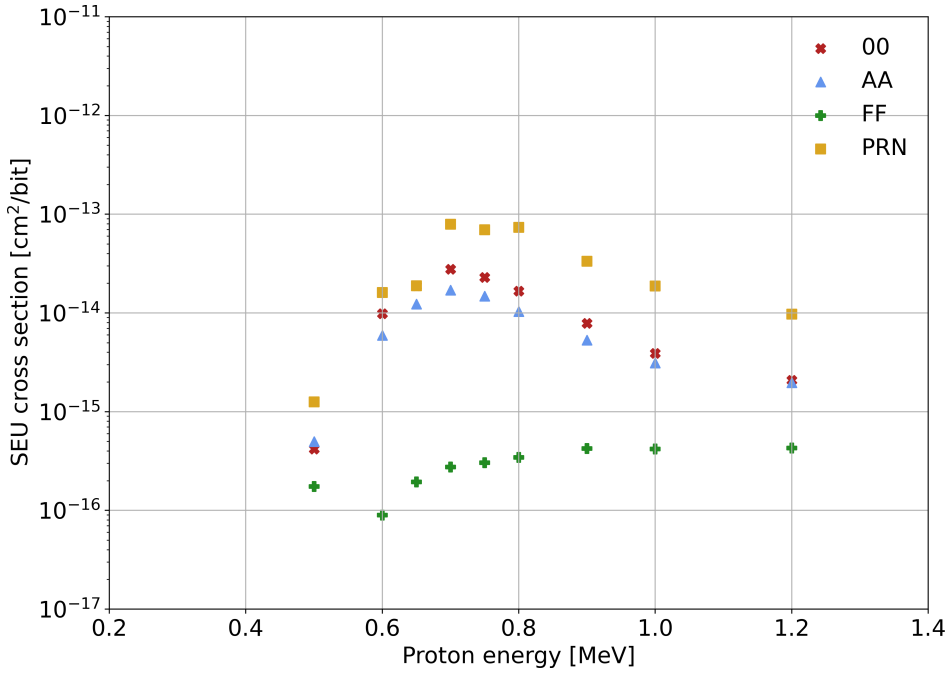


Figure 3.8. SEU cross section as a function of proton energy. Pre-irradiation errors which do not re-appear in the post-irradiation data are not considered.

IV. Initial MRED simulation structure for 3D NAND

Using SEM images of the 3D NAND device as a guideline, the simulation model is constructed as shown in Figure 3.9. The back-end-of-line (BEOL) is defined with thicknesses and materials also taken from the SEM images of the device. It fills the entire 10 µm x 10 µm simulation world size (in the x- and y- dimensions), with the memory array below defined within the center 5 µm x 5 µm region; the material outside this center area is set to SiO_2 (transparent in the figure). Immediately below the BEOL are three layers of tungsten which are 25 nm thick. These represent the three layers of control gates which sit on top of the memory array. The

memory stack itself then contains 79 more layers, each consisting of a transistor layer, followed by a 25 nm layer of SiO_2. Each transistor layer contains a repeated pattern of 3 regions tiled across the array. These material regions are W, TiN, and SiO_2, of (x, y) dimensions (5 μm, 100 nm), (5 μm, 55 nm), and (5 μm, 100 nm) respectively. Four dummy layers are included at the top of the stack, and two more dummy layers are defined at layer number 40. The only difference between the dummy layers and the standard transistor layers is the lack of sensitive detectors in the dummy layers (i.e,. no energy deposition is tracked).

Sensitive detectors are set as SiO_2 regions 30 nm thick and 50 nm in the x- and y-directions, tiled uniformly throughout each transistor layer in the TiN region adjacent to the tungsten word-line. This approximately corresponds to the physical size of the transistors in the memory array (the actual size is slightly smaller). During simulation, the energy deposition events for all sensitive detectors within the same layer (vertically) are tabulated in a single histogram; no (x, y) information about the hits is retained. Each sensitive region has a unique sensitive detector attached so that multiple-bit upsets do not artificially increase the cross section for a single bit.

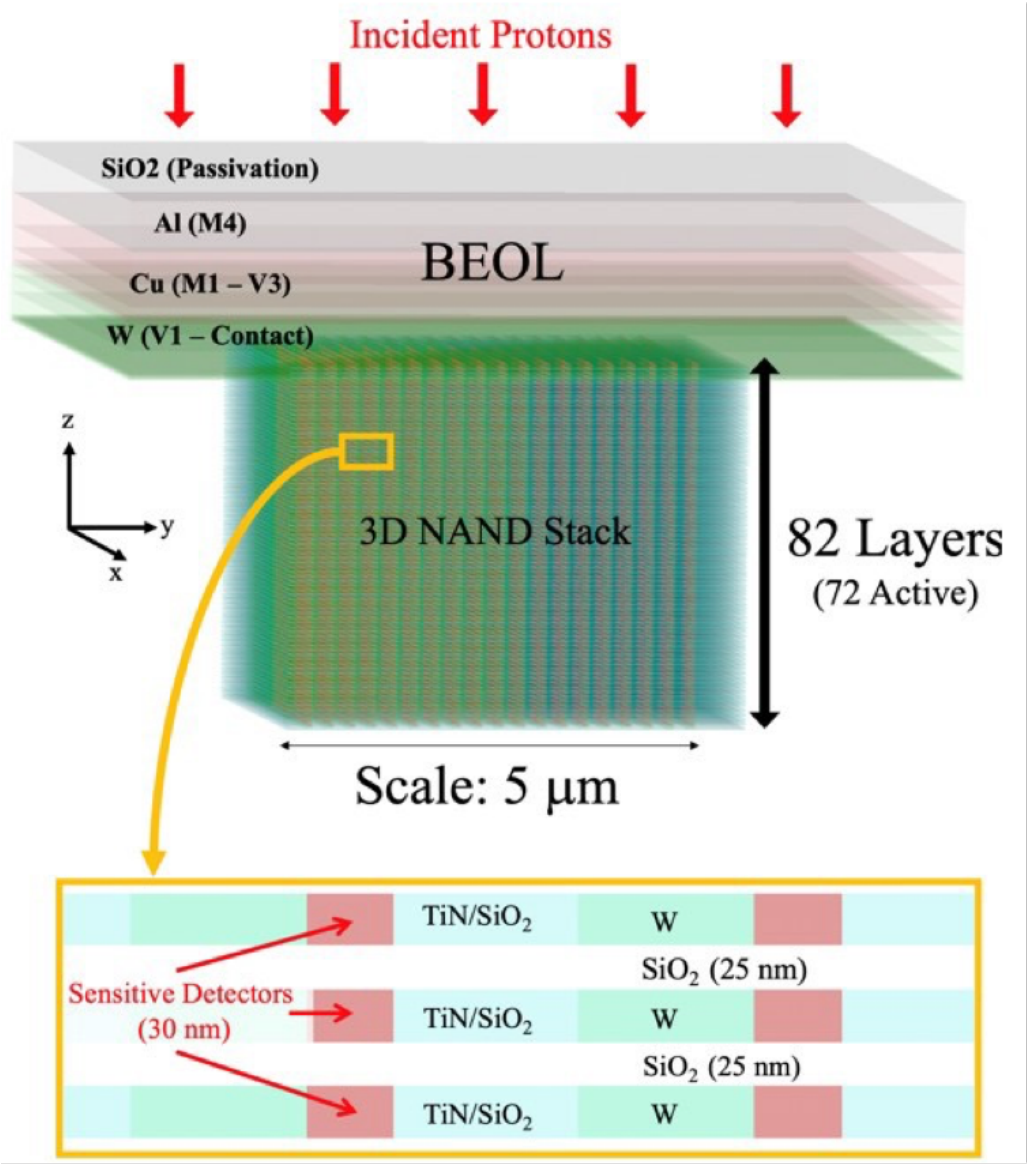


Figure 3.9. Initial simulation structure in MRED.

The material composition of the transistors is of course in reality more complex than simply SiO_2, with each transistor containing silicon nitrides, silicon and aluminum oxides, as well as the polysilicon channel. However, each new physical boundary in MRED forces a stepping action for tracks which traverse the region, and as each of the 72 vertical layers in the simulation already contains 361 sensitive detectors, including these extra boundaries would

significantly increase the computation time required while providing very little if any difference in terms of the overall energy deposition profile as a function of depth in the stack. The key assumptions made in this model are: the densities of the transistor materials are sufficiently similar to that of SiO_2 to allow for full material substitution, that the BEOL may be approximated as full metal layers, and that upsets are caused by charge deposition events which occur in the immediate physical vicinity of the transistors which make up the memory array.

As the number of simulated events accumulate, the resulting distribution for a sensitive detector approaches a limit which ideally has some physical significance. By integrating the event counts over the energy deposition bins of the histogram from high to low energy and normalizing to the simulated fluence, one obtains a cross section (in units of cm^2). The cross section taken at a specific energy, E_{crit}, corresponds to the relative probability that a single event will generate E_{dep} within the volume such that $E_{dep} > E_{crit}$, where E_{crit} is the energy required to induce an upset. Linear energy transfer (LET) is defined as the density-normalized stopping force of an ion ($\rho\ dE/dx$). If the path length l of an ion is small compared to its range, the LET can be approximated as a constant value and the critical energy calculated by:

$$E_{crit} \approx LET\ \rho\ l$$

The LET values from the heavy ion experiments are used to approximate E_{crit} of this device. Using the smallest LET value from the heavy ion testing (1.16 MeV cm^2/mg), taking the thickness of each transistor to be 30 nm (from SEM images), E_{crit} for this device is approximated as 9.2 keV (or Q_{crit} ~ 0.41 fC).

MRED simulations were completed at proton energies of 500 keV, 600 keV, 650 keV, 700 keV, 800 keV, and 1 MeV, with the ion beam randomized over the top surface of the

structure in a unidirectional flux. The total number of ions sampled in each run is at least 14 million, which corresponds to a total fluence of at least 10^{13} cm^2.

Figure 3.10 summarizes the results of the simulations described above. The cross section, which represents the relative probability of generating sufficient charge to induce an upset at a node in the memory array, is given as a function of the proton energy and depth in the stack. The notable difference between simulation and experiment is the lack of events at 700 keV below layer 40 which may be caused by any combination of the approximations used in the model, such as defining the BEOL using full layers of just metal (not including oxide), having range cuts set to 1 μm, or the simplified geometry and materials of the transistor layers. Additionally, the experimental data show a step-like feature between layers 30 and 40 at proton energies greater than 650 keV. There are many possible reasons for this feature, including an alternative material structure in the two dummy layers which exist within this range, or the possible presence of functional circuitry at this depth. The total cross section as a function of proton energy was calculated using differences in the density between silicon and silicon dioxide to account for the intrinsic cell variation within the 3D NAND (see Figure 3.11). The general agreement is good if slightly underestimated compared to the experimental data, but refinements to the MRED model are the next step.

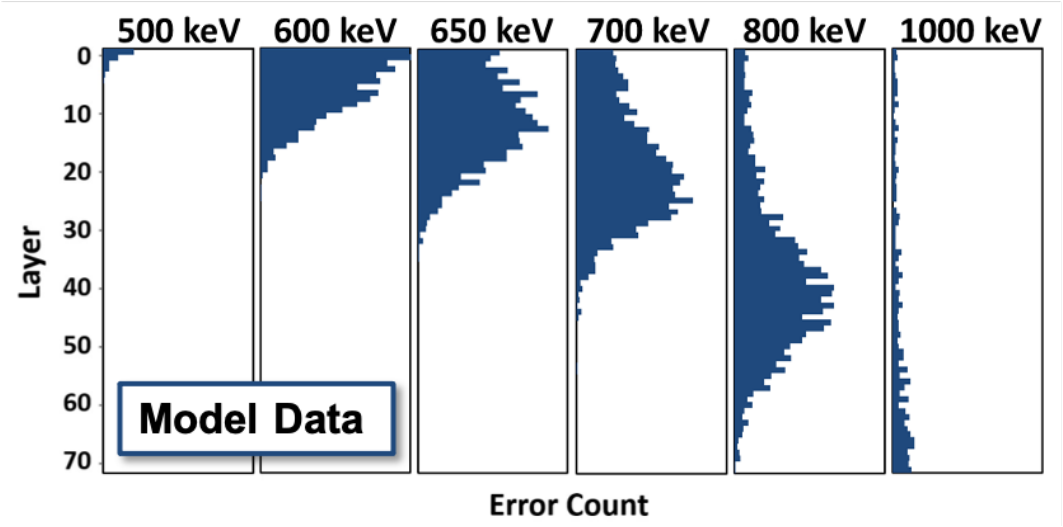


Figure 3.10. Depth profile of simulated proton-induced upsets in 3D NAND memory array using the simplified configuration as shown in Figure 3.9.

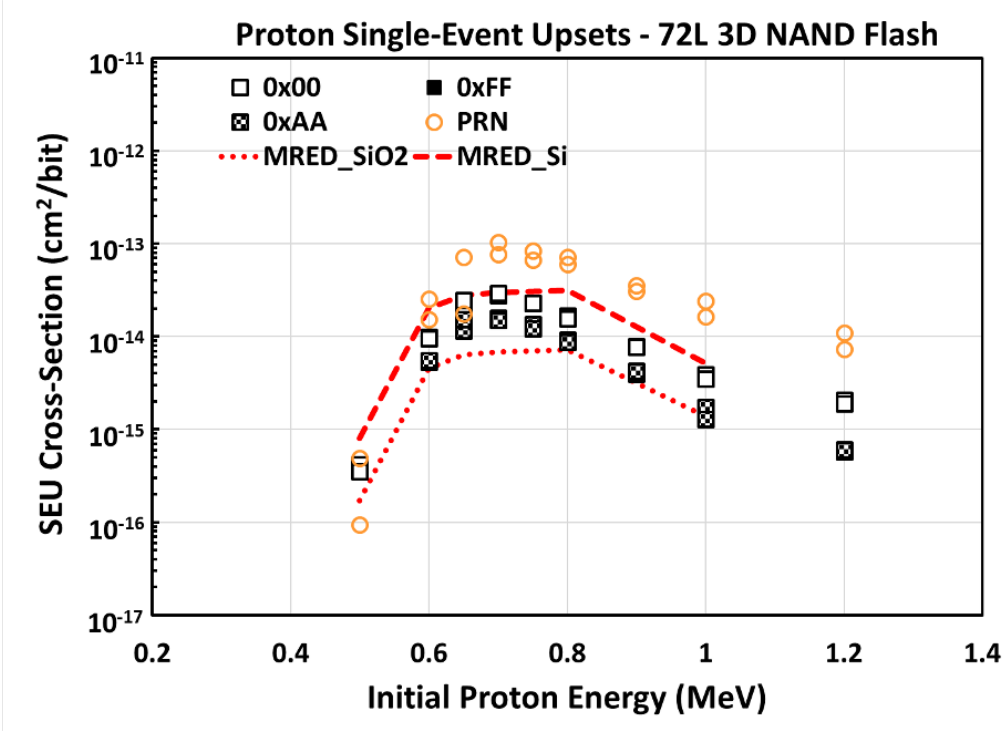


Figure 3.11. Simulation data (in red) shown together with the experimental proton cross section.

Not only are the general shapes in the depth profile consistent with the experimental results (see Fig. 3.6), but the average cross section values over the entire depth of the stack are also on the same order of magnitude as those observed from experimental proton testing. In this model, the transistor area is approximated as pure SiO_2 while the actual transistors have other materials as mentioned above. Using the densities of SiO_2 and pure silicon to bound the range of material densities present in the physical transistor structure, there is strong agreement with the range of cross section values recorded from experiments (Figure 3.6).

V. Resolving differences between experimental data and MRED predictions

To summarize, there are two main discrepancies between experimental and simulated data as presented so far:

1) Seemingly nonphysical jump in event counts between 650 keV and 700 keV proton energies with the corresponding change in the shape of the upset distribution spread as a function of depth in the stack. Whereas upsets in the top layer follow a well-behaved distribution which seems to almost match with the movement of the Bragg peak perfectly, there is significant spreading out at higher energies.
2) The step-like feature roughly halfway through the memory stack that is visible in Figure 3.6 for proton energies at or greater than 700 keV.

The initial simulations performed made several assumptions that might contribute to these differences with the experimental data. These include simplified material configurations, uniformity of the bit line channels, monolithic BEOL configurations, lack of TSV and WL contact materials adjacent to the memory stack, and an approximated general structure with

dimensions based on limited access to detailed SEM images. It is clear that more detailed simulations are required to account for whichever features are actually present in the NAND stack which were not previous captured, or vice versa.

This discussion so far assumes that the discrepancy is the cause of a materials/geometry effect only, and not the result of some experimental, circuit or architectural-level effect. It is worthwhile to explore these alternative possibilities in detail before continuing with more detailed simulations and results.

The first of many explanations is that the proton energies at the test facility were not as tightly tuned as possible, meaning the jump in number of events as a function of depth between 650 keV and 700 keV was actually due to a larger shift in proton energy. This is unlikely however as the shift in the apparent Bragg peak throughout the stack at other energies behaves very much as expected. The facility at NASA-GSFC is routinely used to qualify mission-essential parts and the technical staff confirmed the precision with which the test was performed when asked for verification. Additionally, this would not account for the step-like jump in error counts at halfway through the stack.

It is possible that the errors recorded were "bad bits" in the memory array already, however these were characterized and subtracted out as a filter between each run. If these were responsible, we would expect to see the same bits upsetting over and over again regardless of proton energy. this is not observed.

Another possibility is that there is control or peripheral circuitry which is responsible of the upsets observed at the lower depths/higher energies in question, resulting in a masking effect which makes certain bits “appear” bad in the readout while the bits themselves were not upset (i.e., akin to functional interrupt). This is unlikely however as devices were irradiated in a

powered-off state and there are no known non-volatile memory elements which make up the peripheral circuitry. Effects from upsetting peripherals would also be more likely to result in patterns at the page- or block-level; the experimental errors show none of these suggestive patterns. Circuit level effects would need to be susceptible to upset from direct ionization from protons, which requires relatively low critical charges.

The question then arises of where the control circuitry is located for the NAND array. As a deductive exercise, it turns out not to matter in terms of explanatory power (although the question will be answered). The options are either on top, on the bottom, or, taking into account the step-like jump in events at layer 32, halfway through the stack. If we then assert that control circuitry is upsetting due to the proton irradiation (at this point no matter where such circuitry is located), giving the ultimate appearance of upsets within the memory cells and leading to the apparently nonphysical jump in events at lower depths for 700 keV and higher, then peripherals on the top are safely ruled out as we would expect errors to "appear" throughout the entire stack when such circuits at the top would be subject to proton irradiation at the Bragg peak for the lowest energies tested (i.e., 500 keV). For these lower energies however, the distribution follows a very clean and expected trend with no significant quantity of patterned (or otherwise) errors appearing in the bottom half of the stack.

Similarly, peripheral controls on the bottom of the NAND stack cannot be responsible for the sudden jump in events at 700 keV, considering the range of protons at 600 keV and 650 keV. The distribution of errors tapers off in the 700 keV data set, and assuming that the protons were able to penetrate much further to control circuitry at the bottom of the stack removes the need for such an explanation in the first place, as they would also then have sufficient range to cause the upsets themselves through direct ionization.

The third and final possibility is that the supporting circuitry is somehow halfway through the stack. Focusing on the step feature, we note that it occurs at the same depth of almost exactly halfway through the stack in all energies from 700 keV and up. This suggests a possible architecture-level effect related to the peripheral circuitry. However, because the device was irradiated in a powered-off state, this would require non-volatile elements in some midway layer, for which there is no evidence, nor is there any in the literature for control circuitry being positioned halfway through a memory array in that manner. In fact, TechInsights reported at the Flash Memory Summit in 2019 that Hynix stacked NAND have no additional layers between the top and bottom half of the memory array [38]. Furthermore, published material shows that SK Hynix uses a PUC (peripheral-under-cell) configuration as shown in Figure 3.12 and Figure 3.13. To summarize, control circuitry is ruled out as a likely cause for the upset distributions observed due to the test conditions and the architecture of the device.

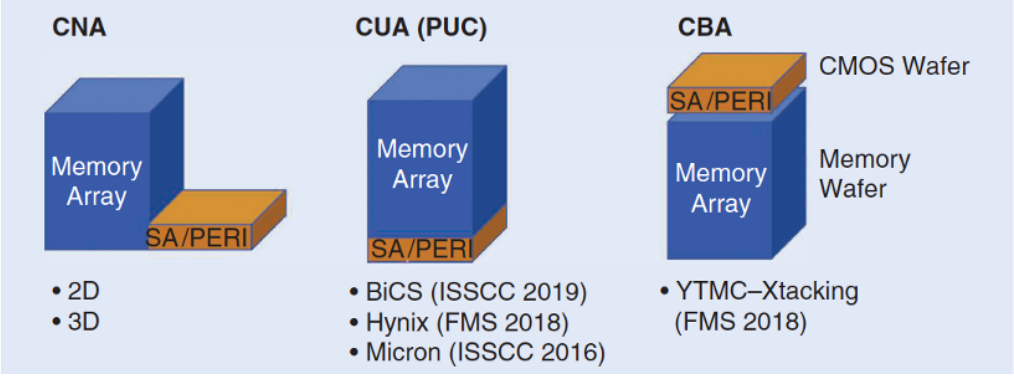


Figure 3.12. Peripheral-under-cell (PUC) architecture is used for Hynix 3D NAND devices [49].

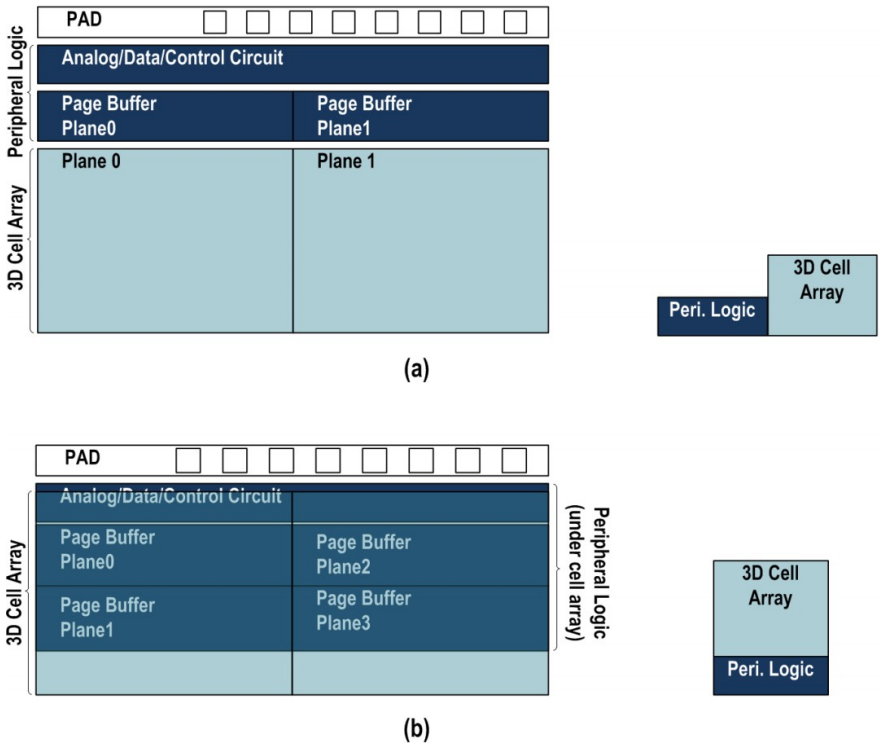


Figure 3.13. Two different peripheral circuitry configurations used in 3D NAND devices [63].

VI. Updated and higher-fidelity simulations of the 3D NAND structure

If the circuit or architecture-level causes mentioned above are effectively ruled out, we should see if the experimental data can be explained through something overlooked in the structure, materials, and geometry of the chip. Figure 3.14 provides a closer and more detailed look at the data; Three separate plots are shown for each energy: 1) the raw events mapped to the logical level in the memory by page number and byte address, 2) the raw events mapped to the physical level in memory by byte address, and 3) the histogram of events over this second physical mapping as previously introduced. The majority of events follow the trends observed in the histogram data, but because the data was checked for bad bits between irradiation, these

errors must be from the protons. This data clearly show that upsets do, in fact, occur throughout the memory array at all depths for all energies, a detail which is easy to miss when focused on the histogram-level data as presented in Figure 3.6.

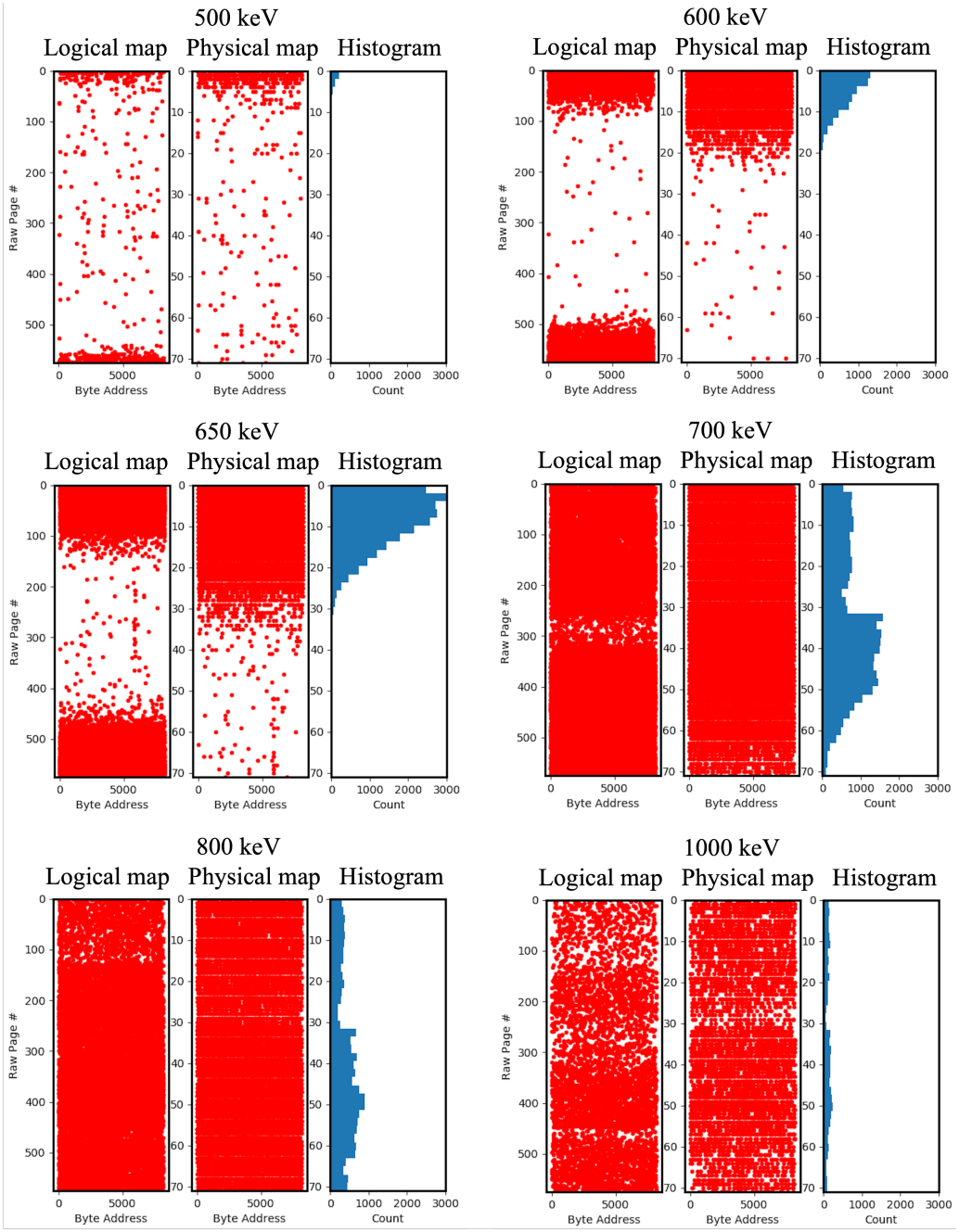


Figure 3.14. Upset distribution as a function of depth for selected energies.

It is possible that some of the variation in the depth profile is due to intrinsic nonuniformities in the cell threshold voltages. Part of the fabrication process for 3D NAND involves very high aspect ratio etches to create the string channels. These etches necessarily taper off as shown in Figure 1.33 at the lower levels. This doesn't account for the events observed deep in the memory stack at low energies however, and while there are certainly other factors which might contribute, the focus here is on explaining how the uniquely 3D structures involved in 3D NAND modulate the radiation response which must be considered for test design and part analysis and qualification.

Several assumptions went in to the original MRED model as presented earlier in this chapter. To account for the 3D structure contributions to the observed depth profile for upsets, a much higher fidelity structure was created in MRED to address these assumptions. The updates to the MRED structure included:

- Gate-all-around transistor geometries (as opposed to block detectors)
- Detailed BEOL from BL contacts to M2 (not monolithic)
- Vertical air-gap voids included between WLs as in the physical devices
- Full range of materials used rather than the silicon/SiO_2 approximation
- Proper p-BiCS configuration with respect to geometry
- Contact fanout and TSVs included adjacent to the structure
- 10 nm offset between top and bottom half of stack accounted for

Figures 3.15 – 3.18 visually show how these various features were included in the MRED simulation structure. Each vertical transistor layer contained 680 sensitive detector collections for a total of over 55,000 cells in the entire NAND stack.

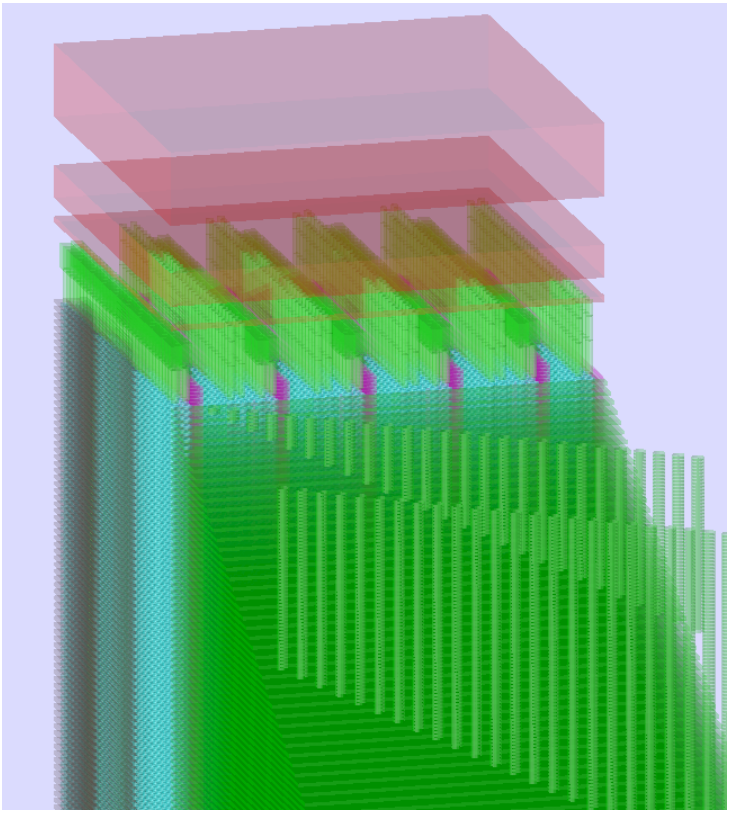


Figure 3.15. Render of the advanced 3D NAND simulation structure in MRED, including the contact fanouts and TSVs.

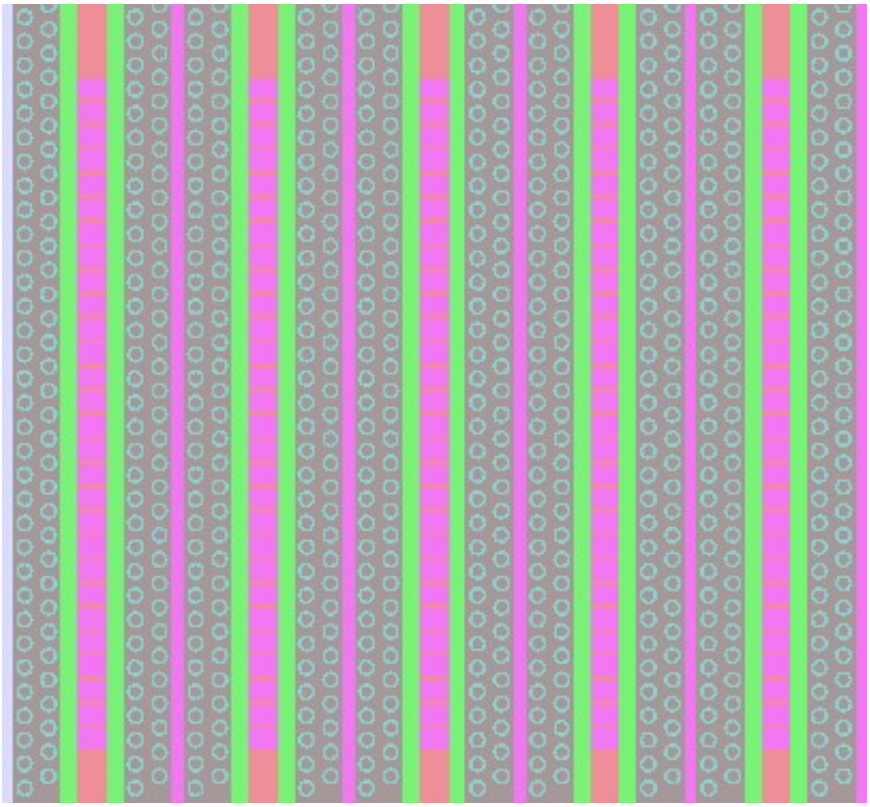


Figure 3.16. Render of the 3D NAND cell geometry viewed in the bit line direction. The WL material is shown in green (tungsten and silicon used), with the pink showing N-filled air gaps. The circle regions correspond to the gate-all-around transistors which make up the cells of the memory array.

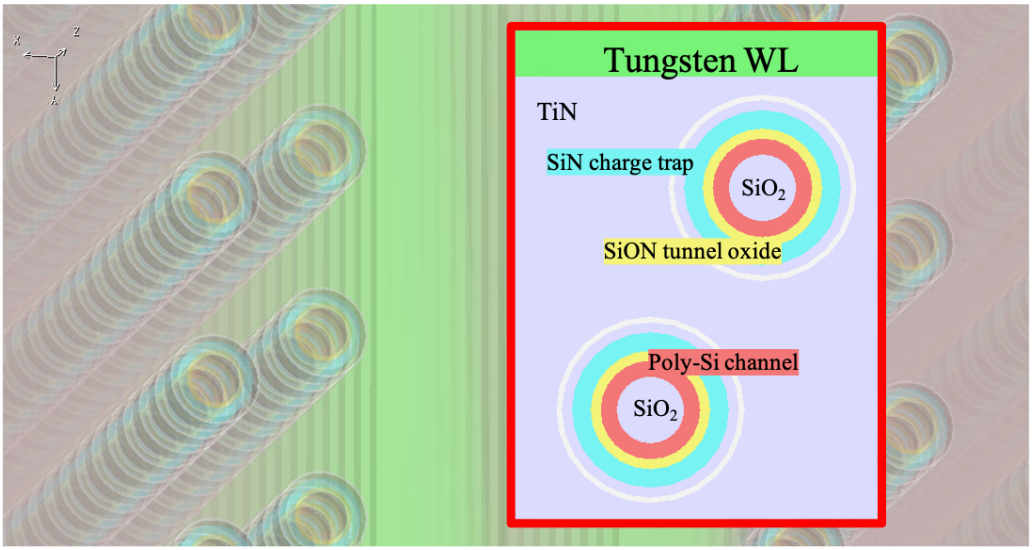


Figure 3.17. MRED 3D NAND model, with focus on the transistor structure.

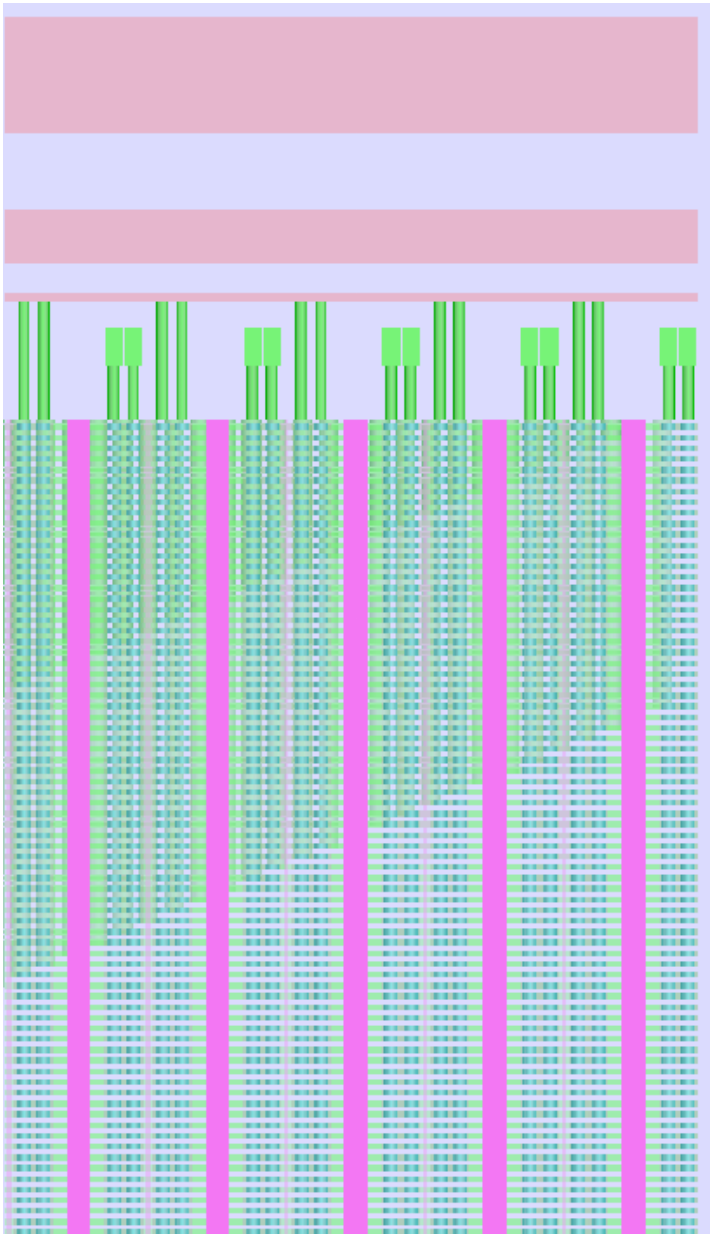


Figure 3.18. Detailed MRED simulation structure as viewed in the WL direction. The BEOL is structure based on the p-BiCS configuration, with monolithic copper placed at level M3 and above in the simulation.

Figure 3.19 shows the simulated physical layer-level SEU cross sections for up to four regions of charge deposition. Unsurprisingly, there were no major shifts in the trends observed due to the materials selection in the four different region, and other work in the literature [52]

suggest that only the charge trap layer and the tunnel oxide contribute significantly to the manifestations of radiation effects in these kinds of devices.

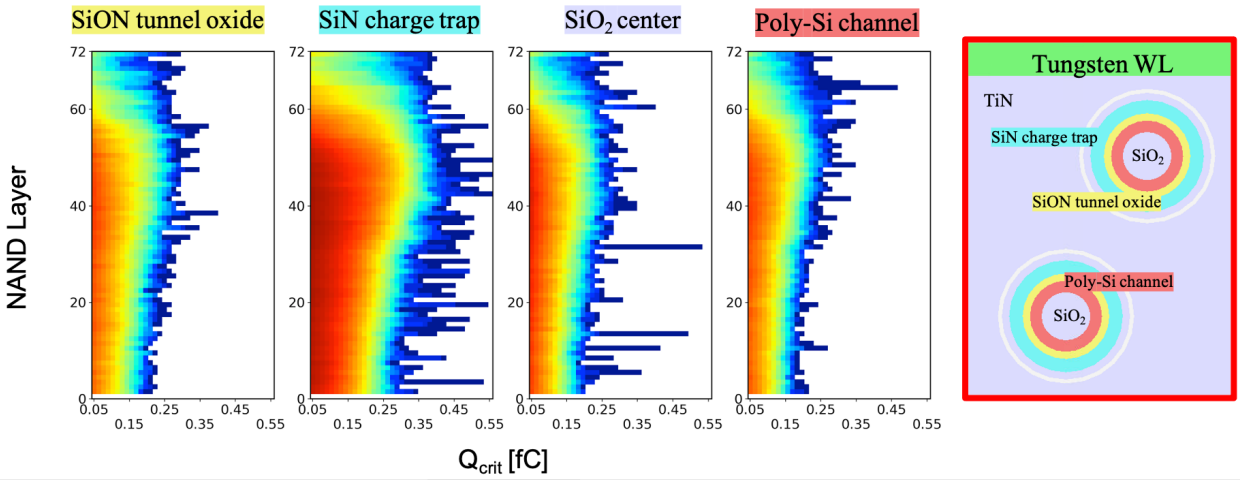


Figure 3.19. Simulated SEU probability distribution per layer for 800 MeV protons.

The updated depth distribution of proton errors using the higher fidelity simulation deck is given in Figure 3.20. While the results still do not align perfectly with the experimental data, there is a marked improvement in terms of general agreement. The distributions shown are more spread out and show that the inclusion of voids allows for upsets deep within the memory stack even at low proton energies (although the probability of these errors is expectedly small). Additionally, there is a slight appearance of the step-like feature at layer 32 due to the misalignment account for in the update MRED deck, as well as the two dummy layers in the middle of the overall stack. While present, this is a small feature in the general trend and suggests that, while the physical features discussed above are at least partially responsible, functional mechanisms likely play a key role as well. Figure 3.21 presents the SEU cross section

as a function of incident proton energy for normal incidence protons. The bands in the plot correspond to Q_{crit} values representative of SLC operating mode, and the cross sections are given in the case of both sensitive detector configurations. The SEU cross sections were calculated over all WL levels, regardless of whether events occurred at that level and energy or not.

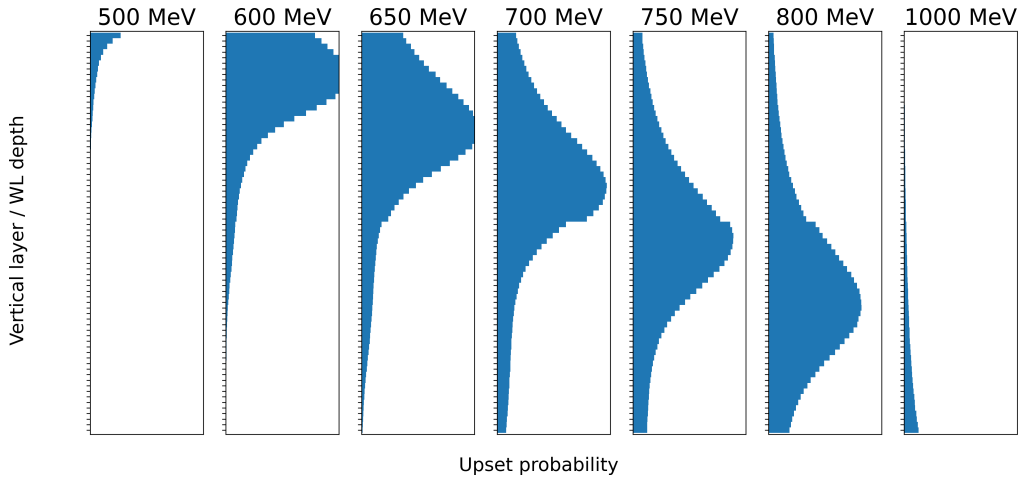


Figure 3.20. Depth distribution of proton-induced upsets from the high fidelity MRED model.

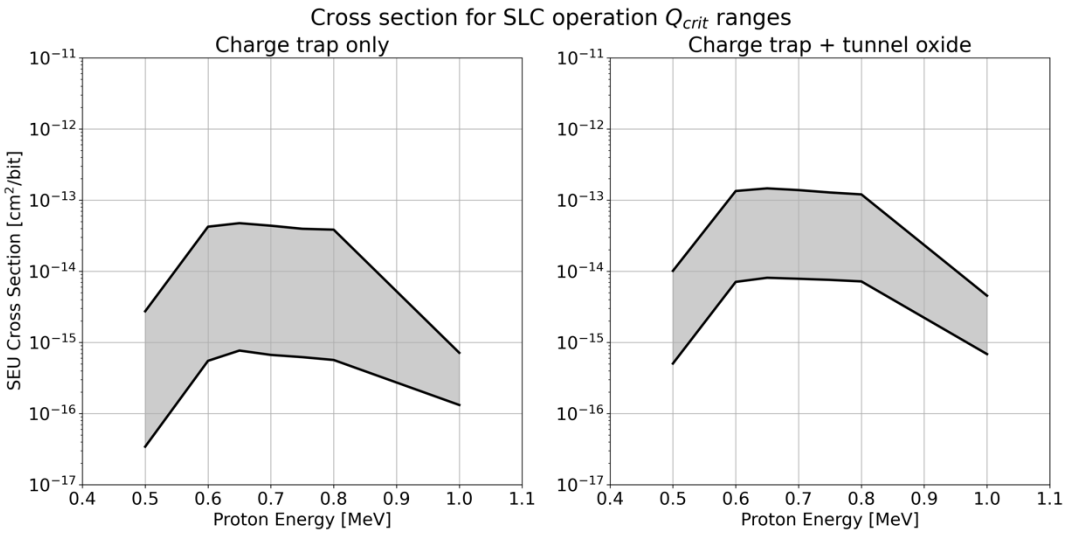


Figure 3.21. Single event upset cross section range based on range of Q_{crit} values corresponding to SLC operation as a function of initial proton energy.

The presence of voids in the vertical dimension through the high aspect ratio slits between WLs suggest an angular study as well. Figures 3.22 to 3.24 show the depth distribution of upsets on the same scale as the previous plot (at normal incidence) for incidences of 15°, 45 °,

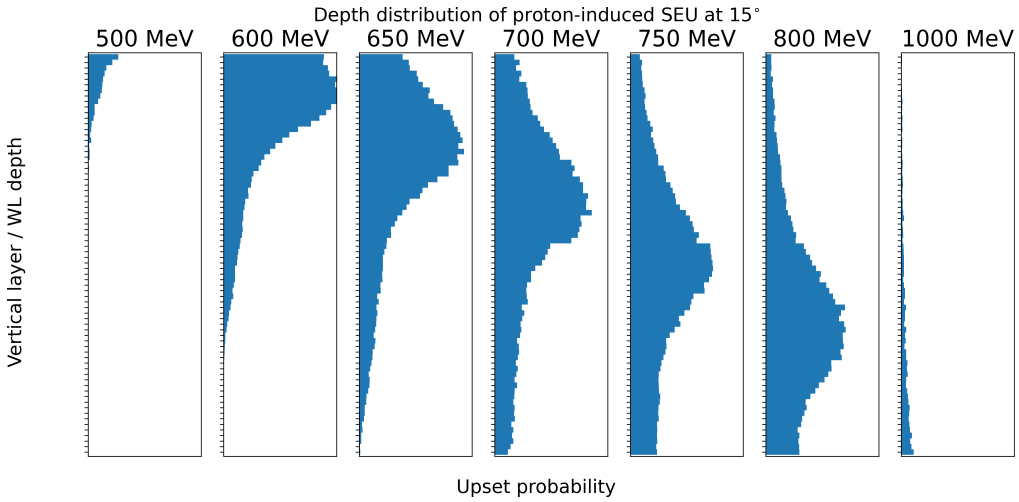


Figure 3.22. Depth distribution of proton-induced upsets from the high fidelity MRED model for proton incidence of 15◦

and 70°. As can be seen, for small angles there is not much change at all in the likelihood of a proton reaching the lower levels of the device. At 45° the lower energy protons begin to be filtered out.

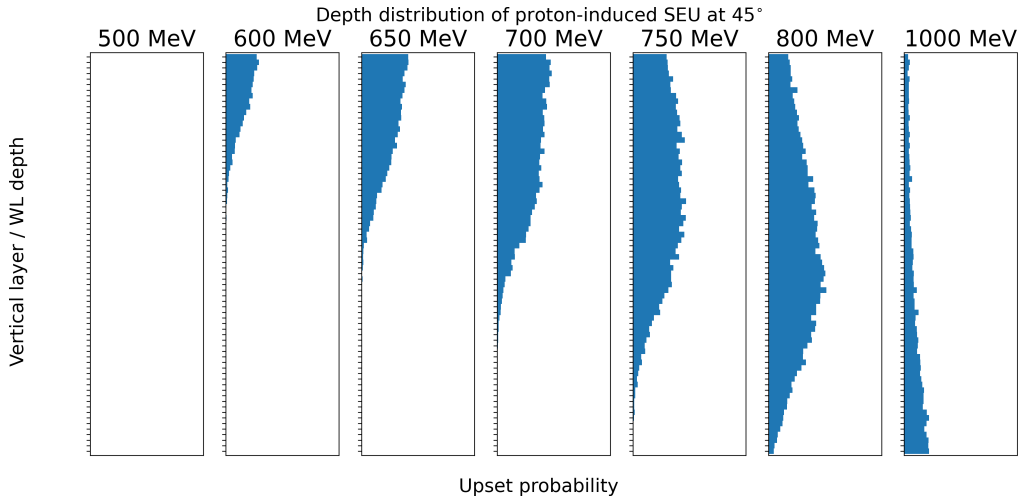


Figure 3.23. Depth distribution of proton-induced upsets from the high fidelity MRED model for proton incidence of 45◦.

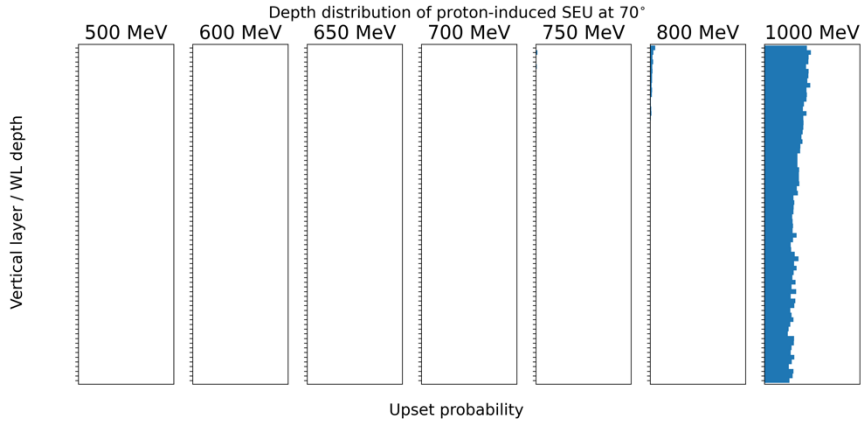


Figure 3.24. Depth distribution of proton-induced upsets from the high fidelity MRED model for proton incidence of 70◦.

Considering the intrinsic ΔV_{th} variability in 3D NAND, we can then plot the total SEU cross section as a function both incident proton energy, proton angle, and cell operating mode. As expected, there is a significant increase in event rate for MLC and TLC operation. The fact that the protons at higher angles of incidence are less likely to cause upsets throughout the whole stack at low incident energies also means that the cross section here is significantly reduced regardless of operation mode. Simulations were performed at normal incidence, and 15, 45, and 70 degrees using the advanced simulation structure which included the air-gap spacings. Figures 3.25 to 3.27 present the SLC, MLC, and TLC distributions of SEU cross section as a function of initial proton energy for all four of these angular configurations. The bounded ranges capture the intrinsic cell variability. There is very little difference in all cases between normal incidence and 15 degrees, while the steeper incident angle mean that fewer to none of the protons are able to penetrate far enough into the chip to cause an upset. Additionally, there is very little difference in terms of the final total cross sections between using just the CT region as sensitive in the simulation vs. using both the CT and tunnel oxide regions.

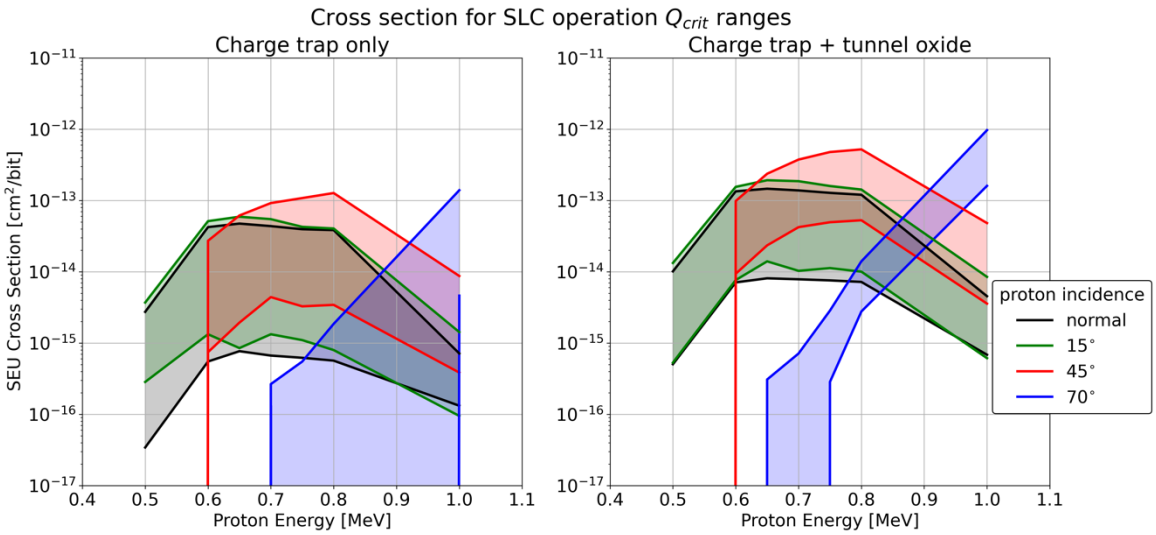


Figure 3.25. Single event upset cross section range based on range of Qcrit values corresponding to SLC operation as a function of initial proton energy for a range of incidence angles.

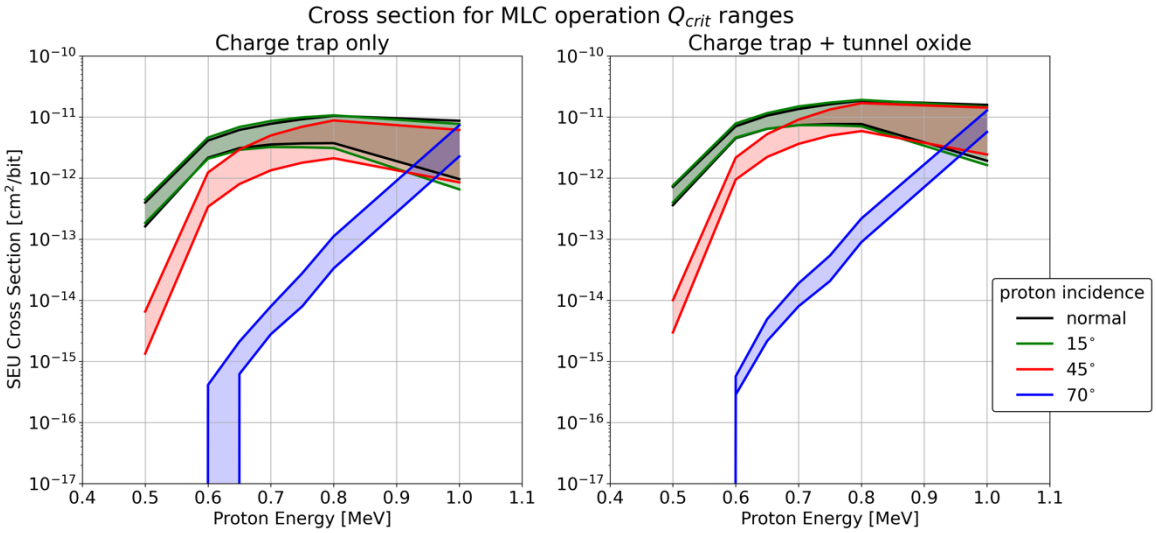


Figure 3.26. Single event upset cross section range based on range of Qcrit values corresponding to MLC operation as a function of initial proton energy for a range of incidence angles.

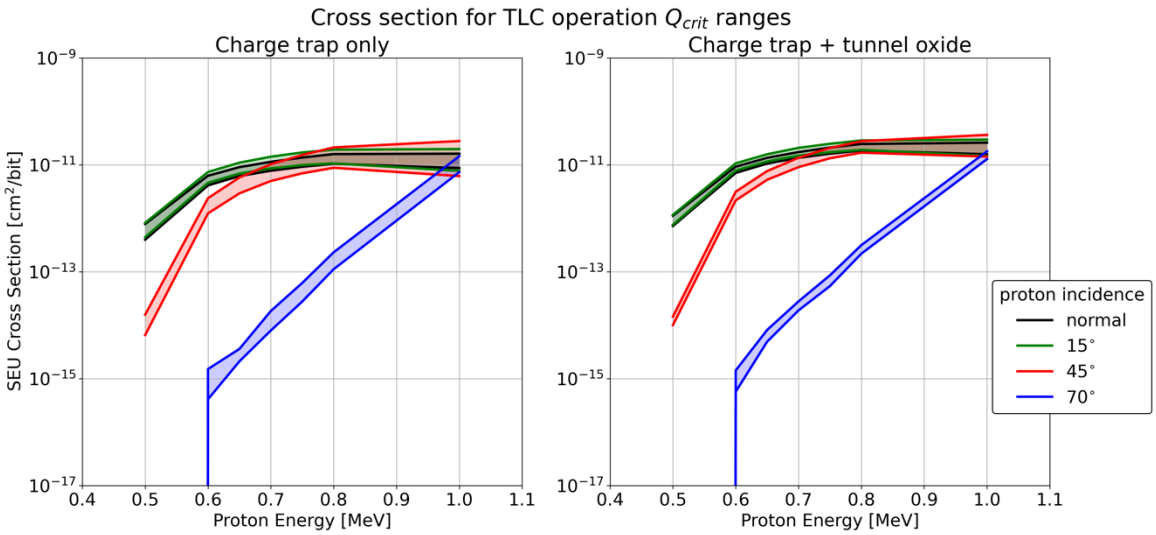


Figure 3.27. Single event upset cross section range based on range of Qcrit values corresponding to TLC operation as a function of initial proton energy for a range of incidence angles.

A final set of simulations was performed to characterize the effects of the air-gap voids between WLs and to determine event rates for the 72-layer Hynix 3D NAND in a trapped proton environment across all 8 die in the package. Figure 3.28 gives the spectrum used in these simulations, which accounted for 100 mils of Al shielding at an altitude of 20,200 km with an inclination of 55 degrees (GPS orbit).

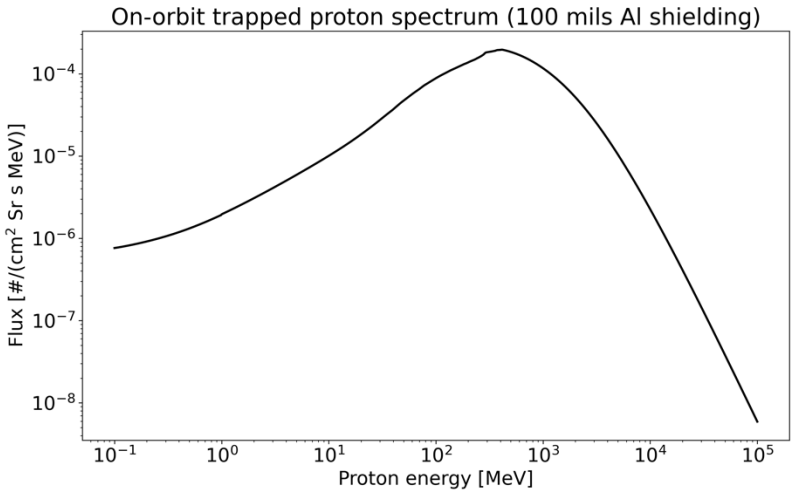


Figure 3.28. Trapped proton spectrum used for event-rate simulations.

Simulations were performed using both an isotropic flux and a directional (normal incidence) flux with the proton energy distribution in Figure 3.28 sampled on a log-basis. While the normal incidence flux doesn't correspond to any real space environment, it is useful in visualizing the cumulative effect of multiple void regions over all 8 die. For critical charges that correspond to the SLC operational mode (approximately 0.3 fC and greater), there is an order of magnitude difference in the event rate that is not accounted for using monolithic simulation structures. A similar increase in event rate is observed in the case of an isotropic flux, although the magnitude of the difference between simulation configurations is much smaller. Because the air-gap is a high aspect ratio region, the probability of the same kind of effect as seen in normal incidence where the effective average range of a proton in the package is extended is less likely to occur in an isotropic environment. For other angles, the air-gap regions mean that there is less material available to stop protons in the memory array, making energy deposition more likely to occur at or near a transistor rather than in the very low-density region of the air-gap.

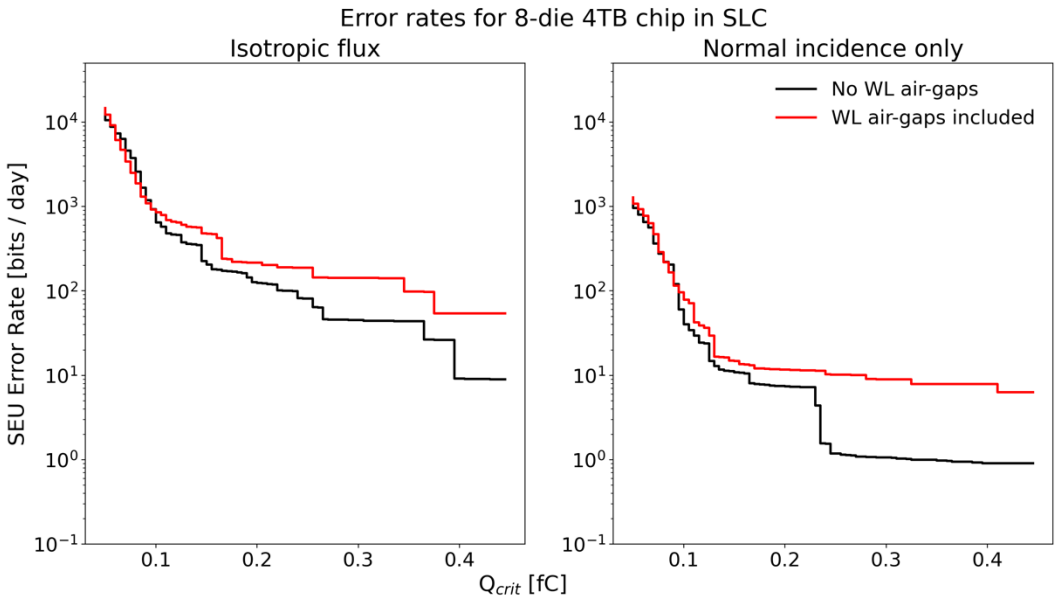


Figure 3.29. SEU error rate for the 512 Gb/die / 8 die 72-layer Hynix 3D NAND as a function of critical charge.

Chapter 4

Heavy-ion-induced upsets in 3D NAND

This chapter introduces single-event effects in heavy-ion environments for the same 72-layer 3D NAND device as the previous chapter. Experimental data from the same GSFC collaborators is included, with validated simulation results using the MRED model which was also previously introduced.

The MRED model is extended to study intra-die upsets and multi-die effects for packaged memories with multiple vertically integrated NAND die in a single chip. Simulations show that there is roughly an order of magnitude increase in the heavy-ion-induced SEU cross section across the range of cell operating modes (SLC, MLC, and TLC), that test conditions which do properly account for a multi-die packaging of the NAND chip can significantly underestimate the actual SEU cross section in high energy environments such as space, and that FG 3D NAND in general exhibits better performance (i.e., lower cross section) when it comes to SEU than CT 3D NAND due to the lack of metallization within the vertical stack near the sensitive regions of the transistors. Little significant intra-die variation in the SEU profile is observed, except in the case of particles near the Bragg peak close to the end of range.

I. Experimental details for SEU in 72-layer CT 3D NAND[3]

Heavy-ion testing was conducted at the Lawrence Berkeley National Laboratory's (LBNL) 88" Cyclotron facility. Most testing was performed at normal incidence, but certain angular irradiations with multiple axes of rotation were performed to evaluate three-dimensional angular dependencies. The test setup at the facility is pictured in Figure 3.2, and the specific ion beams used are listed in the top section of Table 4.1, with surface linear energy transfer (LET) and ranges independently computed with SRIM 2013.

Heavy-ion testing for single-event upsets (SEU) was performed with static, unpowered irradiations to remove any secondary effects caused by control circuitry or page buffer upsets and limit data errors to true bit cell upsets in the memory array. Each run was typically performed to a fluence of $1x10^6/cm^2$ with a flux between $2x10^4/cm^2/s$ and $9x10^4/cm^2/s$. Sixteen blocks each of four data patterns (all zeros, all ones, checkerboard, and pseudorandom), totaling about 85 MB per pattern, were programmed and read back prior to each irradiation. Errors existing prior to the run (common with any advanced flash memory) were masked for analysis. Testing was performed in vacuum at room temperature, and the heavy ion SEU cross-section is shown in log-log format in Fig. 4.1. The trend of SEUs plotted in Fig. 4.1 shows a roughly powerlaw relationship between SEU cross-section and LET down to the lowest LET tested (1.2 MeV cm^2/mg). No error-free LET threshold could be identified with the heavy ion beams available, which was confirmed with the observation of direct ionization from protons as discussed in Chapter 3.

A data pattern dependence is observed in the heavy ion data, with fully erased cells ("all ones" pattern) essentially impossible to upset, "all zero" pattern most susceptible, and mixed data

[3] Partially reprinted with permission from E. P. Wilcox et al., "Observation of Low-Energy Proton Direct Ionization in a 72-Layer 3-D NAND Flash Memory," *IEEE Trans. Nucl. Sci*, vol. 68, no. 5, pp. 835-841, May 2021.

patterns somewhere in between. Such a relationship suggests direct mapping of logical bit value to physical cell program state, and has been shown in previous NAND flash memories, including from the same manufacturer in [39].

Heavy ion	Energy [GeV]	MeV/u	LET [MeV-cm^2/mg]	Range in Si [mm]
^{14}N	0.224	16	1.16	0.470
^{16}O	0.256	16	1.5	0.421
^{40}Ar	0.640	16	7.3	0.254
^{63}Cu	1.008	16	16.5	0.190
^{84}Kr	1.344	16	25.0	0.182
^{28}Si	1.344	48	2.1	1.68
^{56}Fe	3.528	63	5.6	1.69
^{84}Kr	5.040	60	10.5	1.28
^{132}Xe	11.352	86	18.7	1.71
^{48}Ti	28.800	600	0.99	83.37
^{84}Kr	33.600	400	3.2	28.81
^{132}Xe	46.200	350	7.7	16.45
^{181}Ta	61.902	342	13.6	12.76
^{197}Au	19.700	100	32.0	1.79

Table 4.1. Heavy ion selection for simulations. The top corresponds to the test ions used as presented above, while the middle and bottom ions are for ions of similar LETs but much higher energies and ranges.

Experimental and simulation results for the SEU cross section are given in Figure 4.1. The relationship between logical address and physical memory layer was experimentally determined using heavy ion bit upset tracks as a reference, as shown in Fig. 4.2 and previously demonstrated in [54]. Each NAND string within the array is 144 bits long, with a U-shaped tunnel at the base of the structure. The heavy ion SEU data are processed first by grouping the word lines in sets of 144 and then reversing lines 72-143 such that line 0 and 143 both represent the top-most physical layer and lines 71 and 72 are the bottom-most. As an example, 1225 MeV Kr ions created the vertical stripes in Fig. 4.2 when normally-incident ions upset a series of cells in the same address but across multiple layers.

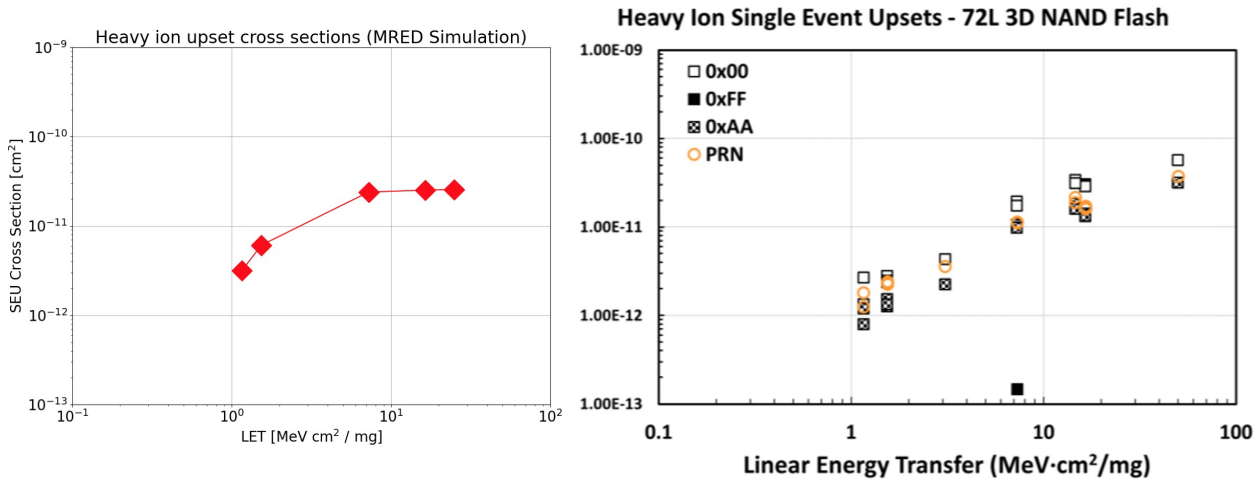


Figure 4.1. Simulation results (left) using the simplified 3D NAND model and experimental data (right) for SEU cross section as a function of LET

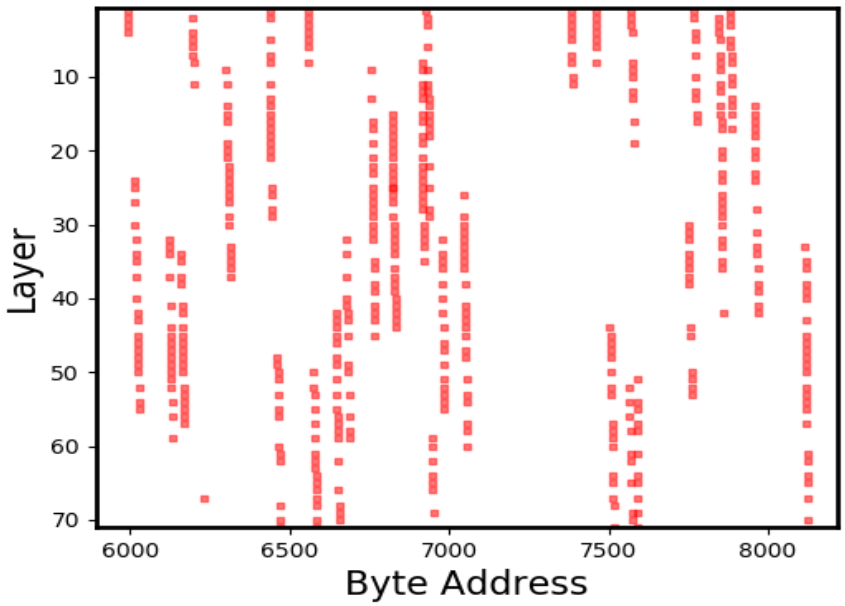


Figure 4.2. Heavy ion tracks in the 3D NAND memory stack.

II. Characterizing heavy-ion SEU in multi-die 3D NAND packages

Table 4.1 contains the ion species, energies, and LETs used in the MRED simulations of heavy-ion SEU in 3D NAND. The three sections correspond to low, medium, and high kinetic energy ions. These energies were selected from available tables for the Texas A&M test facility, the Lawrence Berkeley ion beam lab, and the NASA Space Radiation Laboratory (NSRL). Mass and energy ranges for the ions were chosen such at 16 MeV/u to match experimental data as presented above, while the higher kinetic energy ions were chosen to have roughly the same LET values but much larger range than in the case of 16 MeV/u.

Mechanisms for relative charge generation within the cells are also considered, with the transistors positioned more closely to the tungsten word-lines exhibiting a larger cross section than those further away. This trend holds for both low and high energy ions throughout the entire wire-bonded stack of die.

Figures 4.3 and 4.4 show a cartoon representation and actual image of multi-die stacking in 3D NAND. As mentioned previously, the 72-layer Hynix device is packaged with 8 512 GB TLC dice for a total of 4 Tb. Each die in the simulation has a group of 82 vertical layers of transistors with 680 transistors per layer. For die positioned lower in the stack, the structure is repeated with the only difference being that the die above do not have sensitive detectors attached and thus do not track energy deposition.

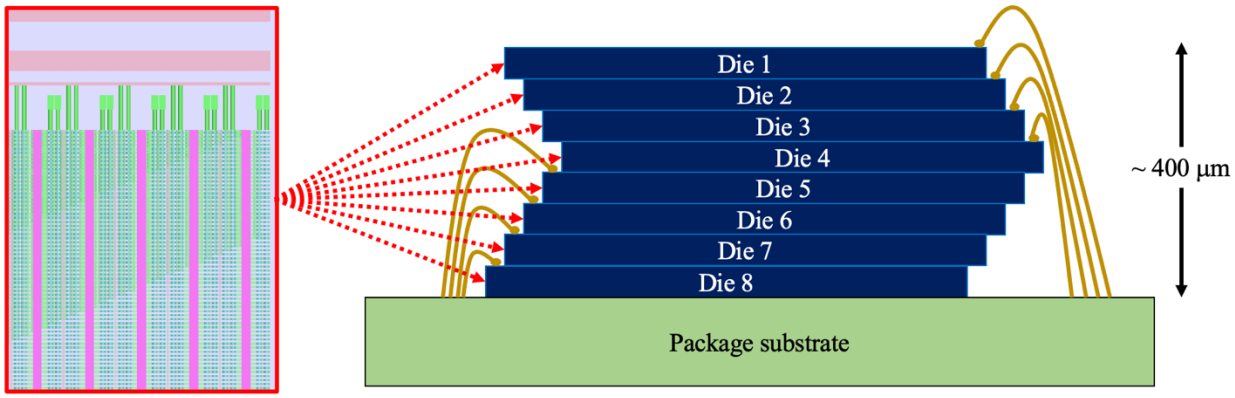


Figure 4.3. Multi-die simulation structure.

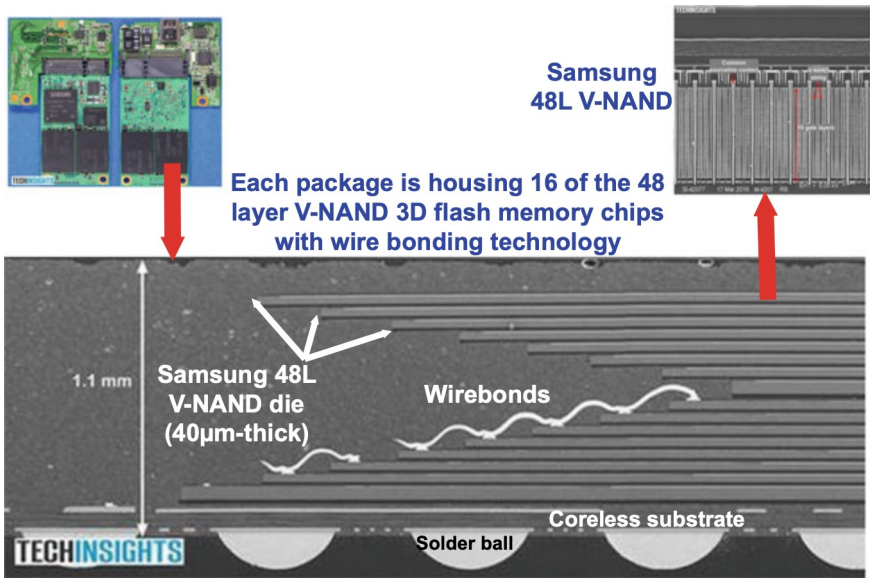


Figure 4.4. Example of multi-die 3D NAND package [64]. Each die for stacks up to and in some cases over 1 mm contain individual stacks of 3D NAND memory.

Experiment and simulation of 3D NANDs have, up until now, only considered the top die in a chip, which, in reality, contains 8 wire-bonded die stacked vertically as shown in Figure 4.4 and 4.5. Other studies in 3D NAND devices have also only focused on the top layer. Extending the MRED simulation model in this manner gives valuable insight into how these chips might responds as a whole in, for example, low vs. high kinetic energy heavy ion environments. The question of testing and qualification is of particular importance here, as heavy ions over the same range of LET values but different kinetic energies result in different cross section profiles both within an individual stack and more drastically across multiple die vertically wire-bonded together.

The heavy-ion data presented in the first section of this chapter characterized the single-event upset cross section as a function of LET using ions at an energy of approximately 16 MeV/u. Because 3D NAND chips are packaged vertically with multiple die in an overall

package spanning upwards of 500 μm in some cases, it becomes important from a parts qualification standpoint to consider whether the upset cross section and event rate for a chip as determined for the top die only accurately describes the overall performance of the entire chip.

The natural space radiation environment consists of particles which exceed common test energies by a wide margin. This section will describe a detailed set of simulations done using the most advanced MRED model of the 3D NAND stack based on a charge-trap configurations to characterize the upset cross section over common LET values over an entire packaged device with up to 8 NAND memory arrays per chip.

Figures 4.5 to 4.7 show MRED predictions of the change in SEU cross section per die per layer with the low-energy (16 MeV/u) ions on top and the high energy ions on bottom. Plotted on the y-axis is the SEU cross section and each plot from left to right represents a 3D NAND die. Within the plot for each die is shown the SEU upset cross sections as a function of physical vertical layer within the die. Ions are shown with a range of LETs of approximately 1 to 30. Within each die plot, the left side corresponds to the top of the die and the right to the bottom.

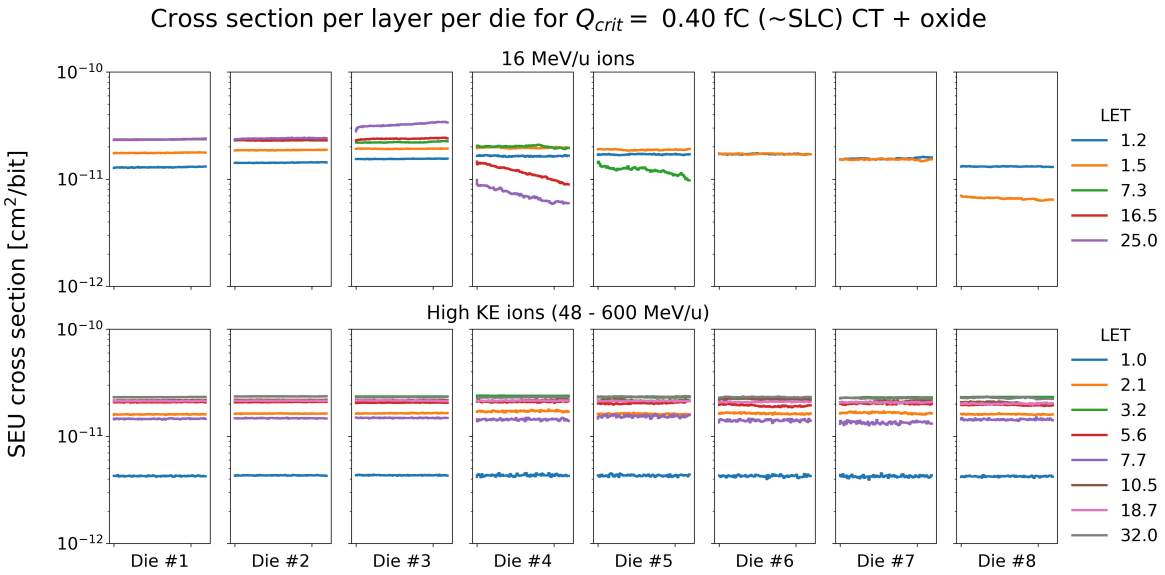


Figure 4.5. SEU cross section, as predicted by MRED, as a function of die in a multi-die package and depth within each die for 3D NAND operating in SLC mode.

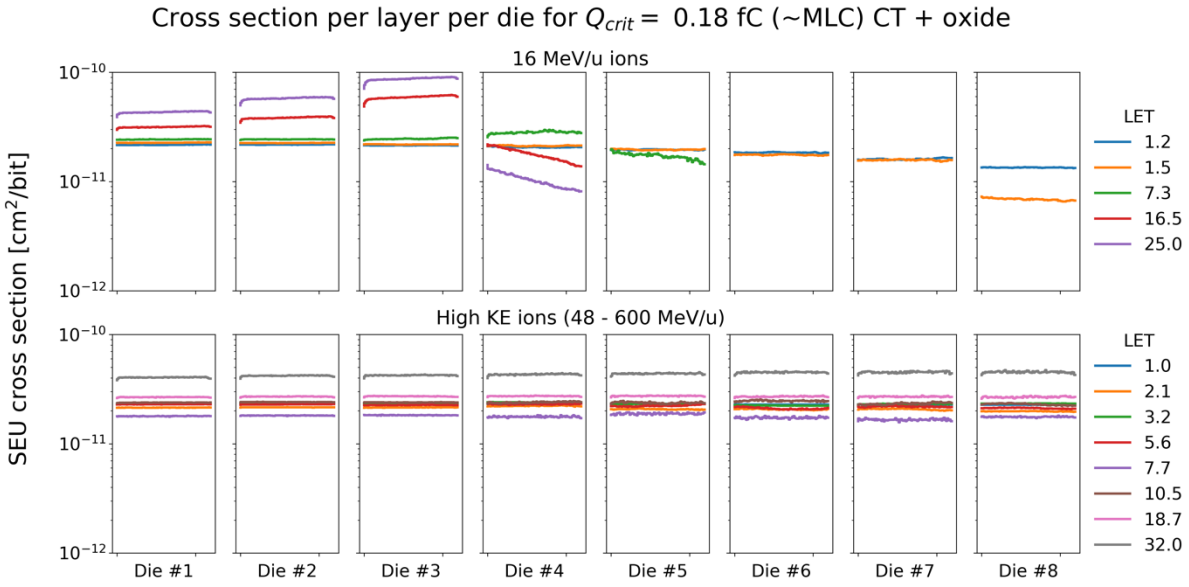


Figure 4.6. SEU cross section, as predicted by MRED, as a function of die in a multi-die package and depth within each die for 3D NAND operating in MLC mode.

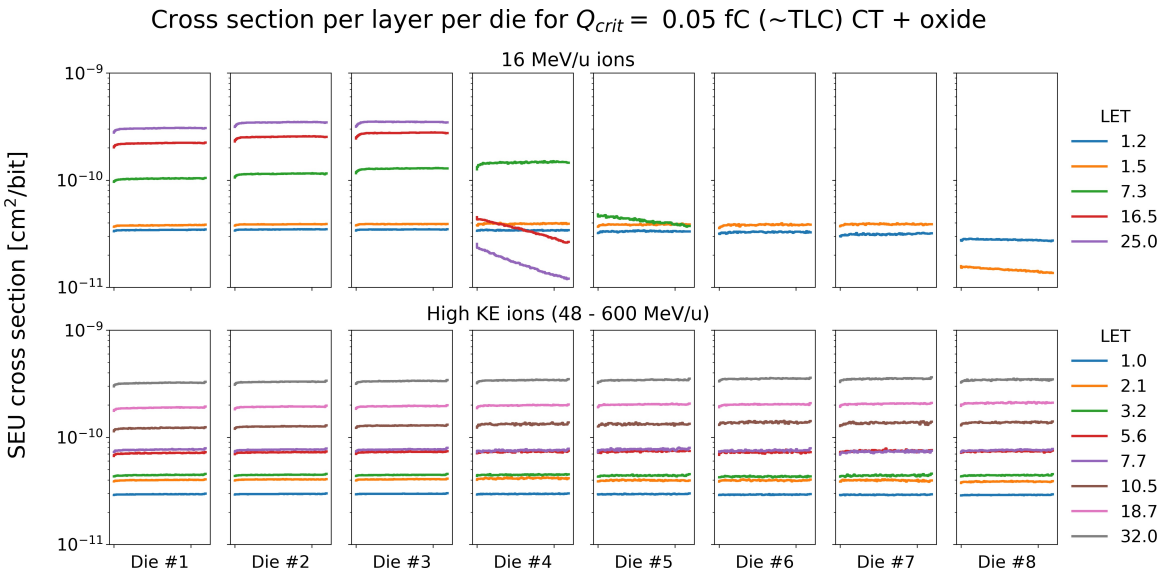


Figure 4.7. SEU cross section, as predicted by MRED, as a function of die in a multi-die package and depth within each die for 3D NAND operating in TLC mode.

From the mostly flat lines within these die plots, there does not appear to be any significant intra-die layer dependence on the heavy-ion response of the devices whether at high or low kinetic energies and regardless of the operating mode of the cell. The more steeply sloping distributions shown in the fourth and fifth die for the 16 MeV/u ions is due to larger fluctuations in LET as ions near the end of their range in the package. In addition to this ranging out of higher LET ions at 16 MeV/u before reaching the bottom chips, there is substantial inter-die variation between the top die, where ions are closer to the Bragg peak at two or three die into the stack than they were at the top level, resulting in SEU cross sections within those die to increase. Little effect is seen in the case of SLC operation, but over an order of magnitude in the case of TLC cells. Related to this result, we notice that the spread of SEU cross section values is much wider in the case of large kinetic energy ions for cells programmed in TLC mode.

To gain more insight into these intra- and inter-die variations, the Figures 4.8 to 4.13 show the total upset cross sections, as predicted by MRED, as a function of LET for all three operational modes (SLC, MLC, TLC), first for sensitive detectors corresponding just to the charge trap region of the transistor (Figs. 4.8 to 4.10), and then with energy deposition collected for both the charge trap and tunnel oxide regions included (Figs. 4.11 to 4.13). Previous results in the literature have shown that charge deposition in both the tunnel oxide and the floating gate of NAND memory cells can contribute to the overall ΔV_{th} [52], and while the specifics will certainly vary between manufacturers (and potentially within manufacturers), these simulations provide reasonable bounds for CT NAND with similar transistor sizes and configurations.

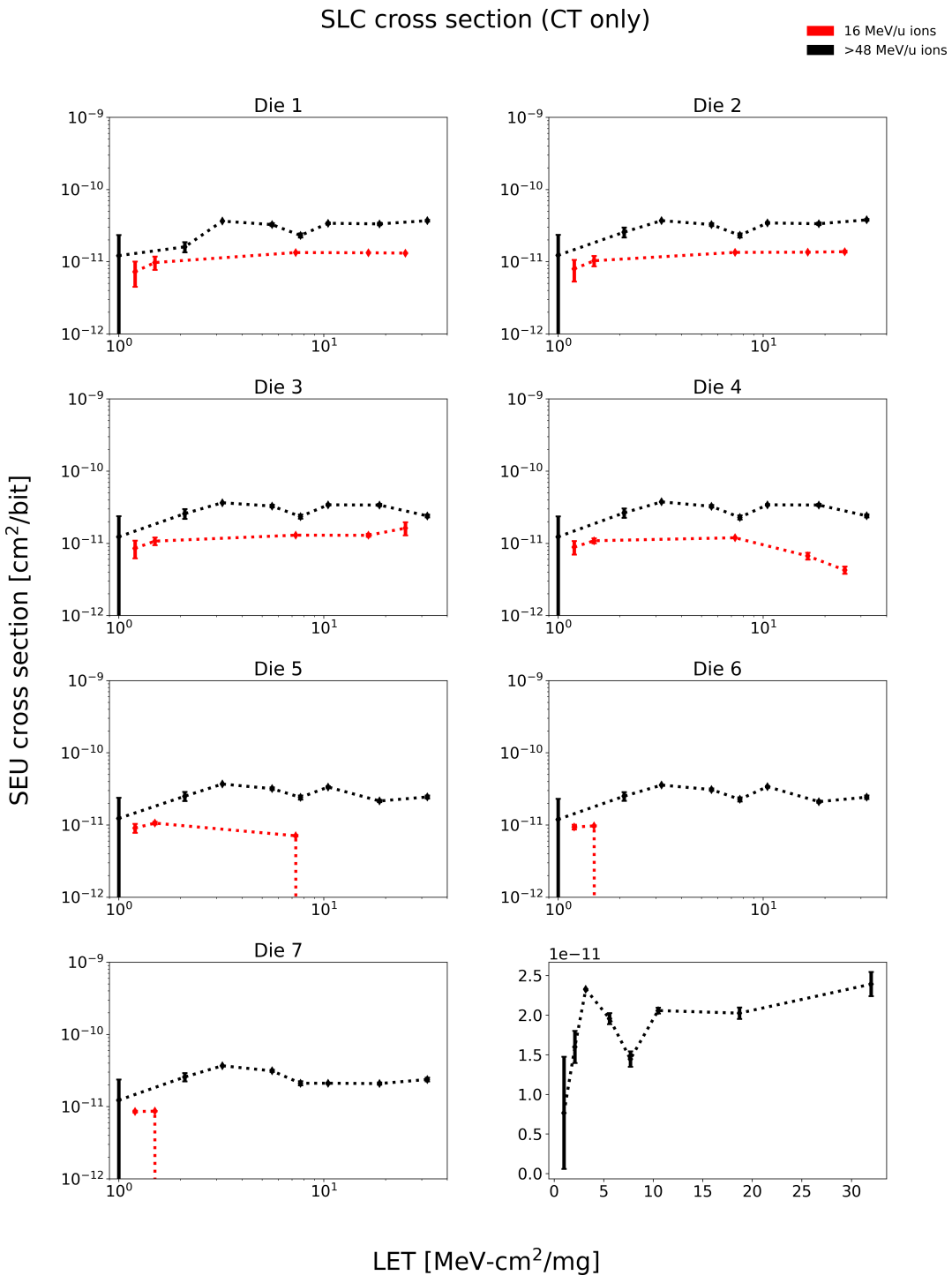


Figure 4.8. Cumulative cross sections, as predicted by MRED, over all NAND arrays in 8-die package for cells operated in SLC mode (CT region sensitive)

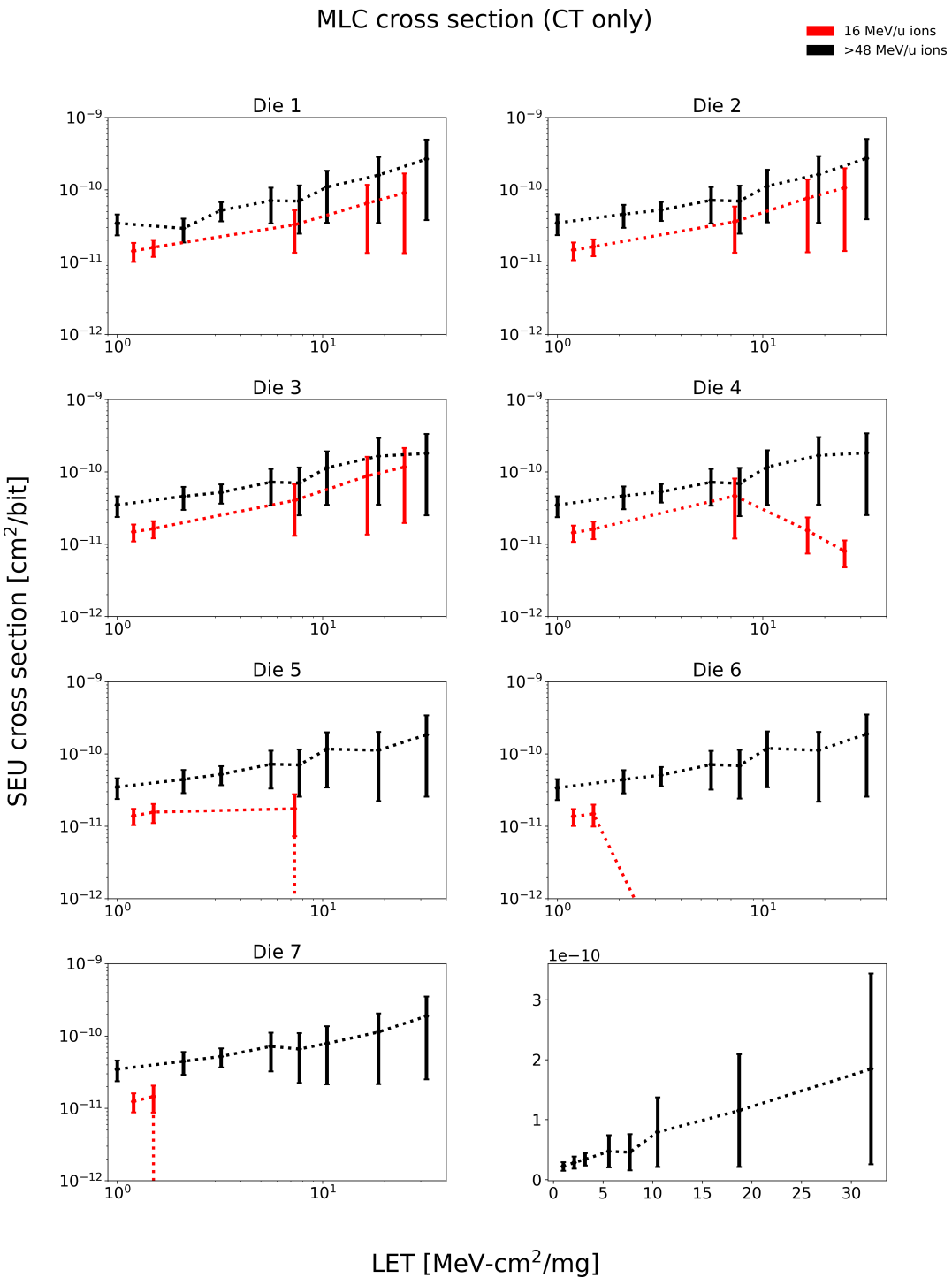


Figure 4.9. Cumulative cross sections, as predicted by MRED, over all NAND arrays in 8-die package for cells operated in MLC mode (CT region sensitive)

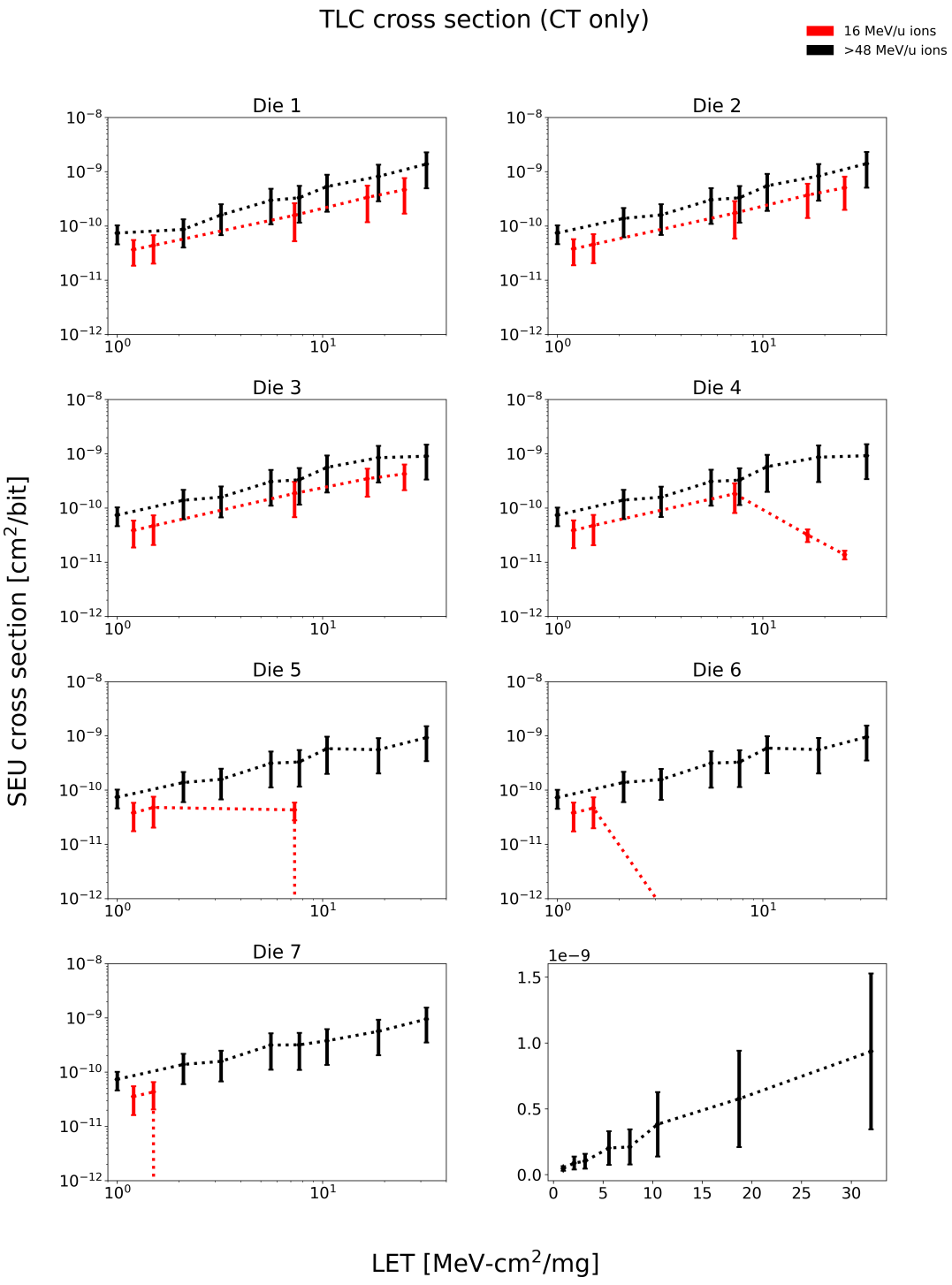


Figure 4.10. Cumulative cross sections, as predicted by MRED, over all NAND arrays in 8-die package for cells operated in TLC mode (CT region sensitive).

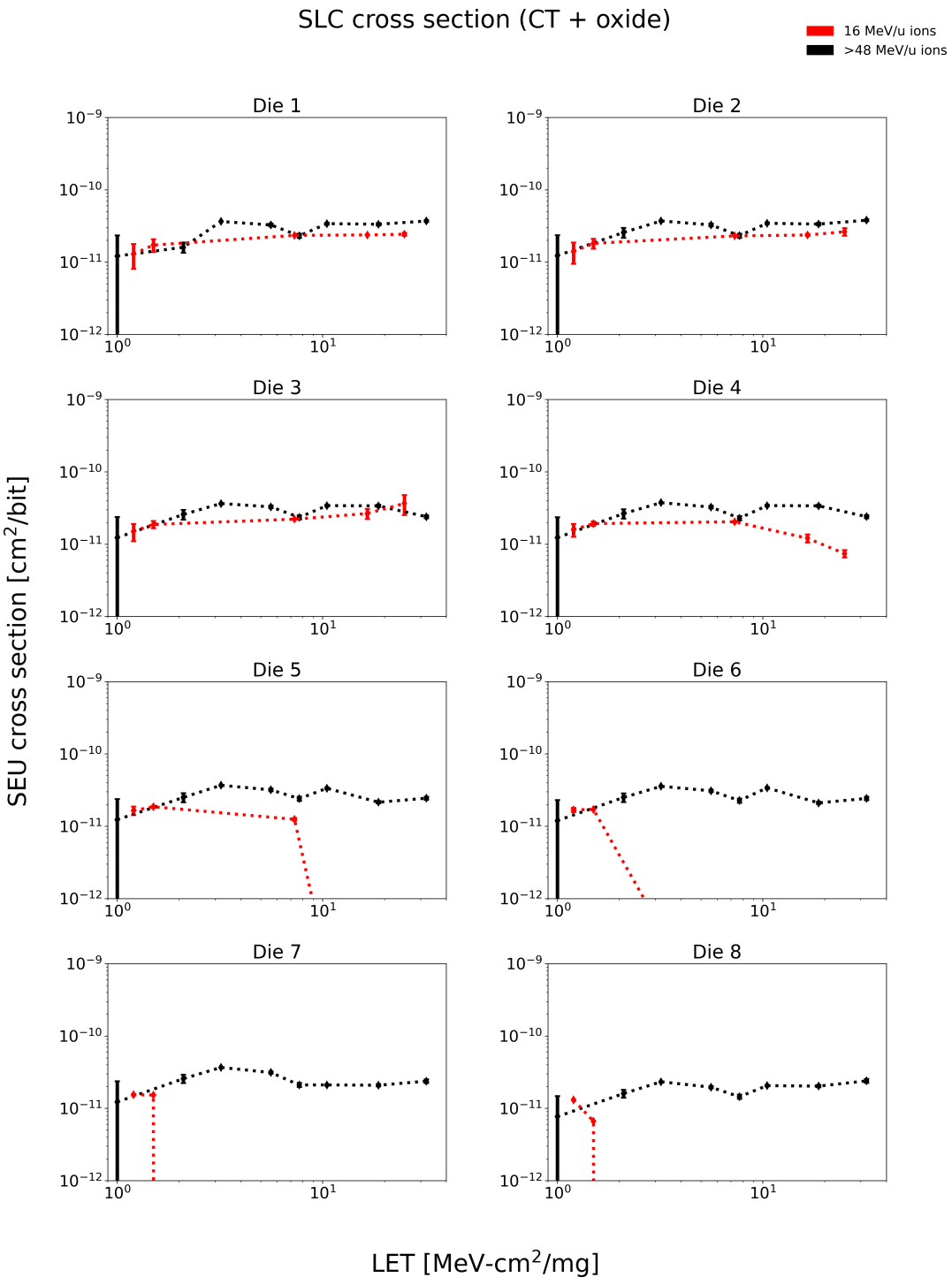


Figure 4.11. Cumulative cross sections, as predicted by MRED, over all NAND arrays in 8-die package for cells operated in SLC mode (CT + tunnel oxide sensitive).

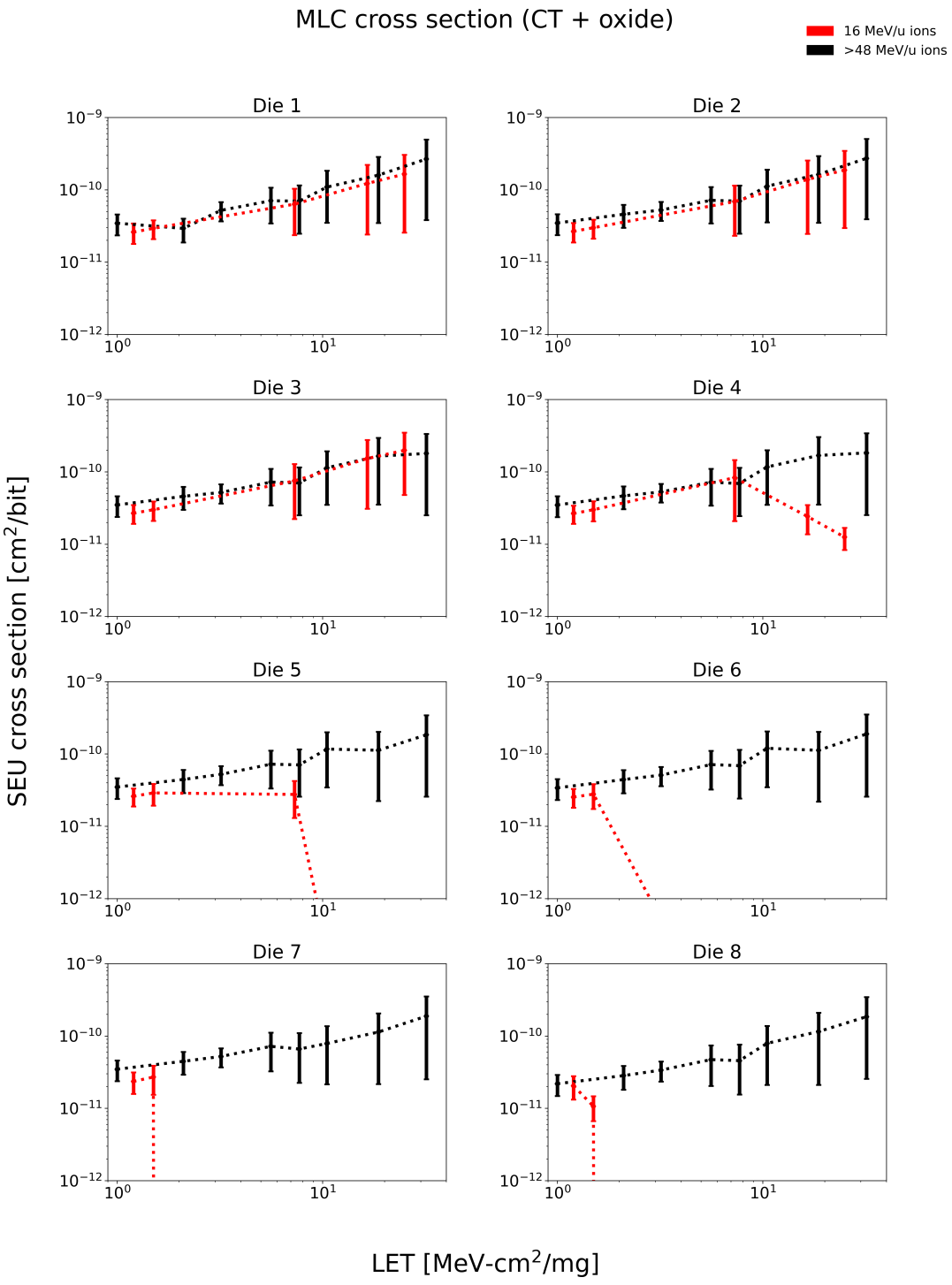


Figure 4.12. Cumulative cross sections, as predicted by MRED, over all NAND arrays in 8-die package for cells operated in MLC mode (CT + tunnel oxide sensitive).

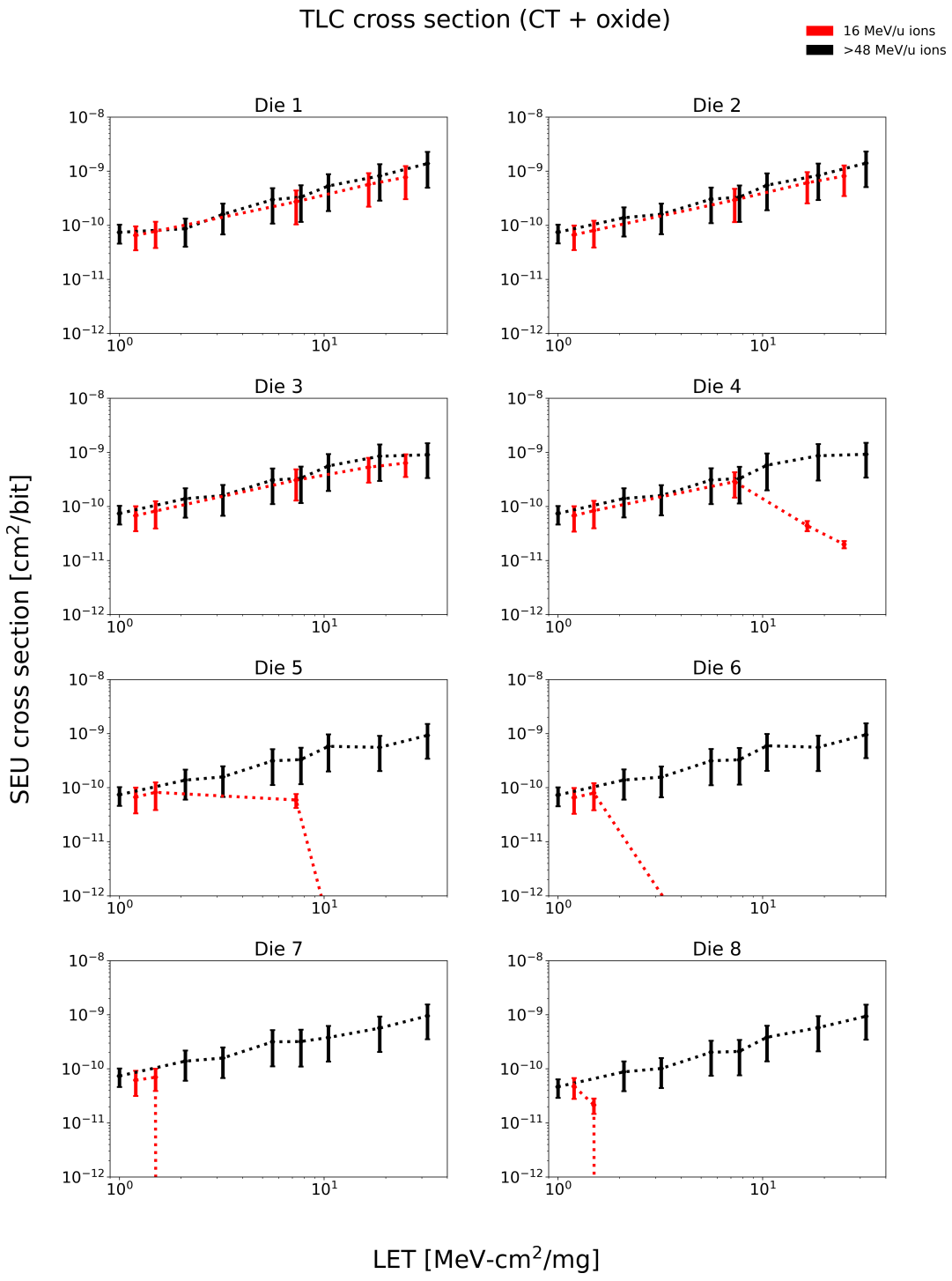


Figure 4.13. Cumulative cross sections, as predicted by MRED, over all NAND arrays in 8-die package for cells operated in TLC mode (CT + tunnel oxide sensitive).

The vertical bars in these plots correspond to approximate ranges of Q_{crit} for the different operating modes, or, equivalently, the average intrinsic in ΔV_{th} variation for the technology. The SEU cross section increases by roughly an order of magnitude between SLC and MLC cell mode.

The difference between the simulations in which the tunnel oxide is present as a region which contributes to upset is noted by a smaller difference between cross section curves for the two testing ion energy ranges over the whole range of LET values. Simulations in which the charge trap only is treated as sensitive show differences of up 3-5x in the SEU cross section as a function of LET, and in both cases the 16 MeV/u ions at higher LETs begin to fall off as the die index increases.

The implications of these results are several. First, designing experiments for part qualification of a 3D NAND for heavy-ion susceptible must take into account the package-level configuration of the chip. While the previous experimental heavy-ion results presented at the beginning of this chapter were obtained using a decapsulated chip within which only the top die was accessed for read/write operations, researchers attempting to characterize or qualify a part without such considerations or testing capabilities must be careful to select a range of heavy-ion beam values which ensure that ions will not range out within the package, or risk significantly underestimating the SEU vulnerability of the DUT especially for high LET values. The other side of this observation is that the significantly longer range of low LET particle even at relatively low kinetic energies ensures very little difference in the obtained SEU cross section regardless of the energy of the beam used. As shown in the previous chapter however, for the 72-layer Hynix NAND device (and most likely others of similar design) a minimum threshold LET below which no single events are observed was not found. This motivated the experiments and

simulations for detecting upsets from directly ionizing protons which have a maximum LET of just over 0.5 MeV-cm^2/mg.

Figures 4.14 and 4.15 show the cumulative SEU cross section for the three operating modes in question over all LET values simulated with low kinetic energy ions (16 MeV/u) in red and higher energy ions in black. In these plots, the total cross section is taken for all die in the package regardless of whether ions ranged out or not. It is worth pointing out that in some commercial chips already on the market, the number of die within the package may be at least double what is presented here, which would certainly increase the discrepancy between event cross sections within the different energy regimes as more low energy ions would range out and not contribute to any events in the buried cells.

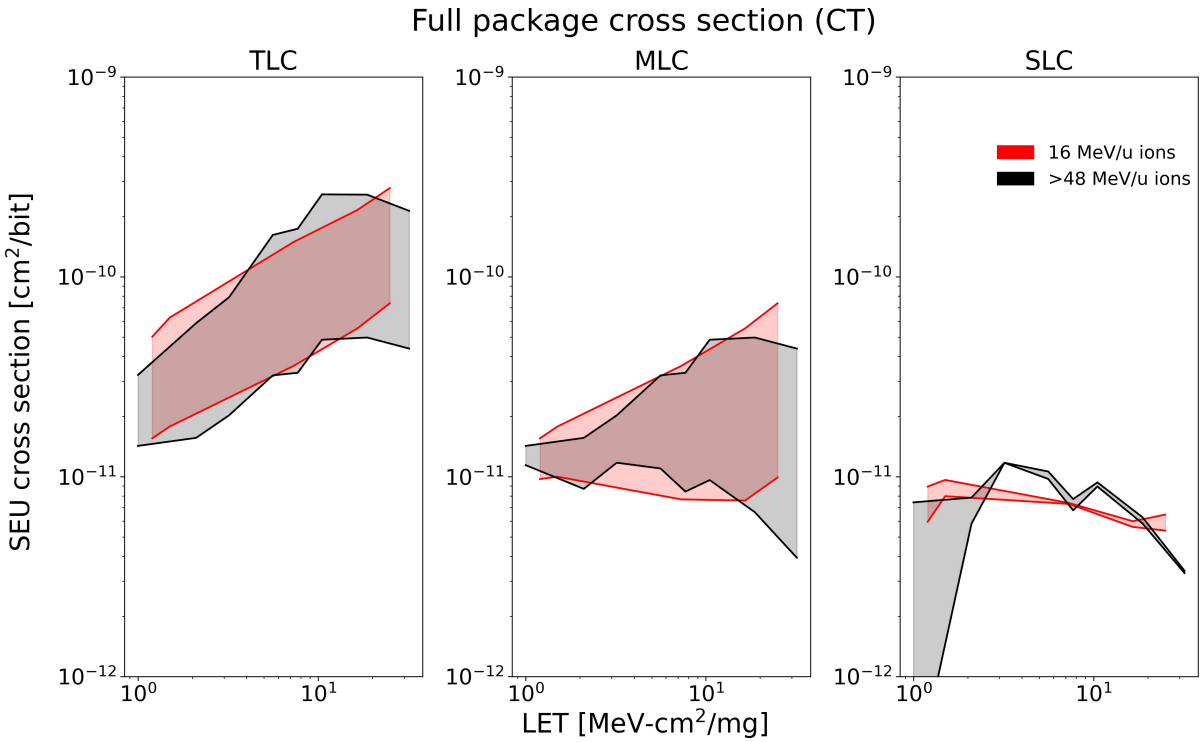


Figure 4.14. Cumulative cross sections, as predicted by MRED, over all CT NAND arrays in 8-die package for cells operated in all three modes. Energy deposition collected for both the charge trap regions for each cell.

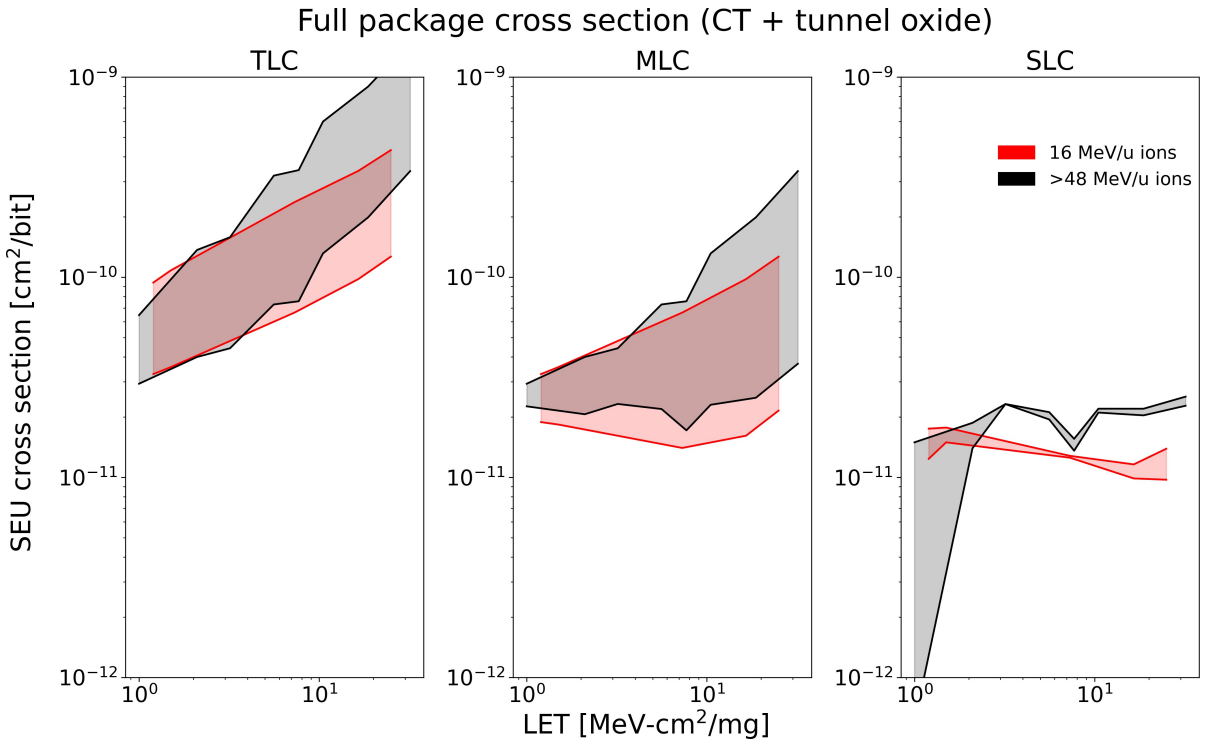


Figure 4.15. Cumulative cross sections, as predicted by MRED, over all CT NAND arrays in 8-die package for cells operated in all three modes. Energy deposition collected for both the charge trap and tunnel oxide regions for each cell.

The above distributions are given for the CT NAND configuration with tungsten WLs. In Figures 4.16 to 4.18, the same simulations were performed with polysilicon WLs to approximate the response of an FG 3D NAND. The same kinds of trends and effects as discussed above apply in the case of FG 3D NAND, with only one slight difference in that the upset cross sections determined by using lower energy ions is slightly larger than those simulated at higher energies. Once again, there is an increase in the SEU cross section for higher bit density operating modes, and the fall-off of particles ranging out in the lower half of the package is clearly seen. Plotting the cumulative upset cross section over all 8 die as done above is shown in Figure 4.19.

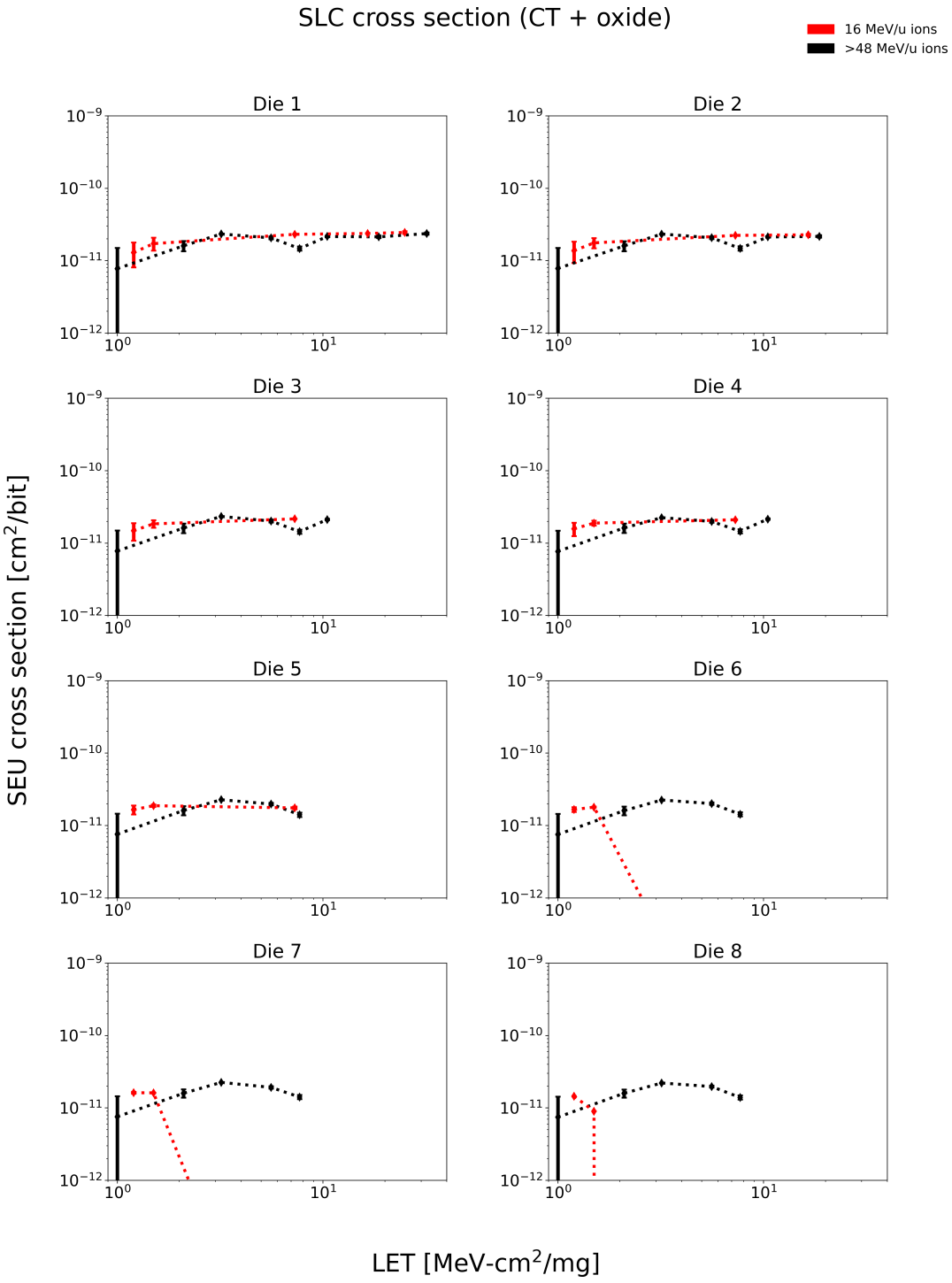


Figure 4.16. Cumulative cross sections, as predicted by MRED, over all FG NAND arrays in 8-die package for cells operated in SLC mode

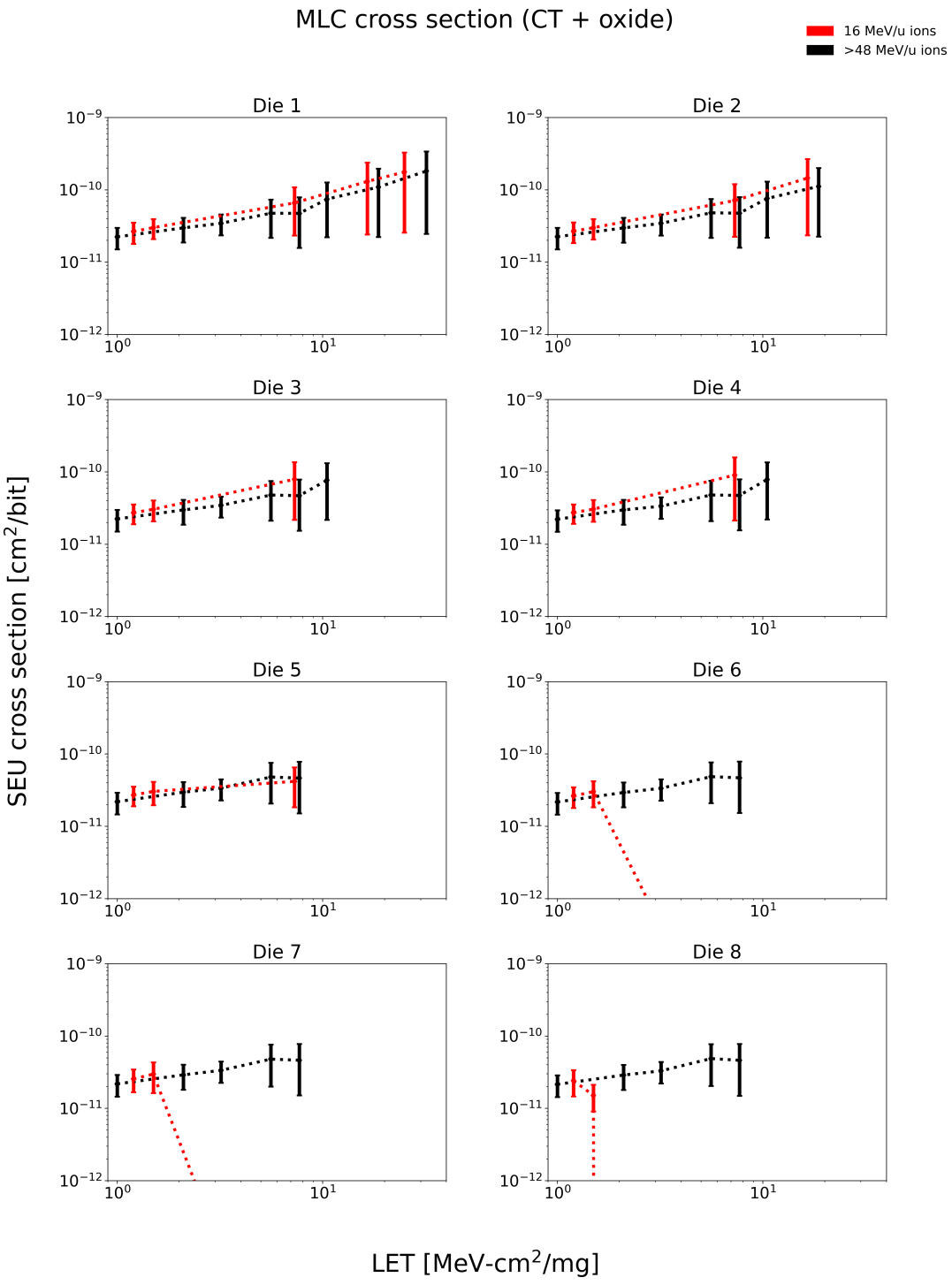


Figure 4.17. Cumulative cross sections, as predicted by MRED, over all FG NAND arrays in 8-die package for cells operated in MLC mode.

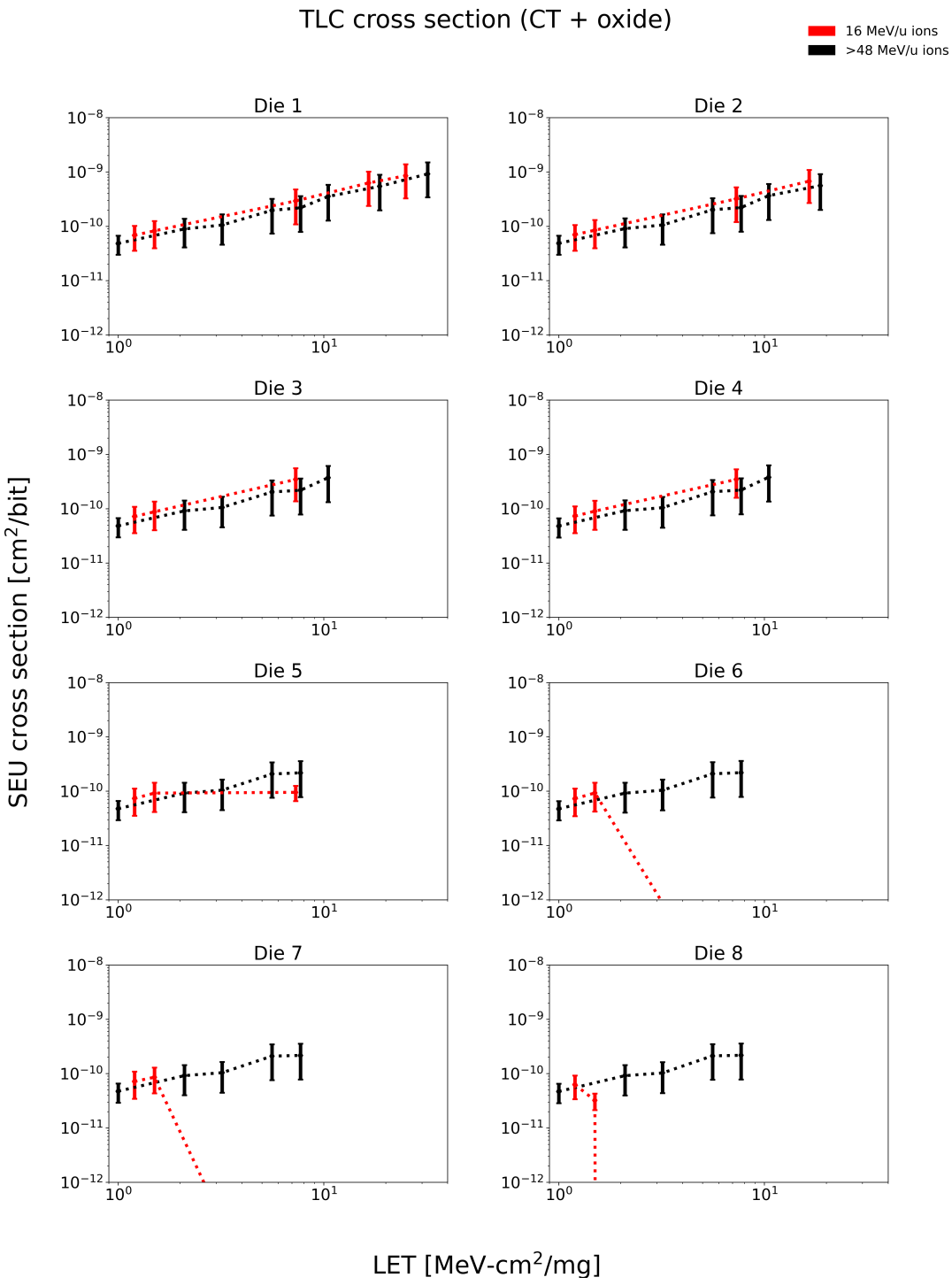


Figure 4.18. Cumulative cross sections, as predicted by MRED, over all FG NAND arrays in 8-die package for cells operated in TLC mode.

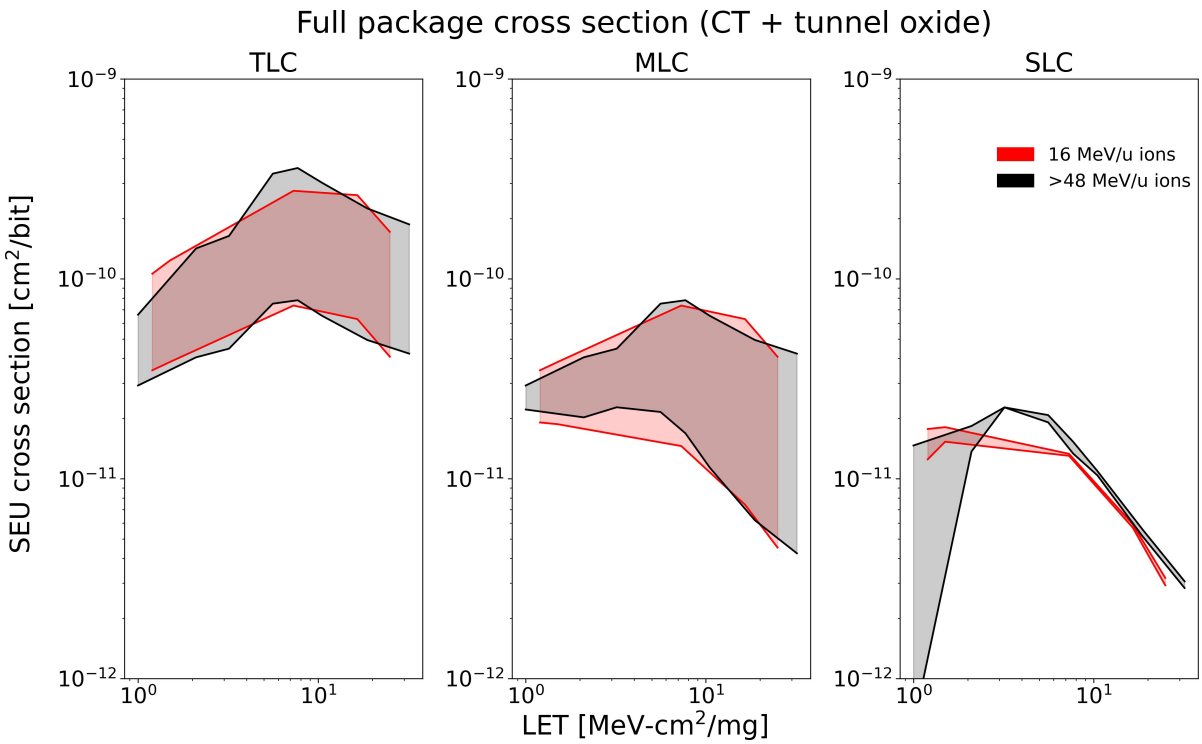


Figure 4.19. Cumulative cross sections, as predicted by MRED, over all FG NAND arrays in 8-die package for cells operated in all three modes. Energy deposition collected for both the charge trap and tunnel oxide regions for each cell.

A key difference between the two configurations in these cumulative plots is that the upset cross sections for all three operating modes are significantly lower at high LET values and for all ion energies than shown in the presence of tungsten WLs (Figure 4.16). Figure 4.20 supports the finding that the presence of high-Z metallization for the WL contacts throughout the NAND array is responsible for an increase in the upset cross section especially at high LETs. In this plot, the contributions to the overall SEU cross section within a CT NAND simulation with tungsten WLs at 16 MeV/u ions is broken into two sets of data which correspond to the relative proximity of each charge collection transistor region to the WL in the p-BiCS structure.

Transistors closer to the tungsten WL have a larger cross section than those further away, although the logical structure of the p-BiCS configuration means that each BL is near the WL on one side of the "U"-shaped string and far away on the other. In FG technologies, the only increased cross section due to high-Z materials comes instead from the BEOL and the TSV structures adjacent to the memory array, which are present in CT NAND as well. As a whole, FG NAND is therefore, from a physical mechanism standpoint, preferable for parts in a heavy-ion environment.

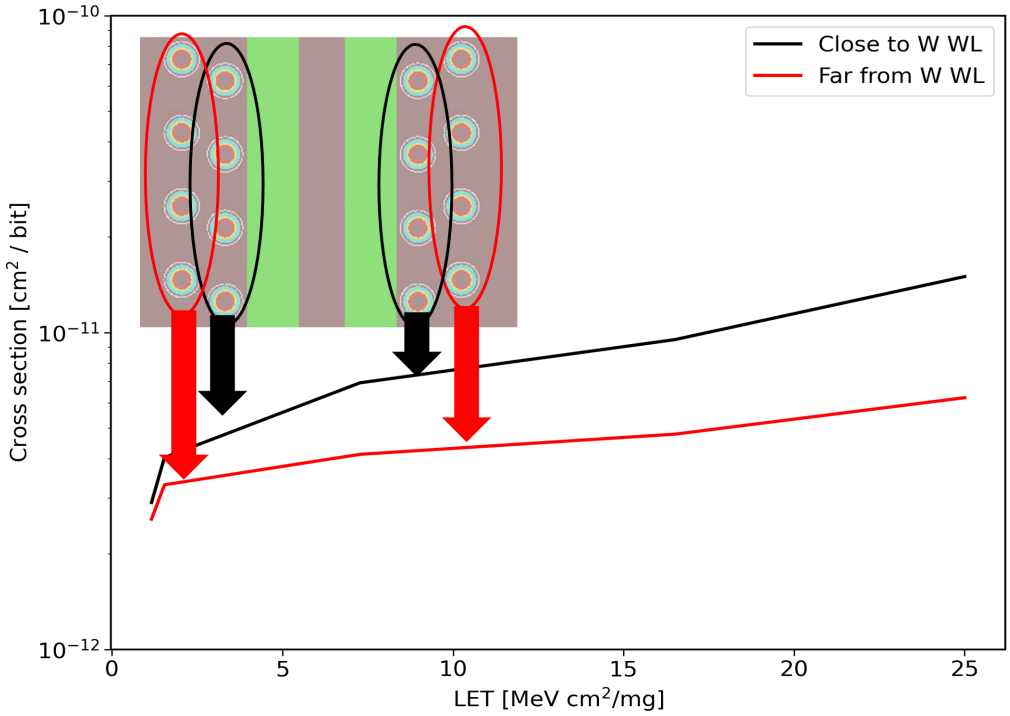


Figure 4.20. The cross section for cells positioned closer to the tungsten word lines is noticeably larger at higher LET than for those positioned farther away.

Chapter 5

Total dose effects and mechanisms in 3D NAND

This final chapter introduces simulations based on the previous MRED model of the 72-layer 3D CT NAND to provide a mechanistic understanding of previously published experimental results which report depth-dependent TID-induced threshold voltage shift (ΔV_{th}) for 3D NAND devices irradiated in X-ray and gamma environments. Floating gate transistors are shown to have significantly less TID-induced ΔV_{th} throughout the memory stack than CT transistors due to the presence of high-Z tungsten WL contacts in the case of CT 3D NAND. Dose enhancement effects from the BEOL metallization are observed in both cases, with a fall-off observed as a function of physical depth in the memory consistent with generalized structure DEF results as already presented in the case of FG 3D NAND where the BEOL is the largest metal region near the cells. This fall off suggests better TID performance for FG 3D NAND especially for newer technologies with more vertical layers. These trends between CT and FG NAND in TID-induced ΔV_{th} are observed in both normal and backside irradiation, with the strongest effects obtained at X-ray energies of 50 keV.

I. Monte Carlo simulation details for dose enhancement in CT and FG 3D NAND

Previous experimental results reported in the literature have suggested [53] and discounted [50] large dose-enhancement effects in 3D NAND, but no studies are found which approach the problem from a mechanistic modeling standpoint while also considering important variations in devices from different manufacturers.

There are a couple key features to consider. The first is that very little experimental data from X-ray and gamma-ray TID testing is available, and some of the reported results do not have the shifts in ΔV_{th} reported at the level of depth/WL [45]. Another important consideration is that variability of ΔV_{th} intrinsically is expected in these devices and accounted for with error-correction code (ECC).

The final key to understanding the distribution of TID in 3D NAND is the structural differences between CT and FG 3D NAND. To recap, FG NAND uses polysilicon WLs as the gate contact for each memory cell in the array; CT NAND uses metal (most commonly tungsten) as the WL contact. This means that for two seemingly similar devices in terms of functionality and application, recalling the results presented in Section 2.3, it is actually likely that the presence of high-Z materials throughout the NAND array results in a significantly stronger dose-enhancement effect for CT NAND than is observed in FG NAND. Characterizing and quantifying this effect due to the relative change in high-Z materials between devices is the focus of this final chapter, in which it is shown that the dose enhancement profile varies significantly for common TID test environments between CT and FG NAND.

Simulations were performed with MRED of the detailed structure as described in Figures 3.15 to 3.19, which takes into account the proper BEOL metallization and the voids within the NAND array between adjacent WLs. Additional simulations were performed with a modified structure which includes a sheet of tungsten situated 1 μm from the bottom of the p-BiCS structure which accounts for the metallization associated with the CMOS peripheral circuitry commonly found in most 3D NAND devices regardless of manufacturers (PUC - peripheral-under-cell).

Environments used for the simulations were a 10-keV X-ray spectrum (as shown in Figure 1.5), monoenergetic X-rays at values of 10 and 50 keV, and ^{60}Co gammas with a mean energy of 1.25 MeV. The incident photons were simulated at normal incidence (through the BEOL, from top to bottom in the NAND array) and backside incidence (through the substrate material and PUC metallization first). These energies were selected due to the common practice of TID testing using ^{60}Co and ARACOR X-rays by researchers.

Based on the results presented in Section 2.3 on the dose-enhancement effect in generalized 3D structures, dose enhancement is more likely to occur for photon energies at 50 keV, with relatively little effect observed at gamma energies. However, as the tungsten word lines are not sheets as in the generalized examples, and are only 10s of nm thick, this effect could very slight or insignificant.

Charge collection for the TID calculation was tracked in larger sensitive detectors than were used in single-event effect investigations, as accumulated dose in both the charge trap layer (or floating gate) and the surrounding oxides contributes to the shift in electrical characteristics for the cell which leads to errors.

II. Contributions to TID-induced ΔV_{th} based on physical memory configurations

Intuitively meaningful comparisons between different simulated energies are not enabled through normalization to fluence, as in the case of single-event effects. This is because the generation of energetic electrons is a function of the initial photon energy and thus two sets of simulations in the same material structure but with different energies of X-rays are not guaranteed to the have the same (or even close to) the same total dose. Because the main question however is focused on the relative change in TID-induced ΔV_{th} distributions for NAND devices with and without tungsten word lines (CT and FG memories), the TID normalization of the distributions as a function of physical depth in the memory array does allow meaningful comparisons to be made.

It is important to emphasize again that different functionality, fabrication methods, feature size, etc. between manufacturers will also potentially contribute to differences in TID response, and that our aim here is not to rigorously characterize the total response over what is, in reality, a large set of parameters. Figure 5.1 shows the simulated TID-normalized dose per vertical layer in the NAND stack for FG NAND without tungsten WLs for both normal and backside photon incidence. These simulations were performed using the 10 keV ARACOR spectrum, and there was no PUC structure positioned underneath the NAND array in these plots. The dose-enhancement profile is clearly visible near the BEOL (which does contain both tungsten and copper), with the relative increase in deposited dose falling off to an approximately equilibrium level around WL 20 (which is roughly 1 μm away from the BEOL). As there is no PUC in this simulation structure, the profile begins to fall off at the bottom of the stack.

Figure 5.2 shows the results of the same simulations performed with tungsten substituted as the WL material. Rather than a sharp fall-off in DEF at depth deeper in the memory stack, there is instead a much more gradual slope observed in the dose enhancement profile. Backside-incident photons produce a dose-enhancement effect due to the presence of tungsten WLs in the bottom of the stack; the preferentially forward directed nature of the dose-enhancement phenomenon in X-ray environments as discussed in Chapter 1 is consistent with these results.

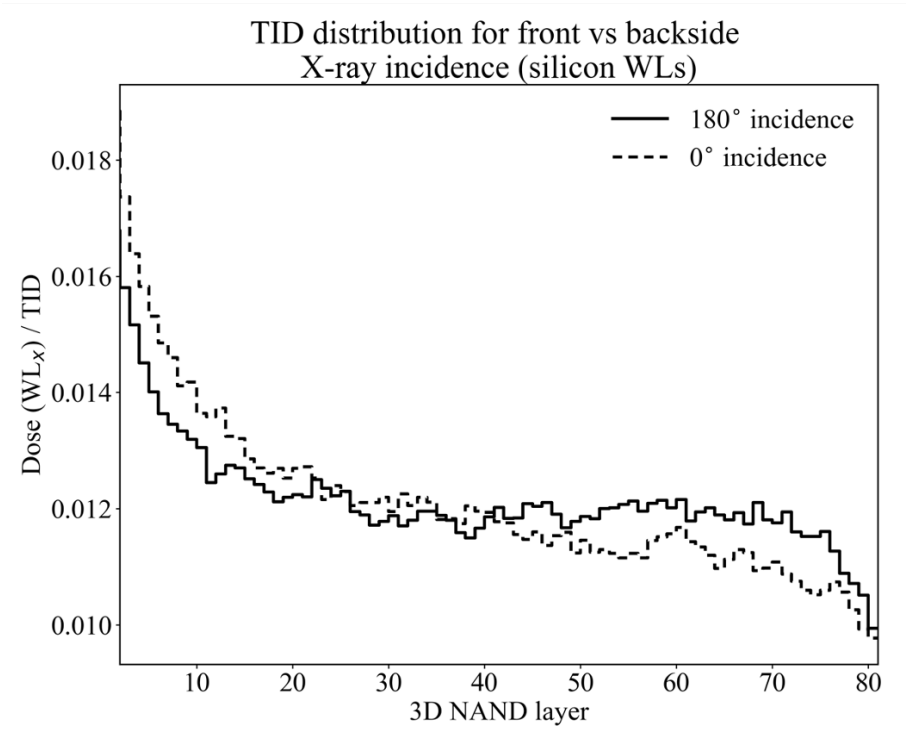


Figure 5.1. Simulated total dose deposited per layer normalized to the dose in the top layer for each specified run. These data correspond to FG NAND (with polysilicon WLs).

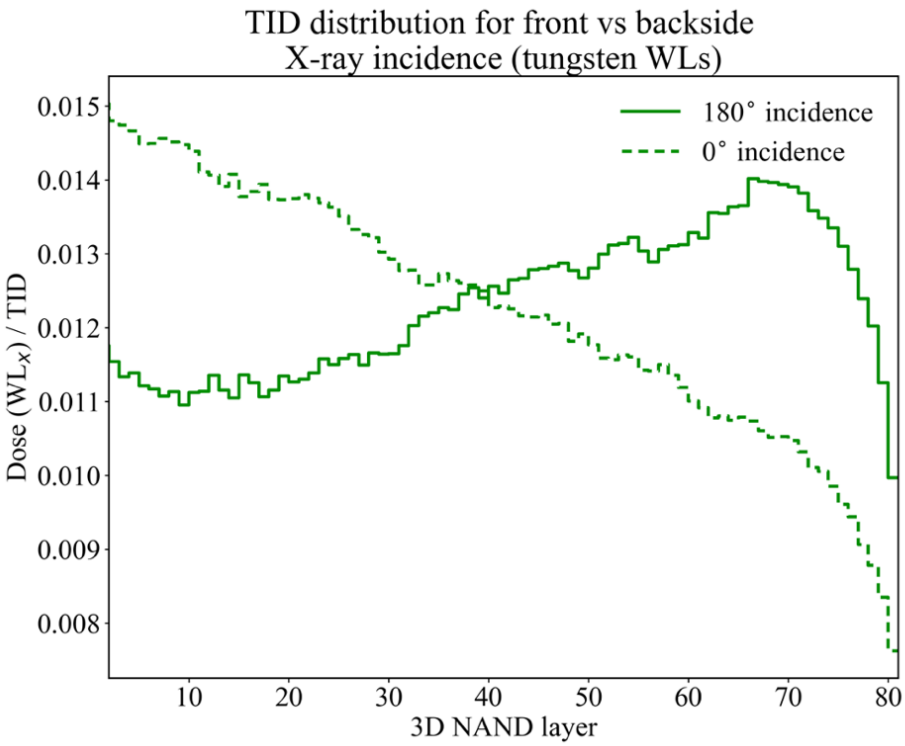


Figure 5.2 Simulated total dose deposited per layer normalized to the dose in the top layer (closest to the BEOL) for CT NAND (tungsten metal WLs).

Since the material composition of the WLs and the incident photon energies are the only parameters adjusted in the simulations, the polysilicon WL results can be used within simulations at the same energies to establish the equilibrium dose to which the data in the tungsten WL simulations is normalized, providing a value for the dose-enhancement factor (DEF). This DEF is not dependent on the BEOL metallization nor the presence or lack of the PUC in the simulation structure, but only on the presence of high-Z WL materials within the NAND array itself. Figure 5.3 and Figure 5.4 presents these distributions of DEF for monoenergetic 50 keV X-rays, 10 keV ARACOR X-rays, and ^{60}Co gammas at both normal and backside incidence. The data for ^{60}Co is rather noisy due to the lack of sufficient computation time. The DEF due to WLs in CT NAND is responsible for anywhere from 2x to 10x the equilibrium dose at lower photon

energies, with some enhancement suggested in the case of gammas as well, as reported in the experimental literature.

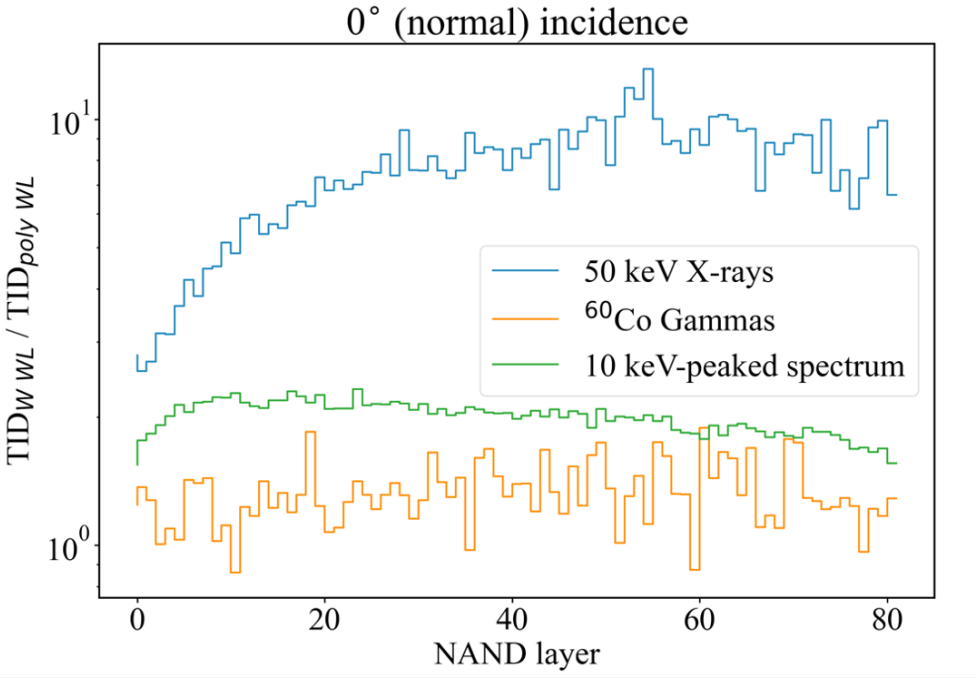


Figure 5.3. The distribution of simulated DEF as a function of depth and testing environment with normal incidence photons (through the BEOL).

While previous experimental studies in TID distribution as a function of depth in 3D NAND point to the narrowing of the channel as a likely cause for the increased TID-induced V_{th} shift at the bottom of the memory array, these results show that there are multiple factors to consider, especially in the case of CT 3D NAND.

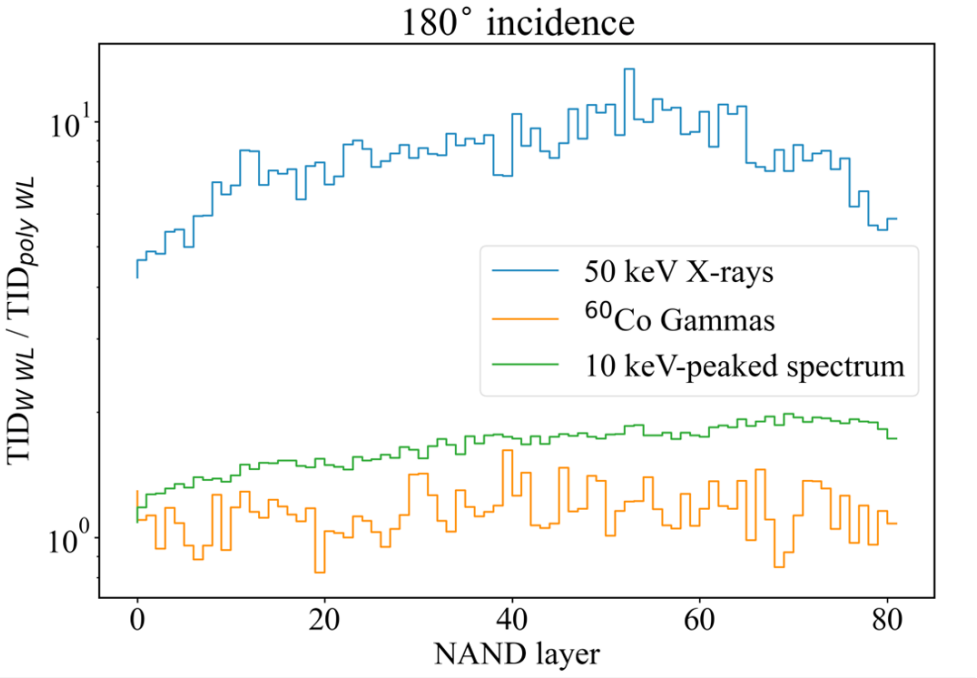


Figure 5.4. The distribution of simulated DEF as a function of depth and testing environment with photons incident through the backside.

Chapter 6

Conclusion

This dissertation described 3DICs with a particular emphasis on both charge-trap and floating gate 3D NAND. General structures representative of 3DICs and relevant to the understanding of radiation effects in 3D NAND memories were introduced and discussed, followed by a detailed treatment of the single-event upset characterization of 72-layer CT NAND in both proton and heavy-ion environments.

Comparisons have been made between experimental data for single-event effect in proton and heavy-ion environments and the MRED model constructed for detailed mechanistic studies. These Monte Carlo simulation models of the NAND stack were extended to provide insight into how SEU cross sections for multi-die packaged NAND memories change depending on the environment of interest, on the multi-bit upset characterization of NAND flash a memory die, and on the importance of fabrication methods and the differences between CT and FG NAND as it pertains to the dose-enhancement effect in X-ray and gamma-ray TID testing. Simulations and modeling of NAND structures also provided insight into the details of NAND structures responsible for a previously unexplained distribution of proton-induced direct-ionization upsets as a function of depth in the stack.

The emphasis in this work has been on the contribution of the physical layout and configuration of 3D NAND stacks to the modulation of the local radiation environments for

multiple external environments of interest, but it is important to consider the functional level of analysis required for a full understanding to be developed. 3D NAND is an incredibly remarkable feat of modern engineering and science, enabling ultra-high memory densities on very small chips. These devices are ideal candidates for future use not just in consumer applications such as cell phones but also in radiation-rich environments such as interplanetary and orbital space missions which require high quality imaging.

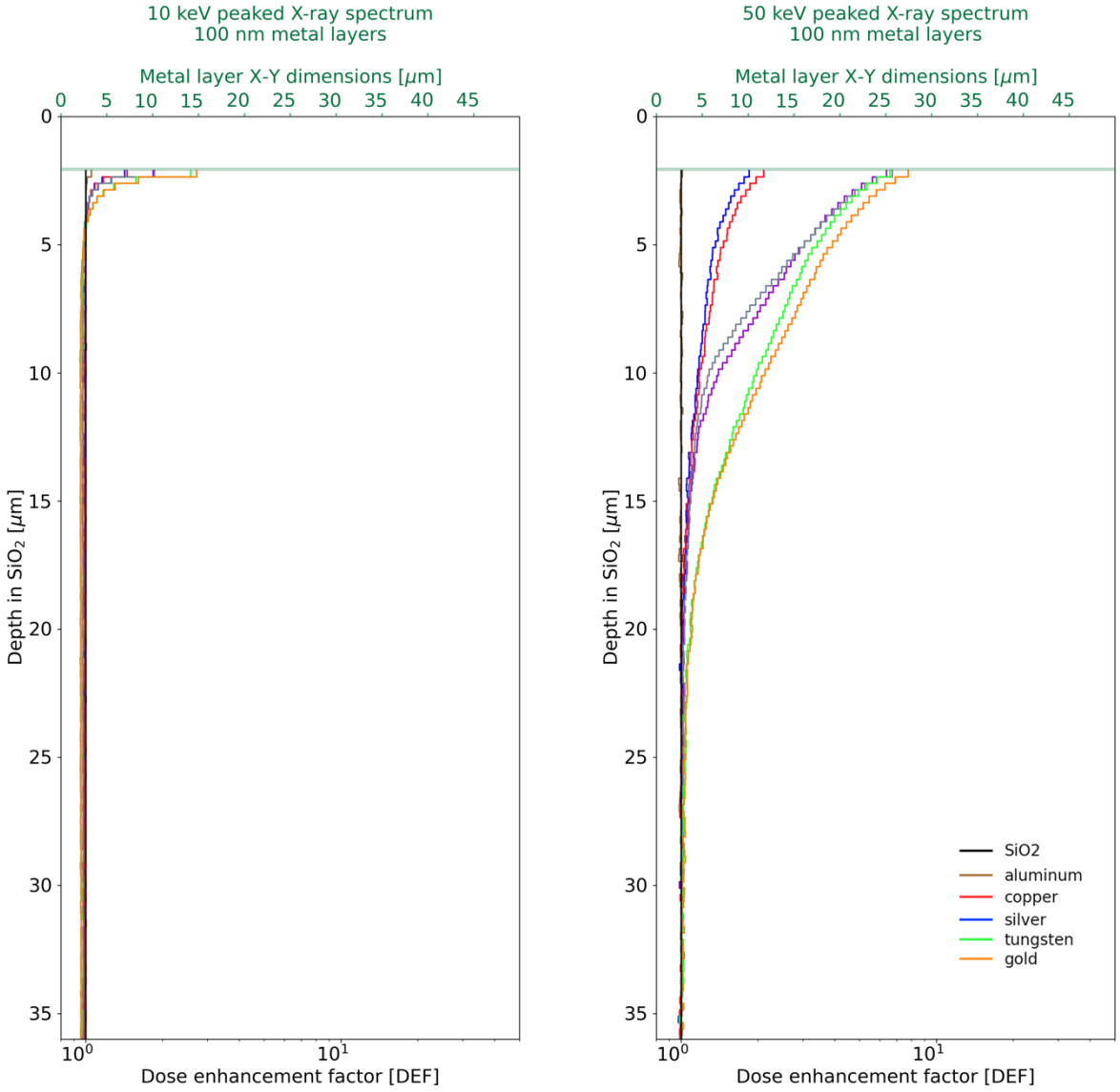
The uniquely 3D structure of these memory chips gives rise to multiple physical mechanisms for modulating local radiation environments and partially determining device performance. These structures include U-shaped bit lines, multiple vertical stacks of high-Z metals, N-filled voids between WLs, channel irregularities including non-uniform radii as a function of depth and misalignments between columns in multi-etch processing, multiple chips included within the same package up to total depths greater than 1 mm, WL contact fan-outs adjacent to rather than above or below active devices, and the gate-all around structure of the individual cells.

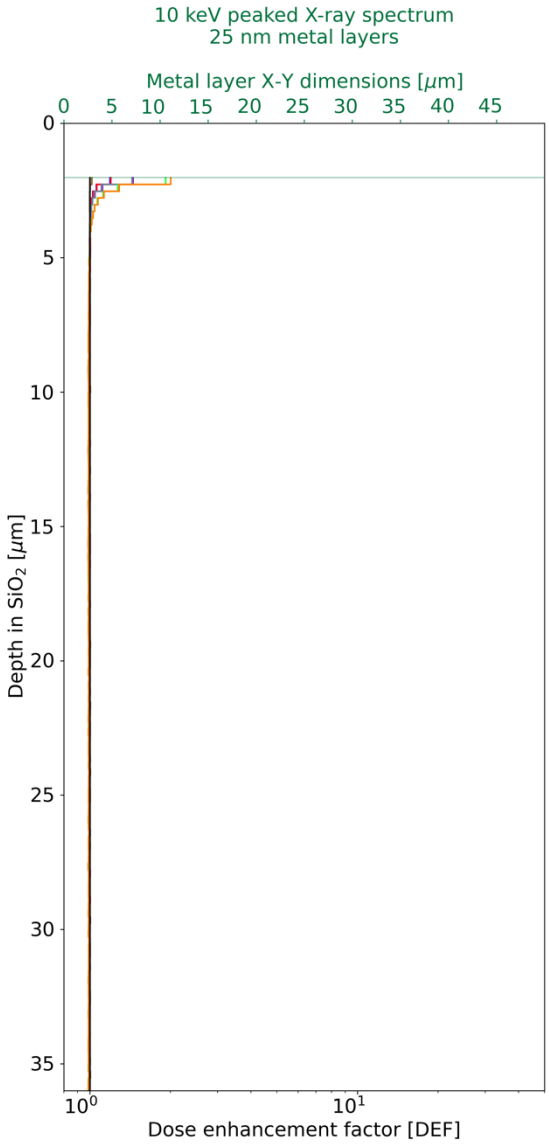
In this work, we have demonstrated how the presence of these structures contributes to the overall response of 3D NAND to radiation exposure for both TID and SEE. The key findings are that there are significant differences in terms of the material- and geometry-dependent effects in charge-trap and floating-gate 3D NAND memories. The central of these differences is the presence of tungsten word lines in CT NAND which introduces a relatively large amount of high-Z material throughout the memory array near the sensitive transistor regions which store data. Using FG NAND devices with polysilicon WLs as a point of comparison, these tungsten WLs are responsible for an increase in the SEU cross section in heavy-ion environments especially at high LETs, and for a much more pronounced dose-enhancement effect throughout

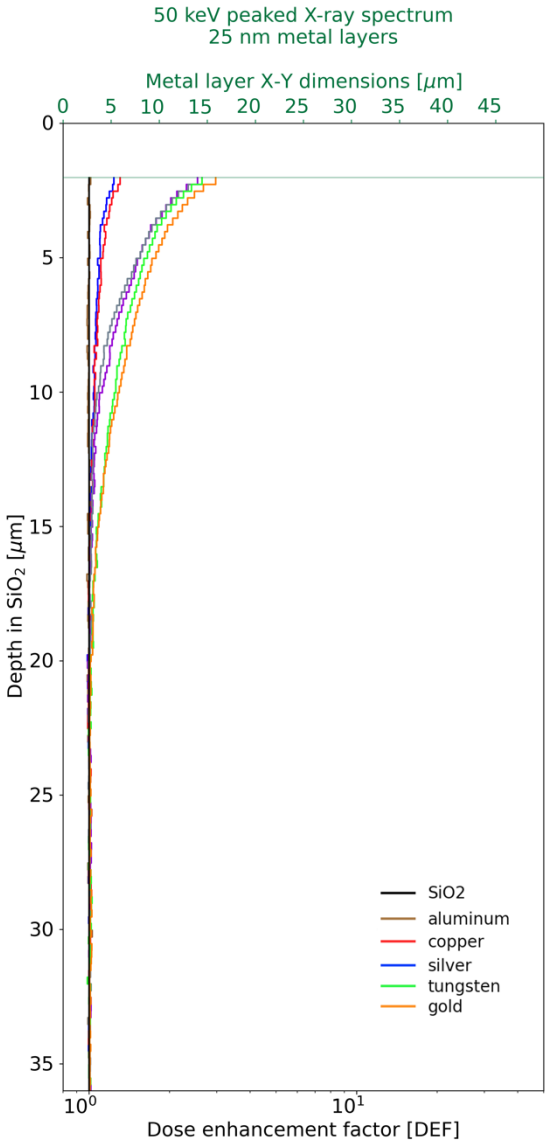
the entire memory array during TID testing with gammas and X-rays, as opposed to effects only caused by the BEOL metallization. The importance of proper test planning for multi-die packages of 3D NAND has been demonstrated, as heavy ions in relatively low-energy ion beams (16 MeV/u) should not be assumed to have sufficient range to penetrate the entire stack of die in the package which can be 1 mm or greater. Extrapolating total SEU cross sections from those measured only in the top die provide a reasonably good estimate of the total performance of the chip in the case of heavy-ion environments, while there are significant angular effects which lead to different event rates in the case of protons. The operating mode of the 3D NAND is also shown to be significant in terms of SEU hardness; greater bit density with MLC or TLC mode results in some cases to over an order of magnitude in SEU cross section.

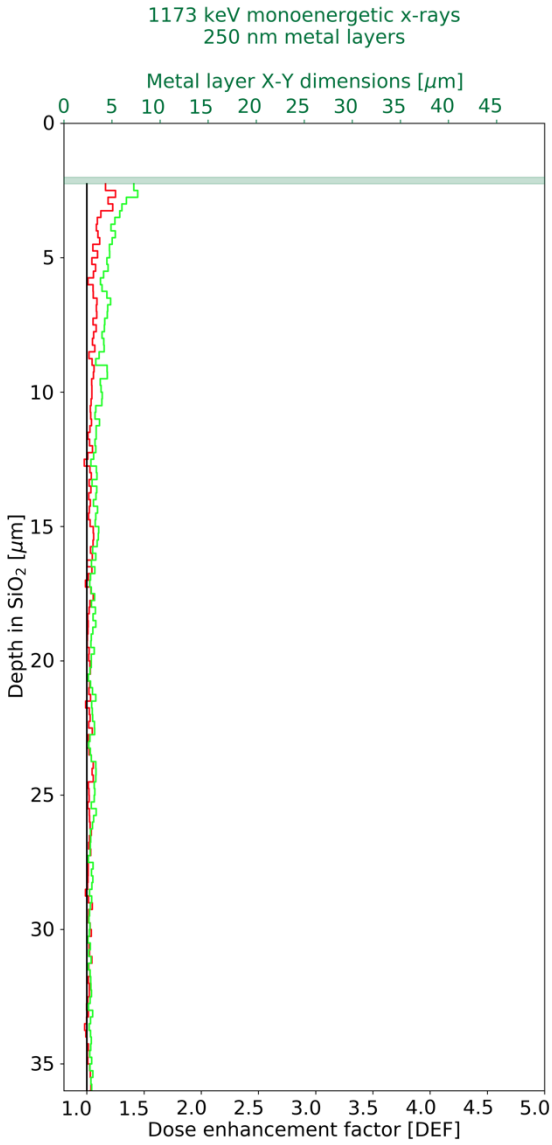
Appendix A

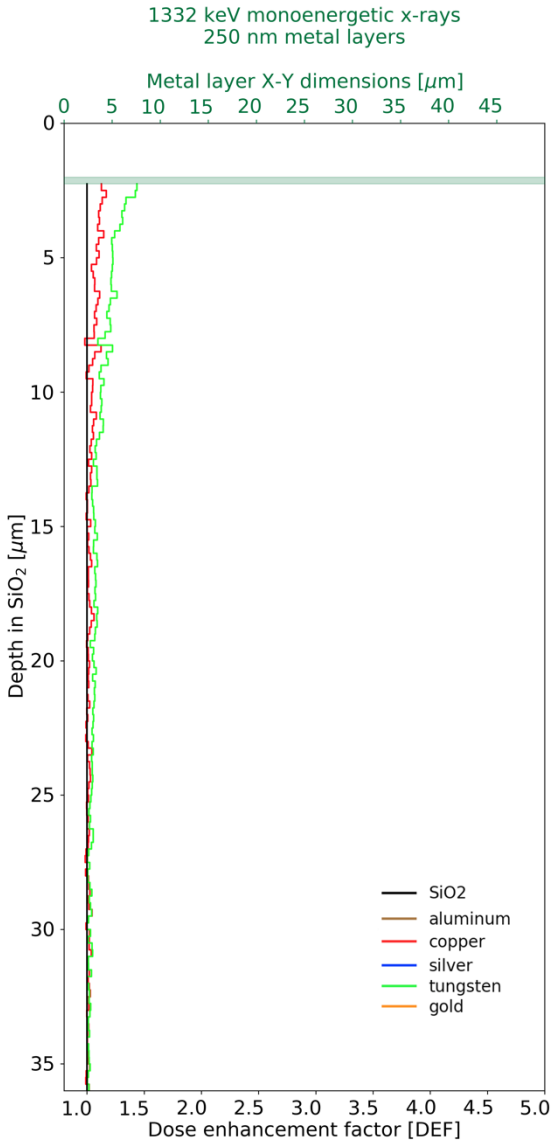
3D Dose Enhancement Effect – Extra Data for monolayers, alternative spacings

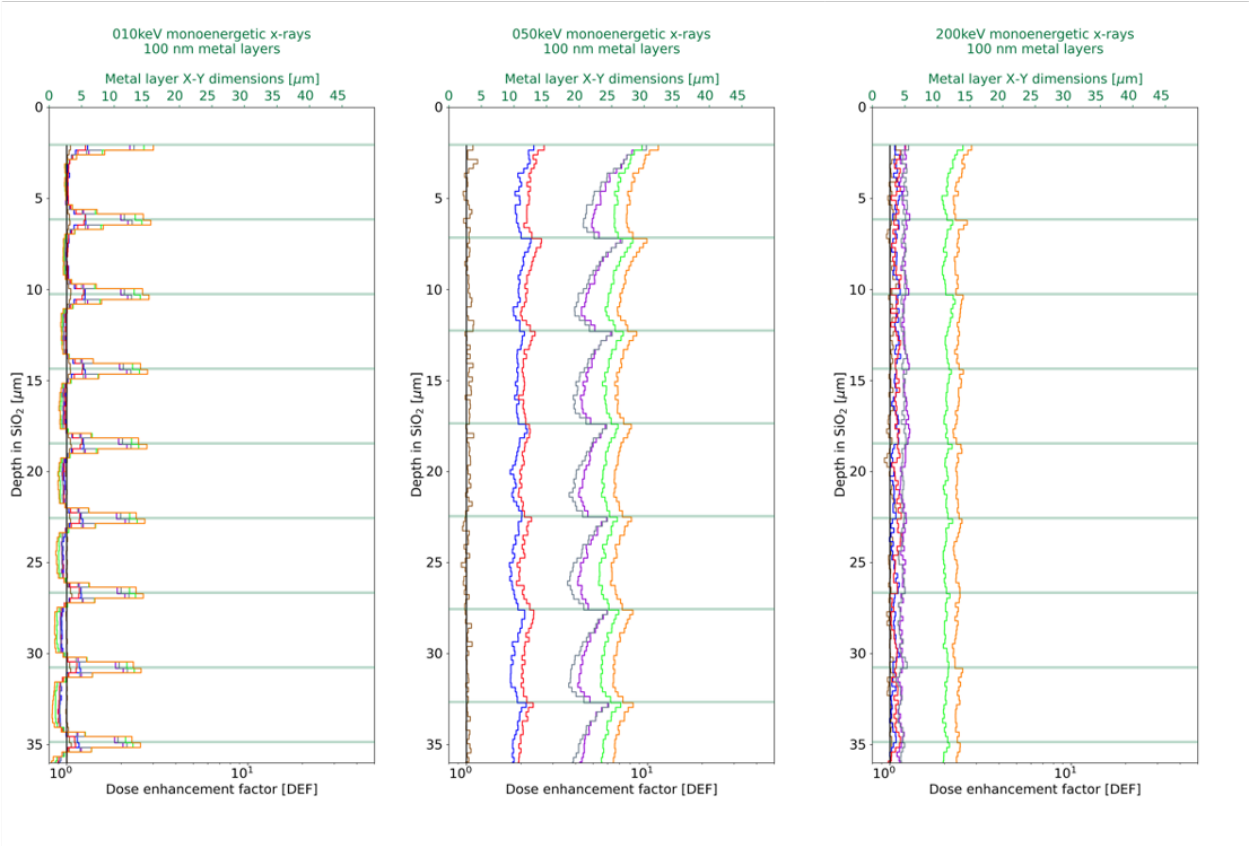


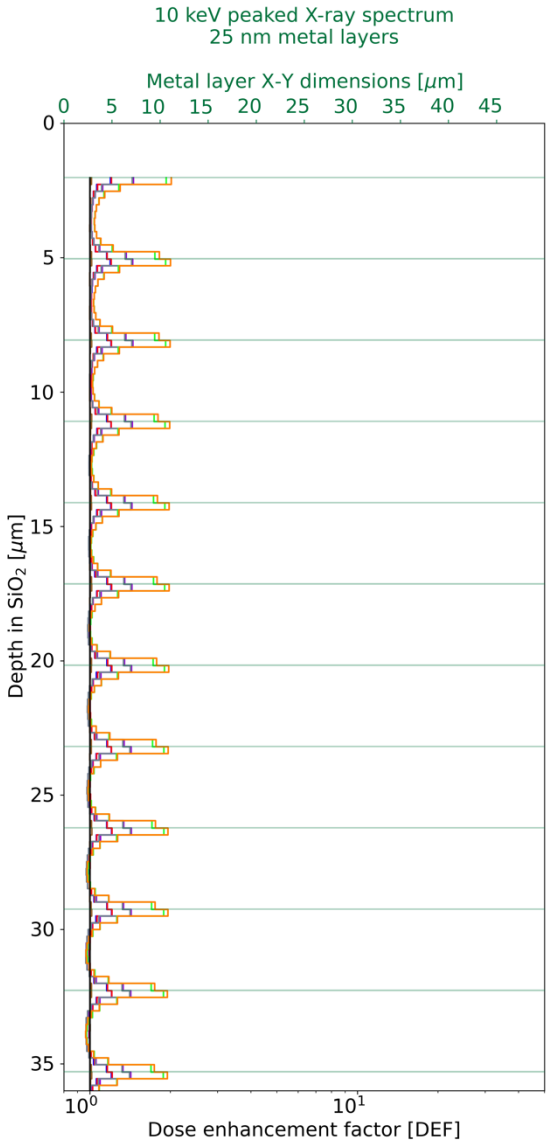

10 keV peaked X-ray spectrum
25 nm metal layers
Metal layer X-Y dimensions [μm]
0 5 10 15 20 25 30 35 40 45
0 5 10 15 20 25 30 35
Depth in SiO2 [μm]
10^0 10^1
Dose enhancement factor [DEF]

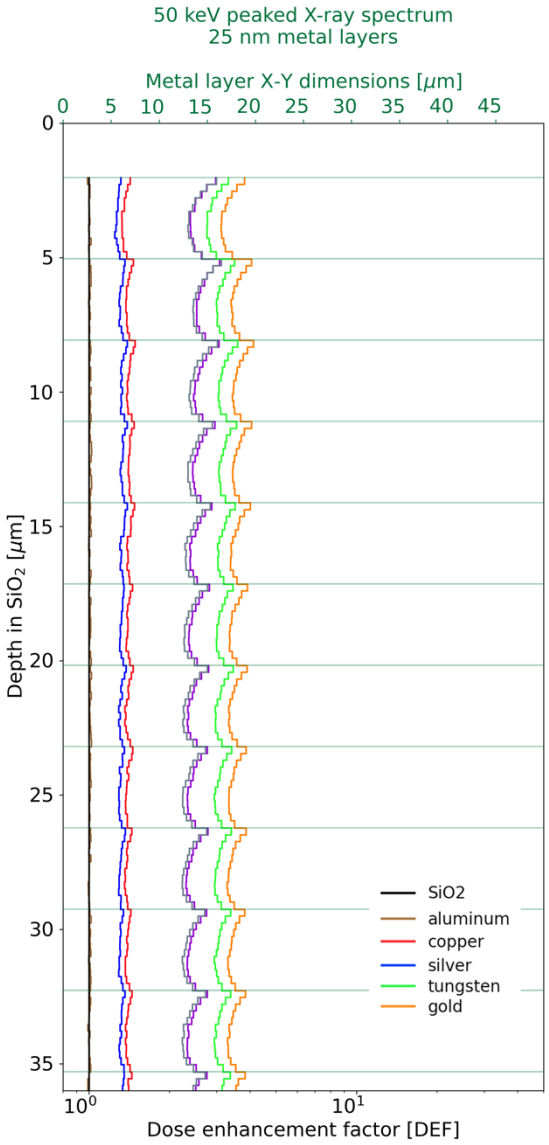

50 keV peaked X-ray spectrum
25 nm metal layers
Metal layer X-Y dimensions [μm]
0 5 10 15 20 25 30 35 40 45
0 5 10 15 20 25 30 35
Depth in SiO2 [μm]
10^0 10^1
Dose enhancement factor [DEF]
SiO2
aluminum
copper
silver
tungsten
gold

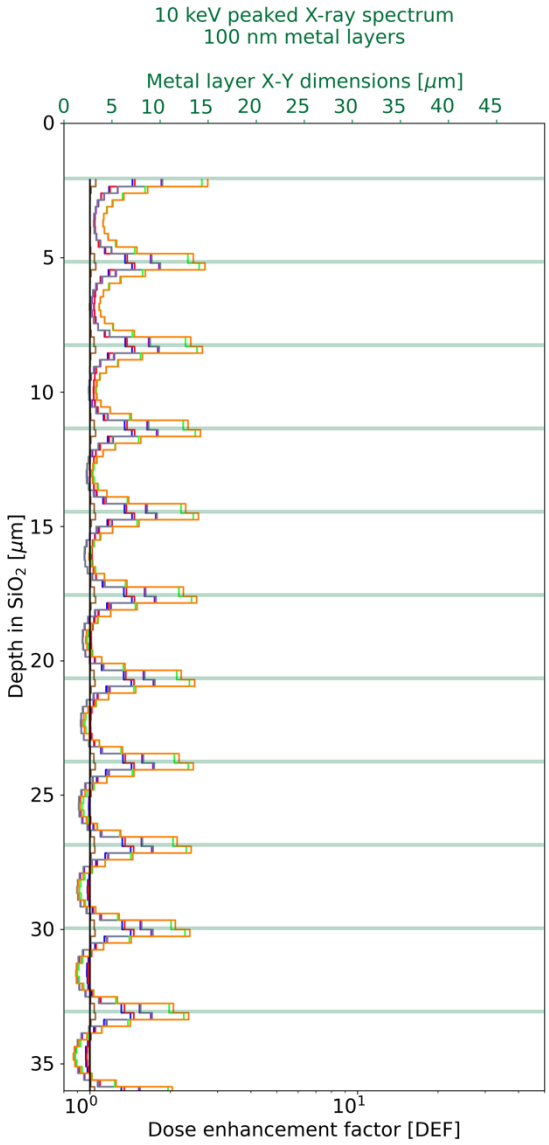

1173 keV monoenergetic x-rays
250 nm metal layers
Metal layer X-Y dimensions [μm]
0 5 10 15 20 25 30 35 40 45
Depth in SiO2 [μm]
0 5 10 15 20 25 30 35
Dose enhancement factor [DEF]
1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0

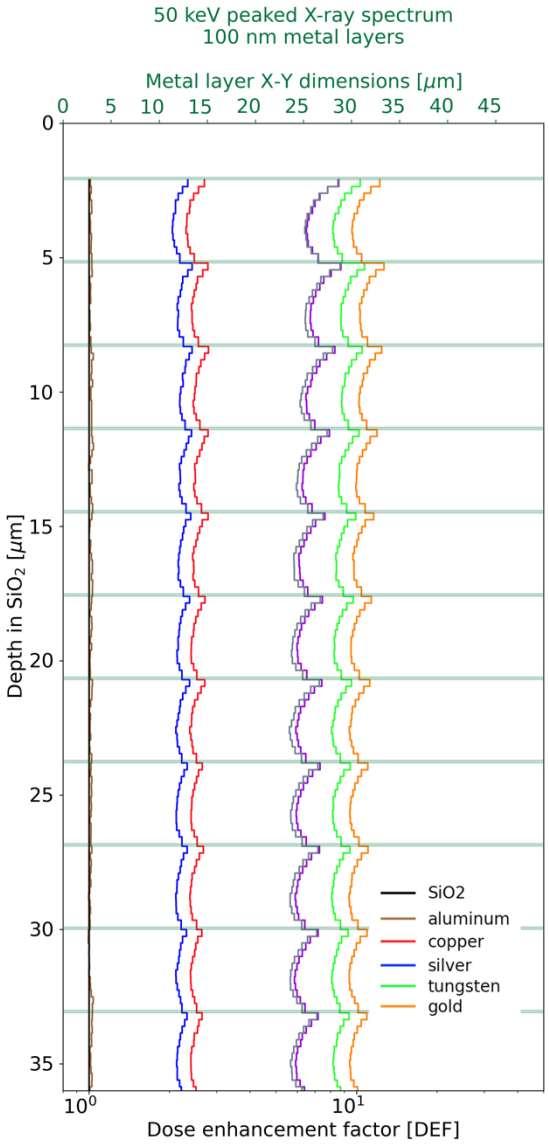

1332 keV monoenergetic x-rays
250 nm metal layers
Metal layer X-Y dimensions [μm]
0 5 10 15 20 25 30 35 40 45
Depth in SiO2 [μm]
0 5 10 15 20 25 30 35
Dose enhancement factor [DEF]
1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
SiO2
aluminum
copper
silver
tungsten
gold

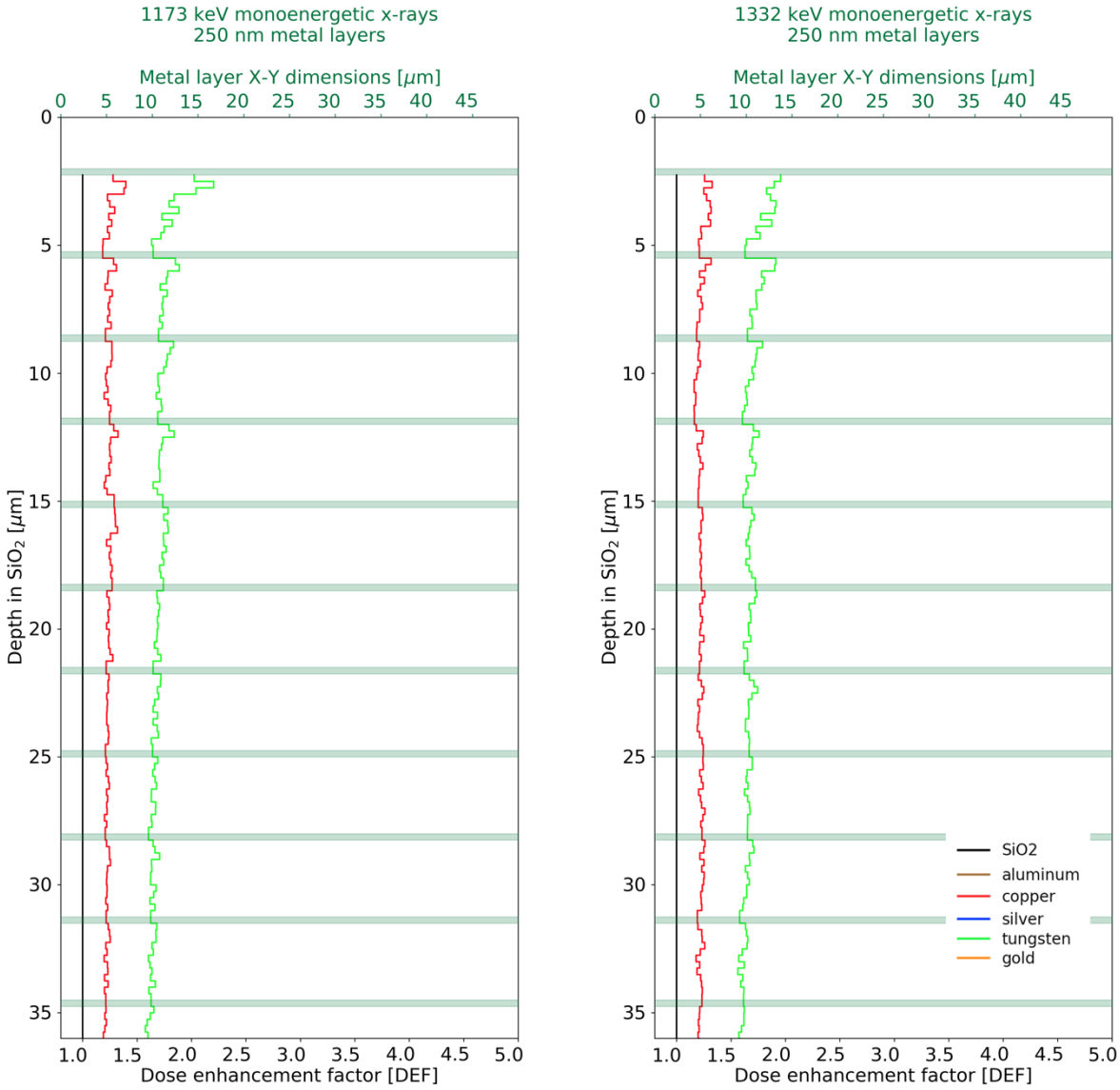

010keV monoenergetic x-rays
100 nm metal layers
Metal layer X-Y dimensions [μm]
Depth in SiO2 [μm]
Dose enhancement factor [DEF]
050keV monoenergetic x-rays
100 nm metal layers
Metal layer X-Y dimensions [μm]
Depth in SiO2 [μm]
Dose enhancement factor [DEF]
200keV monoenergetic x-rays
100 nm metal layers
Metal layer X-Y dimensions [μm]
Depth in SiO2 [μm]
Dose enhancement factor [DEF]


10 keV peaked X-ray spectrum
25 nm metal layers
Metal layer X-Y dimensions [μm]
0 5 10 15 20 25 30 35 40 45
0 5 10 15 20 25 30 35
Depth in SiO2 [μm]
10^0 10^1
Dose enhancement factor [DEF]


50 keV peaked X-ray spectrum
25 nm metal layers
Metal layer X-Y dimensions [μm]
0 5 10 15 20 25 30 35 40 45
0 5 10 15 20 25 30 35
Depth in SiO2 [μm]
10^0 10^1
Dose enhancement factor [DEF]
SiO2
aluminum
copper
silver
tungsten
gold


10 keV peaked X-ray spectrum
100 nm metal layers
Metal layer X-Y dimensions [μm]
0 5 10 15 20 25 30 35 40 45
0 5 10 15 20 25 30 35
Depth in SiO2 [μm]
10^0 10^1
Dose enhancement factor [DEF]


50 keV peaked X-ray spectrum
100 nm metal layers
Metal layer X-Y dimensions [μm]
0 5 10 15 20 25 30 35 40 45
0 5 10 15 20 25 30 35
Depth in SiO2 [μm]
10^0 10^1
Dose enhancement factor [DEF]
SiO2
aluminum
copper
silver
tungsten
gold


1173 keV monoenergetic x-rays
250 nm metal layers
Metal layer X-Y dimensions [μm]
0 5 10 15 20 25 30 35 40 45
Depth in SiO2 [μm]
0 5 10 15 20 25 30 35
Dose enhancement factor [DEF]
1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
1332 keV monoenergetic x-rays
250 nm metal layers
Metal layer X-Y dimensions [μm]
0 5 10 15 20 25 30 35 40 45
Depth in SiO2 [μm]
0 5 10 15 20 25 30 35
Dose enhancement factor [DEF]
1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
SiO2
aluminum
copper
silver
tungsten
gold

References

[1] B. Rossi, High Energy Particles, New York: Prentice-Hall, 1952.

[2] C. A. Klein, “Bandgap Dependence and Related Features of Radiation Ionization Energies in Semiconductors”, *J. Appl. Phys.*, Vol 39, No. 4, 1968, pp. 2029-38.

[3] D. S. Vibbert, M. L. Breeding, et al., "Layout Dependence of Neutron Single-Event Vulnerabilities in Sub-50nm Hardened Latches," *Journal of Radiation Effects and Engineering,* vol. 39, no. 1, pp. 139-146, 2021.

[4] Z. Zhang, "Investigation of Threshold Ion Range for Accurate Single Event Upset Measurements in both SOI and Bulk Technologies," *IEEE Trans. Nucl. Sci.,* vol. 61, no. 3, pp. 1459-1467, 2014.

[5] L. S. Abbott, "Shielding Against Initial Radiation from Nuclear Weapons," Oak Ridge National Laboratory, Knoxville, TN, 1973.

[6] E. Normand and W. Doherty, "Incorporation of ENDF-V Neutron Cross Section Data for Calculating Neutron-Induced Single Event Upsets," *IEEE Trans. Nucl. Sci.,* vol. 36, no. 6, pp. 2349-2355, 1989.

[7] A. Dasgupta, D. M. Fleetwood, R. A. Reed, R. A. Weller and M. H. Mendenhall, "Effects of Metal Gates and Back-end-of-line Materials on X-ray Dose in HfO2 Gate Oxide," *IEEE Trans. Nucl. Sci.,* vol. 58, no. 6, pp. 3139-3144, 2011.

[8] A. Dasgupta, D. M. Fleetwood, R. A. Reed, R. A. Weller, M. H. Mendenhall and B. D. Sierawski, "Dose Enhancement and Reduction in SiO2 and High-κ MOS Insulators," *IEEE Trans. Nucl. Sci.,* vol. 57, no. 6, pp. 3463-3469, 2010.

[9] M. P. King and a. et, "Electron-Induced Single-Event Upsets in Static Random Access Memory," *IEEE Trans. Nucl. Sci,* vol. 60, no. 6, pp. 4122-4129, 2013.

[10] D. M. Fleetwood, et al., "Comparison of enhanced device response and predicted X-ray dose enhancement effects on MOS oxides," *IEEE Trans. Nucl. Sci.,* vol. 35, no. 6, pp. 1265-1271, Dec. 1988.

[11] D. E. Beutler, D. M. Fleetwood, W. Beezhold, D. Knott, L. J. Lorence and B. L. Draper, "Variations in Semiconductor Device Response in a Medium-Energy X-Ray Dose-Enhancing Environment," *IEEE Trans. Nucl. Sci.,* vol. 34, no. 6, pp. 1544-1550, Dec. 1987.

[12] D. M. Long, D. G. Millward and J. Wallace, "Dose Enhancement Effects in Semiconductor Devices,," *IEEE Trans. Nucl. Sci.,* vol. 29, no. 6, pp. 1980-1984, 1982.

[13] G. F. Knoll, Radiation Detection and Measurement, Wiley, 2000.

[14] E. Burke, L. Lowe, D. Snowden, J. Capelli and S. Mittleman, "The Direct Measurement of Dose Enhancement in Gamma Test Facilities," *IEEE Trans. Nucl. Sci.,* vol. 36, no. 6, pp. 1890-1895, 1989.

[15] R. A. Weller, et. al., "Monte Carlo Simulation of Single Event Effects," *IEEE Trans. Nucl. Sci.,* vol. 57, no. 4, pp. 1726-1746, Aug 2010.

[16] R. A. Reed, et al., "Physical Processes and Applications of the Monte Carlo Radiative Energy Deposition (MRED) Code," *IEEE Trans. Nucl. Sci.,* vol. 62, no. 4, pp. 1441-1461, Aug. 2015.

[17] S. Agostinelli, et al., "Geant4 -- A Simulation Toolkit," *Nucl. Instr. Methods Phys.*

Res. Sec. A, vol. 506, no. 3, pp. 250-303, Jul. 2003.

[18] P. M. Gouker, et al., "Radiation effects in 3d integrated SOI SRAM circuits," *IEEE Trans. Nucl. Sci.,* vol. 58, no. 6, pp. 2845-2854, 2011.

[19] M. L. Breeding, R. A. Reed, K. M. Warren and M. L. Alles, "Exploration of the Impact of Physical Integration Schemes on Soft Errors in 3D ICs Using Monte Carlo Simulation," in *IEEE IRPS*, Monterey, CA, 2019.

[20] K. M. Warren, et al., "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, 2005.

[21] K. M. Warren, et al., "Predicting Neutron Induced Soft Error Rates: Evaluation of Accelerated Ground Based Test Methods," in *2008 International Reliability Physics Symposium (IRPS)*, 2008.

[22] A. D. Tipton, et al., "Multiple-Bit Upset In 130 nm CMOS Technology," *IEEE Trans. Nucl. Sci.,* vol. 53, no. 6, pp. 3259-3264, 2006.

[23] K. M. Warren, et al., "Monte-Carlo Based On-Orbit Single Event Upset Rate Prediction for a Radiation Hardened by Design Latch," *IEEE Trans. Nucl. Sci.,* vol. 54, no. 6, pp. 2419-2425, 2007.

[24] M. Hirose, et al., "Three-Dimensional Integrated Circuit with Optically Coupled Shared Memories". Patent US4672577A, 1987.

[25] G. Zhang, M. Graef and F. van Roosmalen, "The Rationale and Paradigm of 'More than Moore'," in *56th Electronic Components and Technology Conference*, 2006.

[26] E. P. DeBenedictis, M. Badaroglu, A. Chen, T. M. Conte and P. Gargini, "Sustaining

Moore's Law with 3D Chips," *IEEE Computer,* vol. 50, no. 8, p. 2017, 2017.

[27] R. R. Tummala, "Moore's Law for Packaging to Replace Moore's Law for ICs," in *2019 Pan Pacific Microelectronics Symposium*, 2019.

[28] C. Yu, "The 3rd Dimension—More Life for Moore's Law," in *2006 International Microsystems, Package, Assembly Conference*, Taiwan, 2006.

[29] E. Beyne, "The 3-D Interconnect Technology Landscape," *IEEE Design Test,* vol. 33, no. 3, pp. 8-20, 2016.

[30] S. K. Lim, "Bringing 3D ICs to Aerospace: Needs for Design Tools and Methodologies," *J. Info. Comm. Convergence Eng.,* vol. 15, no. 2, pp. 117-122, 2017.

[31] Y. Xie, "Processor Architecture Design Using 3D Integration Technology," in *23rd International Conference on VLSI Design*, 2010.

[32] C. Liu and S. K. Lim, "A Design Tradeoff Study With Monolithic 3d Integration," in *13th International Symposium on Quality Electronic Design*, 2012.

[33] B. Mattis, L. Soirez, C. Bullock, D. Martini, S. Jensen, J. Levy, and A. Jones, "Front-side Mid-level Tungsten TSV Integration for High-density 3D Applications," in *2016 IEEE International 3D Systems Integration Conference (3DIC)*, 2016, pp. 1–4.

[34] J.-M. Yook, J. C. Kim, S.-H. Park, J.-I. Ryu, and J. C. Park, "High Density and Low-Cost Silicon Interposer Using Thin-Film and Organic Lamination Processes," in *2012 IEEE 62nd Electronic Components and Technology Conference*, 2012, pp. 274–278.

[35] V. Sukumaran, G. Kumar, K. Ramachandran, Y. Suzuki, K. Demir, Y. Sato, T. Seki, V. Sundaram, and R. R. Tummala, "Design, Fabrication, And Characterization Of Ultrathin 3-D Glass Interposers With Through-Package-Vias At Same Pitch As TSVs In Silicon,"

IEEE Trans. Components, Packaging and Manufacturing Tech., vol. 4, no. 5, pp. 786–795, 2014.

[36] M. Fernndez-Bolaos, W. A. Vitale, M. M. Lpez, A. M. Ionescu, A. Klumpp, R. Merkel, J. Weber, P. Ramm, I. Ocket, W. De Raedt, and A. Enayati, "3D TSV Based High Frequency Components For RF IC And RF MEMS Applications," in *2016 IEEE International 3D Systems Integration Conference (3DIC),* 2016, pp. 1–4.

[37] G. Pars, N. Bresson, S. Minoret, V. Lapras, P. Brianceau, J. F. Lugand, R. Anciant, and N. Sillon, "Through Silicon Via Technology Using Tungsten Metallization," in *2011 IEEE International Conference on IC Design Technology*, 2011, pp. 1–4.

[38] J. Choe, "Comparison of Current 3D NAND Chip and Cell Architecture.", in *Flash Memory Summit,* 2019

[39] D. Chen, E. Wilcox, R. L. Ladbury, C. Seidleck, H. Kim, A. Phan, and K. A. LaBel, "Heavy Ion and Proton-Induced Single Event Upset Characteristics of a 3-D NAND Flash Memory," *IEEE Trans. Nucl. Sci.*, vol. 65, no. 1, pp. 19–26, 2018.

[40] J. Seo, et al., "Highly Reliable m1x MLC NAND Flash Memory Cell With Novel Active Air-Gap And P+ Poly Process Integration Technologies," in 2013 IEEE International Electron Devices Meeting, 2013, pp. 3.6.1–3.6.4.

[41] P. M. Conway, M. J. Gadlage, J. D. Ingalls, A. M. Williams, D. I. Bruce, and D. P. Bossev, "Impact of the Elemental Makeup of an IC in Generating Single-Event Upsets from Low-Energy (<10 MeV) Neutrons: A 3-d NAND Flash Case Study," *IEEE Trans. Nucl. Sci.*, vol. 66, no. 1, pp. 466–473, 2019.

[42] Y. Li, “3d NAND Memory And Its Application In Solid-State Drives: Architecture, Reliability, Flash Management Techniques, And Current Trends,” *IEEE Solid-State Circuits Magazine*, vol. 12, no. 4, pp. 56–65, 2020.

[43] R. N. Advani, “The Business of NAND,” in *3D Flash Memories,* R. Micheloni, Ed. Springer Scientific, 2004, ch. 1, p. 24.

[44] K. Parat and C. Dennison, “A Floating Gate Based 3d NAND Technology With CMOS Under Array,” in 2015 IEEE International Electron Devices Meeting (IEDM), 2015, pp. 3.3.1–3.3.4.

[45] H. Tanaka, et al., “Bit Cost Scalable Technology With Punch And Plug Process For Ultra High Density Flash Memory,” in 2007 IEEE Symposium on VLSI Technology, 2007, pp. 14–15.

[46] M. J. Gadlage, D. I. Bruce, J. D. Ingalls, D. P. Bossev, M. Mckinney, and M. J. Kay, “Directional Dependence of Co-60 Irradiation On The Total Dose Response Of Flash Memories,” *IEEE Trans. Nucl. Sci.*, vol. 66, no. 1, pp. 148–154, 2019.

[47] “SK Hynix Unveils The Industry’s Most Multilayered 176-Layer 4d NAND Flash,” accessed: May 2021, url: https://news.skhynix.com/sk-hynix-unveils-the-industrys-highest- layer-176-layer-4d-nand-flash/.

[48] J. Jang, et al., “Vertical Cell Array Using TCAT (Terabit Cell Array Transistor) Technology For Ultra High Density NAND Flash Memory,” in 2009 Symposium on VLSI Technology, 2009, pp. 192–193.

[49] C. Monzio Compagnoni, A. Goda, A. S. Spinelli, P. Feeley, A. L. Lacaita, and A. Visconti, “Reviewing The Evolution Of The NAND Flash Technology,” *Proc. IEEE*, vol. 105, no. 9, pp. 1609–1633, 2017.

[50] R. Katsumata, et al., “Pipe-Shaped BiCS Flash Memory With 16 Stacked Layers And Multi- Level-Cell Operation For Ultra High Density Storage Devices,” in *2009 Symposium on VLSI Technology*, 2009, pp. 136–137.

[51] M. Bagatin, et al., “Effects Of Heavy-Ion Irradiation On Vertical 3-D NAND Flash Memories,” *IEEE Trans. Nucl. Sci.*, vol. 65, no. 1, pp. 318–325, 2018.

[52] M. Bagatin, S. Gerardin, A. Paccagnella, and S. Beltrami, “Depth Dependence Of Threshold Voltage Shift In 3-D Flash Memories Exposed To X-Rays,” *IEEE Trans. Nucl. Sci.*, vol. 68, no. 5, pp. 659–664, 2021.

[53] M. Bagatin, et al., “Sample-To-Sample Variability And Bit Errors Induced By Total Dose In Advanced NAND Flash Memories,” *IEEE Trans. Nucl. Sci.*, vol. 61, no. 6, pp. 2889–2895, 2014.

[54] E. P. Wilcox and M. J. Campola, “A TID and SEE Characterization Of Multi-Terabit Cots 3d NAND Flash,” in *2019 IEEE Radiation Effects Data Workshop*, 2019, pp. 1–7.

[55] P. Kumari, S. Huang, M. Wasiolek, K. Hattar, and B. Ray, “Layer-Dependent Bit Error Variation In 3-D Nand Flash Under Ionizing Radiation,” *IEEE Trans. Nucl. Sci.*, vol. 67, no. 9, pp. 2021–2027, 2020.

[56] U. Surendranathan, P. Kumari, M. Wasiolek, K. Hattar, T. Boykin, and B. Ray, “Gamma-Ray-Induced Error Pattern Analysis For MLC 3-D NAND Flash Memories,” *IEEE Trans. Nucl. Sci.*, vol. 68, no. 5, pp. 733–739, 2021.

[57] C. Howe, et al., “Role of Heavy-Ion Nuclear Reactions In Determining On-Orbit Single Event Error Rates,” *IEEE Trans. Nucl. Sci.*, vol. 52, no. 6, pp. 2182–2188, 2005.

[58] N. A. Dodds, et al., “Charge Generation By Secondary Particles From Nuclear Reactions In BEOL Materials,” *IEEE Trans. Nucl. Sci.*, vol. 56, no. 6, pp. 3172–3179, 2009.

[59] M. A. Clemens, et al., “The Effect Of High-Z Materials On Proton-Induced Charge Collection,” *IEEE Trans. Nucl. Sci.*, vol. 57, no. 6, pp. 3212–3218, 2010.

[60] M. A. Clemens, et al., “The Effects Of Neutron Energy And High-Z Materials On Single Event Upsets And Multiple Cell Upsets,” *IEEE Trans. Nucl. Sci.*, vol. 58, no. 6, pp. 2591–2598, 2011.

[61] N. Seifert, S. Jahinuzzaman, J. Velamala, R. Ascazubi, N. Patel, B. Gill, J. Basile, and J. Hicks, “Soft Error Rate Improvements In 14-Nm Technology Featuring Second-Generation 3d Tri-Gate Transistors,” *IEEE Trans. Nucl. Sci.*, vol. 62, no. 6, pp. 2570–2577, 2015.

[62] B. D. Sierawski, et al., “Impact Of Low-Energy Proton Induced Upsets On Test Methods And Rate Predictions,” *IEEE Trans. Nucl. Sci.*, vol. 56, no. 6, pp. 3085–3092, 2009.

[63] H. Huh, et al., “13.2 a 1tb 4b/cell 96-stacked-wl 3d NAND Flash Memory With 30mb/S Program Throughput Using Peripheral Circuit Under Memory Cell Array Technique,” in *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*, 2020, pp. 220–221.

[64] F. Irom, D. N. Nguyen, and G. R. Allen, “Single Event Effect And Total Ionizing Dose Results Of Highly Scaled Flash Memories,” in *2013 IEEE Radiation Effects Data Workshop (REDW)*, 2013, pp. 1–4.