UNIVERSITY OF MILANO-BICOCCA

Department of Materials Science

PhD School in

Nanostructures and Nanotechnologies

PhD Thesis

ATMOSPHERIC NEUTRON INDUCED SOFT ERRORS ON ELECTRONIC DEVICES

Supervisor:

Prof. GIUSEPPE GORINI

PhD Thesis of:

ALBERTO FERRARIO

CoSupervisor:

Prof. ALESSANDRO PACCAGNELLA

XXIV CYCLE

Abstract

Soft errors at sea level, originating from scattered particles in the atmosphere or alpha-emitting contaminants in chip materials, are a known source of disturbances in SRAM's and, to a lesser extent, in DRAM's. Relatively less is known about the sensitivity of Floating Gate memories, one of the most pervasive type of memory. An extensive literature covers the effects of heavy ions on floating gate cells, but little data obtained with particles matching the terrestrial neutron environment are available.

The purpose of this thesis is to *investigate atmospheric neutron effects on floating gate cells in NAND Flash memories.*

Experimental data were obtained with neutron irradiation on commercial devices from different vendors. Irradiation was performed at the VESUVIO line of the ISIS facility at the Rutherford Appleton Laboratory, in Didcot, UK, using a wide-energy neutron beam. The VESUVIO neutron spectrum reasonably reproduces the terrestrial environment, with several orders of magnitude of acceleration. A GEANT4-based code has been developed to simulate the irradiation of three technological floating gate cells. With MATLAB data post processing, simulations have provided physical informations about the interactions between incident neutrons and chip materials. Finally, a relation between experimental data and simulations has been studied.

The thesis is organized as follows:

Chapter 1 gives an overview of radiations effects. Different sources of neutrons, like space environments, terrestrial environments, but also man-made radiation facilities are introduced. The second part of Chapter 1 is focused on radiation effects on CMOS technology, with attention on *single event effects.*

Chapter 2 is a brief survey of Flash technology. The Chapter explains the basic principles and operation of non-volatile memories, with attention on the cell

array architecture. The multi level cell memories are also introduced. Then, the reliability problems of industry-standard Flash Cell and scaling issues of the latest Flash structures are presented.

Chapter 3 contains the experimental work of this thesis. In the first part, general published data of ionizing radiation on floating gate cell are given. Hence, experimental data of neutron radiation on commercial Flash memories at the Isis facility are presented, focusing on the *Cross Section* and the *Threshold Voltage* shift.

Chapter 4 presents the simulation work. After an introduction of the general concepts of the Geant4toolkit, the chapter explains the developed simulation tool in terms of *3D Geometry*, *Materials*, *Sensitive Detector* and the simulated *Particles Spectrum*. Data obtained from simulations and analysed, in all their components, are exposed, suggesting a relation with the experimental data.

Appendices gives the reader details of the Geant4 source code implemented for this work. In particular, the descriptions of the *Sensitive Detector* and the *Physic List* are reported.

Sommario

I *Soft errors* in ambiente terrestre, causati da particelle generate negli strati esterni dell'atmospera, o dall'emissione di partcelle alpha dai contaminanti radioattivi dei materiali dei chip, sono una fonte ben nota di disturbi nelle memorie di tipo SRAM e, in misura minore, nelle memorie DRAM. Ancora poco è stato studiato sulla sensibilità delle memorie a Floating Gate, ad oggi uno dei più diffusi tipi di memoria. Mentre un'ampia letteratura copre gli effetti di ioni pesanti nelle celle di tipo Floating Gate, pochi dati sono al momento disponibili per quanto riguarda gli effetti causati da neutroni atmosferici.

Lo scopo di questa tesi è di *studiare gli effetti di neutroni atmosferici sulle celle floating gate di memorie NAND Flash.*

I dati sperimentali sono stati ottenuti irraggiando memorie commerciali da diversi fornitori. Gli irraggiamenti sono stati eseguiti nella linea Vesuvio della sorgente di neutroni Isis (Rutherford Appleton Laboratory), presso Didcot (UK), utilizzando una sorgente di neutroni ad ampio spettro. Lo spettro di neutroni della line Vesuvio riproduce bene lo spettro terrestre con un'accelarazione di diversi ordini di grandezza. Per il lavoro di tesi è stato sviluppato un software basato sul toolkit Geant4 per simulare l'irraggiamento di celle di memoria appartenenti a tre diversi nodi tecnologici. Le simulazioni, insieme all'analisi dei dati con Matlab, hanno fornito informazioni a livello fisico sulle interazioni tra i neutroni incidenti e il materiale dei chip. Con queste informazioni, stato infine studiata una relazione tra i dati sperimentali e le simulazioni.

La tesi è organizzata come segue:

Capitolo 1 offre una panoramica degli effetti da radiazione. Le diverse tipologie di sorgenti di neutroni, come l'ambiente spaziale, l'ambiente terrestre, e anche le sorgenti di radiazione artificali, sono introdotte. La seconda parte del Capitolo 1 è focalizzata sugli effetti da radiazione nella tecnologia CMOS,

con attenzione particolare effetti da evento singolo.

Capitolo 2 è una rassegna della tecnologia Flash. Il capitolo spiega i principi base e il funzionamento delle memorie non volatili, con attenzione all'architettura dell'array di celle. Inoltre, vengono introdotte le memorie con celle a multi livello. Quindi sono presentati i problemi di affidabilit delle celle Flash commerciali, i problemi di scaling e le più recenti strutture Flash.

Capitolo 3 contiene il lavoro sperimentale della tesi. Nella prima parte sono forniti i dati attualmente noti sugli effetti dovuti a radiazioni ionizzanti nelle celle floating gate. Quindi, sono presentati i dati sperimentali ottenuti dagli irraggiamenti con neutroni di memorie Flash commerciali, eseguiti presso i laboratori di Isis, con rilievo per la *Cross Section* e la variazione della *Tensione di Soglia*.

Capitolo 4 presenta il lavoro di simulazione. Dopo un'introduzione sui concetti generali del toolkit Geant4, il capitolo spiega lo sviluppo del software di simulazione in termini di *Geometria 3D*, dei *Materiali*, *Volume Sensibile* e dello *Spetro di Radiazione* simulato. Sono quindi esposti i dati ottenuti dalle simulazioni e analizzati in tutte le loro componenti, indicando una relazione con i dati sperimentali.

Appendici danno dettagli al lettore sul codice sorgente basato su Geant4 implementato in questa tesi. In particolare, sono riportati la descrizione del *Sensitive Detector* e della *Physic List*.

Contents

List of Figures

List of Tables

Nomenclature

CMOS Complementary metal oxide semiconductor is a technology for constructing integrated circuits.

COLLECTED CHARGE The charge collected by a particular device node during the passage of a particle.

CRITICAL CHARGE The minimum amount of collected charge that will cause a device node to change state.

DIFFERENTIAL FLUX The time rate of fluence per unit energy, the rate of the quantity of radiation, particle fluence, per unit area incident on a surface per unit energy.

DUT Device under test.

ECC Error correction code, sometimes called error detection and correction (EDAC).

FG Floating gate of a Flash cell. Often indicates the cell itself.

FITs Failures in time; the number of failures per 10^9 device-hour.

FLUENCE The total amount of particle radiant energy incident on a surface in a given period of ime, divided by the area of the surface. The fluence is usually expressed in particles per unit area.

FLUX The time rate of flow of particle radiant energy incident on a surface, divided by the area of that surface.

HARD ERROR An irreversible change in operation that is typically associated with permanent damage to one or more elements of a device or circuit (e.g., gate oxide rupture, destructive latch-up events).

LET Linear Energy Transfer.

MLC Multi level floating gate cell.

MULTIPLE-BIT UPSET (MBU) A multiple-cell upset in which two or more error bits occur in the same word.

MULTIPLE-CELL UPSET (MCU) A single event that induces several bits in an IC to fail at one time.

ONO Oxide-Nitride-Oxide Structure of the FG cell.

RADIATION Energy emitted in the form of electromagnetic waves or moving nuclear particles.

SENSITIVE VOLUME A region, or multiple regions, containing nodes whose states can be changed by incident radiation.

SINGLE EVENT TRANSIENT (SET) A momentary voltage excursion (voltage spike) at a node in an integrated circuit caused by a single energetic particle strike.

SINGLE-EVENT EFFECT (SEE) Any measurable or observable change in state or performance of a microelectronic device, component, subsystem, or system (digital or analog) resulting from a single energetic particle strike.
NOTE Single-event effects include single-event upset (SEU), multiple-bit upset (MBU), multiple-cell upset (MCU), single-event functional interrupt (SEFI), single-event latch-up (SEL), single-event hard error (SHE) and single-event transient (SET), single-event burnout (SEB), and single-event gate rupture (SEGR).

SINGLE-EVENT FUNCTIONAL INTERRUPT (SEFI) A soft error that causes the component to reset, lock-up, or otherwise malfunction in a detectable way, but does not require power cycling of the device (off and back on) to restore operability, unlike single-event latch-up (SEL), or result in permanent damage as in singleevent burnout (SEB).

SINGLE-EVENT LATCH-UP (SEL) An abnormal high-current state in a device caused by the passage of a single energetic particle through sensitive regions of the device structure and resulting in the loss of device functionality.

SINGLE-EVENT UPSET (SEU) CROSS-SECTION the number of events per unit fluence. For device SEU crosssection,the dimensions are area per device. For bit SEU cross-section, the dimensions are area per bit.

SINGLE-EVENT UPSET (SEU) RATE the rate at which single event upsets occur.

SINGLE-EVENT UPSET (SEU) A soft error caused by the transient signal induced by a single energetic particle strike.

SLC Single level floating gate cell.

SOFT ERROR RATE (SER) The rate at which soft errors occur.

SRAM Static rando access memory.

TID Total ionizing dose.

Chapter 1

Introduction to ionizing radiation effects

Electronics chips operating at sea level are constantly bombarded by a shower of high- and low-energy neutrons, which originate from the interactions of cosmic rays with the outer layers of the atmosphere. The neutron flux changes with altitude, reaching a peak very close to the cruise altitude of airplanes, posing an even more serious threat to avionics. In addition, inevitable radioactive contaminants in the package and solder materials emit alpha particles, which may reach sensitive device areas and produce errors. Spacecraft and satellite electronics must operate reliably in a much harsher environment, characterized by a significant presence of ionizing radiation, in the form of protons, electrons, and heavy-ions coming from various sources.

Ionizing radiation can cause either permanent or temporary damage to electronic chips, generating a plethora of effects, from flipping an SRAM memory bit from one to zero or vice versa, to burning-out a power MOSFET.

NAND Flash memories are not immune to radiation effects. On the contrary, due to their complexity and large number of diverse building blocks, they exhibit quite complex failure signatures when exposed to ionizing particles. This is especially true for the space environment, where mitigation strategies are mandatory due to the severity of the effects. Yet, ionizing radiation issues are becoming a growing concern also at ground level, due to the ever-decreasing feature size, which is making floating gate cells sensitive even to small external disturbances, such as those caused by atmospheric neutrons and alpha particles. Sensitivity to radiation is related both to the floating

gate cells, and the peripheral circuitry.

This chapter will give a brief overview of radiation effects. Initially, differents sources of neutrons are introduced, like space environments, terrestrial environments, but also man-made radiation facilities. Then the effects of radiation on electronic devices are investigated, focusing on CMOS technology, with attention on both long-term effects and single event effects.

1.1 Environments

Radiation environments are encountered in military applications, nuclear power stations, nuclear waste disposal sites, high-altitude avionics, medical and space applications. Radiation type, energy, dose rate and total dose may be very different in each of these application areas and require in many cases radiation-tolerant electronic systems.

In this section, the main sources of radiation are exposed. First of all, space is analized, as it is a very hard environment, from a radiation point of view, having serious effects on electronic behaviour of spacecrafts and satellites. The, the terrestrial one, which can have important and unexpected effects in everyday electronic devices. Finally, the high energy physics experiments, which are man-made radiation sources, are briefly introduced.

1.1.1 Space

Our planet is surrounded by a radiation rich environment, consisting of mainly energetic charged particles (electrons, protons, heavy ions). They can either be trapped particles, bound to trajectories dictated by the earths magnetic field, or free, transiting particles originating from the sun or from galactic sources. The particles associated with ionizing radiation are categorized into three main groups relating to the source of the radiation:

i trapped radiation belt particles (Van Allen Belts);

ii solar flare particles;

iii cosmic rays.

1.1.1.1 Van Allen Belts

The natural space environments in which most of the satellites operate is an orbit ranging in altitudes from low earth orbits ($150 - 600$ km) to geosynchronous orbits (roughly 35.880 km). Most of the particles in interplanetary space come from the sun in the form of a hot ionized gas called the solar wind; it flows radially from the sun with a speed that in proximity of the Earth varies from about 300 to 1000 km/s, and represents a solar mass loss of about 10^{14} kg per day.

The radiation environment of greatest interest is the near earth region, about $1 - 12$ earth radii R_e (where $R_e = 6380$ km), which is mainly dominated by electrically charged particles trapped in the earth's magnetosphere, and to a lesser extent by the heavy ions from cosmic rays (solar and galactic) [6]. As the earth sweeps through the solar wind, a geomagnetic cavityis formed by the earth's magnetic field (Figure 1.1[1]).

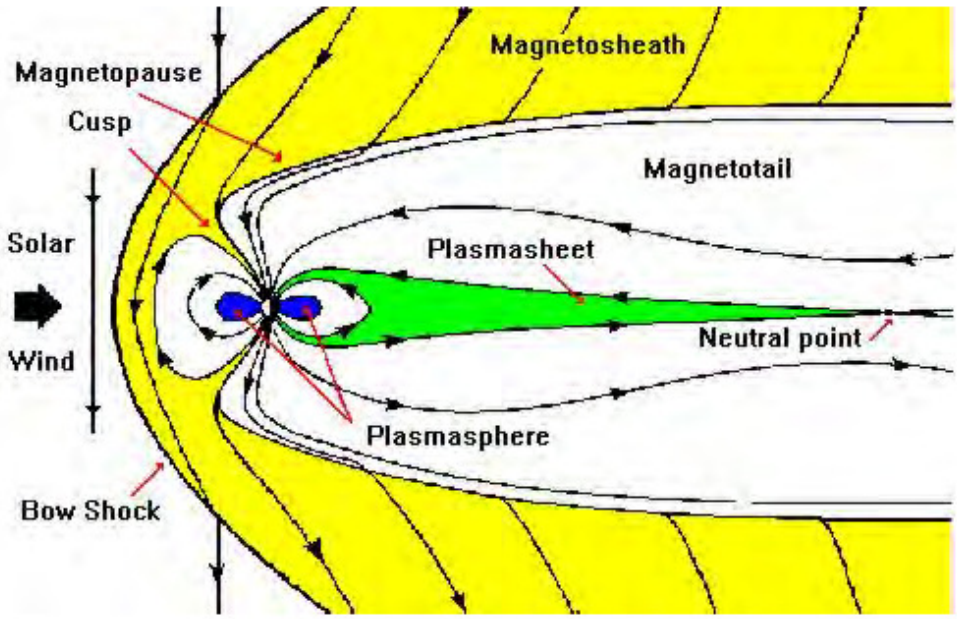


Figure 1.1: Interactions between Earth magnetosphere and the solar wind.

The motion of the trapped charge particles is complex, as they gyrate and bounce along the magnetic field lines, and are reflected back and forth between the pairs of conjugate mirror points (regions of maximum magnetic field strength along their trajectories) in the opposite hemispheres. Also, because of the charge, the electrons drift in an easterly direction around earth, whereas protons and heavy ions drift westward. Interplanetary space probes such as the Voyager (and

[1] http://helios.gsfc.nasa.gov/magnet.html

Galileo to Jupiter) encounter ionizing particles trapped in the magnetosphere of other planets, as well as the solar flares and heavy ions from cosmic rays.

Electrons in the earth's magnetosphere have energies ranging from low keV to about 7 MeV, and are trapped in the roughly toroidal region which is centered on the geomagnetic equator and extends to about 1–12 earth radii. These trapped electrons are differentiated by "inner zone" (< 5 MeV) and "outer zone" (~ 7 MeV) electron populations. The trapped protons originating mostly from the solar and galactic cosmic rays have energies ranging from a few MeV to about 800 MeV. They occupy generally the same region as the electrons, although the region of highest proton flux for energies $E_{\mathrm{p}} > 30$ MeV is concentrated in a relatively small area at roughly $1.5R_{\mathrm{e}}$. The actual electron and proton flux encountered by a satellite is strongly dependent upon the orbital parameters, mission launch time, and duration. Electrons and protons from the trapped radiation belts on interacting with spacecraft materials produce secondary radiation (e.g., "bremsstrahlung" or braking radiation from the deceleration of electrons). This secondary radiation can extend the penetration range of primary radiation and lead to an increase in dose deposition. Incident electron and proton fluxes are typically calculated from the trapped radiation environmental models developed by the U.S. National Space Sciences Data Center (NSSDC). The trapped particle fluxes responding to changes in the geomagnetic field induced by the solar activity exhibit dynamic behavior.

1.1.1.2 Solar flares

A solar flare is an enormous explosion in the solar atmosphere. It results in sudden bursts of particle acceleration, heating of plasma to tens of millions of degrees, and the eruption of large amounts of solar mass. Flares are believed to result from the abrupt release of the energy stored in magnetic fields in the zone around sunspots. There are two types of flares: impulsive and gradual. Impulsive flares accelerate mostly electrons, with some protons. They last minutes or hours and the majority appear near the solar equator. Impulsive flares occur at a rate of about 1000 per year during solar maximum [55, 35].

Gradual flares accelerate electrons, protons, and heavy ions to near the speed of light, and the events tend to last for days. They occur mainly near the poles of the Sun and happen about 100 times per year.

This acceleration of solar flare particles to extremely high energies involves all

the different elements in the solar atmosphere. Ions of elements such as carbon, nitrogen, oxygen, neon, magnesium, silicon, and iron, excited in this way, end up in solar cosmic rays, also called solar energetic particles (SEPs).

A small percentage of solar flares are accompanied by the ejection of significant numbers of protons. Solar proton events occur sporadically, but are most likely near solar maximum. Events may last hours or up to more than a week, but typically the effects last 2 to 3 days. Solar protons add to the total dose and may also cause single event effects in some cases.

The high-energy particles comprising the radiation environment can travel through spacecraft material and deposit kinetic energy. This process causes atomic displacement or leaves a stream of charged atoms in the incident particles wake. Spacecraft damage includes decreased power production by solar arrays, failure of sensitive electronics, increased background noise in sensors, and radiation exposure of the spacecraft crew.

1.1.1.3 Cosmic rays

Cosmic rays are high energy charged particles, originating in outer space, that travel at nearly the speed of light and strike the Earth from all directions. Most cosmic rays are the nuclei of atoms, ranging from the lightest to the heaviest elements in the periodic table. Cosmic rays also include high energy electrons, positrons, and other subatomic particles. The term *cosmic rays* usually refers to galactic cosmic rays, which originate in sources outside the solar system, distributed throughout our Milky Way galaxy. However, this term has also come to include other classes of energetic particles in space, including nuclei and electrons accelerated in association with energetic events on the Sun (called solar energetic particles), and particles accelerated in interplanetary space.

Cosmic rays were discovered in 1912 by Victor Hess, when he found that an electroscope discharged more rapidly as he ascended in a balloon. He attributed this to a source of radiation entering the atmosphere from above, and in 1936 was awarded the Nobel prize for his discovery. For some time it was believed that the radiation was electromagnetic in nature (hence the name cosmic "rays"), and some textbooks still incorrectly include cosmic rays as part of the electromagnetic spectrum. However, during the 1930's it was found that cosmic rays must be electrically charged because they are affected by the Earth's magnetic field [72, 34].

The energy of cosmic rays is usually measured in units of MeV, for mega-electron volts, or GeV, for giga-electron volts. (One electron volt is the energy gained when an electron is accelerated through a potential difference of 1 V). Most galactic cosmic rays have energies between 100 MeV (corresponding to a velocity for protons of 43% of the speed of light) and 10 GeV (corresponding to 99.6% of the speed of light). The number of cosmic rays with energies beyond 1 GeV decreases by about a factor of 50 for every factor of 10 increase in energy. Over a wide energy range the number of particles per m^2 per steradian per second with energy greater than E (measured in GeV) is given approximately by $N(> E) = k(E + 1)^{-a}$, where $k \sim 5000$ per m^2 per steradian per second and $a \sim 1.6$. The highest energy cosmic rays measured to date have had more than 10^{20} eV, equivalent to the kinetic energy of a baseball traveling at approximately 100 mph [34].

It is believed that most galactic cosmic rays derive their energy from supernova explosions, which occur approximately once every 50 years in our Galaxy. To maintain the observed intensity of cosmic rays over millions of years requires that a few percent of the more than 10^{51} ergs released in a typical supernova explosion be converted to cosmic rays. There is considerable evidence that cosmic rays are accelerated as the shock waves from these explosions travel through the surrounding interstellar gas. The energy contributed to the Galaxy by cosmic rays (about $1\,\mathrm{eV} \cdot \mathrm{cm}^{-3}$) is about equal to that contained in galactic magnetic fields, and in the thermal energy of the gas that pervades the space between the stars.

Cosmic rays include essentially all of the elements in the periodic table; about 89% of the nuclei are hydrogen (protons), 10% helium, and about 1% heavier elements. The common heavier elements (such as carbon, oxygen, magnesium, silicon, and iron) are present in about the same relative abundances as in the solar system, but there are important differences in elemental and isotopic composition that provide information on the origin and history of galactic cosmic rays. For example there is a significant overabundance of the rare elements Li, Be, and B produced when heavier cosmic rays such as carbon, nitrogen, and oxygen fragment into lighter nuclei during collisions with the interstellar gas. The isotope ^{22}Ne is also overabundant, showing that the nucleosynthesis of cosmic rays and solar system material have differed. Electrons constitute about 1% of galactic cosmic rays. It is not known why electrons are apparently less efficiently accelerated than nuclei.

Because cosmic rays are electrically charged they are deflected by magnetic

fields, and their directions have been randomized, making it impossible to tell where they originated. However, cosmic rays in other regions of the Galaxy can be traced by the electromagnetic radiation they produce. Supernova remnants such as the Crab Nebula are known to be a source of cosmic rays from the radio synchrotron radiation emitted by cosmic ray electrons spiraling in the magnetic fields of the remnant. In addition, observations of high energy (10 MeV - 1000 MeV) gamma rays resulting from cosmic ray collisions with interstellar gas show that most cosmic rays are confined to the disk of the Galaxy, presumably by its magnetic field [34]. Similar collisions of cosmic ray nuclei produce lighter nuclear fragments, including radioactive isotopes such as ^{10}Be, which has a half-life of 1.6 million years. The measured amount of ^{10}Be in cosmic rays implies that, on average, cosmic rays spend about 10 million years in the Galaxy before escaping into inter-galactic space.

Just as cosmic rays are deflected by the magnetic fields in interstellar space, they are also affected by the interplanetary magnetic field embedded in the solar wind (the plasma of ions and electrons blowing from the solar corona at about $400\,\mathrm{km}\cdot\mathrm{s}^{-1}$), and therefore have difficulty reaching the inner solar system. Spacecraft venturing out towards the boundary of the solar system they have found that the intensity of galactic comic rays increases with distance from the Sun. As solar activity varies over the 11 year solar cycle the intensity of cosmic rays at Earth also varies, in anti-correlation with the sunspot number.

The Sun is also a sporadic source of cosmic ray nuclei and electrons that are accelerated by shock waves traveling through the corona, and by magnetic energy released in solar flares. During such occurrences the intensity of energetic particles in space can increase by a factor of 10^2 to 10^6 for hours to days. Such solar particle events are much more frequent during the active phase of the solar cycle. The maximum energy reached in solar particle events is typically 10 to 100 MeV, occasionally reaching 1 GeV (roughly once a year) to 10 GeV (roughly once a decade). Solar energetic particles can be used to measure the elemental and isotopic composition of the Sun, thereby complementing spectroscopic studies of solar material.

A third component of cosmic rays, comprised of only those elements that are difficult to ionize, including He, N, O, Ne, and Ar, was given the name *anomalous cosmic rays* because of its unusual composition. Anomalous cosmic rays originate from electrically-neutral interstellar particles that have entered the solar system unaffected by the magnetic field of the solar wind, been ionized, and

then accelerated at the shock wave formed when the solar wind slows as a result of plowing into the interstellar gas, presently thought to occur somewhere between 75 and 100 AU from the Sun (one AU is the distance from the Sun to the Earth).

1.1.2 Terrestrial environments

Neutrons, originating from the interaction of cosmic rays with the outer layers of the atmosphere, and alpha particles, emitted by radioactive contaminants in the package and solder materials are the main sources of radiation-induced effects at sea level.

As cosmic rays penetrate in the atmosphere they interact with Nitrogen and Oxygen atoms, generating a cascade of secondary particles, as illustrated in Figure 1.2. Several reactions are possible, giving rise to many different particles (protons, pions, muons, neutrons) and an electromagnetic component. In turn the generated particles can have enough energy to create even more particles. As a result, as cosmic rays penetrate into the atmosphere, the number of particles initially increases and then decreases, when the shielding effect of the atmosphere becomes dominant.

From the standpoint of electronics, atmospheric neutrons are the most important particles, because, even though they come in a smaller number than muons, they are able to trigger nuclear reactions inside the chips, giving rise to charged secondary byproducts (something which muons cannot do), which deposit charge and disturb the operation of electronic chips.

The neutron flux has some peculiar characteristics. It increases with altitude, peaking (more than two orders of magnitude larger than at sea level) very close to the altitudes of commercial airplanes. As a result, avionics is much more threatened by atmospheric neutrons, than electronics operating at sea level. The energy spectrum is not constant, but features a characteristic 1/Energy dependence. The energy range of interest for radiation effects includes thermal neutrons (around 25 meV) and neutrons with energy above about 10 MeV. Thermal neutrons are important, because of their large cross section of interaction with Boron 10. ^{10}B is contained in Borophosphosilicate layers for inter-metal isolation [6, 37].

Many semiconductor manufactures (but not all) have eliminated ^{10}B, because of concern over radiation effects. High-energy neutrons (> 10 MeV) are of concern, because they are capable of triggering nuclear reactions, for example with Silicon

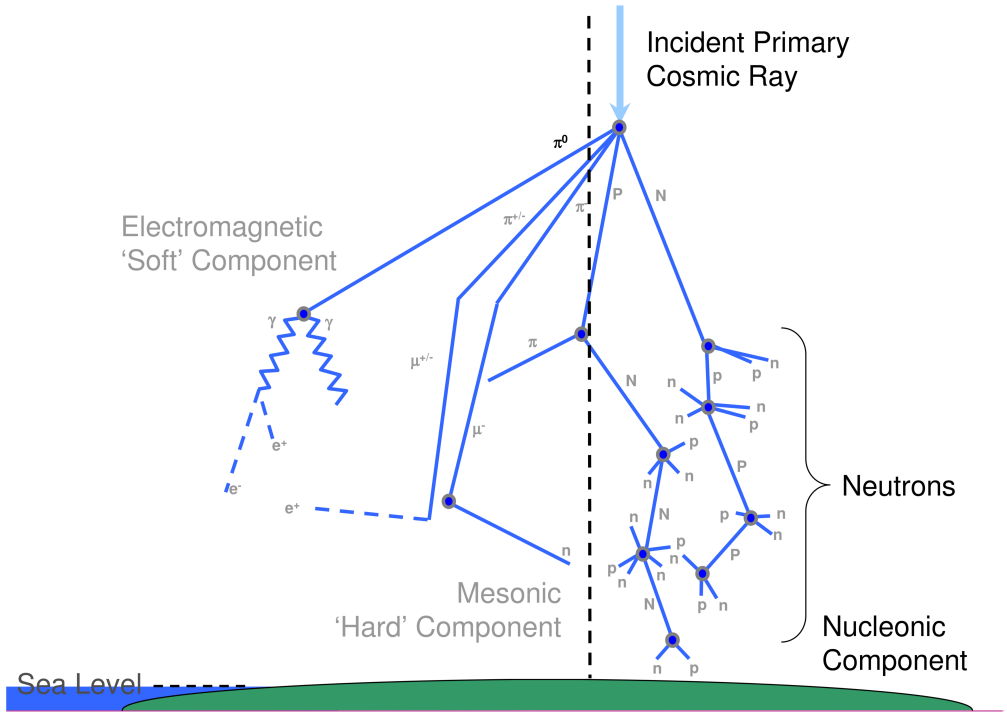


Figure 1.2: Cascade of byproducts produced by cosmic-ray interaction. Generated particles include protons (p), neutrons (n), pions (π), electrons (e), muons (μ) and photons (γ) [45].

atoms, giving rise to charge secondary byproducts. The neutron flux changes also with latitude, atmospheric pressure, and solar activity. As a reference, the atmospheric neutron flux (thermal and high-energy components) at New York City is about 14 neutrons/cm^2/hour.

Radioactive contaminants inside chip materials are the second important source of radiation effects at sea-level. Their effects tend to be less important at high altitudes (where the neutron flux is much higher, and therefore dominates the error rate), but can be responsible for a large part of radiation-induced errors on the ground, especially on deeply scaled technologies. ^{238}U, ^{234}U, ^{232}Th, ^{190}Pt, ^{144}Nd, ^{152}Gd, ^{148}Sm, ^{187}Re, ^{186}Os, and ^{174}Hf are the naturally occurring alpha emitters (i.e. the part of alpha emitters with sufficiently long half lives, which is defined as the time needed to reduce the emission by a factor 2). Some of them can be used as materials for the fabrication of integrated circuits. Others, such as Uranium and Thorium are unwanted impurities, that unfortunately find their way in raw materials and process flows. They are actually at the beginning of a chain of alpha decays [51], which end up with Pb. Typical emission levels in a chip are on the order of 10^{-3} alphas/cm^2/h.

1.1.3 Artificial radiation environments

On Earth, there are some specific man-made environments, which have very high requirements in terms of radiation hardness. High-energy physics experiments, such as the large hadron collider at CERN, Geneve, require the detector read-out electronics to work very close to the beam, where high fluences of fast hadrons are present. Because of this, a dedicated library of rad-hard components has been developed, where radiation hardness has been attained by using a standard CMOS process, with some layout modifications, that imposes a speed and area penalty, but afforded a considerable increase in the tolerable radiation levels.

Nuclear fission power plants also require the use of radiation-hard electronics in certain applications, where exposure to radiation is inevitable. Although it is unlikely that electronic devices are required in the reactor core, controllers and imager may be placed within the bio-shield [6].

Future nuclear fusion power plants are also expected to have issues with radiation, due, for instance, to fast neutrons coming from Deuterium-Tritium reactions.

1.1.3.1 ITER

Radiation effects on components and materials will be one of the most serious technological issues in fusion systems realizing burning plasmas. Especially, diagnostic components, which should play crucial roles to control plasmas and to understand physics of burning plasmas, will be exposed to high-flux neutrons and gamma-rays.

The International Thermonuclear Experimental Reactor (ITER)[2] is a large-scale scientific experiment intended to prove the viability of fusion as an energy source, and to collect the data necessary for the design and subsequent operation of the first electricity-producing fusion power plant. ITER is the first theater, where diagnostic components will be exposed to intense irradiation environments associated with high-flux high-energy neutrons. Radiation effects will influence performance of diagnostic components substantially at the onset of fusion nuclear reactions, and successful control and operation of burning plasmas will strongly depend on development of radiation-hardened diagnostic components, and quantitative understandings of radiation effects [78].

The Japan Materials Testing Reactor (JMTR), having a neutron fluxes and gamma-ray dose rates, similar to those expected near burning plasma regions in ITER, is currently used to test diagnostic components. Several radiation effects, such as radiation induced electrical conductivity (RIC) and radiation induced electromotive force (RIEMF) have been demonstrated. For example, some results showed that effects of the RIEMF may cause serious problems in magnetic-field measurements for a long plasma discharge duration, because it will generate a substantial drift voltage in some occasions.

JMTR results indicate that the ITER-relevant diagnostic components could be developed in time, though there are still some radiation effects issues to overcome.

1.2 Overview of radiation effects

1.2.1 Radiation-matter interaction

Ionizing particles (electrons, protons, heavy ions) impinging on a semiconductor or insulating material lose energy as they are stopped by the target material.

[2]For more informations see: http://www.iter.org/mach.

Different energy loss mechanisms exist: Coulomb interactions, either with electrons or with nuclei, and nuclear reactions. As a result, electron-hole pairs are created (ionizing component) and lattice atoms are displaced from their lattice position (non-ionizing loss). The ionizing loss is by far the most important in CMOS circuits and will be treated here in more detail. It can be characterized by the *Linear Energy Transfer* (LET), which is the energy deposited per unit length by the impinging particle in the target material, through ionization processes. 3.6 eV are needed to produce an $e-h$ pair in Silicon, while 17 eV to 18 eV are required for an $e-h$ pair in Silicon Dioxide [7]. The initially liberated electrons can further ionize the material as long as they possess the necessary energy. The LET is customarily normalized by the density of the target material, and is measured in units of

$$Linear\ Energy\ Transfer = [\mathrm{MeV} \cdot \mathrm{cm}^{-2} \cdot \mathrm{mg}^{-1}\,].$$

The LET depends on the particle energy, displaying a maximum for intermediate energies, called the Bragg peak.

Multiplying the LET by the particle fluence of a monochromatic beam and by a proper constant, one obtains the total amount of energy deposited per unit of mass through ionization, which is called *Total Ionizing Dose* (TID), and is measured in rad, 1 rad = 100 erg/g or Gray, 1 Gy = 1 J/kg in SI units. LET and TID are the fundamental metrics in the study of radiation effects in CMOS circuits.

A large part of the produced electron-hole pairs quickly recombine, both in semiconductors and insulators, soon after generation. The percentage of carriers that survive recombination, called charge yield, depends on the characteristics of the impinging particle, on the target material, and on the magnitude of the local electric field [31] [7]. The charge yield decreases for increasing LET, because the pairs are more dense and escape mutual recombination with more difficulty with higher LET. On the contrary increasing the electric field separates electrons from holes more efficiently and reduce recombination.

Neutral particles such as neutrons, or lowly-charged penetrating ones such as protons, can trigger nuclear reactions, generating secondary charged by-products (heavy ions), thus indirectly ionizing the target material.

Energetic heavy particles ($Z \geq 2$) follow a straight trajectory in matter, being only rarely deflected. For a given energy and species, the distance travelled in the

target material is well defined and almost constant for nominally identical ions, and is called range. The release of energy associated with the passage of an ion gives rise to a cylindrical electron-hole pair track, whose size is on the order of tens of nanometers in Silicon, and considerably less in Silicon dioxide [60]. This is something to remember when deep-submicron technologies, whose feature size is on the same order of magnitude, are studied.

For heavy energetic particles, the loss of energy occurs mostly through ionization, but as the particles slow down, the contribution of non-ionizing processes becomes increasingly important, and it is the dominant contribution near the end of the particle range (i.e., when the particle is about to stop).

Light particles, such as electrons, experience frequent scattering events and describe a zig-zag trajectory. No range can be defined for such particles.

Photons are not a concern in space, but can be used for x-ray PCB inspections at significant levels (doses used at airports check points are very low), and in addition γ- and x-rays sources are routinely used for ground testing of electronic components. Photons interact in different ways with matter depending on their energy. For low energy photons (such as those produced by 10 keV x-rays; a standard test source) the photoelectric effect is the dominant mechanism. For higher-energy photons (such as those produced by ^{60}Co gamma rays; another standard test source), Compton scattering is the dominant mechanism. As a result, the impact of photons, or some peculiar phenomena such as dose-enhancement effects [31], are strongly influenced by their energy. However, regardless of the mechanism, secondary electrons are produced at all energies, which are responsible for a large part of the energy deposition.

1.2.2 Categories of radiation effects

The electron-hole pairs generated by the interaction of radiation and chip materials may induce a variety of effects. They cause charge trapping and defect generation in insulators. They generate spurious currents in reverse-biased pn junctions, which may upset memory elements or give rise to voltage transients in combinational logic and analog circuits. As we will see, in floating gate cells they generate discharge currents that may corrupt the stored information.

Depending on the type and characteristics of the impinging radiation, both irreversible (i.e., causing permanent damage to the exposed device) and reversible

(loss of information) effects may arise.

In some cases, a single particle, with highly ionizing power, deposits enough charge to cause malfunctions, which range from the corruption of the information stored in a single memory bit, to the burn-out of power MOSFETs. This type of events are called *single event effects* (SEE). SEE's affect memory, combinatorial, and analog circuits with destructive or non-destructive effects.

In other cases, damage builds up strike after strike, during a prolonged exposure to ionizing radiation, that generates drifts in components parameters (such as threshold voltage shift, power consumption increase, gain decrease, etc.). This second type of effects can involve both damage to dielectrics layers (gate oxide, lateral isolation, passivation layers, etc.), resulting in so-called *total ionizing dose effects* (TID), and displacement damage (DD) to bulk semiconductor materials. TID is responsible for parametric degradation and functional failure of MOSFETs (e.g., threshold voltage shift) and BJTs (e.g., current gain decrease). On the contrary, Displacement Damage (DD) effects are due to the generation of point or extended defects in the crystalline semiconductor lattice, degrading devices which rely on bulk semiconductor conduction, such as BJTs, solar cells, etc., but not MOSFETs. Annealing effects can take place, even at room temperature, causing total ionizing dose and displacement damage effects to change over time, resulting in more or less severe degradation.

Low-LET particles, such as electrons, generate TID and DD. Protons give rise to TID, DD, and SEE. Protons SEE's are usually indirectly generated, i.e., a charged secondary product, coming from a proton nuclear reaction is responsible for the SEE. Recently, bit-flips resulting from proton direct ionization (without the generation of a secondary particle) have been observed in very scaled SRAMs as well [69]. Alpha particles (He ions) cause TID and SEE. Heavier ions ($Z \geq 2$) generate SEE's and microdose effects, i.e., effects similar to TID, but localized to the ion track. Given the fluxes and types of present particles (see radiation environments on Section 1.1), TID and DD effects occur primarily in space, HEP experiments, and nuclear power plants, whereas SEE's take place also on the ground. TID effects can degrade circuit parameters also when high doses of radiation are used for inspection purposes.

1.3 Total ionizing dose effects

Total Ionizing Dose effects in CMOS devices are due to charge trapping and defect generation in insulating layers. Historically, the main problems of MOSFETs with respect to radiation were shifts in the threshold voltage and degradation in trasconductance. Thanks to the thinning of the gate oxide, radiation effects in the gate stack are now much less severe. Nevertheless, damage may still occur in the thick isolation oxides, which surround the conductive channel.

1.3.1 Basic mechanisms

Figure 1.3 depicts the fundamental mechanisms underlying charge trapping and defect generation in an oxide layer. Radiation generates defects in insulating layers through indirect processes, i.e., by releasing species, such as holes and hydrogen ions, which in turn are responsible for the radiation response of the exposed devices. The first step in the process is the generation of energetic electron-hole pairs. After thermalization, the electrons, which have a higher mobility, are quickly swept towards the anode by the applied bias, leaving the heavier and slower holes to move inside the oxide in the opposite direction. But before they do that, a large part of the $e - h$ pairs will recombine (charge yield).

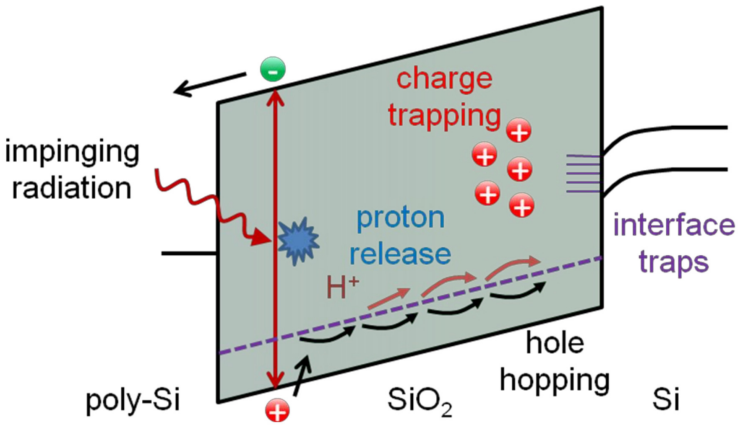


Figure 1.3: Basic Mechanisms of Total Ionizing Dose Effects.

The surviving holes may be trapped in pre-existing deep traps while migrating towards the cathode under the influence of the applied field. The transport of holes occurs by hopping through localized states, in a rather complicated and

dispersive way (i.e., occurring over many decades in time). Holes arriving at the Si/SiO_2 interface, where the density of defect sites is higher, can be trapped. These trapping sites have been related to the E' center, a trivalent silicon associated to an oxygen vacancy [49]. Vacancies are related to the out-diffusion of oxygen in the oxide and lattice mismatch at the surface. The amount of trapped charge depends on the number of holes that survive recombination, on the number of O vacancies, and on the field-dependent capture cross-section of the traps [49] . It is also very dependent on the quality of the oxide, with "hardened" ones showing orders of magnitude less radiation-induced charge trapping than "soft" oxides. Processing conditions strongly influence oxide hardness [49]: high temperature anneals, for instance, can increase the number of oxygen vacancies. Increasing the amount of hydrogen during processing also decreases oxide hardness, as is discussed below.

During the transport and trapping of holes, hydrogen ions (protons) are likely released. Hydrogen ions arriving at the interface can generate interface traps [71], by reacting with hydrogen-passivated dangling bonds at the interface trap. Interface traps may readily exchange carriers with the channel, and are full or empty depending on the position of the relevant quasi Fermi level. Interface states are usually donor (positive when empty, neutral when charged) when their energy position is above midgap or acceptor (neutral when empty, negative when charged) when below [74]. The creation of interface traps is much slower (up to thousands of seconds) than the build-up of trapped charge. Measurements have shown that interface traps are related to Pb centers, a trivalent center at the Si/SiO_2 interface bonded to three Si atoms with a dangling orbital perpendicular to the interface.

Annealing of charge in oxide-traps may start immediately and occurs due to tunneling or thermal processes [60]. Indeed, the trapped charge can be neutralized by electrons, either being thermally excited from the valence band [74], or tunneling through the oxide barrier [60, 74]. In the first case, annealing increases with higher temperatures and for shallower traps; in the second case, annealing depends on the tunneling distance and trap energy spatial position. In contrast, high temperatures are needed to recover interface traps [60]. As a result, in low-dose rate environments (such as space), interface traps may play a predominant role.

1.3.2 Scaling effects

TID issues in CMOS devices have been strongly mitigated by scaling, and the associated thinning of the gate oxide. A simple model (valid for low doses

and relatively thick oxides) shows that the amount of radiation-induced threshold voltage shift quadratically decreases for decreasing oxide thickness [71]. This decrease is even sharper for ultra-thin gate oxides ($< 10\,\mathrm{nm}$), due to the short tunneling distance from the gate and the channel to the oxide bulk [71]. State-of-the-art silicon oxides for low-voltage MOSFETs are only 1–2 nanometers thick, and neither the build up of oxide trap charge, nor interface trap generation is significant even after high radiation doses.

SOI devices, once confined to niche applications (e.g, in the rad-hard arena, due to their low sensitive volume to single event effects and the inherent suppression of single event latchup), have become mainstream in the microelectronic industry in their partially depleted version. In this case, the presence of a thick Buried Oxide (BOX) is detrimental for the TID response [71]. Indeed, charge trapping in the BOX leads to leakage currents in partially depleted devices, and alterations of the front gate characteristics in fully depleted MOSFETs, where coupling exists between the front and the back channels.

1.4 Single event effects

A single ionizing particle crossing a sensitive device area can trigger a spurious response in the circuit or damage it. These events are classified as *soft errors*, if they do not induce any physical damage, but only loss of information (e.g., a bit flip in a memory array); or as *hard errors* if they do induce permanent damage (e.g., the gate oxide rupture following the strike of a heavy ion). Some other effects may be destructive or not depending on the intervention of protection structures, such as the latch-up induced by ionizing particles. The most significant SEE's are listed below:

Non-destructive (soft) effects

Single Event Upset, SEU the corruption of a single bit in a memory array.

Multiple Bit Upset, MBU the corruption of multiple bits due to a single particle.

Single Event Transient, SET a transient induced by an ionizing particle in a combinatorial or analog part of a circuit.

Single Event Functional Interrupt, SEFI the corruption in the controlling

state machine of a chip, leading to functional interruption.

Destructive (hard) effects

Single Event Gate Rupture, SEGR rupture of the gate oxide occurring especially in power MOSFETs.

Single Event Burnout, SEB burnout of a power device (IGBT, MOSFETs, etc.).

Effects that may be destructive or not

Single Event Latch-up, SEL the activation of parasitic bipolar structures, leading to a sudden increase in the supply current.

Single Event Snapback, SES a regenerative feedback mechanism sustained by impact ionization in SOI devices.

The cross section is used to express the sensitivity of a chip to a particular type of event, and is defined as

$$\sigma_{SEE} = \frac{number\ of\ events}{particle\ fluence}$$

The cross section increases for increasing LET, and the cross section versus LET curve is usually fitted with a Weibull cumulative probability distribution. Of great importance are the threshold LET, i.e., the minimum LET required to experimentally observe an event, and the saturation LET, i.e., the LET above which the cross section does not increase anymore, or at least is considered to saturate.

The concept of critical charge is often introduced in the study of single event effects. The critical charge is the minimum charge that must be deposited to trigger a certain event, and it is in close relation with the threshold LET.

The single event upset in static RAMs is probably the best known single event effects and occurs both in space and also on Earth. To cause disturbance in a circuit, the charge generated by a heavy ion strike must be collected by a sensitive node. Reverse-biased pn junctions feature a large depletion region and a relatively strong electric field, so they are one of the most likely candidate to collect charge [29]. The process leading to a bit-flip in an SRAM can be summarised as follow:

an ionizing particle strikes the drain of the off nMOSFET in the cross-coupled inverters, which form an SRAM cell.

i The particle generates a cascade of electron-hole pairs both inside and outside the depletion region of the reverse-biased drain pn junction.

ii The released charge is collected by the junction, both by drift and diffusion, giving rise to a transient current. This current is sourced by the on pMOSFET in the same inverter, which tends to restore the initial state.

iii Since the pMOSFET has a finite output conductance, the voltage at the struck node decreases, turning the radiation-induced current into a voltage transient.

iv The current decreases the potential of the drain node, possibly below the cell switching voltage.

v If the radiation-induced transient is long enough, the feedback may cause the cell to flip, changing its initial state.

Many factors come into play to determine the probably of an upset. The higher the LET of the impinging particle, the larger the deposited charge. Charge collection is also of fundamental importance. The drift component is given by the $e - h$ pairs separated by the electric field in the junction of the depletion region and collected by the drain node. In addition, particles generated in the neutral region may diffuse towards the junction. Simulations have shown an even more complex picture, involving the distortion of the junction potential by the ion track (*funnel effect*) [49]. The funnel effect greatly extends the region over which charge can be collected by drift, and increases SEU sensitivity, even though its impact may be minor in circuits, like SRAM cells, where the junction bias is allowed to change [29]. Indeed, a junction connected to a pMOSFET may go from reverse bias to zero bias during a strike, limiting the influence of drift collection.

1.4.1 Scaling effects

Single event effects have become more complex to study, as the feature size has been scaled into the deep submicron realm. Indeed, nowadays the size of the ion track has become comparable to feature size of modern chips. Phenomena that were once confined to a single circuit node, can now involve multiple nodes and

charge sharing can occur [85]. This has caused a significant rise in the number of multiple bit upsets as technologies have been scaled.

As will be shown in Chapter 2.3, in particular for Flash memories, all storage devices have been the subject of aggressive scaling and, as a result, have improved their density at an exponential rate. As memory packing density increases, the probability that a neutron or alpha strike will upset more than one cell also increases. Based on simulations from Ibe and coworkers [39], the ratio of high energy neutron multi-cell upsets to single cell upsets grows from a few percent at 250 nm to almost 50% at 22 nm. This indicates that more powerful ECC codes and increasing the interleaving distance will be required.

Many soft error mitigation techniques rely on spatial separation of sensitive devices – e.g. cell interleaving for memory, dual interlocked storage cell (DICE) and triple modular redundancy (TMR) for logic, etc. As devices scale below 100 nm, the assumption that a neutron or alpha strike will only impact one node or device is no longer valid. New layout techniques to ensure isolation or suppression of charge collection will be required [76].

Chapter 2

Overview on Flash memories

Complementary metal-oxide-semiconductor (CMOS) memories can be divided into two main categories: random access memories (RAM's), which are volatile, i.e., they lose stored information once the power supply is switched off, and read-only memories (ROM's), which are nonvolatile, i.e., they keep stored information also when the power supply is switched off. Nonvolatile memory market share has been continuously growing in the past few years, and further growth in the near future is foreseen, especially for Flash memories (in which a single cell can be electrically programmable and a large number of cells-called a block, sector, or page are electrically erasable at the same time) due to their enhanced flexibility against electrically programmable read-only memories (EPROM's). Flash memories combine the capability of nonvolatile storage with an access time comparable to DRAM's, which allows direct execution of microcodes. Hence, the reason for the great success of Flash memories in present data storage market [8].

There are two major applications for Flash memories that should be pointed out. One application is the possibility of nonvolatile memory integration in logic systems (mainly, but not only, microprocessors) to allow software updates, store identification codes, reconfigure the system on the field, or simply have smart cards. The other application is to create storing elements, like memory boards or solid state hard disks, made by Flash memory arrays which are configured to create large-size memories to compete with miniaturized hard disks.

This section will give an overview of Flash technology, as Flash memories (in particular NAND) are the main devices used in neutron irradiation experiments (see

Chapter 3.2) and simulations (see Chapter 4). Basic operating principles of a schematic cell will be introduced in Section 2.1, while the main array organizations are given in Section 2.1.5. Section 2.2 will point out the reliability problems of industry-standard Flash Cell. Scaling issue and the latest Flash structures are presented in Section 2.3.

2.1 Basic principle and operations of NVM memories

There is a widespread variety of NVM's, and they all show different characteristics according to the structure of the selected cell and the complexity of the array organization. They all have performance that can go from those of ROM memories, which cannot be reconfigured, to those of information alterability with almost the same flexibility of RAM memories.

2.1.1 Information storage and access

The need for information alterability always contrasts with the need for good data retention. Cells with different characteristics have different applications according to the relevance that the device functional parameter has (absorbed power, programming/erasing speed and selectivity, capacity, and so forth). To have a memory cell that can commute from one state to the other and that can store the information independently of external conditions, the storing element needs to be a device whose conductivity can be changed in a non destructive way.

One solution is to have a transistor with a threshold voltage that can change repetitively from a high to a low state, corresponding to the two states of the memory cell, i.e., the binary values ("1" and "0") of the stored bit. Cells can be "written" into either state "1" or "0" by either *programming* or *erasing* methods. One of the two states is called "programmed", the other "erased". In some kinds of cells, the low-threshold state is called "programmed"; in others, it is called "erased". Though this may induce some confusion, the different terms are related to the different organizations of the memory array. In fact, if a datum has to be stored in a bit of the memory, there are different procedures.

1. The whole memory is erased (i.e., all the cells are driven in the same conductive or nonconductive state) and, after this, the information is programmed

in the selected bit; the rest of the array is reprogrammed.

2. Only the byte that includes the bit to change is erased and then reprogrammed with the new information.

3. Only the bit that has to be changed is addressed; the value to be stored is compared with the already stored one and written only if different.

When memories are organized as in cases 1) and 2), there is only one operation that can be performed bit by bit, called "program". The other operation, which is performed on the whole array or on a part of it, is the "erase" operation. When memories are organized as in case 3), both operations can be performed bit by bit but "program" needs a much more complicated array organization.

The "read" operation is performed by applying to the cell a gate voltage that is between the values of the thresholds of the erased and programmed cells and senses the current flowing through the device.

2.1.2 FG device

A Flash cell is basically a floating-gate MOS transistor (see Figure 2.1), i.e., a transistor with a gate completely surrounded by dielectrics, the floating gate (FG), and electrically governed by a capacitively coupled control gate (CG). Being electrically isolated, the FG acts as the storing electrode for the cell device; charge injected in the FG is maintained there, allowing modulation of the "apparent" threshold voltage (i.e., V_{th} seen from the CG) of the cell transistor. Obviously the quality of the dielectrics guarantees the non-volatility, while the thickness allows the possibility to program or erase the cell by electrical pulses. Usually the gate dielectric, i.e., the one between the transistor channel and the FG, is an oxide in the range of $9 - 10$ nm and is called "tunnel oxide" since FN electron tunneling occurs through it. The dielectric that separates the FG from the CG is formed by a triple layer of oxide-nitride-oxide (ONO). The ONO thickness is in the range of $15 - 20$ nm of equivalent oxide thickness. The ONO layer as interpoly dielectric has been introduced in order to improve the tunnel oxide quality. In fact, the use of thermal oxide over polysilicon implies growth temperature higher than 1100 °C, impacting the underneath tunnel oxide. High-temperature postannealing is known to damage the thin oxide quality.

If the tunnel oxide and the ONO behave as ideal dielectrics, then it is possible

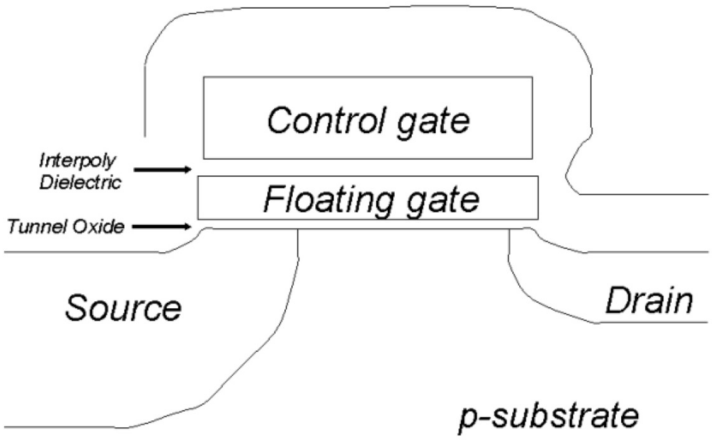


Figure 2.1: Schematic cross section of a Flash cell. The floating-gate structure is common to all the nonvolatile memory cells based on the floating-gate MOS transistor.

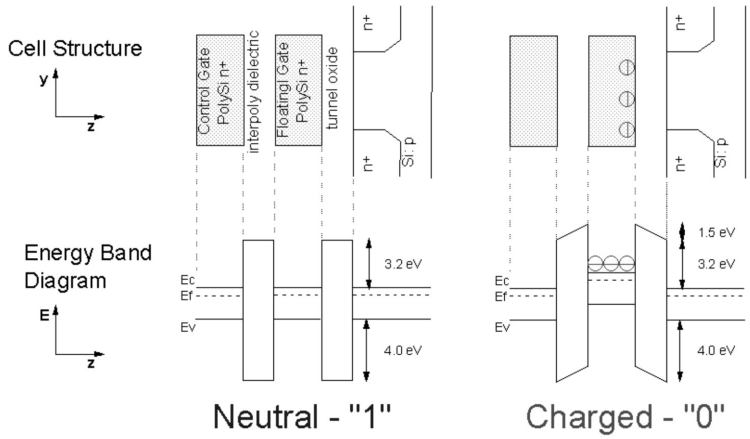


Figure 2.2: Schematic energy band diagram (lower part) as referred to a floating gate MOSFET structure (upper part). The left side of the figure is related to a neutral cell, while the right side to a negatively charged cell.

to schematically represent the energy band diagram of the FG MOS transistor as reported in Figure 2.2. It can be seen that the FG acts as a potential well for the charge. Once the charge is in the FG, the tunnel and ONO dielectrics form potential barriers.

The neutral (or positively charged) state is associated with the logical state "1" and the negatively charged state, corresponding to electrons stored in the FG, is associated with the logical "0".

2.1.3 Read operation

The data stored in a Flash cell can be determined measuring the threshold voltage of the FG MOS transistor. The best and fastest way to do that is by reading the current driven by the cell at a fixed gate bias. In fact, as schematically reported in Figure 2.3, in the current-voltage plane two cells, respectively, logic "1" and "0" exhibit the same transconductance curve but are shifted by a quantity, voltage shift ΔV_{th}, that is proportional to the stored electron charge Q. In particular, the threshold voltage shift ΔV_{th} is derived as

$$\Delta V_{\text{th}} = V_{\text{th}} - V_{\text{th},0} = -Q/C_{FC}$$

where $V_{\text{th},0}$ is the thershold voltage when $Q = 0$ and C_{FC} is the capacitance the FG and the control gate.

Hence, once a proper charge amount and a corresponding ΔV_{th} is defined, it is possible to fix a reading voltage in such a way that the current of the "1" cell is very high (in the range of tens of microamperes), while the current of the "0" cell is zero, in the microampere scale. In this way, it is possible to define the logical state "1" from a microscopic point of view as no electron charge (or positive charge) stored in the FG and from a macroscopic point of view as large reading current. Vice versa, the logical state "0" is defined, respectively, by electron charge stored in the FG and zero reading current.

2.1.4 Charge injection mechanisms

Considering Figure 2.2, the problem of writing an FG cell corresponds to the physical problem of forcing an electron above or across an energy barrier. The problem can be solved exploiting different physical effects [77]. In Figure 2.4, the

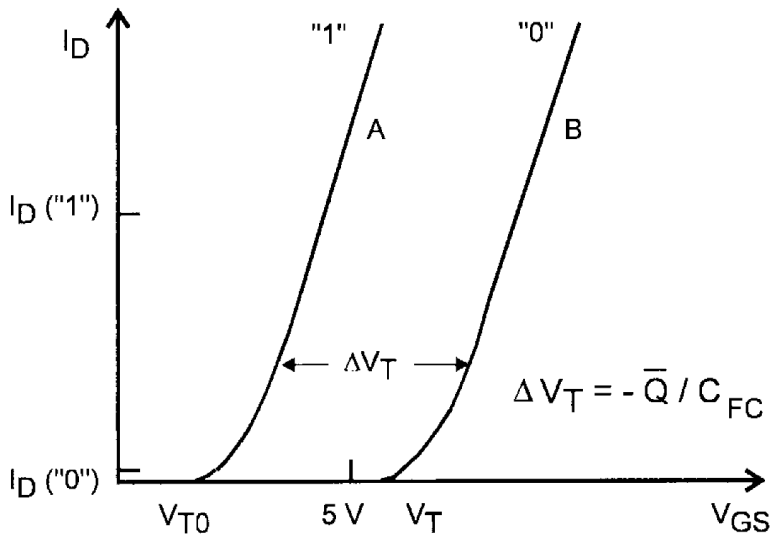


Figure 2.3: $I - V$ curves of an FG device when there is no charge stored in the FG (curve A) and when a negative charge Q is stored in the FG (curve B) [33].

three main physical mechanisms used to write an FG memory cell are sketched.

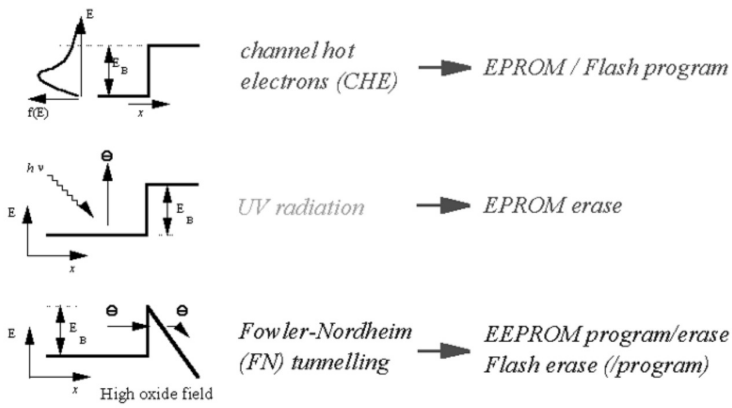


Figure 2.4: Writing mechanism in floating-gate devices.

CHE The *channel hot electron* mechanism, where electrons gain enough energy to pass the oxide-silicon energy barrier, thanks to the electric field in the transistor channel between source and drain. In fact, the electron energy distribution presents a tail in the high energy side that can be modulated by the longitudinal electric field.

UV The *photoelectric effect*, where electrons gain enough energy to surmount the

barrier thanks to larger the interaction with a photon with energy $h\nu$ larger than the barrier itself. For silicon-dioxide, this corresponds to UV radiation. This mechanism is the one originally used in EPROM's products to erase the entire device.

FN The *Fowler-Nordheim* electron tunneling mechanism is a quantum-mechanical tunnel induced by an electric field. Applying a strong electric field (in the range of $8 - 10\,\mathrm{MV} \cdot \mathrm{cm}^{-1}$) across a thin oxide, it is possible to force a large electron tunneling current through it without destroying its dielectric properties.

A NOR Flash memory cell is programmed by CHE injection in the FG at the drain side and it is erased by means of the FN electron tunneling through the tunnel oxide from the FG to the silicon surface.

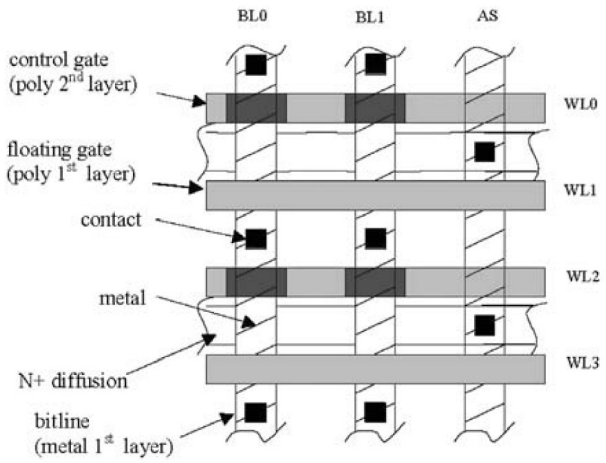
2.1.5 Memory Cell Array

Different type of Flash cells and architectures have been developed since 1990's. They can be divided in terms of access type, parallel or serial, and in terms of the utilized programming and erasing mechanism, Fowler-Nordheim tunneling (FN), channel hot electron (CHE), hot-holes (HH), and source-side hot electron (SSHE).

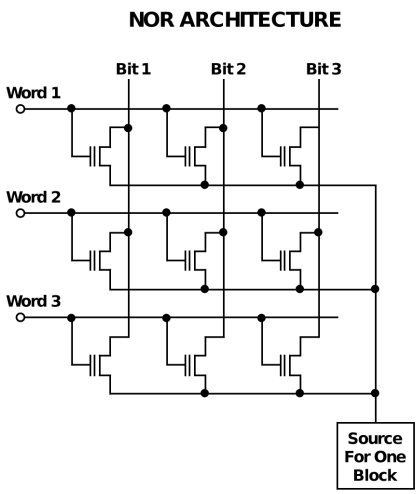
Among all of these architectures, the most widely adopted Flash memory cell array can be group into four categories: NOR type, divided bitline NOR (DiNOR) type, NAND type, and AND type. In the following sections, the basic concepts of these Flash cell array will be presented.

2.1.5.1 NOR

NOR-type Flash [84] is characterized by high-speed random access and very weel suited to program storage applications. Typical sector size is 64 kbytes. Nowadays, NOR-type Flash chips are available in densities ranging from 16 to 256 Mbits (source: *Samsung*). The Figure 2.5(a) represents the typical layout topology of a NOR memory cell array. Figure 2.5(b) shows the equivalent circuit schematic of the Flash array.



(a)



(b)

Figure 2.5: NOR-type Flash: array layout topology (a) and equivalent schematic (b).

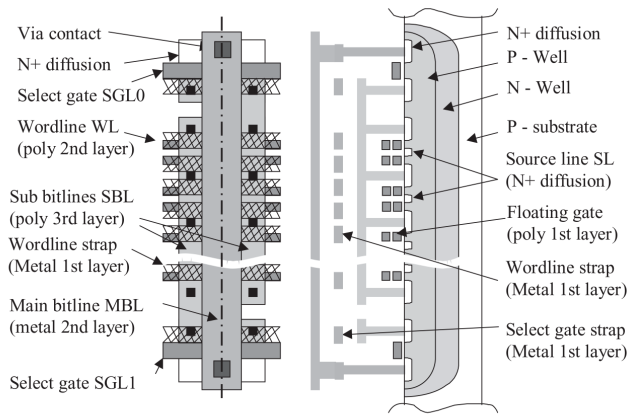
2.1.5.2 Divided Bit line NOR

DiNOR-type (divided bitline NOR) was introduced by Mitsubishi [44]. An outline of DiNOR Flash is reported in Figure 2.6.

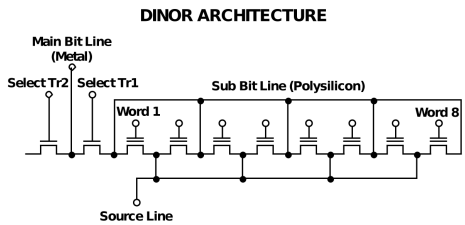
DiNOR-type Flash make use of triple-well process technology. In the elementary DiNOR array structure two polysilicon subbitlines, SBL0 and SBL1, are connected through select gates to an aluminium main bitline (MBL). Typically 64 cells (one block) are connected to each subbitline in parallel similar to a NOR-type array. With this arrangment a roughly 70% cell area reduction can be achieved with respect to the NOR-type configuration. Writing data to the cell is performed by applying about 8 V to the select gateline SGL (deselected SGLs are kept at 0 V). THe selected wordline (WL) in the selected block is biased a −11 V (with deselected WLs at 0 V). The main bitline (MBL) potential is raised to about 8 V. With the above biasing voltages, the cell at the crossing of the selected worldline and the selected subbitline sees −11 V on the gate and 8 V on the drain, and, therefore, electrons stored in its floating gate are discharged to the drain by tunneling across the gate oxide. Erase is achieved, again by FN tunneling, by applying about 12 V to the word lines of the selected block, −11 V to the P-well, the source line (SL) and the select gate. Since worldlines are at 12 V and P-well is at −11 V, a high electric field is applied between the floating gate and the P-well, which pushes electrons inside the floating gate. Different than the NOR-type Flash, for the DiNOR technology, erase causes an increase in the cell threshold voltage and is done at block level, while writing causes a decrease in threshold and is performed at the bit level.

2.1.5.3 NAND

The NAND-type Flash [83, 48] array features are summarized in Figure 2.7. Figure 2.7(a) shows the basic NAND-type cell layout and the array cross section, while Figure 2.7(b) outlines the equivalent NAND string circuit schematic. The NAND Flash technology outlined in Figure 2.7 is based on a triple-well process. There are, however, implementations of NAND-type Flash in which the array is built inside P-well regions on a N-type substrate. NAND Flash exploit Fowler-Nordheim tunneling for both programming and erase. In erase mode all worldlines (Wl) and select lines (SSL) are at 0 V. Select GSL, common-source line (CS) and bitline (BL) are floating; and the P-well is at 20 V. In programming mode the selected WL is at 19 V, deselected WLs are at 12 V, SSL and CS are at 5 V, and



(a)

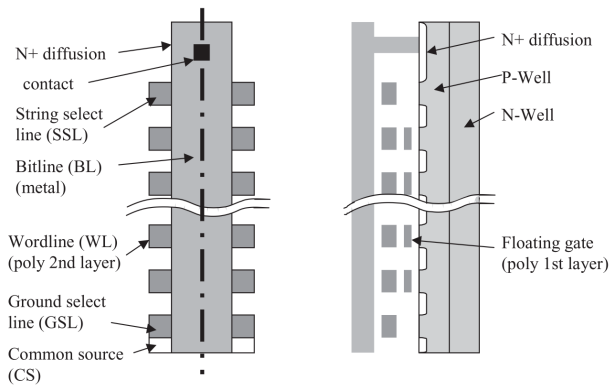


(b)

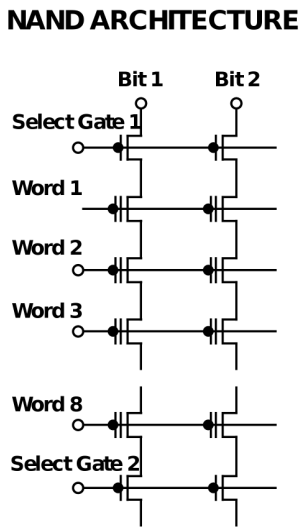
Figure 2.6: DiNOR-type Flash: array layout topology (a) and equivalent schematic (b).

P-well and GSl are at 0 V. The BL is at 0 V if the data to be programmed is logic 0 or at V_{cc} if the data is logic 1. Read is performed with BL at 3 V, CS and P-well at 0 V, SSl and GSL at V_{cc}. All WLs are at V_{cc} except the one selected, which is set at 0 V. Erased threshold voltage V_{th} is about -2 V (overerased state), and programmed V_{th} is in the $0.6 - 1.7$ V range.

The NAND-type Flash allows high integration densities due to the small number of contacts in the array. The NAND cell size is roughly half that of the NOR-type. NAND Flash is accessed in a serial mode and is widely used in data storage applications, such as Flash memory cards.



(a)



(b)

Figure 2.7: NAND-type Flash: array layout topology (a) and equivalent schematic (b).

2.1.5.4 AND

The AND-type Flash is a serial architecture that combines some features of NOR and NAND technologies. It is optimized for data storage applications such as Flash cards and fast silicon disk. AND technology features are summarized in Figure 2.8 which shows the array equivalent circuit schematic. The cells are connected in parallel to local diffusion lines. The local diffusion lines can be connected to metal bitlines (BL) through select transistors. Similar to the NAND technology, thanks to the small number of contacts (cells connect to diffusion lines), high integration densities are achieved. The programming and erase mechanisms are, respectively, floating-gate channel and floating-gate drain FN tunneling.

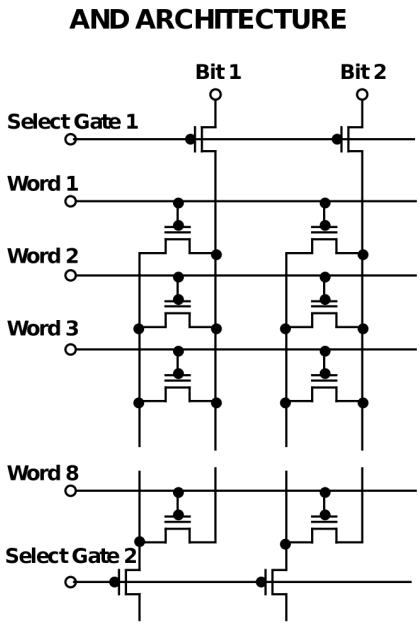


Figure 2.8: AND-type Flash: array equivalent schematic.

2.1.6 Multi Level Cell

An attractive way to speed up the scaling of Flash memory is offered by the multilevel (ML) concept [68]. The idea is based on the ability to precisely control the amount of charge stored into the floating gate in order to set the threshold voltage V_{th} of a memory cell within any of a number of different voltage ranges, corresponding to different logical levels. A cell operated with 2^n different levels is capable of storing n bits, the case $n = 1$ being the conventional single-bit cell.

Three main issues must be afforded when going from conventional to ML Flash [53]. A high programming accuracy is required to obtain narrow V_{th} distributions; reading operation implies multiple, either serial or parallel, comparison with suitable references to determine the cell status, requiring accurate and fast current sensing; V_{th} window and read voltage are larger while read margins are smaller than the single-bit case, this for allocating all levels, requiring improved reliability and/or error-correction circuitry. These key points will be discussed with reference to a common-ground NOR architecture.

Figure 2.9 shows the threshold voltage distribution of a 2-b/cell memory. The 11, 10, and 01 cell distribution will give rise to a different current distribution, measured at fixed V_{read}, while the 00 cell distribution does not drain current as well as the programmed level of a standard 1-b/cell device. High read data rate, via page or burst mode, is normally supported by large internal read parallelism.

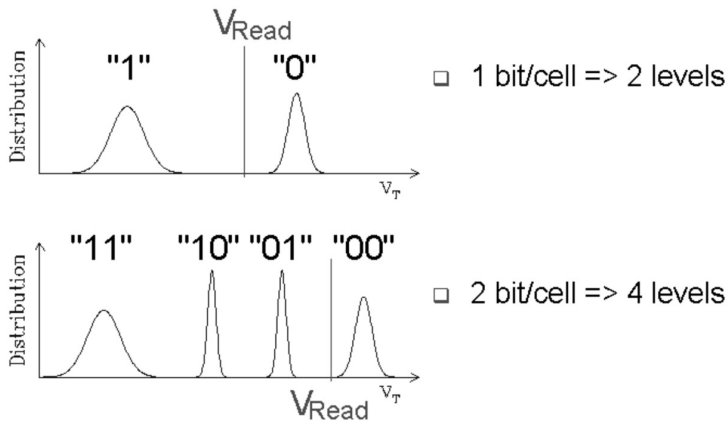


Figure 2.9: Threshold voltage distribution for 2-b/cell compared with the standard 1-b/cell.

2.2 Reliability

Many issues have to be addressed when, from the theoretical model of a single cell, a Flash product has to be realized, integrating millions of cells in an array. Nonvolatility implies at least ten years of charge retention, and the data must be stored in a cell after many read/program/erase cycles. The confidence in Flash memory reliability has grown together with the understanding of the single memory cell failure mechanisms.

The high degree of testability [12] allows the detection at wafer level of latent defects which may cause single cell failures related to programming disturbs, data retention, and oxide defects [11, 41], thus making Flash one of the most reliable nonvolatile memories.

2.2.1 Threshold voltage distribution

When dealing with a large array of cells, e.g., from tens of thousands to one million, it is very important to understand the type of dispersion given by the large set of cells. The best way to do it is to compare the threshold voltage distribution of the whole array, considering it after UV erasure (that can be considered as the reference state) after Channel Hot Electron programming and after FN erasing.

Figure 2.10 shows typical distributions of cell threshold voltages in a large memory array. The UV-erased distributionis pretty narrow and symmetrical. A more accurate analysis would reveal a Gaussian distribution due to random variations of critical dimensions, thickness, and doping which contribute to cause a dispersion of threshold voltages, either directly or through coupling ratios.

The programmed distribution is wider than the UV-erased one, but it is still symmetrical. The enlargement occurs because most of the parameters that cause V_{th} dispersion of UV-erased cells also impact the threshold shift of programmed cells.

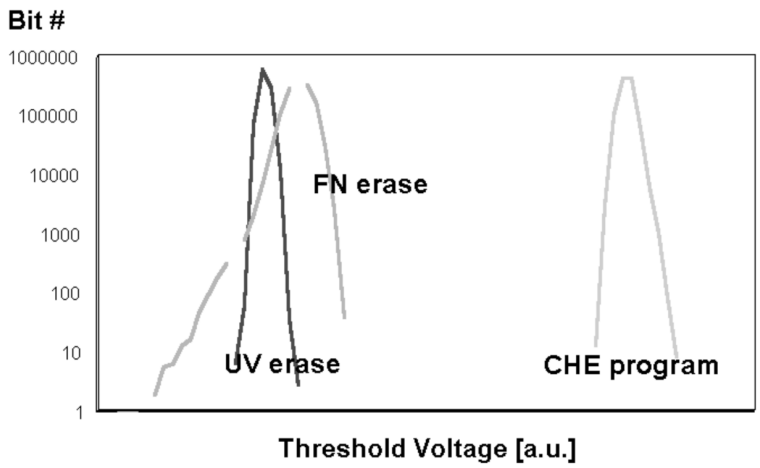


Figure 2.10: Threshold voltage distribution of a 1 Mb Flash array after UV erasure, after CHE programming, and after FN erasure [11].

The distribution of threshold voltages after electrical erase is much wider and heavily asymmetrical. A more detailed analysis would show that the bulk of the distribution is again a Gaussian with a standard deviation larger than the one of programmed cells. Cells in this part of the distribution are referred to as "normal" cells. But there is also an exponential tail at low V_{th}, composed of cells that erase faster than the average, also called "tail" cells.

The dispersion of threshold voltages of normal cells is due to coupling ratio variations, and it has been accurately modeled [90]. Instead, the understanding of the tail cells, although of key importance, is more difficult. In fact, as these cells erase faster than normal cells with the same applied voltage, one should assume that they are somehow "defective". However, they are just too numerous for being associated with extrinsic defects.

Different models have been presented with the aim to explain the tail cells. For example, a model proposed by Dunn *et al.* [32] (1994) explains the tail cells as due to randomly distributed positive charges in the tunnel oxide. This model is solidly based on the well-known existence of donor-like bulk oxide traps and on calculations that show the huge increase of the tunnel current density caused by the presence of an elementary positive charge closed to injecting electrode.

Independently from a consolidated model, it can be stated that the exponential tail of the erased distribution is mostly related to structural imperfections, i.e., intrinsic defects, and it can be minimized by process optimization (for example, working on silicon surface preparation, tunnel oxidation, FG polysilicon optimization) but not eliminated. Flash products must be designed taking into account the existence of such a tail.

2.2.2 Program disturb

The failure mechanisms referred to as "program disturbs" concern data corruption of written cells caused by the electrical stress applied to these cells while programming other cells in the memory array. Two types of program disturbs must be taken into account: row and column disturbs, also referred as gate and drain stress, as schematically reported in Figure 2.11, representing a portion of a cell array.

Row disturbs are due to gate stress applied to a cell while programming other cells on the same wordline. If a high voltage is applied to the selected row, all the

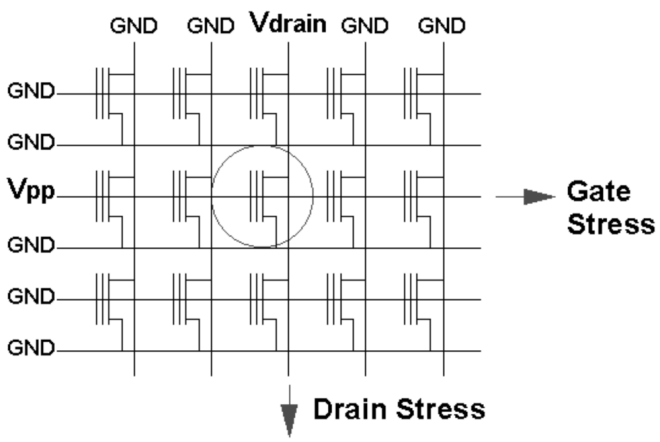


Figure 2.11: Schematic of a Flash array, showing row and column disturbs occurring when the cycled cell is programmed [11].

other cells ofthat row must withstand the gate stress without losing their data. Depending on the data stored in the cells, data can be lost either by a leakage in the gate oxide or by a leakage in the interpoly dielectric.

Column disturbs are due to drain stress applied to a cell while programming other cells on the same bitline. Under this condition, programmed cells can lose charge by FN tunneling from the FG to the drain (soft erasing). The program disturb depends on the number of cells along bitline and wordline and then depends strongly on the sector organization. The most effective way to prevent disturb propagation is to use block select transistor in a divided bitline and wordline organization to completely isolate each sector. Program disturb really could be a critical issue in Flash memory, and cells and circuits must be designed with safety margins versus the stress sensitivity.

2.2.3 Data retention

As in any nonvolatile memory technology, Flash memories are specified to retain data for over ten years. This means the loss of charge stored in the FG must be as minimal as possible. In updated Flash technology, due to the small cell size, the capacitance is very small and at an operative programmed threshold shift, about 2 V, corresponds a number of electrons in the order of 10^3 to 10^4. A loss of 20% in this number (around 2–20 electrons lost per month) can lead to a wrong read of the cell and then to a data loss.

Possible causes of charge loss are:

1. defects in the tunnel oxide;
2. defects in the interpoly dielectric;
3. mobile ion contamination;
4. detrapping of charge from insulating layers surrounding the FG.

The generation of defects in the tunnel oxide can be divided into an extrinsic and an intrinsic one. The former is due to defects in the device structure; the latter to the physical mechanisms that are used to program and erase the cell. The tunnel oxidation technology as well as the Flash cell architecture is a key factor for mastering a reliable Flash technology.

The best interpoly dielectric considering both intrinsic properties and process integration issues has been demonstrated to be a triple layer composed of ONO. For several generations, all Flash technologies have used ONO as their interpoly dielectric.

The problem of mobile ion contamination has been already solved on the EPROM technology, taking particular care with the process control, but in particular using high phosphorus content in intermediate dielectric as a gettering element [28]. The process control and the intermediate dielectric technology have also been implemented in the Flash process, obtaining the same good results.

Electrons can be trapped in the insulating layers surrounding the floating gate during wafer processing, as a result of the so called plasma damage, or even during the UV exposure normally used to bring the cell in a well defined state at the end of the process. The electrons can subsequently detrap with time, especially at high temperature. The charge variation results in a variation of the floating gate potential and thus in cell V_{th} decrease, even if no leakage has actually occurred. This apparent charge loss disappears if the process ends with a thermal treatment able to remove the trapped charge.

The retention capability of Flash memories has to be checked by using accelerated tests that usually adopt screening electric fields and hostile environments at high temperature.

2.3 Evolution and scaling trend

Integrated circuit technology has historically improved functional density at an exponential rate. This exponential trend was first pointed of by Gordon Moore in a 1965 article [54] where he observed that the number of transistors per chip was doubling every year. This trend has since slowed, but an exponential rate has been maintained. This so-called *Moore's law* has been the guiding rule for the International Technology Roadmap for Semiconductors (ITRS). The industry progresses by means of development of technology such that a new generation, referred to as a technology node, with minimum feature (F) sizes 0.7 x of the previous generation emerges at 3-year intervals. The 2007 ITRS includes data for five nodes: 2004, 90 nm; 2007, 65 nm; 2010, 45 nm; 2013, 32 nm, and 2016, 22 nm.[1]

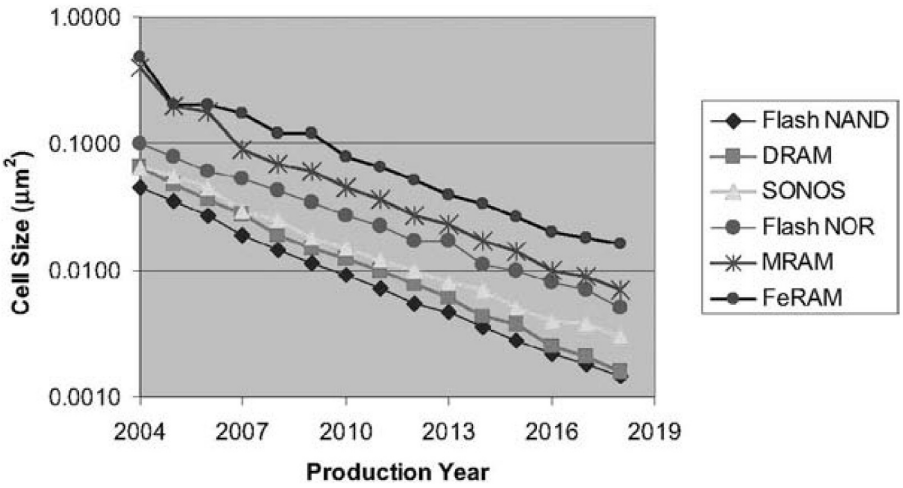


Figure 2.12: Plot of 2007 ITRS memory cell size requirements.

Figure 2.12 provides a graphical view of the cell size requirements given for DRAM and five NVM technologies. The expected density advantage of Flash NAND is evident.

Up to this point in time the industry has been successful in maintaining the frantic pace of development set out in the ITRS. Every year at every major semiconductor meeting and in private discussions among technologists there has been someone who argues that the problems on the horizon are insurmountable and scaling will end. Usually the proponent of doom has one or two specific issues

[1] The most recent version of the ITRS is available at http://public.itrs.net.

that he or she feels are going to be fatal. As the minimum feature size progresses closer to atomic dimensions, no doubt the nature of density improvement will shift from simple scaling to alternative ways of advancing functional integration. The challenge for nonvolatile memory is perhaps even more difficult than for CMOS logic because of the need to manage both dielectric stresses and maintain retention characteristics.

2.3.1 Market scenario

The Flash memories were commercially introduced in the early 1990's and since that time they have been able to follow the Moore law or, better, the scaling rules imposed by the market.

In last decade, the NAND flash market has witnessed tremendous growth with revenues increasing from $370 million in 2000 to over $20 billion in 2010 (see Figure 2.13). Continuous price declines for NAND flash memory have enabled it diversify beyond removable storage applications to consumer, industrial and computing applications, each with its own specific requirements in terms of performance, features and density. Since 2006, MLC technology has become mainstream with SLC technology occupying a sliver of the total market. SLC technology is favored in applications requiring high performance, high reliability and wide operating temperatures. Other than the solid state computing and mobile application markets, SLC devices can be found in embedded applications.

Above all, the development of Flash memory technology over the past decade has been one of the driving forces in memory market, such as digital consumer applications (digital cameras, handheld music players and mobile phones). NAND Flash has been leading the way from 16 Mb in 1994 to 8 Gb today.

2.3.2 NAND Flash scaling

NAND Flash memory has been the subject of aggressive scaling during the past decade, as shown in Figure 2.14, moving from the trailing position to the leading-edge of Flash technology. Besides 30 percent lithographic shrinkage for each generation, there have been three outstanding innovations in device-scaling.

1. The cell area has been scaled down by about 50 percent at each generation with a new floating gate structure and device isolation.

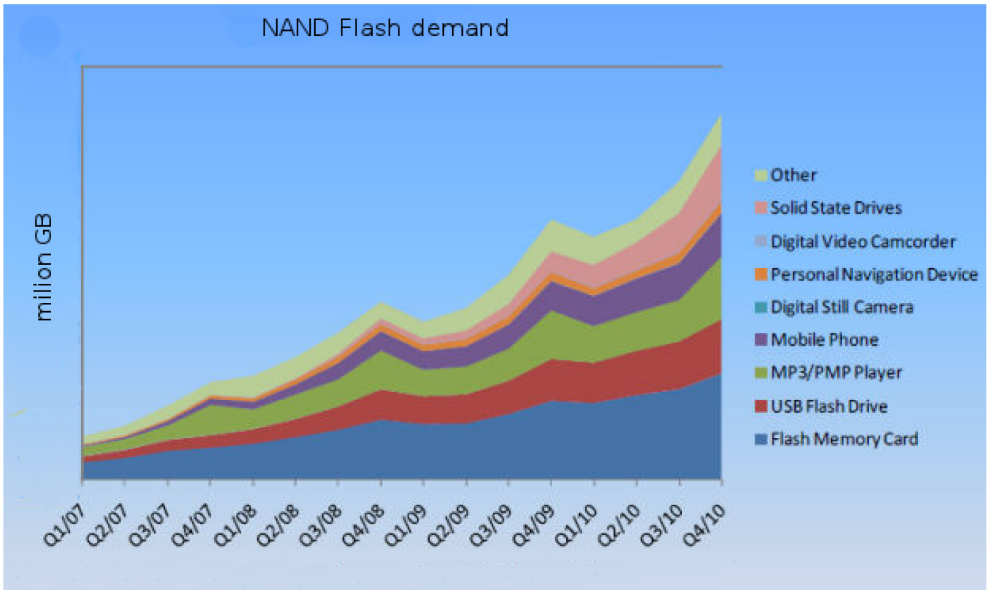


Figure 2.13: NAND Flash memory demand for different applications and/or devices.

2. Cell overhead was reduced 15 percent by increasing the number of cells in a NAND string from 16 to 32.

3. The number of bits in a cell doubled when using MLC technology.

Recently, an 8 Gb MLC NAND Flash, with these improvements, has been developed as the highest memory density with the smallest cell size of 0.0093 μm^{-2} per bit.

Given the current technology evolution trend, the minimum design rules of NAND Flash memory will move beyond the 30-nm region by the next decade (2010's), and single chip NAND Flash density will reach to 32/64 Gb or above. From the standpoint of device scaling, however, NAND Flash will face tough challenges, such as much more severe floating gate interference, a lower coupling ratio and less tolerant charge loss. As word-line space drops below 30 nm, the capacitance coupling among floating gates is increased as much, which shifts and widens distribution of the cell threshold voltage (V_{th}). The height of the floating gate must be decreased to reduce the coupling. Figure 2.15 shows the necessary height of the floating gate for suppressing a V_{th} shift induced by coupling below 0.2 V. While low-k dielectric materials help improve the floating gate coupling, the capacity to scale it down will be increasingly limited.

Another structural limit at around 35 nm NAND technology node is that the utilization of floating gate sidewall ONO (Oxide Nitride Oxide) capacitance will no longer be possible because the inter-poly dielectrics will be thicker than the

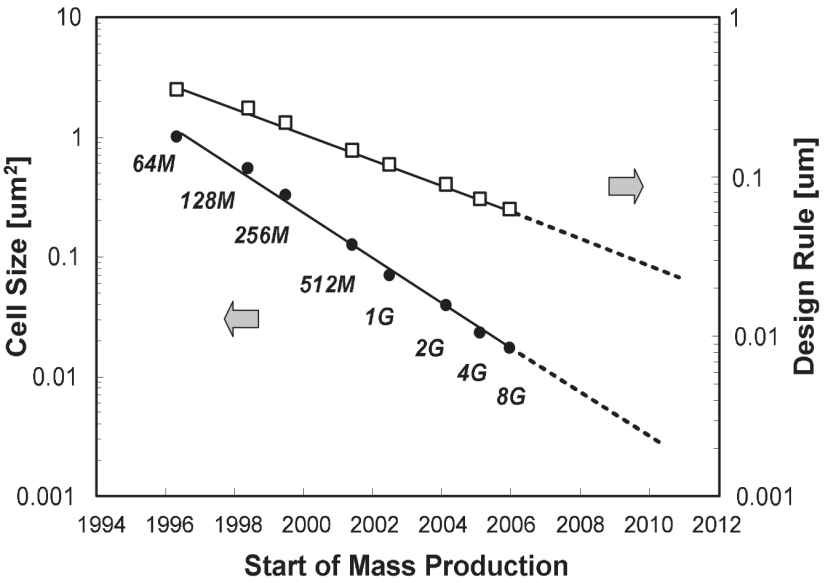


Figure 2.14: Historical trend of NAND Flash memory densities and technology.

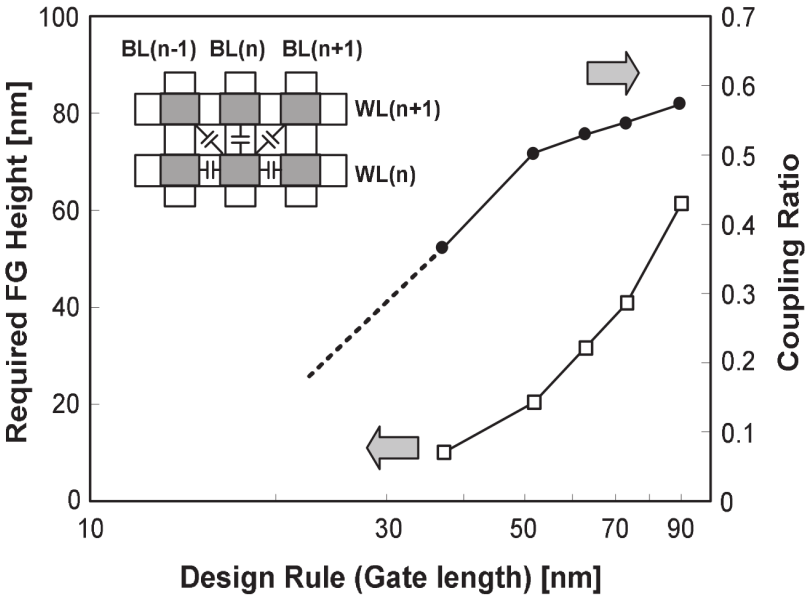


Figure 2.15: Required floating gate height for the effective suppression of inter-poly coupling interferences and coupling ratio by design rule.

space between floating gates. Since the sidewall's contribution to the coupling ratio from control gate to floating gate is reduced, the coupling ratio will drop drastically to below 0.4 at 35-nm node as shown in Figure 2.15. To enhance the coupling ratio, inter-poly ONO dielectrics needs to be scaled down from around 15 nm today. The development of ultra low leakage high-k dielectric materials for the NAND Flash inter-poly coupling capacitor application will be one of the most important and difficult tasks for the further scaling of NAND Flash at around and beyond 35-nm node.

The number of electrons on the floating gate is significantly decreased due to the decrease of interpoly ONO capacitance, as shown in Figure 2.16. It is expected that less than 100 electrons for a V_{th} shift of 6 V will be stored following 30 nm design rule. Considering that a MLC will fail after a 5 percent loss of the charges stored in the floating gate, loss of no more than 10 electrons is allowed over a 10-year period of operation for the device. In addition, increasing a portion of the cell edge, which is prone to generate interface states and charge traps, further undermines the tolerance for endurance and retention characteristics. Intensive efforts to incorporate nitrogen or deuterium into the tunnel oxide and substrate interface are required to improve endurance and charge loss (or gain) characteristics via the tunnel oxide. Other fundamental approaches are the use

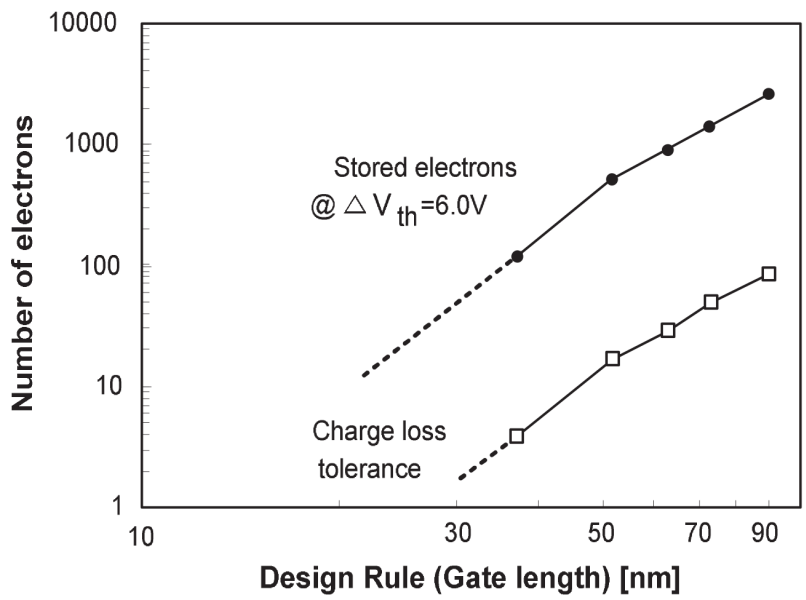


Figure 2.16: Number of stored electrons in a NAND Flash cell and the amount tolerable for charge loss by technology.

of non-floating gate structures, such as SONOS or nano-crystal dots. These cell

structures, which represent a new type of charge trapping, will be one of the future candidates for a technology node below 30 nm. When applied to NAND Flash, the use of low leakage high-k dielectric materials on top of nitride charge trap layer and corresponding integration technology will be a key to the success of these cell architectures. A SONOS cell provides nearly planar cell array with drastically reduced inter-cell coupling, thanks to the absence of floating gates. But, when Fowler-Nordheim tunneling is used for NAND type SONOS cell, the slow erase speed and more apparent charge loss need to be resolved. High-k coupling dielectrics which has good band gap matching characteristics with the charge trap layer improves the coupling ratio onto the tunnel oxide, allowing it to be thicker for improved charge loss while maintaining fast erase and program speed.

2.3.3 NOR Flash scaling

The density of NOR Flash has doubled every two years, complying with Moore's Law. While volume production of 90 nm MLC NOR technology is expected to start later this year, several critical limitations in NOR cell technology pose tough challenges for future scaling. For example, drain program voltage, theoretically, cannot be scaled below 3.2 V because the electrons to be programmed onto the floating gate should be activated to high enough energy states to jump over the potential barrier from silicon to oxide. The cell drain should be able to sustain voltage significantly higher than 3.2 V to compensate for the voltage drop along the channel. For the continuous cell size scaling, cell transistor gate length should also be scaled down to about two times minimum design rule and suppressing drain to source punch-through requires increasing channel doping level, which results in lower junction breakdown voltage at cell drain. Figure 2.17 shows the historical and projected changes in cell gate length and associated junction breakdown voltage by technology. It indicates that junction breakdown voltage will be dropping to the critical boundary around 4 V at the 65 nm technology node.

Scaling also will present major challenges in NOR MLC technology. Maintaining narrow erase cell threshold voltage distribution, narrow program cell threshold voltage distribution and data retention characteristics after the endurance cycles over multiple charge states are the main technical challenges. For example, a reliable means of controlling tunnel oxide thickness in the active edge area is one of the critical factors for the narrow distributions of erase and program cell threshold

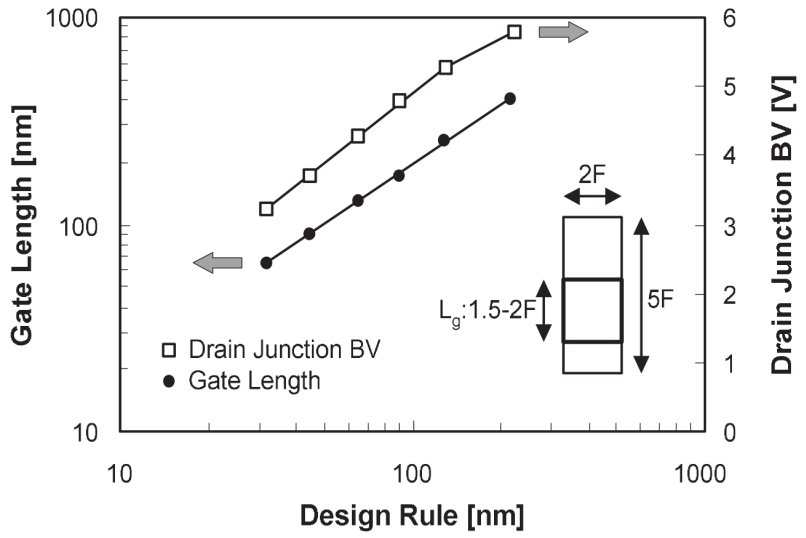


Figure 2.17: NOR cell transistor Gate Length estimation and drain junction breakdown voltage vs. Design Rule.

voltages. From these considerations, together with the structural changes needed for gate length scaling, the additional process solution and considerable design investigations are required for further binary and multi-level cell scaling at any technology node under 45 nm.

2.3.4 New non-volatile memory (NVM) technologies

Relatively mature, new material memories such as FRAM, MRAM and PRAM can offer very fast write speed with random accessibility and have potentially high endurance and retention characteristics as shown in Table 2.1. Using a polarization mechanism, FRAM is hampered by the relatively large cell size of $10 - 20\,\mathrm{F}^2$. MRAM, using the magneto-resistance effect, could provide sufficiently faster read/write performance, but it also suffers from a large cell size, as well as high write current and the disturbance inherent in magnetic fields. PRAM, using a reversible phase-change mechanism, has a relatively small cell size and good retention/endurance characteristics with a simple process integration scheme.

The primary requirements for new non-volatile memory can be summarized as good scalability and superior functionality. PRAM offers RAM-compatible fast read access and a reasonably fast, random write speed of faster than 500 ns with the potential of $4\,\mathrm{F}^2$ scalability. While its write speed is not fast enough for it

	FRAM	MRAM	PRAM
Cell size	15 F^2	40 F^2	20/10 F^2
Operational mechanism	polarization	MR change	phase change
Technology issue	cell size scalability	cell size write current disturbance	write current thermal crosstalk
Speed (read/write)	50/50 ns	50/50 ns	4 50/ $<$ 500 ns
Endurance/Retention	10^{12}/10 years	10^{15}/10 years	10^{11}/10 years

Table 2.1: Characteristics of new NVM technologies.

to be used as the working RAM, it provides a RAM-like simple write and read operations.

The cell size of PRAM using a MOS switch device is relatively large (currently 20 F^2). Until now, the ability to reduce the cell size of MOS switch PRAM has been limited by the amount of write current to melt the phase change storage material ($Ge_2Sb_2Te_5$: GST). Three dimensional FinFET for cell access devices, which can deliver two to three times the current of the present planar MOS transistor, may need to be introduced to resolve this problem. On the ther hand, a Diode switch PRAM can draw a larger amount of write current than an MOS switch PRAM, effectively permitting the design of a smaller cell. Regardless of the current driving capability of the switching device, the ideal cell size of MOS switch PRAM is 6 F^2, whereas a Diode switch PRAM cell could be reduced to the lithographic limit of 4 F^2. However, the complexity in array design, such as isolated base-active and the need for a contact per cell, and the need for dedicated process integration steps for diode formation can be major inhibitors for a Diode switch PRAM to achieve 10 F^2 for replacing NOR Flash.

As Flash memory continues to evolve from binary to MLC types, the resistance of GST storage material of PRAM can be modulated to multiple resistance states by applying appropriate current pulses. This intrinsic property points to the possibility of developing multi-bit PRAM cell technology, even though the means in which it distributes resistance states needs to be successfully designed to adequately identify multiple resistance states at readout circuits.

Chapter 3

Radiation effects on Flash memories

As mentioned in Chapter 2, nonvolatile Flash memories represent the most attractive storage device, due to their high density and low cost, for space engineers and terrestrial application. Accordingly to increasing request of performance and availability, several studies have been performed on the radiation sensitivity of this technology, concerning both total ionizing dose and single event effects. Unfortunately, these devices showed an high sensitivity to radiation in space environment, so as to limit their use, while a few is known in the terrestrial environments.

While SRAM devices can be made insensible to soft errors in many different ways (by design as in Liaw [50] (2003) or by software as in Klein [43] (2005) and in Huang [38] (2010), NVMs are susceptible to all three categories of errors above. The lack of any refresh cycle of the stored information make flash memories vulnerable to data loss at each exposure to ionizing radiation. Considering that Flash memories standards impose a retention time for the data stored of 10 years at least and a minimum 10^6 write/erase operations before performance degradation starts, is clear that non-volatile memory cells are in a passive state for most of their lifetime.

As been shown in Chapter 2, Flash memories contains different peripheral circuits (charge pumps, microcontrollers, page buffer, etc. . .) which can be seriously corrupted by radiations. In the 90's, the floating gate cells were considered a minor problem, while the control circuitry was recognized as the most sensitive part of commercial devices [57, 58, 75, 70]. Recent experiments on current generation Flash memories

have however shown that significant amount of radiation effects can be observed at the sea level or terrestrial environments. As the technology evolution has developed more and more advanced devices, increasing the architectural density while decreasing the featuring size of cells, the sensitivity of the floating gate cell seems to be an important issue.

Single event effects and total ionizing dose effects are recently reported in literature for both the peripheral circuitry and the floating gate cells, depending on the type of memory and the conditions during exposure [21, 36]. However, is quite difficult to have an accurate comprehension of all radiation effects, and in particular to identify the fundamental causes, in the state-of-the-art Flash memories. Radiation sensitivity can exhibit large variation from manufacturer to manufacturer and under different operating conditions.

In first sections of this chapter (Section 3.1), an overview of the effects of ionizing radiation on NAND Flash are presented, focusing on the floating gate cell. Hence, the second part of the chapter (Section 3.2) is focused in particular on the effects of neutrons radiation. After an explanation of used devices and facilities, this section presents experimental results and related consequences obtained exposing Flash devices to neutrons.

3.1 Floating gate cells

The effects of ionizing radiation on the FG array is for the largest part related to the corruption of the stored information. While the "user mode" information is visible in digital term, i.e. "0" or "1" in a SLC device, or in a multi bit configuration for a MLC, the information by itself is an analog quantity, being its granularity linked to the electron charge [27], represented by the threshold voltage V_{th}. Accessing the V_{th} permits to investigate the complex behaviour of many recent technologies. The results of radiation studies in FG cell seems to be strictly related to the physics of the cell more than the memory operating condition.

3.1.1 Total ionizing dose effects

The effects of total ionizing dose (TID) are typical of space environment or high energy experiments, where the amount of radiation is very high.

The main effect of a TID irradiation on a FG array is to progressively discharge the cell as dose accumulates. This is actually experimentally found and

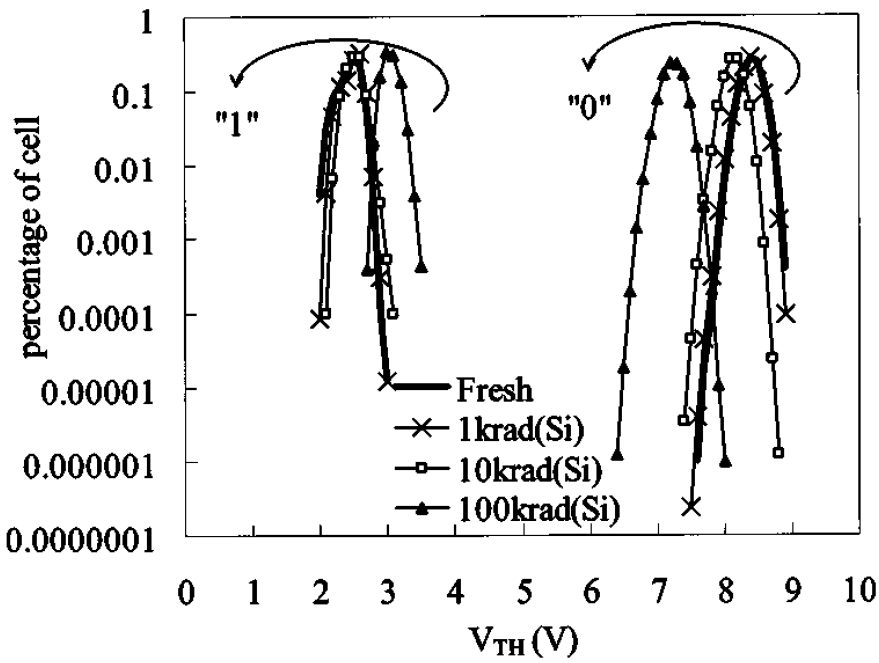


Figure 3.1: Probability densities of FG MOSFET threshold voltage V_{th} before and after irradiation at 1 krad(Si), 10 krad(Si), 100 krad(Si), with 10 keV x-rays [15].

is illustrated in Figure 3.1, reporting the V_{th} probability densities for FG arrays programmed in the "0" and "1" state before and after different 10 krad x-rays TID levels [15]. Before irradiation, V_{th} distributions closely resemble the expected Gaussian shape. After irradiation, the threshold voltage of all FGs programmed in the "0" state ("high" V_{th}, corresponding to electron stored in the FG) uniformly moves toward lower V_{th}: the FGs progressively lose electrons because of irradiation. Conversely, devices programmed in the "1" state ("low" V_{th}, holes stored in the FG) experience a V_{th} increase, that is, a positive charge loss from the FG. The V_{th} shift (V_{th}) is lower for the "1" than that for the "0" state, because of the much lower net charge stored in the FG in that condition. Of great importance is the fact that distributions rigidly shift during irradiation since TID radiation effects are uniformly distributed over the chip area.

A physics-based model to account for these effects has been developed at first by Snyder et al. in 1989 [81] and further developed in recent times to account for the developments in the technology happened in almost 20 years [16, 18]. Briefly, charge is generated in all oxides surrounding the FG. Depending on the oxide electric field part of these carriers recombine; after recombination, carriers thermalize, then electrons are quickly swept away from the oxide thanks to their

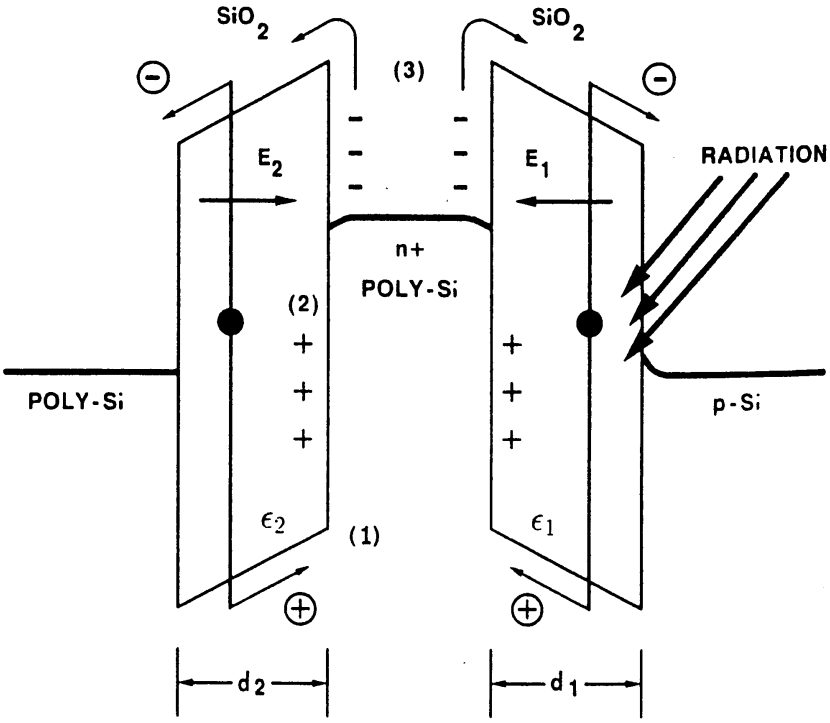


Figure 3.2: Main contribution to threshold voltage degradation in a FG cell during TID experiments [81].

high mobility, while holes slowly move toward the FG and may be either reach the FG, where they recombine part of the stored negative charge (mechanism (1) in Figure 3.2), or they are trapped in the oxide (mechanism (2) in Figure 3.2). In both cases, the net effect is a reduction of the cell V_{th}. Further, electrons may acquire enough energy form the incoming radiation [42] to jump over the oxide barrier (mechanism (2) in Figure 3.2). There are several aspects which need to be considered when adapting this model to modern devices, including:

i recombination, thermalization, and electron transit times are of the same order of magnitude in the range of the tunnel oxide thickness (10 nm or less) proper of modern devices: the different phenomena, above described, are not strictly sequential;

ii in all modern devices the interpoly dielectric is a ONO sandwich, thus the different dynamics of charge generation and trapping in the nitride layer has to be carefully considered [67];

iii the above description do not consider the lateral dimensions of FG cell.

Cellere *et al.* [18] (2005a) decribed accurately the average V_{th} shift as a function

of dose, comparing data obtained with ^{60}Co γ-rays and with 10 keV x-rays. In this study is evident that the degradation induced by x-rays is by far larger, likely due to dose enhancements effects, while results from ^{60}Co perfectly superimpose to results obtained after protons irradiation [18, 24].

Finally, Scheick *et al.* [73] (1998) measured the number of a non-volatile memory cells which flip from "1" to "0" a function of the duration of exposure to UV radiation at an intensity of 50 μWcm^{-2}. When subjected to ionizing radiation, the array progressively discharge, resulting in a shift of the average erasure time [73]. in general single bit errors can appear during TID exposure due to large V_{th} shifts [58, 59, 61].

3.1.2 Single event effects

The effect of single ions on the FG array is, as expected, different from that of TID since the energy is released by each ion in a small volume. Hence, every single ion can result in a large but localized effect (SEE). As an example, a chip with all FG programmed at "1" was irradiated with $2 \cdot 10^7$ Iodine ions cm^{-2} [13]. The FGs with corrupted information are randomly distributed across the chip surface, and their number roughly corresponds to the number of FGs being hit by ions, meaning that every time a single FG is hit by a single ion, it experiences a net charge loss, resulting in a V_{th} shift [21, 13, 14, 17]. A more informative picture of these same data can be obtained by looking at the statistical description of V_{th} (Figure 3.3): before irradiation, V_{th} distribution resembles the expected Gaussian shape, but after irradiation, a secondary peak appears, corresponding to the FGs hit by ions. Hence, the distance between the main and the secondary peak can be taken as the average V_{th} shift experienced by hit cells, $< \Delta V_{\text{th}} >$. This parameter, as a function of the pre-rad tunnel oxide electric field and of ion LET, At least for ions having a relatively low energy, shows an almost perfectly linear relationship. Such a linear relationship allows excluding a model similar to that developed for TID effects, that is, one based on charge generation, recombination, and transport. In fact, all these phenomena have a strongly non-linear dependence on the electric field [21]. Further, the amount of charge lost for FGs Figure 3.3 exceeds that calculated for those same ions with a generation/recombination/transport model by a factor of more than 20 [17, 21].

One of the model currently available to explain these results is based on the idea that the dense electron/hole plasma generated in the tunnel oxide by the

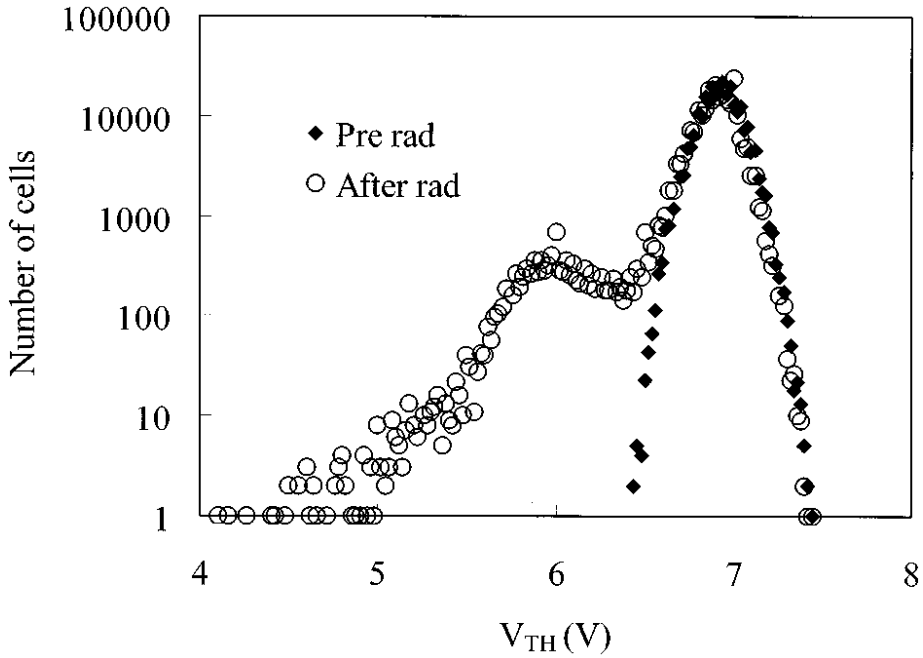


Figure 3.3: Probability density of threshold voltages for before and after irradiation with $2 \cdot 10^7$ Iodine ions cm^{-2} [13].

incoming ion acts like a resistance, thus promptly discharging the FG [17, 21]. While value of this resistance is linked to the ion LET (heavier ions result in denser track, hence in lower resistance and in larger discharge), the duration of the path is linked to the presence of electrons themselves in the FG. In fact, since $< \Delta V_{\text{th}} >$ linearly depends on the number of generated electrons/holes, not on this surviving recombination, it follows that charge loss has to happen before recombination. Moreover, as discussed in Cellere *et al.* [21], the average $< \Delta V_{\text{th}} >$ strongly depends on the FG area, being larger for smaller FGs. This could explain why user mode errors in the array were not detected in old generation Flash memories ($< \Delta V_{\text{th}} >$ was too small to result in errors). Also, one of the key points which derive from this complex picture is that phenomena are going on in the nano scale at a pace which is not always compatible with models developed in many years at the micron or sub-micron scale. Actually, things get even more complicated when dealing with aggressively scaled devices.

In most modern technologies, some charge loss happen even from FGs which are just grazed, not hit by ions [20]. This phenomenon is not covered in details here due to space constrains and since a full model is not yet available; however, it is worth noting that these "grazed" FGs have peculiar characteristics which allow clearly distinguish them from their "hit" counterparts.

Also, in any real world environment (be it space, high energy physics, or simply the atmosphere) particles do not cross the devices normally to the silicon surface. It is not totally surprising but of great interest the fact that, at grazing angles, a single ion can lead to the formation a "tail" of FGs with corrupted information [23].

3.1.3 Long-term effects

Another important problem of radiation is the possibility of long-term, or even permanent, damages in FG devices. FGs being hit by single ions can be easily identified thanks to their large post- irradiation ΔV_{th}. If they have been programmed once again the cumulative probability of V_{th} increased, as shown in Figure 3.4 [19]. Just one hour and half after programming the memory array,

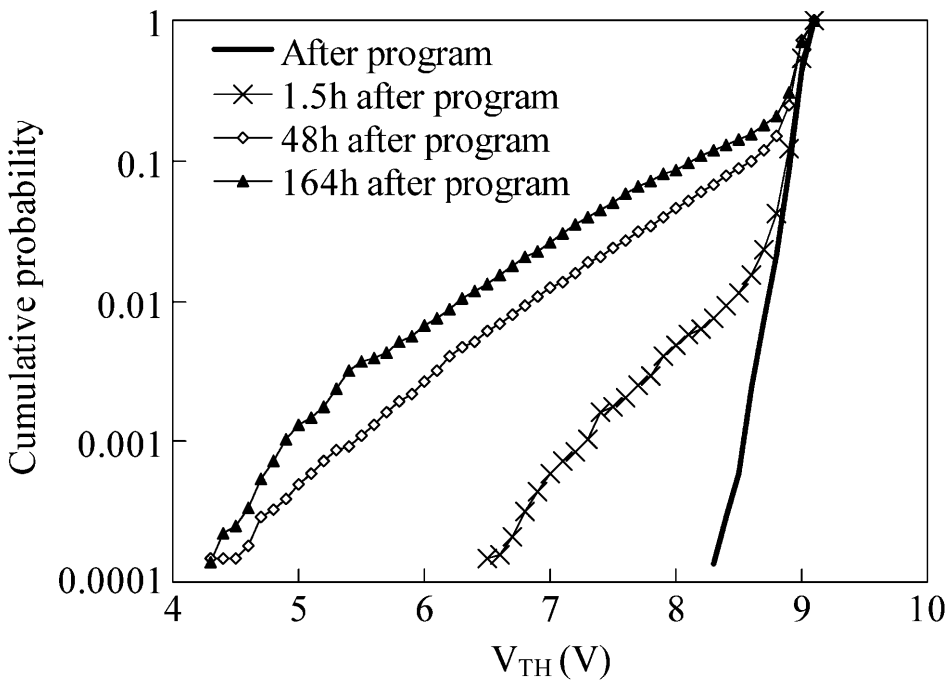


Figure 3.4: Cumulative distribution of V_{th} for cells of FG device hit by iodine ions, after being re-programmed [19].

a tail appears extending down to less than 7 V due to FGs changing their V_{th} and leaking charge. This charge loss should be related to the ion impact, since no similar effect was observed in non-hit FGs [19]. The tail increases with time, thanks to the progressive discharge of those FGs which were already leaking charge 90 minutes after programming, and to new FGs featuring lower leakage current, which progressively add to the tail of the distribution. The tail increases with

time, thanks to the progressive discharge of those FGs which were already leaking charge 90 minutes after programming, and to new FGs featuring lower leakage current, which progressively add to the tail of the distribution.

This phenomenon derives from the onset of the so called Radiation Induced Leakage Current (RILC) [47] and is basically a multi-Trap Assisted Tunneling, m-TAT, conduction mechanism through electrically active defects created, directly or indirectly, by the ion impact. The damage does not scale with the pre-irradiation electric field [19], so that the actual physical mechanism leading to the creation of defects is not clear. On the other side, the resulting defects have a close affinity with those generated by electrical stress, such as those responsible for SILC, as evidenced by their recovery after being exposed to Forming Gas Anneal (FGA) after irradiation and before the electron retention experiment [19].

3.2 Neutron radiation

The heavy, energetic galactic and solar ions discussed in 1.1 bombard the Earth's upper atmosphere, collide with oxygen and nitrogen atoms, and generate a cascade of fragments through the atmosphere. The results at ground level are particles including neutrons, protons, electrons, muons, and gamma rays. As exposed in the previous sections, the effects of radiation in Flash memories have mainly been a concern with heavy ions or high energetic radiation at the terrestrial environments. Recent experiments on current generation Flash memories have however shown that significant amount of radiation effects can be observed with the flux of neutrons at the sea level.

Besides, atmospheric neutrons are a known source of Soft Errors (SE) in static [4] and dynamic [80] CMOS memories (the majority of studies have been conducted on SRAM). The experiments presented in this work show that atmospheric neutrons are able to induce SE in Flash memories as well. Working with a spallation beam source, which mimics the energy spectrum of atmospheric neutrons, several single bit error were identified in different devices, indicating that the number of excess electron/hole stored in the FG can be corrupted by neutrons. Despite the low failure probability of a single FG cell, neutrons effects on these devices are important if we consider a mainstream integrated circuits of everyday applications where (1) the total chip are is very large, usually several cm^{-2}; (2) there is a huge amount of high density devices, in particular memories and microprocessors; (3)

the device scaling tends to make the critical charge very small.

3.2.1 Experimental

The terrestrial neutron spectrum used for soft error measurements is documented in JEDEC Standard "JESD89A—Measurement and Reporting of Alpha Particle and Terrestrial Cosmic Ray-Induced Soft Errors in Semiconductor Device" [40]. Several facilities throughout the world offer high energy broad spectrum neutron sources for accelerated soft error testing, but none accurately replicate the terrestrial neutron spectrum across the range of 1 MeV to 1 GeV.

The Isis neutron beam features an energy spectrum that is similar to the terrestrial one with an acceleration factor between 10^7 and 10^8, making it appropriate for accelerated tests of sea-level electronic applications.

Flash memories of structure, technological nodes and from different vendors were used in the experiments.

3.2.1.1 The Isis facility

The Isis neutron source is located at the CCLRC Rutherford Appleton Laboratory (Didcot, U.K.) and has been used so far for condensed matter studies. Neutrons are produced at Isis by the spallation process [88]: a heavy-metal target (tungsten) is bombarded with pulses of highly energetic protons, generating neutrons from the nuclei of the target atoms. The acceleration process is composed of two steps: first H^- are injected into a linear accelerator (LINAC). The beam is converted to protons by a 0.3 µm thick aluminium oxide stripping foil and then accelerated in a synchrotron. The high-energy proton pulses finally strike the tungsten target and corresponding pulses of neutrons are freed by spallation. The energy of the produced neutrons is reduced through a moderator, which can be of different types. More information about Isis can be found in [1]. The resulting neutron beam reaches 26 different lines (Figure 3.5), including the Vesuvio line where our experiments were performed. Vesuvio is commonly employed for condensed matter studies, exploiting neutrons above 1 eV, the so-called epithermal neutrons. As shown in Figure 3.6, the sample S is located at a distance $L_0 = 11.055$ m from the water moderator. The Aluminium sample tank is of cylindrical form, with an internal diameter of 50 cm. The height is 65 cm and the beam center is located at about 30 cm from the top. The spectrum in the

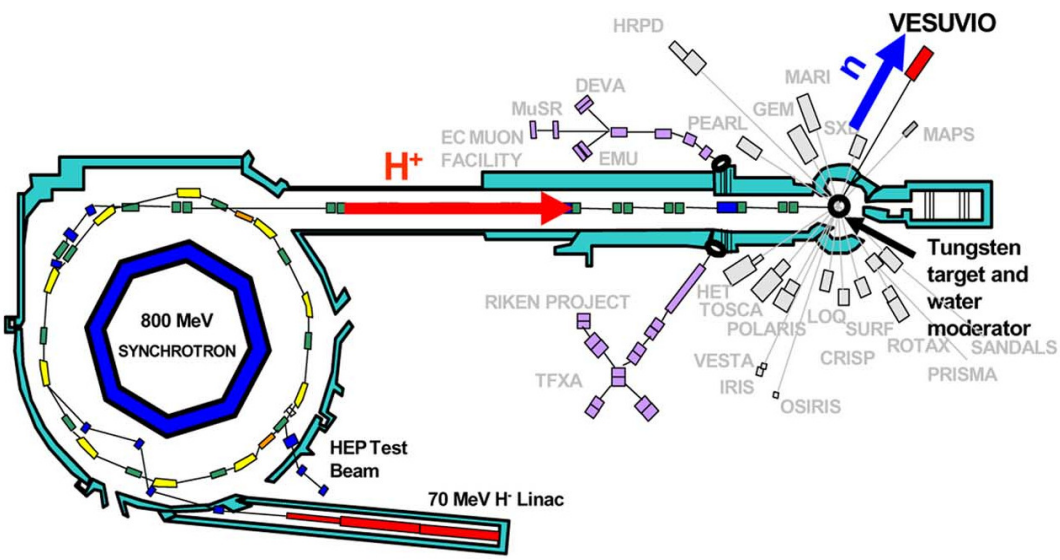


Figure 3.5: Schematic of the Isis neutron facility including the LINAC, the synchrotron and 26 lines [1].

high-energy region has been measured through the threshold activation target technique [62]. The Isis spectrum is illustrated in Figure 3.7, where it is compared with two of the most widely used neutron sources (LANSCE and TRIUMF) and with the terrestrial flux multiplied by a factor 10^7 and 10^8. As seen, the Isis spectrum features a $1/E^\alpha$ characteristic, with α larger than one, and provides a flux similar to the terrestrial one with acceleration between 10^7 and 10^8 in the energy range $10 - 100$ MeV. The Isis spectrum integrated above 10 MeV yields $7.86 \cdot 10^4\ \mathrm{n \cdot cm^{-2} \cdot s^{-1}}$ on the irradiated device.

Studies have been made to compare Isis to other spallation neutron sources through the use of the charge-coupled device (CCD) sensor described in [66]. Differences in the neutron spectrum of the Isis source with respect to LANSCE result in different charge collection spectra in the CCD. A "LANSCE equivalent flux" [65], i.e., the effectiveness of Isis neutrons in producing events in the CCD compared to LANSCE ones, has been measured yielding a value of $6.7 \cdot 10^4\ \mathrm{n \cdot cm^{-2} \cdot s^{-1}}$. The equivalence is based on the number of events measured with the CCD above a certain threshold (417 ke). In other words, if we divide the number of events in the CCD sensor collected in the Isis and LANSCE beams by the LANSCE-equivalent flux, we obtain the same event cross section. Thus, the LANSCE-equivalent flux can be used to compare the experimental data obtained at the Isis facility with those at LANSCE [64], the facility that comes the closest to matching the JEDEC JESD89A [40] reference spectra. This correlation is

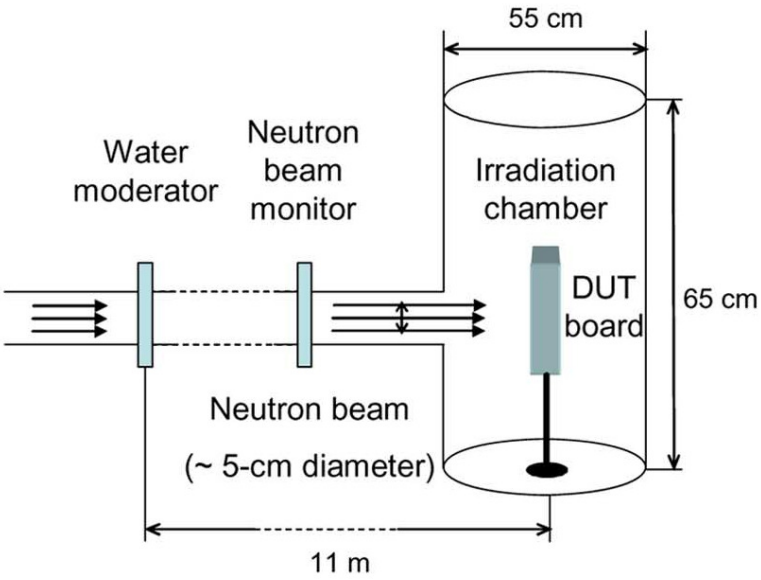


Figure 3.6: Schematic of the VESUVIO irradiation chamber.

necessary, because ISIS-VESUVIO features a different spectrum than LANSCE.

3.2.1.2 Devices

Used devices are Single Level Cell (SLC) and Multi Level Cell (MLC) commercial NAND and NOR Flash memories and some test chip. The feature size of the tested devices ranges from 90 nm to 48 nm, with a capacity, i.e. the array size, between 0.5 Gbit and 32 Gbit. Vendors of used memories are indicated only with capital letters (A, B, C), as many informations in this work are restricted. In the case of MLC devices, the cell levels are numbered from the lowest to highest V_{th} as L0, L1, L2, L3. Main characteristics of devices are summarized in Table 3.1.

In one set of samples only raw digital errors was measured, that means the flip of floating gate from its original logic state, "0" or "1", reading the memory before and after irradiation. In these kind of measurements non error code corrections (ECC) was applied. During irradiation devices were left unbiased. Unirradiated references devices were kept alongside the irradiated ones and measured at the same time. On another set of samples, in order to have cleaner and more physically significant information, the threshold voltage V_{th} of the floating gate was analysed through proprietary Test Modes.

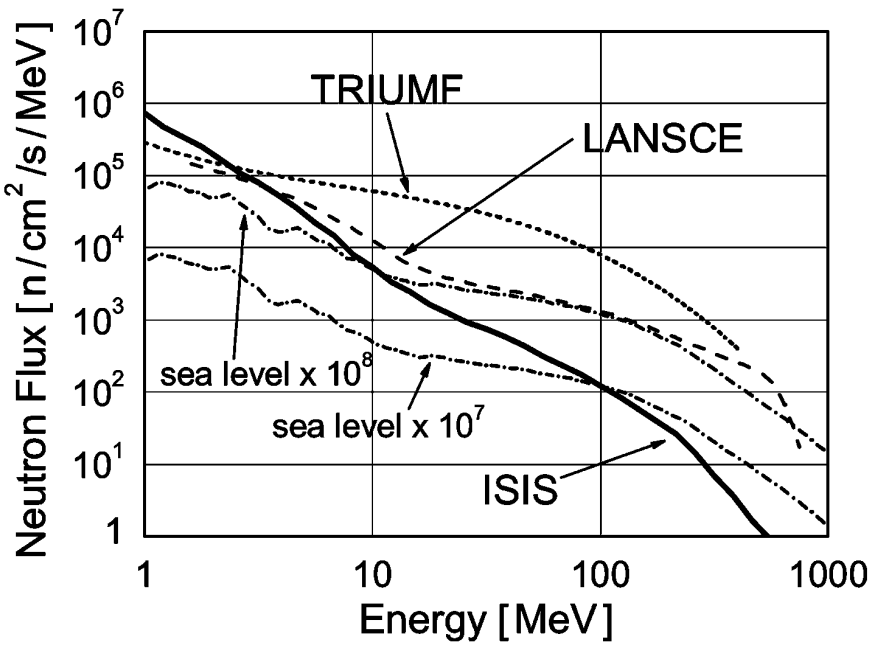


Figure 3.7: ISIS spectrum compared to those of the LANSCE and TRIUMF facilities and to the terrestrial one at the sea level multiplied by 10^7 and 10^8.

Vendor	Architecture	Feature Size [nm]	Capacity [Gbit]
A	NAND MLC	90	8
A	NAND MLC	65	4
A	NAND MLC	51	32
A	NAND SLC	65	4
B	NAND MLC	65	8
B	NAND MLC	50	32
B	NAND SLC	90	4
C	NAND MLC	60	8
C	NAND MLC	48	16
C	NAND SLC	48	1
C	NOR MLC	90	0.5
C	NOR MLC	65	0.5

Table 3.1: Main characteristics of devices used in this work.

3.2.2 Cross-section

Figure 3.8 shows the raw bit cross section (i.e. with no error correction code), σ, of neutrons induced errors, as a function of feature size for bits programmed at the highest level (L3) in MLC Flash memories. The bit cross section is defined as the number of errors divided by the fluence and by the total number of bits. Figure 3.9 illustrates the dependence of the cross section on the level. The most affected bits are those belonging to the two highest levels, L2 and L3.

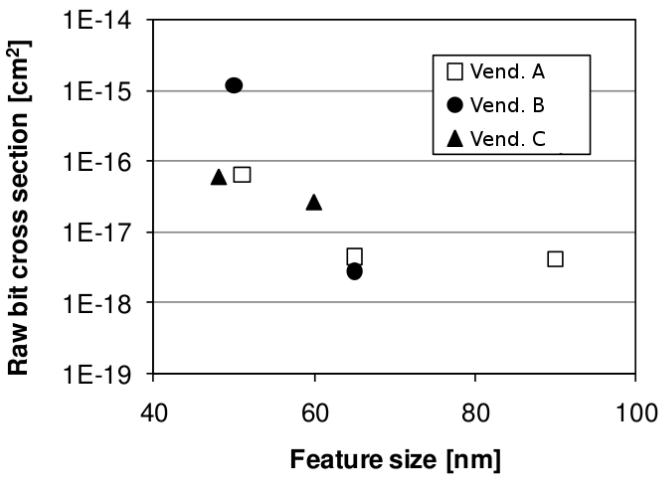


Figure 3.8: Cross section for raw bit errors induced by wide-spectrum neutrons in MLC NAND cells programmed at L3 in Flash memories from different manufacturers. Thousands of events were collected for each point. Error bars are smaller than the symbols.

σ increases more than one order of magnitude in two generations for vendor A. The cross section for the smallest feature size is less than 10^{-16} cm^2, a factor of 100 smaller than that for a typical SRAM cell [79]. The increase from one generation to the other is stronger in vendor B, more than three orders of magnitude. The absolute value of σ is higher in vendor B, as well, at 10^{-15} cm^2, for the worst-case pattern and smallest feature size. In all cases, the upsets are from one level to the adjacent one with lower V_{th}.

σ can be used to calculate the raw bit error rate on the field for a certain location, using data on the neutron flux. Assuming a value of $14\,\text{n}\cdot\text{cm}^{-2}\cdot\text{h}^{-1}$ at NYC [40], the raw bit error rate due to atmospheric neutrons over a period of ten years in the worst case (Vendor B, L3) is $1.14\cdot10^9$. This value is comparable to other mechanisms of bit errors [52]. The uncorrectable bit error rate is the probability that more errors will be generated in an ECC codeword than can

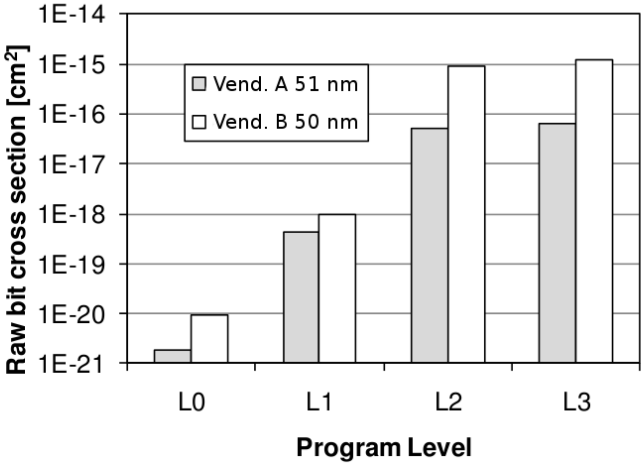


Figure 3.9: Cross section for raw bit errors as a function of the program level for two of the tested memories.

be corrected. Assuming an ECC scheme capable of correcting 8 bits and an ECC codeword of 539 bytes, as prescribed by the manufacturer datasheet, the uncorrectable bit error rate will be, according to the binomial distribution, much less than $1 \cdot 10^{18}$. At avionics altitudes, the neutron flux can be 300 times higher, resulting in a worst-case raw bit error rate of $3.4 \cdot 10^{7}$, which, translated into uncorrectable bit errors, is still well below $1 \cdot 10^{-18}$. The other tested memories exhibit lower raw bit error rates.

Figure 3.10 shows the sensitivity of SLC NAND Flash memories to neutrons. The three memories are produced different vendors (Table 3.1), which makes the comparison between different feature sizes somewhat less precise, and only the program state has been considered (the erase level is not sensitive). At a feature size of 65 nm, the raw bit cross section is more than three orders of magnitude below that of MLC devices. An increasing trend with scaling is visible, but we collected just a handful of events in the 65-nm parts, so no conclusion can be drawn about the real magnitude of the increase due to this large uncertainty.

Figure 3.11 illustrates the raw bit cross section for NOR MLC memories. Errors appear first at the 65 nm node, whereas no fails are detected in 90-nm cells. σ is two or three orders of magnitude less than that of NAND MLC, but larger than that of NAND SLC. An increasing trend with decreasing feature size is clearly visible in Figure 3.11.

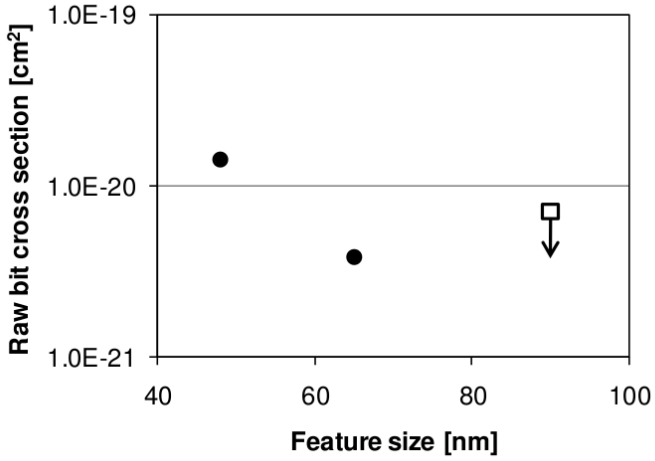


Figure 3.10: Cross section for raw bit errors induced by wide-spectrum neutrons in SLC NAND. The memories are produced by different vendors (Table 3.1). No errors have been observed on 90-nm parts, the reported value is an upper bound based on the number of tested cells and neutron fluence. Only one error has been measured on the 65-nm parts, which gives a large uncertainty on the result. On the contrary hundreds of events have been measured on 48-nm parts, with an error of about 5%.

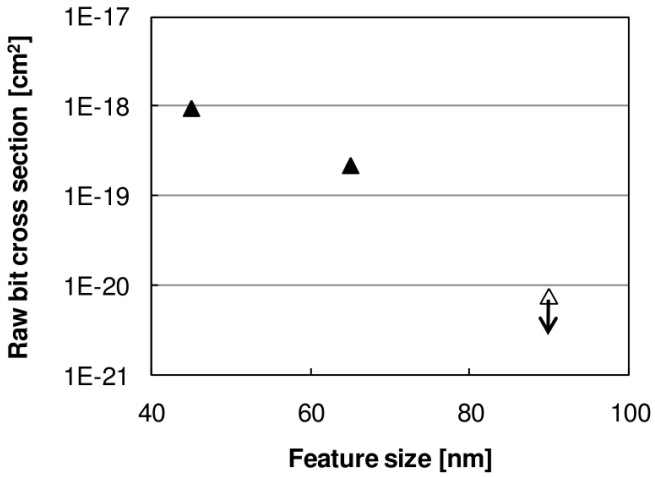


Figure 3.11: Cross section for raw bit errors (disabled internal ECC) induced by wide-spectrum neutrons in MLC NOR. No errors have been observed in 90-nm parts, the reported value is an upper bound based on the number of tested cells and neutron fluence. Tens of errors have been measured for 65 and 45-nm parts. No error is measured with internal ECC on.

3.2.3 V_{th} distribution

In addition to the number of errors, the voltage threshold, V_{th}, distributions were measured, before and after neutron irradiation in the set of Multi Level Cell NAND samples from Vendor C, where we have full access to the threshold voltage distribution. Figure 3.12 and Figure 3.13 show that in NAND technologies neutron irradiation results in the formation of low-V_{th} FGs. Increasingly longer tails appear in the distributions (Figure 3.12), for increasing cell V_{th}. Since ΔV_{th} is larger in devices programmed at higher average V_{th}, the error probability is larger for the logic states programmed at higher V_{th}. For this reason, the cell FIT for MLC devices depends on the status. SLC devices are less sensitive than MLC because of their larger noise margin.

A closer look at the V_{th} distribution gives precious insights into the physical mechanism leading to such errors (Figure 3.12). Since the main distribution does not move because of neutron irradiation, but a tail forms, from the previous discussion it is clear that neutron interactions with the device nuclei leads to recoils or nuclear reaction byproducts which, being ionizing, can cross a FG thus partially discharging it.

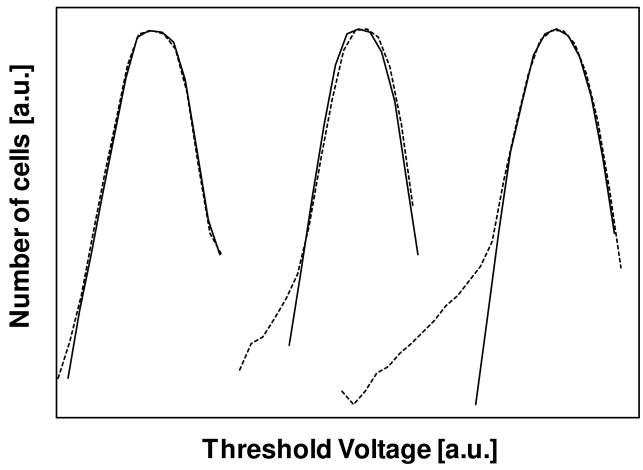


Figure 3.12: Threshold voltage distribution for a 60-nm MLC NAND Flash memory (Vendor C) irradiated with wide-spectrum neutrons at Isis facility.

Figure 3.13 compares the V_{th} distributions before and after irradiation for two different technology nodes. The two devices were irradiated with the same fluence ($1.68 \cdot 10^{10}$ n $\cdot$ cm^{-2}), so it is apparent how both the number of cells in the tail and the maximum V_{th} shift increase with decreasing feature size.

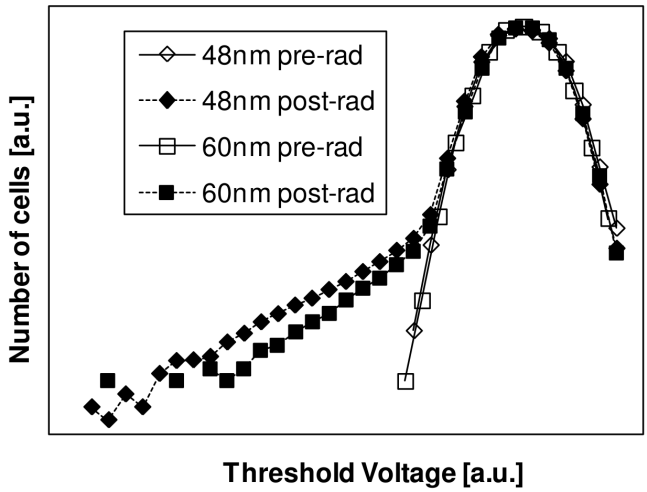


Figure 3.13: Comparison between the threshold voltage distributions before and after wide-spectrum neutron irradiation for two MLC NAND Flash memories (Vendor C) with different feature size.

Based on previous discussion on heavy ion effects, several peculiar aspects of neutron-induced errors can be explained, including [26] the larger sensitivity of MLC devices with respect to SLC, the larger sensitivity of program states corresponding to larger V_{th}, and the increasing sensitivity with technology scaling.

3.2.4 Scaling effects

Figure 3.14 shows the raw cross section, of errors attributable to neutrons as a function of the feature size for MLC NAND Flash memories manufactured by vendor A (see Table 3.1). The most affected bits are those belonging to L2 and L3. For these bits, increases more than one order of magnitude in two generations for the floating gate size from 65 nm to 51 nm. The value for the smallest feature size is less than 10^{-16} cm^{-2}, a factor of 100 smaller than that for a typical SRAM cell. Figure 3.15 shows a similar figure for the MLC NAND Flash memories manufactured by vendor B where the floating technology nodes are 65 and 50 nm. The increase from one generation to the other is stronger, more than three orders of magnitude. The absolute value of σ is higher as well at 10^{-15} cm^{-2}, for the worst-case pattern and smallest feature size.

Similar results are obtained with devices from vendor C (Figure 3.16). The variation of the raw cross section per bit from 60 nm devices to 48 nm devices

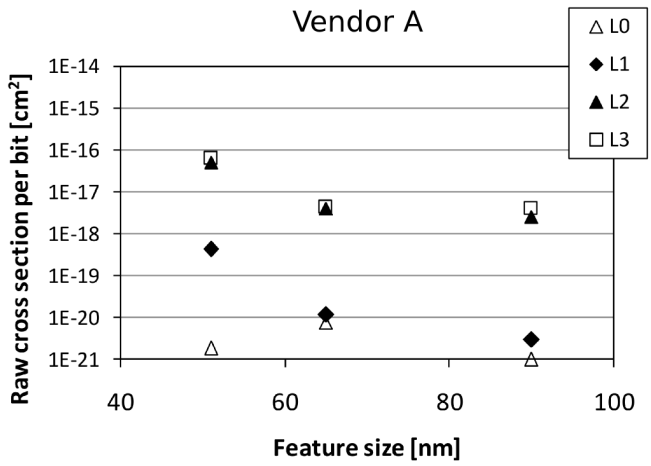


Figure 3.14: Cross section for raw bit errors induced by wide-spectrum neutrons in MLC NAND Flash memories of different features size from manufacturer A.

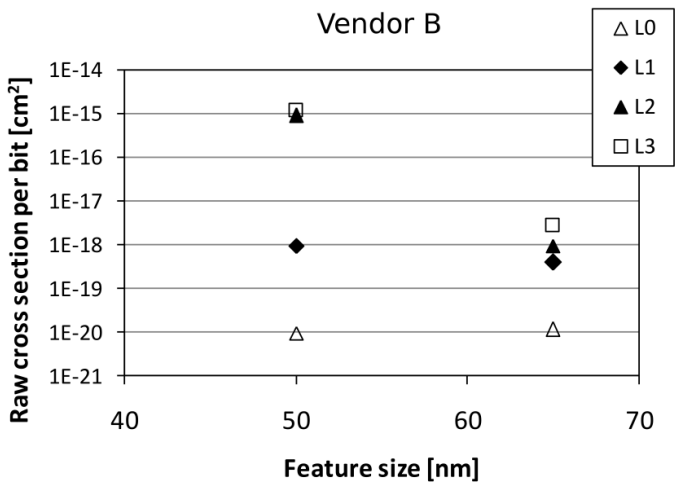


Figure 3.15: Cross section for raw bit errors induced by wide-spectrum neutrons in MLC NAND Flash memories of different features size from manufacturer B.

are lower than other vendors. In this case, the highest program level L3 shows a significant effect.

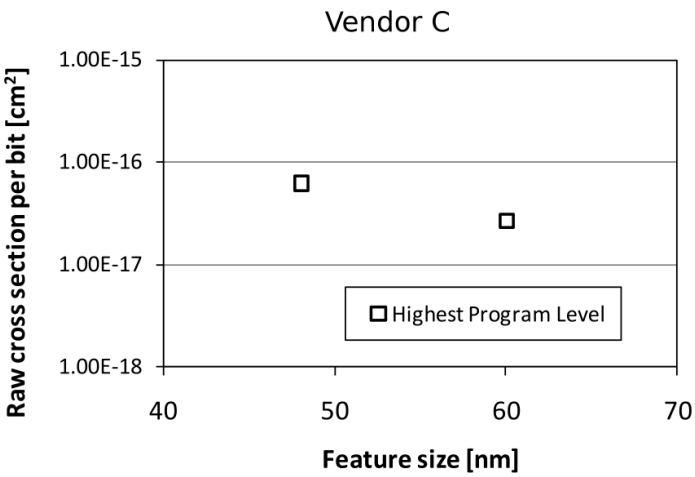


Figure 3.16: Cross section for raw bit errors induced by wide-spectrum neutrons in MLC NAND Flash memories of different features size from manufacturer C. Only the highest program level L3 is shown.

The scaling trend illustrated in Figures 3.14-3.16 and 3.13 of the previous section clearly points to an increased sensitivity towards neutron effects as the feature size is scaled, as predicted in [9]. The raw bit error rate increases with decreasing feature size, with large differences from vendor to vendor. This reflects the simple fact that the charge used to store a bit decreases with each new generation, whereas the ion track stays always the same. In one of the tested devices, the raw bit error for the most sensitive level increased by three orders of magnitude with just a generation. Hence, *the effect of scaling is to reduce the stored charge, and so, to increase* ΔV_{th}. That's the reason why neutron-induced errors are present in modern technologies, while they were not in older ones.

However, changes in materials, for instance the replacement of Tungsten with lighter elements, like Cobalt and Copper, could offset this trend. Considering the technology nodes of FG devices used in this work, the numbers are reassuring, and only in some conditions the neutron threat appears to be of some significance, but nowhere large enough to defy current ECC schemes.

Chapter 4

Geant4 simulation of neutron irradiation of Flash memories

The effects of neutrons radiation on modern electronic devices is currently recognized as a real problem both in space and in terrestrial environment. From this arises the need to understand neutron induced single event effects and to predict the reliability of new devices.

Much has been studied on devices such as SRAM and FPGA [79, 5, 46], or heavy ions radiation on floating gate memory [15, 56, 22], while little is known about neutrons effects in Flash memory. The previous chapter (see Chapter 3.2) shows the experimental proof that these effects are presents in modern Flash devices.

To better understand the physic behind measured effects, i.e. the bit errors or the V_{th} shift, a gate level simulation is required. Thus, simulations can provide important informations about the number, species and energy about the secondary particles crossing the floating gate; moreover can give us information about the regions of memory where these secondary particles are generated and about the materials mostly involved in nuclear process.

Nowadays, GEANT4 is one of the most powerful simulation toolkit for the interaction between radiation and matter and a lot GEANT4-based codes have been worked out in many research fields: like MULASSIS, MRED, CREAM and M^{2}EDUSA just to name a few [89, 87, 86]. Pia *et al.* [63] (2009) highlighted the growing role played by GEANT4 Monte Carlo simulations both in fundamental nuclear and technological research, in the last ten years. In this work a *Geant4-based code for the simulation*

neutrons radiation on Flash memory has been developed.

An overview of GEANT4 is given in Section 4.1 while the Section 4.2 explains how the code has been developed for the aim of this thesis. Results and data analysis obtained from simulations are exposed in Section 4.3.

4.1 Overview of Geant4

GEANT4 is a state-of-the-art radiation toolkit for the comprehensive Monte Carlo simulation of energetic particles from PeV energies to 100's eV (and thermal energies for neutrons)[2]. It is based on a collection of C++ classes, which contain a variety of utilities and physics processes to model the interaction of radiation with matter. It includes both electromagnetic and hadronic interactions.

The toolkit was initially developed for the high-energy physics community, primarily to support experiments for the CERN's Large Hadron Collider, but has found increasing use within other communities, such as medical physics, high-energy astrophysics and the space radiation effects community. Its object-oriented design and implementation in C++ means that the code is flexible and can be easily enhanced through class inheritance.

4.1.1 Structure

GEANT4 adopts object-oriented technology with a clear hierarchical structure of categories and sub-categories, linked by a uni-directional flow of dependencies. Domains represented in GEANT4 have the structured shown in Figure 4.1. The functions of the different shown categories may be briefly summarized as:[1]

Run and Event concern the simulation of the cinematic of the *event* with the primary and secondary trajectories.

Tracking manages the transport of the particle described by the physical interactions.

Hits and Digit provide the capability to reproduce the structure of the detector read-out electronics and its response.

[1]For a complete description see Agostinelli e *et al.* [2] (2003) or the GEANT4 documentation on `http://geant4.cern.ch/support/userdocuments.shtml`.

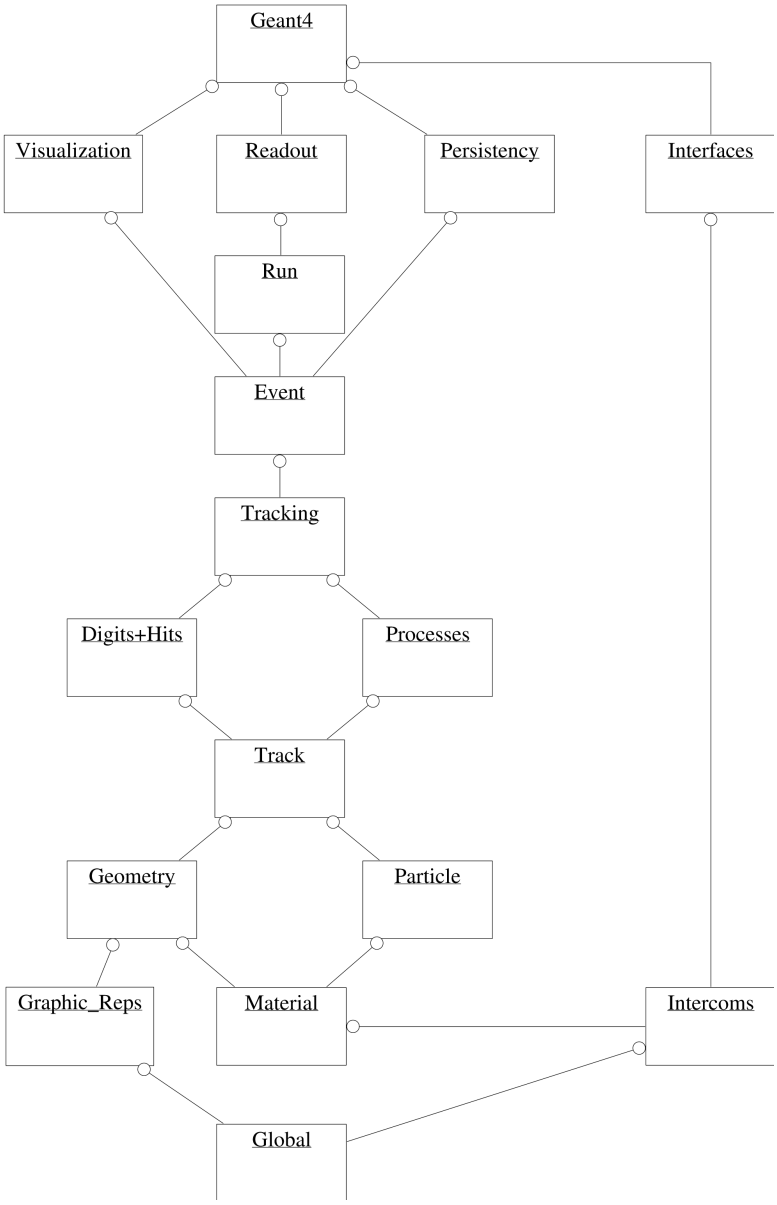


Figure 4.1: Diagram of the logical categories structure of GEANT4 structure [2].

Geometry allows the user to model the experimental set-up in detail. Representations of complex solid geometries and Boolean operations between objects are possible.

Particle Definition and Material manage the definition of materials and particles.

Processes is the domain that manages all the physical processes involved in interactions between particles and matter. Contains the implementation of all the physical models, divided according to the range, type of particle and material that crosses. Cross sections of the physical process involved in the interaction are transparently calculated or through analytical calculation or by using tabulated experimental data.

4.1.2 Physics

The user is able to apply a wide variety of different physics models to treat, for example, energetic nuclear-nuclear collisions, radioactive decay, low-energy electron-photon transport, and low-energy neutron transport. The following is a summary of the physics currently covered or being implemented in GEANT4:

- Hadron-Parameterized; parton-string (5 GeV); kinetic (10 MeV–10 GeV); QMD models; precompound ($>$ 100 MeV); Low-energy neutron (thermal–20 MeV); isotope production;
- Nuclear De-excitation; evaporation ($A > 16$); Fermi break-up ($A \leq 16$); Fission ($A \geq 65$); photo-evaporation (ENSDF); radioactive decay (ENSDF);
- Electromagnetic-ionization; multiple scattering; δ-ray production; Bremsstrahlung; annihilation; photo-electric effect; Compton scattering; Rayleigh scattering; pair-production; atomic relaxation;
- Others-Cerenkov radiation production; scintillation; transport of optical photons.

In GEANT4, neutron transport at energies below 20 MeV is based on empirical elastic and inelastic neutron cross-section data drawn from many of the standard database sources. Models that are based on theory can typically be applied to any material. Those that utilize databases, such as the low-energy neutron transport

and photo-evaporation processes, are limited by the availability of data for specific materials. These include commonly used spacecraft and electronic materials (H, C, N, O, Al, Si, Ti, Fe, Ni, Cu, Ga, As, Ta, and W).

Ionization due to the recoil nucleus and secondary charged nucleons and nuclear fragments includes accurate treatment of low-energy electromagnetic processes using:

1. Electronic stopping powers for protons and α-particles based on ICRU-49.
2. Multiple scattering.
3. δ-ray production from ions.

Recent enhancements to the electron-photon physics mean that the lower-energy threshold for electron transport is $\approx 100\,\mathrm{eV}$.

4.2 Geant4 toolkit for neutron irradiation of Flash memories

In this thesis, an application based on GEANT4 toolkit to simulate the interaction of neutron particles with the chip materials has been developed. The output of this tool are the number, species, energy and trajectory of the particles that cross the floating gate.

The GEANT4 version used in this work is *geant4.9.4* installed on a *Linux* system.

4.2.1 Geometry and materials

The geometry and materials of the different Flash memories samples were modeled in detail in three dimensions. Three technological nodes devices from Vendor C have been chosen and are listed below:[2]

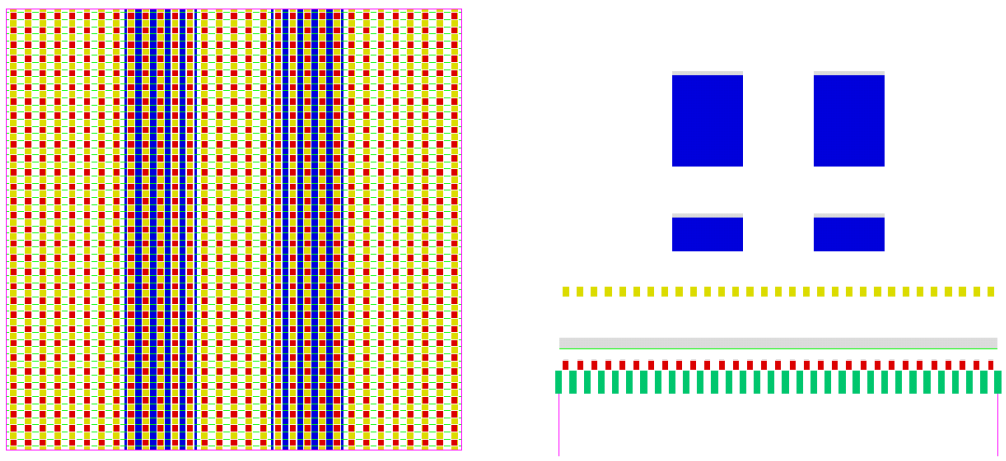
FG60 60 nm NAND SLC Flash memory

FG48 48 nm NAND SLC Flash memory

[2]Devices are different from those reported in Table 3.1.

FG32 32 nm NAND SLC Flash memory

Each model consists of a matrix of 31 × 31 floating gate cell. A 2D drawing of the examined structure (not to scale) is shown from the top in Figure 4.2(a) and from a side in Figure 4.2(b). The different part of the modeled structure are



(a) (b)

Figure 4.2: 2D drawing of the 31 × 31 structure (not to scale) of the modeled NAND Flash memory shown from the top (a) and from the side (b).

claer in Figure 4.2. The shallow trench isolation (STI) is coloured in green while the red squares indicate the Floating Gate cells. Over the gate level, the wordline is represented in grey and the bitline in yellow, while the other metal layers are coloured in blue. The materials that surround the lines are not represented for simplicity. Figure 4.3 show the cross section model, not to scale, of the simulated Flash memory devices. All the components of a complete Flash memory chip as been modelled so as to take into account the secondary particles generated from all layer and materials, i.e. substrate, wordline, bit line, isolation, metal layers and passivation. The dimensions and materials of the model reported in Figure 4.3 can be summarised as follow:[3]

Dimensions

- Floating gate size: $32 - 60$ nm
- Tunnel oxide thickness: ≈ 10 nm
- Floating Gate thickness: $\approx 50 - 150$ nm

[3]Shown values are approximate as they are confidential.

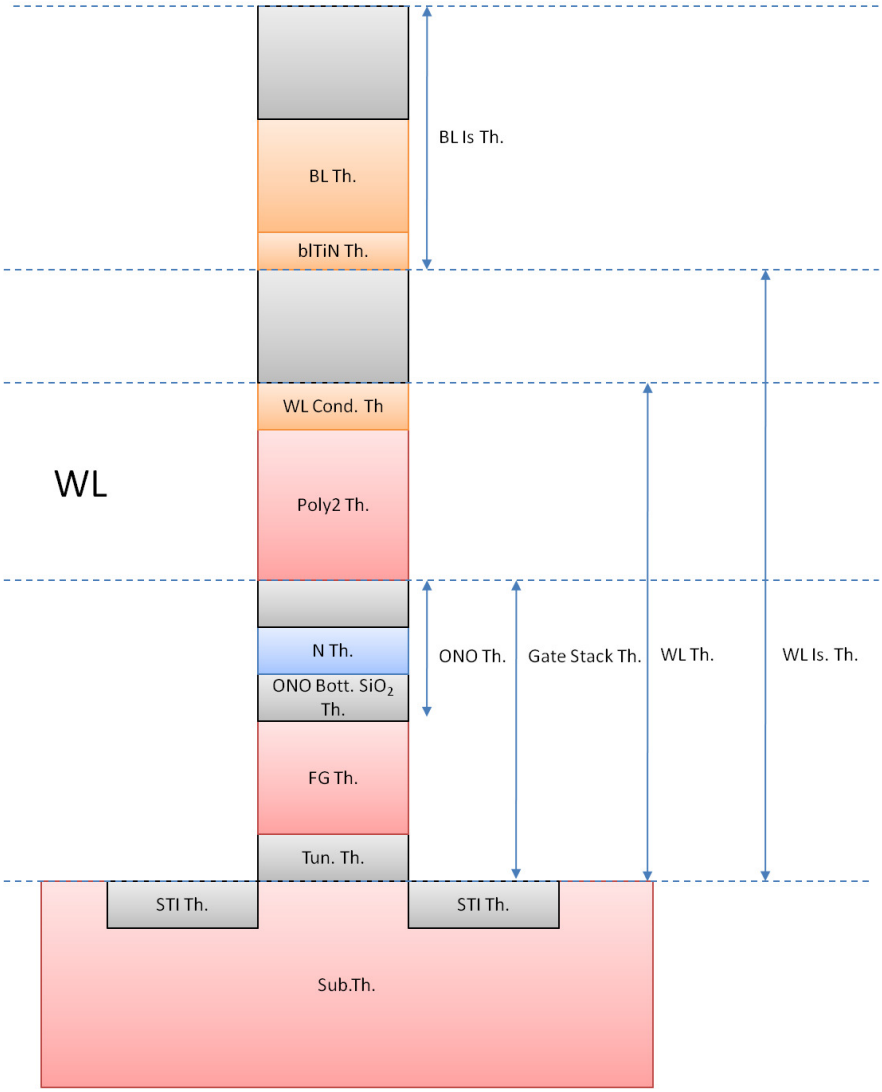


Figure 4.3: Cross section model, not to scale, of the simulated Flash memory devices. All the components of a complete Flash memory chip as been modelled so as to take into account the secondary particles generated from all layer and materials, i.e. substrate, wordline, bit line, isolation, metal layers and passivation.

- Metal layers: $\approx 50 - 1000\,\text{nm}$ (depends on layer)
- Substrate thickness: $\approx 500\,\mu\text{m}$
- Passivation: $> 5\,\mu\text{m}$

Materials

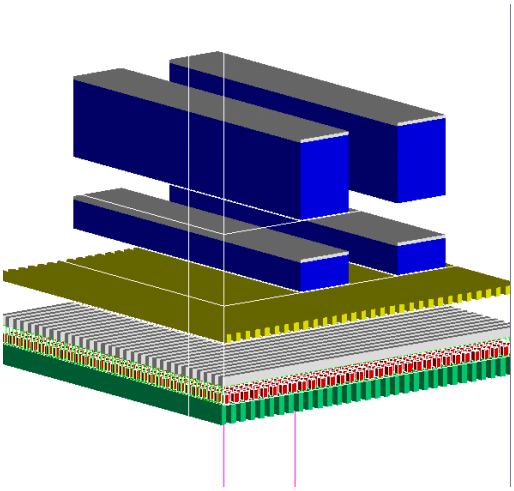
- Isolation: SiO_2
- Substrate: Si
- Floating gate: $\text{Si}, \text{SiO}_2, \text{N}$
- Wordline: polySi
- Bitline, metal layers: $\text{AlCu}, \text{Ti}, \text{TiN}, \text{TiW}$

Three dimensional views of the Flash memory structure modeled with GEANT4 geometry are shown in Figure 4.4. It is clear that the differences between the three memories *FG32*, *FG48* and *FG60*, are not only the floating gate dimensions, but also the layout of the metal layers. In particular, the *FG60*device has one more metal layer with respect to the other technological nodes.

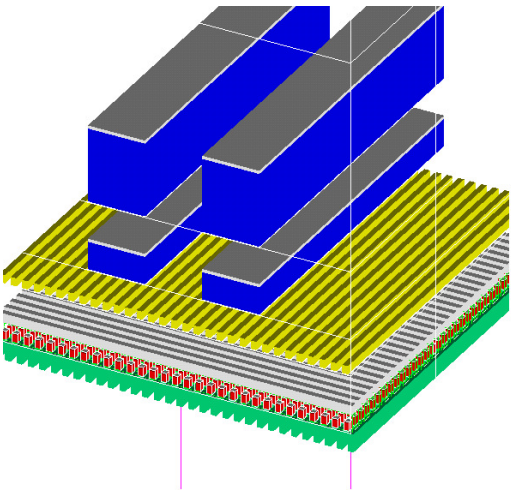
4.2.2 Sensitive detector

Once the 3D model of the simulated devices is defined, is necessary to indicate the *sensitive detector* and the *cut range*. A sensitive detector either constructs one or more hit objects or accumulates values to existing hits. For the simulation of neutron radiation on Flash memories, the floating gate (FG) has been set as the sensitive detector. Thus, an event is defined as the deposition of energy inside the floating gate. For each event in the floating gate, the species, the kinetic energy, the deposited energy and the trjectories of all particles generated are recorded (for details see Appendix A.2). Figure 4.5 shows a zoom of the 3D structure, where the is clear the red cubes indicate the *sensitive detector*, i.e. the floating gate.

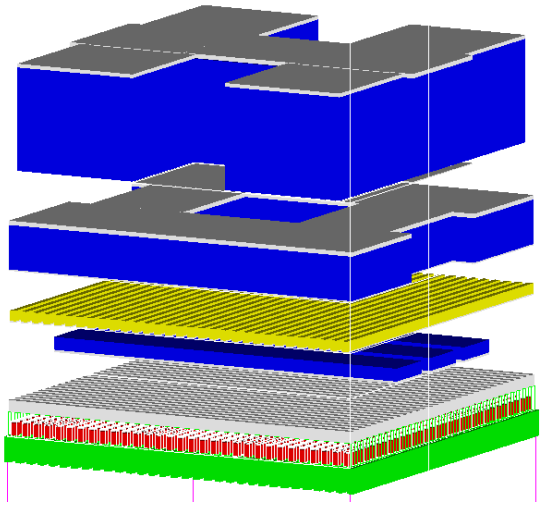
GEANT4 treats energy deposition by ionization as the sum of a continuous deposition and discrete events in the sensitive detector, whose production depends on the range cut set by the user. When the ionization process produces electrons whose range is above the cut, these electrons are tracked, otherwise the energy



(a)



(b)



(c)

Figure 4.4: 3D drawing of the modeled NAND Flash memory: 60 nm Floating gate (*a* and *b*) and 48 nm Floating gate (*c*).

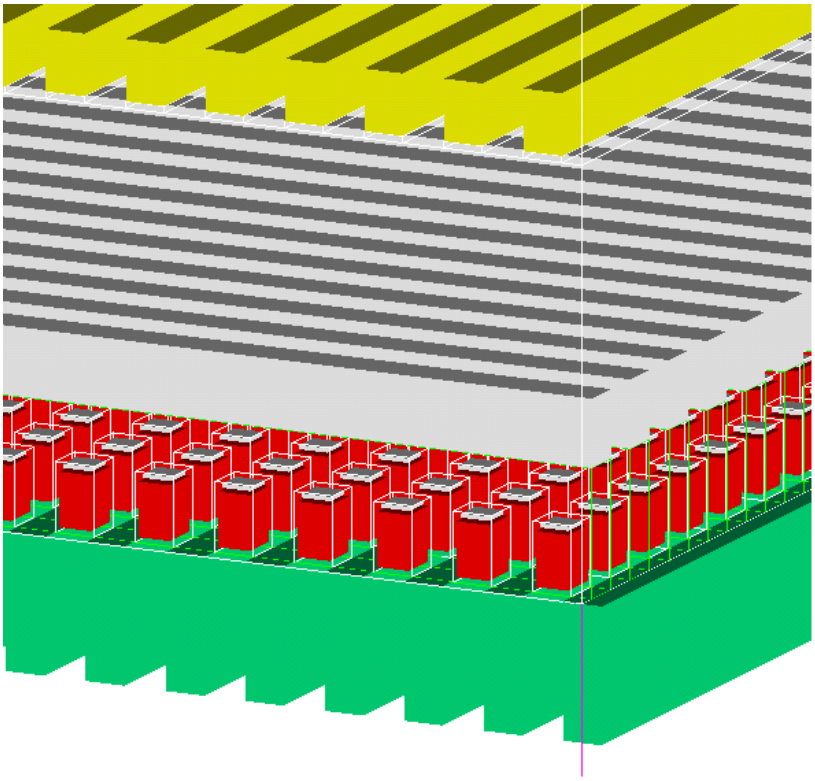


Figure 4.5: Detail of the 3D model showing the sensitive detector as the floating gate (red).

deposition associated with them becomes part of the continuous component. The range cut for electrons was set to 1 nm, so that energy deposition in the FG is possible not only for ions directly crossing the FG, but also for particles passing close enough to generate electrons that can reach the FG.

4.2.3 Particles spectrum

The ISIS and the terrestrial neutron energy spectra were carefully reproduced in the simulations. $2.8 \cdot 10^9$ incident neutrons were simulated in *normal* incident irradiation in two condition: *front* irradiation, i.e. from the top of the device, and *back* irradiation, i.e. from the bottom (the substrate) of the memory device.

To reduce the total amount of events to be simulated, while producing enough high energy neutrons events, the ISIS and terrestrial spectra were divided in three parts:

- low: $2\,\text{MeV} < Energy < 10\,\text{MeV}$
- medium: $10\,\text{MeV} < Energy < 100\,\text{MeV}$

- high: $Energy > 100$ MeV

Results were combined afterwards applying the proper weight to the results coming from different simulations.

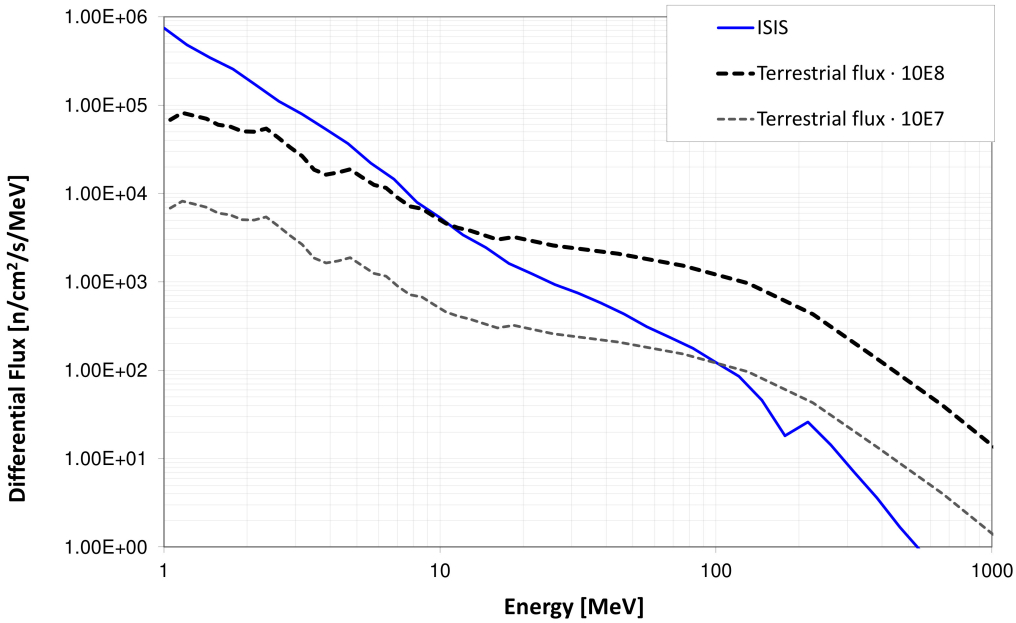


Figure 4.6: Neutron Isis spectrum as compared to the IEC reference terrestrial spectrum multiplied by a factor 10^7 and 10^8.

In Figure 4.6 the neutron Isis and the IEC reference terrestrial spectrum multiplied by a factor 10^7 and 10^8, which are used in the simulations, are compared.

4.3 Neutron radiation simulations

In this section, the main results obtained with the developed Geant4-based software (described in the previous section) are reported. The following graphs will point out the number of events as function of the effective LET, the energy distribution, the species and the origin location of the secondary particles produced by neutrons for three memories *FG60*, *FG48* and *FG32*. Moreover, a comparison between the Isis and terrestrial spectrum, the dependence of incident angle radiation and the effect of floating gate dimension are shown.

Simulations have been performed on an Intel ® Core™2 Duo Processor E6400 (2M Cache, 2.13 GHz) with 2MB RAM. Each complete simulation have been

sectioned into shorter *runs* of $2 \cdot 10^8$ incident neutrons, so as to simplify the analysis and to improve the performance[4]: for the lower part of spectrum 14 short simulation were executed while for the high energy part of the spectrum, only 8 simulations were enough.

Each simulation can be indicated with a kind of device (*FG60*, *FG48* or *FG32*), the kind of neutron spectrum (ISIS or terrestrial) and the incident angle ($0°(front)$ and $180°(back)$).

Data analysis have been carried out with MATLAB.

4.3.1 Data analysis

The result of simulation is a list events, which are recorded when an amount of energy is deposited inside the floating gate (FG) after a particle crossing. The extrapolated data for each event are

i the *species* of generated particle,

ii the *kinetic energy* of the particle,

iii the *ionizing deposited energy*,

iv the *non ionizing deposited energy*,

v the *volume* where the particle is generated,

vi the in and out *trajectory* of generated particle.

Then, to filter and elaborate the data, the parameters listed in Table4.1 have been used. The *threshold energy*, E_{th} indicates the LET above which an events is considered. Hence, the first step of data processing is the filtering of all stored events with this *threshold energy*. Then, filtered data are compressed: when different particles are generated from the same neutron, or when different events happened in the same cell, their deposited energy are summed in a single event.

From the parameters in Table 4.1, the following quantities are calculated for the statistical analysis. In order to normalize the effects in different devices we need the total amount of simulated particles

$$particles_{sim} = runs \times 2.0 \cdot 10^8,$$

[4] A single *run* takes from 5 to 8 hours, depending of the part of the spectrum

and the overall irradiated area

$$area_{irr} = (FG \times 2 \times FG \times 2) \times cells \times 10^{-14}\text{cm}.$$

To compare the simulated data of the terrestrial spectrum and the ISIS spectrum with the LANSCE one, the equivalent fluence, and the equivalent number of particles have been calculated

$$\Phi_{eq} = \frac{5.82}{7.44} \times \Phi_{LANSCE}$$

$$particles_{eq} = area_{irr} \times \Phi_{eq}$$

giving the scaling factor

$$scale = \frac{Particles_{eq}}{Particles_{sim}}.$$

To recombine the spectrum with the right weights we need the following ratios

$$R_{\text{high/medium}} = \frac{flux_{\text{MeV}>10}}{flux_{10<\text{MeV}<100}},$$

$$R_{\text{low/medium}} = \frac{flux_{2<\text{MeV}<10}}{flux_{10<\text{MeV}<100}}$$

and

$$W_{\text{medium}} = 1.$$

Finally with can combine the overall number of events and the energy as a function of the effective LET, LET_{eff}, from the different simulations

$$events = (e_{\text{low}} \cdot R_{\text{low/medium}} + e_{\text{low}} \cdot W_{\text{medium}} + e_{\text{high}} \cdot R_{\text{high/medium}}) \cdot scale$$

and

$$energies = (E_{\text{low}} \cdot R_{\text{low/medium}} + E_{\text{low}} \cdot Weigth_{\text{medium}} + E_{\text{high}} \cdot R_{\text{high/medium}}) \cdot scale$$

where e_{low}, e_{medium} and e_{high} (E_{low}, E_{medium} and E_{high}) are the recorded *events* (*energies*) from the three different part of the spectrum, and the effective LET is calculated as the deposited energy E_{dep} divided by the density of Silicon and by the FG thickness

parameter	value	description
FG	(depends on device)	floating gate size
T_{FG}	95 nm	FG thiskness
W_{cell}	(depends on device)	cell width
L_{cell}	(depends on device)	cell length
$cells$	4539480000	number of cells
ρ_{Si}	$2.329\,\mathrm{g}\cdot\mathrm{cm}^{-3}$	Silicon density
$runs$	$8-14$	number of performed runs
Φ_{LANSCE}	$1.68\cdot 10^{10}\,\mathrm{n}\cdot\mathrm{cm}^{-2}$	LANSCE Fluence
$E_{\mathrm{dep,th}}$	0.02 ($LET = 1\,\mathrm{MeVcm}^{-2}\mathrm{mg}^{-1}$)	threshold energy for event definition

Table 4.1: Parameters used in the analysis of simulation data.

$$LET_{\mathrm{eff}} = E_{\mathrm{dep}} \cdot \rho_{\mathrm{Si}} \cdot T_{\mathrm{FG}}$$

and E_{dep} is deposited energy of each event after the filtering and compressing process.

4.3.2 Events and energy

Figure 4.7 shows the number of the secondary particles crossing the FG as a function of the effective LET for the 32 nm Flash memory device. The effective LET was calculated as the deposited energy divided by the density of silicon and by the FG thickness. The two curves represent the case of particles with energy above 2 MeV (solid blue) and above 10 MeV (dotted gray). Is important to note that even neutrons with energy below 10 MeV give an important contribution on total events with effective LET above $1\,\mathrm{MeV}\cdot\mathrm{cm}^{-2}\mathrm{mg}^{-1}$. The difference is nearly one order of magnitude. Figure 4.8 shows the same behaviour for the 48 nm and 60 nm.

As seen in Figure 4.7, the distributions are approximately linear in a log-lin scale and the maximum LET is around $1\,\mathrm{MeV}\cdot\mathrm{cm}^{-2}\mathrm{mg}^{-1}$ for all technologies. The smallest LET bin, below $1\,\mathrm{MeV}\cdot\mathrm{cm}^{-2}\mathrm{mg}^{-1}$, contains a lot more particles than the other bins, due to the large amount of generated protons and alpha particles. These lighter particles can travel a longer distance than those with higher LET, which have lower initial energy and lose their energy faster along the way.

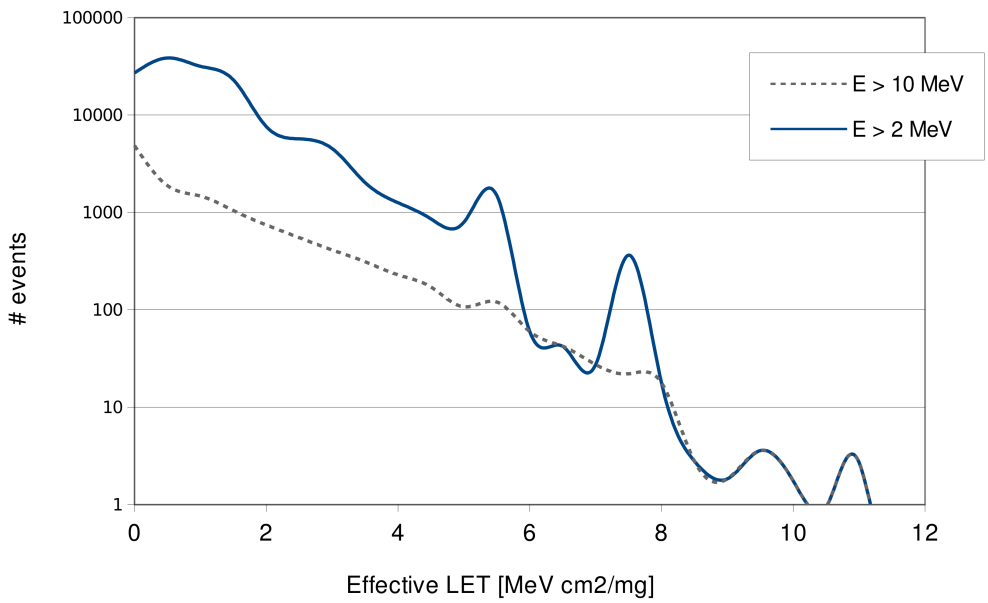
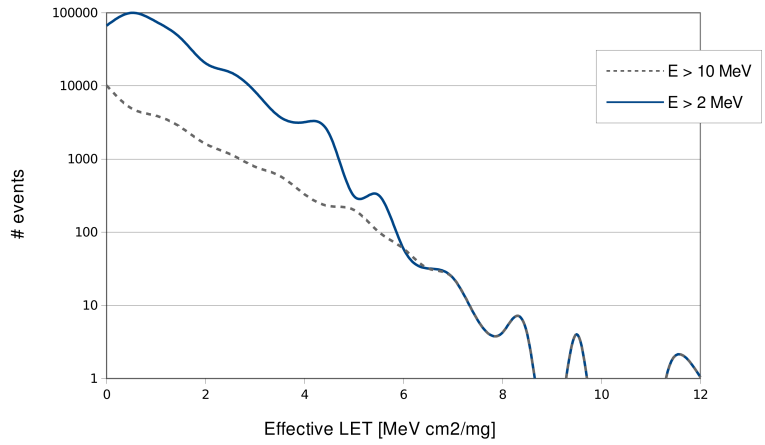


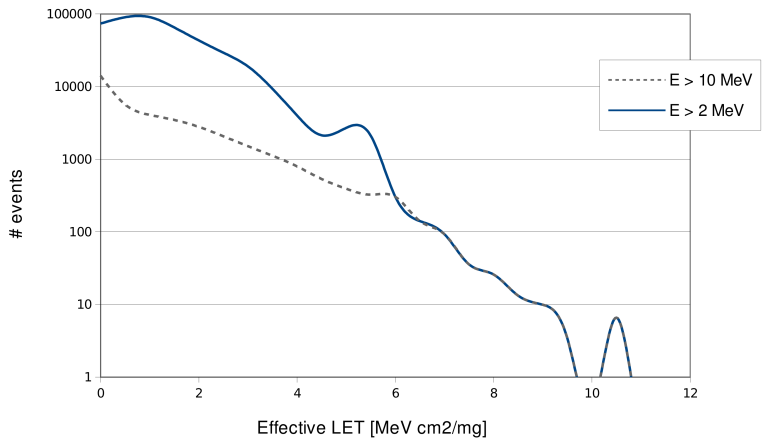
Figure 4.7: Simulated distribution of neutron-induced secondary particles LET in 32 nm Flash memory device. The two curves represent the case of particles with energy above 2 MeV (solid blue) and above 10 MeV (dotted gray).

Figure 4.9 shows a comparison of simulated distribution of neutron-induced secondary particles LET between different Flash memory technological nodes. Because of the larger geometrical size, the total number of events (the integral of the curves in Figure 4.9) is larger in the 60 nm device than in the 48 nm and 32 nm ones. However, the effect of an event with LET above $1\,\mathrm{MeV}\cdot\mathrm{cm}^{-2}\mathrm{mg}^{-1}$ can be higher in smaller device. The shapes of the two curves are different because of the different thickness of the materials the particles cross before reaching the floating gate.

Figure 4.10 shows the energy distribution of the particles with LET above $1\,\mathrm{MeV}\cdot\mathrm{cm}^{-2}\mathrm{mg}^{-1}$ produced by the neutrons in all the simulated devices, with the Isis spectrum. As seen, the energies are fairly limited, most of the particles are below 10 MeV. This is an important aspect, because the structure of the $e-h$ track generated by the ions is dependent on the energy, in addition to LET [25]. Furthermore, low energies mean that the neutron-triggered events are not likely to affect many bits, and that long tracks of adjacent FG cells experiencing charge loss (such as those observed in Cellere *et al.* [23] (2007a)) are unlikely with neutrons.



(a)



(b)

Figure 4.8: Simulated distribution of neutron-induced secondary particles LET in 48 nm (a) and 60 nm (b) Flash memory devices. The two curves represent the case of particles with energy above 2 MeV (solid blue) and above 10 MeV (dotted gray).

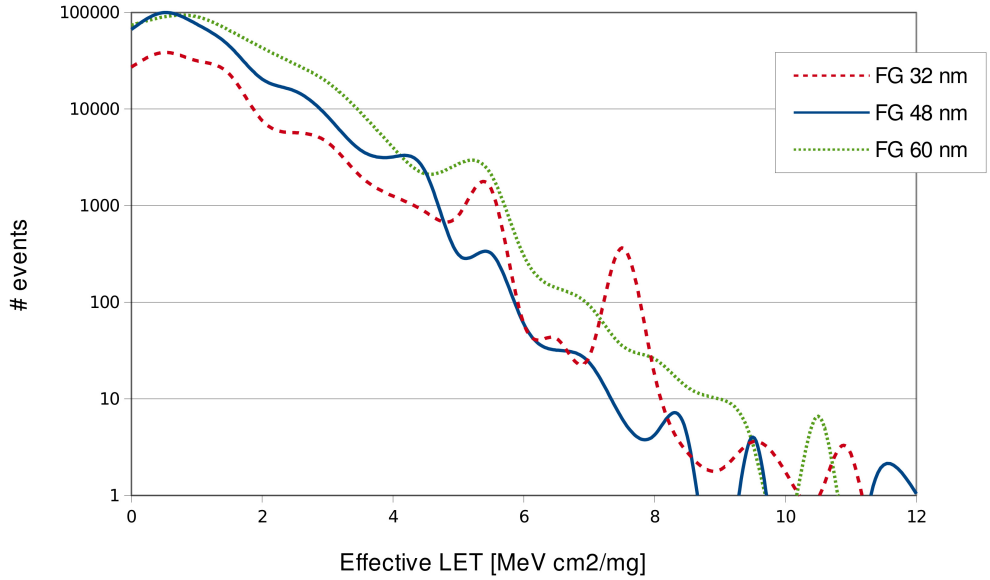


Figure 4.9: Comparison of simulated distribution of neutron-induced secondary particles *LET* between different Flash memory technological nodes.

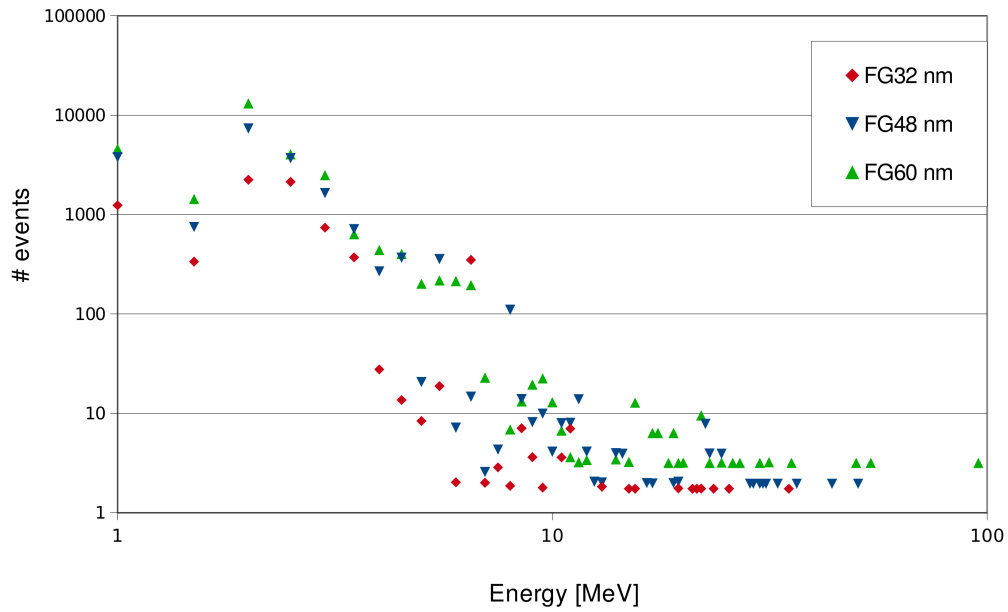


Figure 4.10: Simulated distribution of neutron-induced secondary particles energy in 32 nm, 48 nm and unit60nm Flash memory devices. ISIS neutron spectrum is used.

4.3.3 Particles

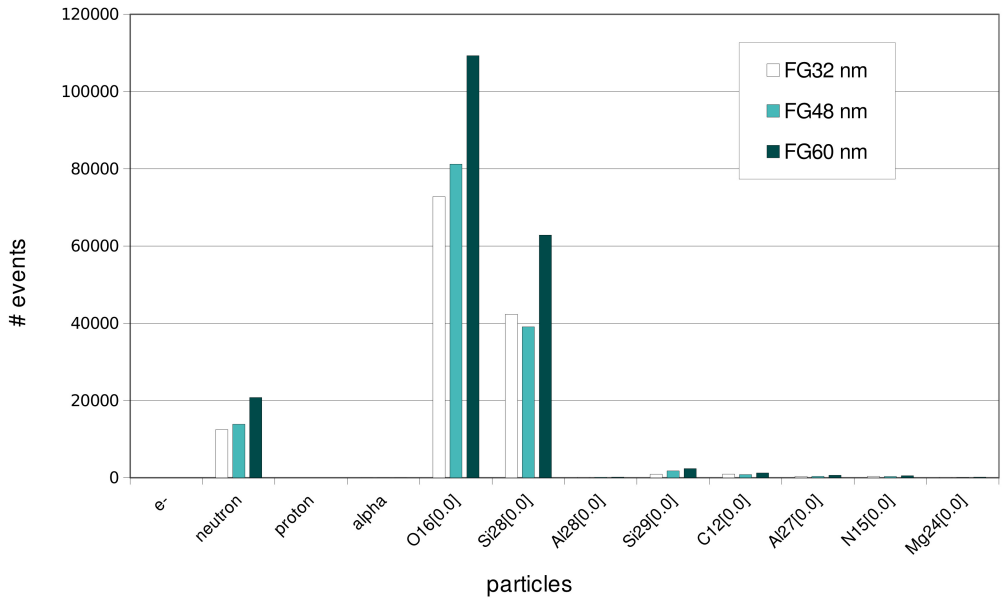


Figure 4.11: Simulated secondary particles distribution with effective $LET > 1\,\mathrm{MeV} \cdot \mathrm{cm}^{-2}\mathrm{mg}^{-1}$ crossing floating gate in 32 nm, 48 nm and 60 nm Flash memory devices.

As shown in Figure 4.11, simulation results show that the largest part of the events is due to reactions between neutrons and Silicon or Oxygen atoms. Although neutron nuclear interactions with Si and O are by far the most common, byproducts originating from reactions with high-Z materials employed in the CMOS flow (such as Tungsten) can be very important in certain situations [30], because they generate high-LET particles. However, presented simulations show that vents with Tungsten and other heavier materials in the back-end are much more rare, and contribute to the total neutron-induced number of events for less than 1%. Interactions with heavier materials might have been significant a few generations ago, when the threshold LET for an error was far higher, so that these events were the only ones able to produce an error. Given the relatively low threshold LET necessary to produce an error (see Section 4.3.5), changes in materials (for instance the replacement of Tungsten) in future technology nodes are not expected to play a major role on the sensitivity to neutrons (as long as Silicon and Oxygen will be the two most common materials).

The three simulated devices show a similar behaviour in byproducts generation.

The overall number of interactions with same material is larger in 60 nm device because of the large floating gate size.

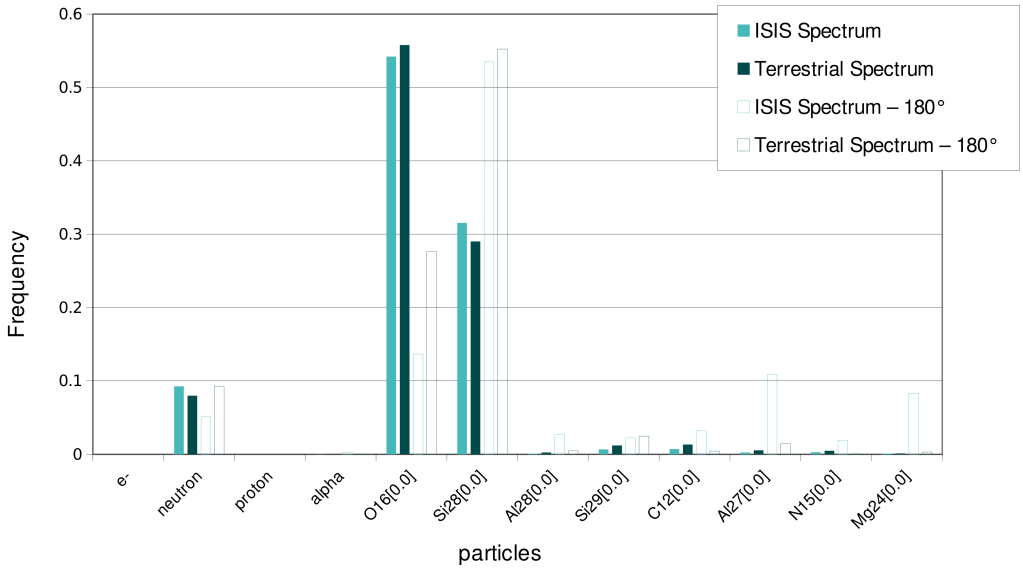
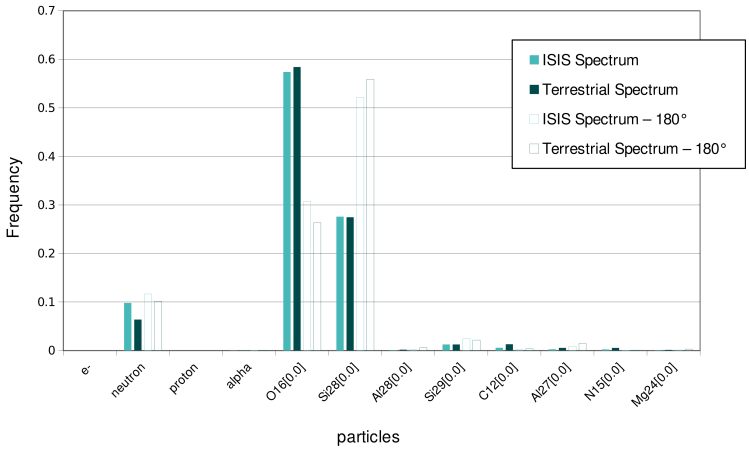


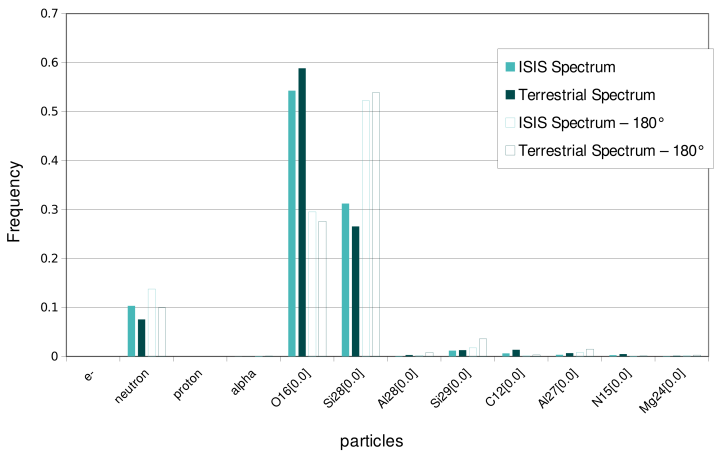
Figure 4.12: Simulated secondary particles distribution with effective $LET > 1\,\mathrm{MeV} \cdot \mathrm{cm}^{-2}\mathrm{mg}^{-1}$ crossing floating gate in 32 nm Flash memory device. The graph compares interaction with materials between Isis and terrestrial spectra, with a *front* and *back* radiation.

Figure 4.12 compares the distribution between the Isis spectrum and the terrestrial spectrum in the *FG32*. The histograms compares also the effect of a *front*, i.e normal incident, and *back* radiation. The species of particle are the same above indicated. With the *front* radiation the results from Isis and terrestrial spectra are nearly the same, also in terms of the numbers of events, supporting the fact that the Isis source reproduces the terrestrial one very well.

A clear difference can be seen, with both Isis and terrestrial spectra, between the *front* and *back* radiation. Most of events concern ^{28}Si particles as compared to ^{16}O particles. This fact can be explain considering that in the *back* radiation neutrons cross the very thick Silicon substrate, and thus interact more with Silicon atoms. Instead, Oxygen atoms are present only in area above the substrate, near the floating gate, and interact easily with neutrons coming from the top. Moreover, considering also the results presented in Section 4.3.4, events come for the most part from the oxide isolating the wordlines, rather than the wordlines themselves. Figure 4.13 show similar results in 48 nm (a) and 60 nm (b) Flash memory devices.



(a)



(b)

Figure 4.13: Simulated secondary particles distribution with effective $LET > 1\,\mathrm{MeV} \cdot \mathrm{cm}^{-2}\mathrm{mg}^{-1}$ crossing floating gate in 48 nm (a) and 60 nm (b) Flash memory devices. The graph compares interaction with materials between ISIS and terrestrial spectra, with a *front* and *back* radiation.

4.3.4 Byproducts origin locations

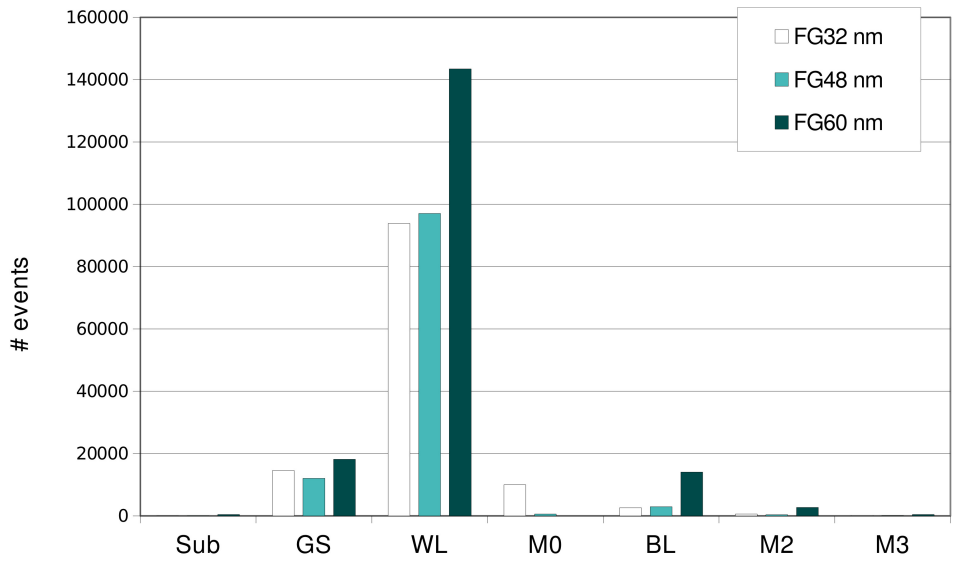
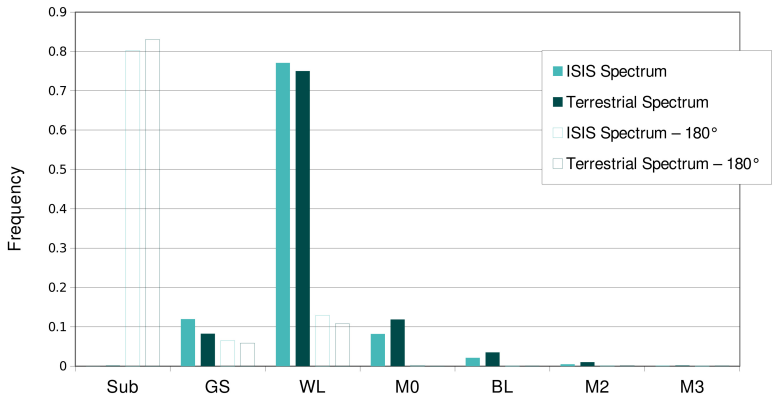


Figure 4.14: Simulated origin location distributions of the neutron byproducts with effective $LET > 1\,\mathrm{MeV} \cdot \mathrm{cm}^{-2}\mathrm{mg}^{-1}$ crossing floating gates in 32 nm, 48 nm and 60 nm Flash memory devices. Sub = Substrate, GS = Gate stack, WL = Wordline, BL = Bitline, Mx = Metal layer x. Each layer contains also the isolation up to the next layer going from the substrate to the passivation. Note that M0 is not present in 60 nm device.

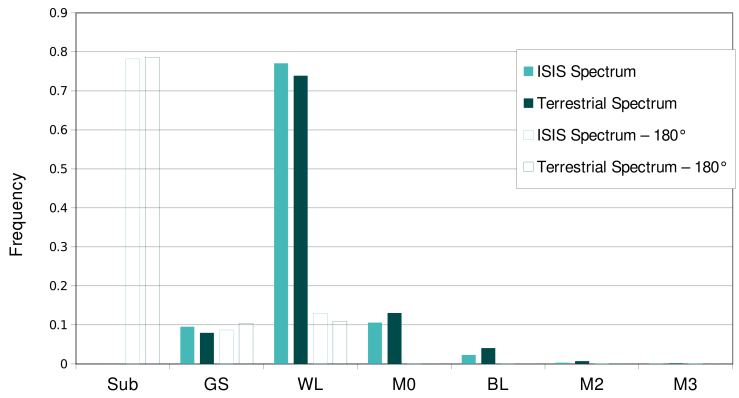
Another interesting correlated aspect is shown Figure 4.14, which reports the locations in the chip where ionizing secondaries originate during exposure to neutrons. Most of the events are due to particles generated in the proximity of the worldines, whereas, as expected, the other layers give a decreasing contribution as their distance from the floating gate increases [10]. As seen in the previous section (Section 4.3.3), where the *front* radiation shows an higher generation of Oxygen byproducts compared to other materials, events come for the most part supposedly from the oxide isolating the wordlines, rather than from the wordlines themselves.

4.3.5 ΔV_{th} and LET relation

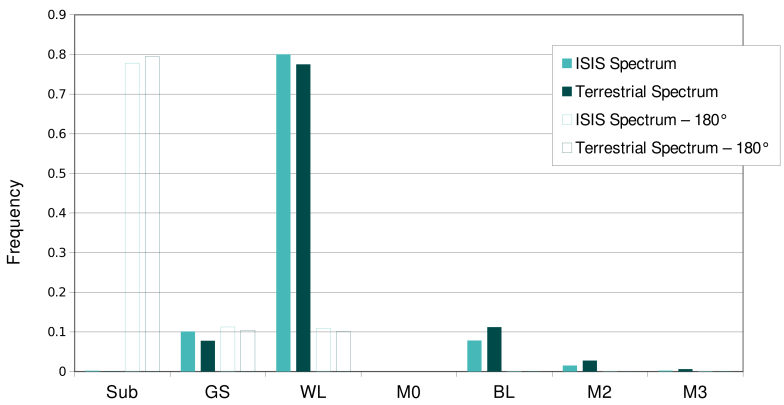
As we have discussed in the previous section, neutrons interact with chip materials through elastic or inelastic nuclear reactions that generate charged



(a)



(b)



(c)

Figure 4.15: Simulated origin location distributions of the neutron byproducts with effective $LET > 1\,\mathrm{MeV} \cdot \mathrm{cm}^{-2}\mathrm{mg}^{-1}$ crossing floating gates in 32 nm Flash memory device. The ISIS and terrestrial spectra, and the *front* and *back* radiation are reported. Sub = Substrate, GS = Gate stack, WL = Wordline, BL = Bitline, Mx = Metal layer x. Each layer contains also the isolation up to the next layer going from the substrate to the passivation.

secondary particles (heavy ions). These byproducts can cross (or pass nearby) floating gates, inducing charge loss and, to a lesser extent, charge trapping in the oxides surrounding the FG. This translates into a shift in the threshold voltage of the hit cells, which, when large enough to take the cell beyond the reference voltage, causes a raw bit error. Obviously this is not enough to cause an error visible to the user, because mandatory ECC schemes have to be implemented by the application in the case of NAND devices.

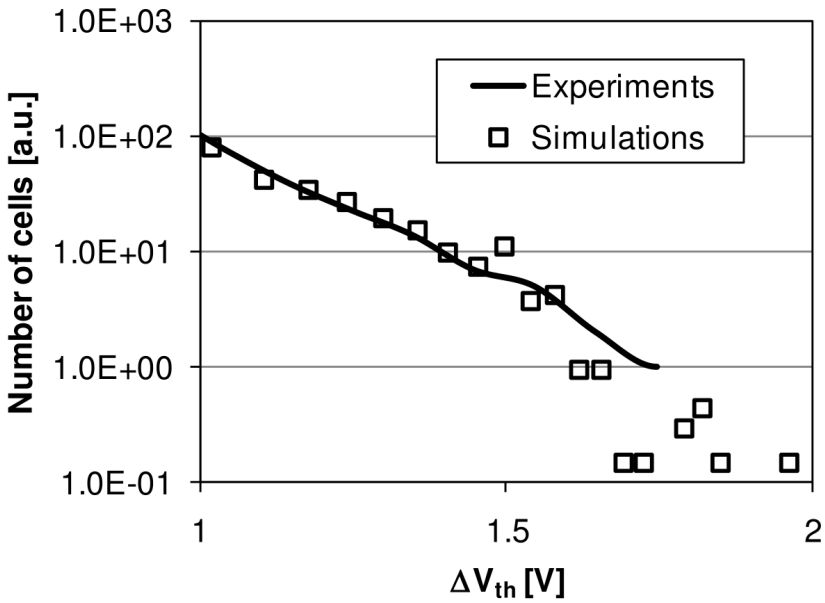


Figure 4.16: Comparison between experimental neutron-induced tails and weighted simulations of energy deposition in 60 nm Flash memory device.

The tails in Figures 3.12 and 3.13 (see Section 3.2) are approximately linear in a log-lin scale, meaning that the distribution of threshold voltage shifts induced by neutron byproducts is exponential. This distribution arises from the fact that the generated secondaries have a wide spread in energy, range, and LET. To match the experimental data (Figures 3.12 and 3.13) with the simulated events (Figure 4.9)[5], we need to establish a relation between the LET of the generated particles and the threshold voltage shift experienced by the cells. In doing so, we neglect the impact of the impinging particle energy, which is anyway a second-order factor. The correspondence between ΔV_{th} and LET is done as follows: if there are n cells experiencing a threshold voltage ΔV_{th} and n cells in our simulations hit by

[5]Fitting of experimental data has been applied only on *FG60* and *FG48* devices, as there are not Isis experimental data on *FG32* device.

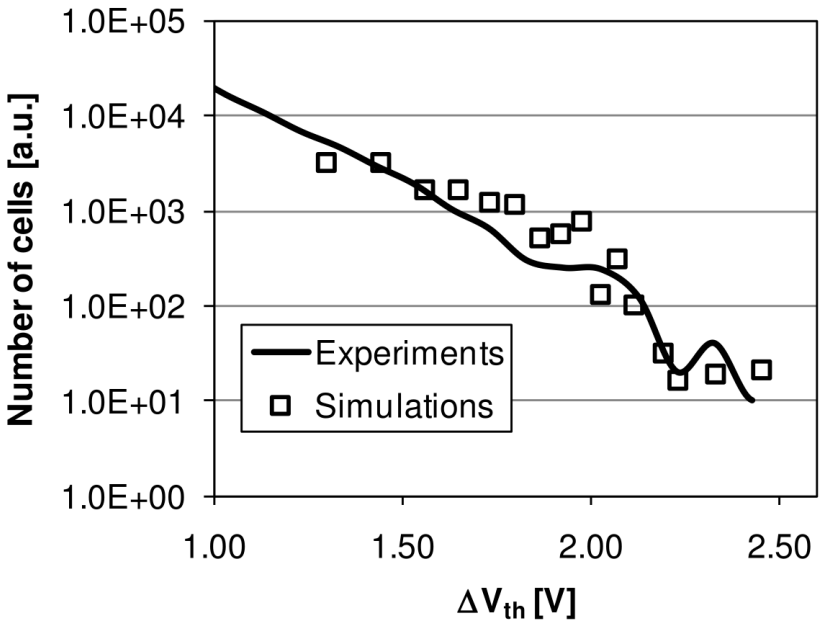


Figure 4.17: Comparison between experimental neutron-induced tails and weighted simulations of energy deposition in 48 nm Flash memory device.

particles having a linear energy transfer LET, we assume that LET produces ΔV_{th}. For this purpose we used a power law, choosing the pre-factor (A) and the exponent (B), fitting the data on the basis of the number of events:

$$\Delta V_{\text{th}} = A \cdot LET^{B} \tag{4.3.1}$$

The results are shown in Figures 4.16 and 4.17, where A is equal to 0.8 and 1.2 and B to 0.35 and 0.3 for 60 nm and 48 nm devices, respectively. A linear relation, which was assumed in previous works [21], between ΔV_{th} and LET does not provide a good fit for the data. Yet, equation 4.3.1 is only an approximation, because, as shown in the past, LET is not the only parameter that determines ΔV_{th}. In fact, the impinging particle energy plays a fundamental role as well. A lower energy ion produces a track with a smaller diameter, because of the lower energy of the delta electrons emitted normal to the ion trajectory, and this translates into a higher threshold voltage shift, because of the higher conductivity of the transient conductive path [17].

4.3.6 Scaling trends

A clear trend towards an increasing error rate for decreasing feature size is visible in experimental results. NAND MLC are the most sensitive among the studied devices (Figures 3.8 and 3.10). In fact, the smaller read margin make MLC more sensitive than SLC.

In Chapter 3.2 has been mentioned that errors take place only if the threshold voltage shift is large enough, that is to say, only if the LET of the impinging particle is above a certain threshold LET (LET_{th}). With the relation 4.3.1 is possible to estimate the LET_{th} for the two devices, by simply looking at the number of errors. The results are

$$LET_{\mathrm{th}} \approx 2\,\mathrm{MeV} \cdot \mathrm{cm}^{-2}\mathrm{mg}^{-1}$$

for 60 nm floating gate cell, and

$$LET_{\mathrm{th}} \approx 0.5\,\mathrm{MeV} \cdot \mathrm{cm}^{-2}\mathrm{mg}^{-1}$$

for 48 nm devices.

These considerations justify the observed scaling trend: the beneficial impact of the reduction in cell area is more than offset by the decrease in threshold LET. The rapid increase in cross section with decreasing feature size can be explained considering as follows:

- the total number of events depends on the cell area, i.e. it is quadratically dependent on the feature size (Figure 4.9),
- the number of events vs LET is exponentially distributed (Figure 4.9),
- to a first order, the threshold LET is a function of the ratio between the stored charge and the charge deposited by the heavy ion.

Chapter 5

Conclusions

Soft Error (SE) mechanisms in commercial electronics have been studied and implemented for over three decades. The most hard radiation environment is certainly space, where radiation is mainly composed of energetic charged particles (electrons, protons, heavy ions). However, atmosphere is not free from radiation, as a shower of particles (in particular, *neutrons*) is generated by protons from cosmic rays.

Neutrons are nonionizing, but they can interact with matter through nuclear reactions and/or generate nuclear recoils. Reaction byproducts are ionizing particles, i.e., nuclei that will lose energy through the generation of electron/hole pairs. Such nuclei can be generated in biased regions where the deposition of tiny amounts of charge can be enough to change the logic status of a node, i.e. a Soft Error (SE). The probability of interaction between neutrons and matter is very small, and not all ionizing byproducts will cross the device sensitive area; however, a huge amount of devices are unavoidably exposed to the average flux of about $14\,\text{neutrons} \cdot \text{cm}^{-2} \cdot \text{hours}$.

Soft Errors due to *atmospheric neutrons* are well known in CMOS technologies and in particular in SRAM memories, while few is known in non-volatile devices.

In this thesis the effects of *atmospheric neutrons on advanced Flash memory devices* have been studied throught experiments and simulations of neutron radiation.

The experiments presented in this work, carried out at the Isis neutron source, indicate that the number of excess electron/hole stored in the FG can be corrupted by neutrons. The analysis of cross section and of threshold voltage V_{th} distribution of the FG cells show that:

- the neutron raw bit error rate increases with decreasing feature size for all type of FG.
- the cross section for raw bit errors increases more than three orders of magnitude in one generation (Vendor B).
- MLC NAND are the most sensitive among the studied devices.
- in the MLC devices, the most affected bits are those belonging to the highest level.
- the raw bit cross section for NOR MLC memories is two orders of magnitude less than that of NAND MLC.
- the FG cell is partially discharged after neutrons irradiation, leading to the formation of a tail in the threshold voltage distribution.
- both the number of cells in the tail and the maximum V_{th} shift increase with decreasing feature size.

An application based on th Geant4 toolkit has been developed to simulate the interaction of neutron with the chip materials. The output of the simulation are the number, species, energy and trajectory of particles that cross the floating gate depositing energy along their paths. Simulations have been performed with Isis and terrestrial spectra on three technology nodes devices. Results can be summarized as follows:

- the effects of Isis and terrestrial scpectra are comparable.
- neutrons in the spectral range of $1 - 10\,\mathrm{MeV}$ appear to be significant for number of secondary particles generated.
- tthe effective LET distribution is different in th three devices because of the differences thicknesses of material the particles cross before reaching the floating gate.

- the energies of secondary particles are fairly limited as most of the particles are below 10 MeV.
- the largest parts of events is due to reactions between neutrons and Silicon or Oxygen atoms.
- events come for the most part from the oxide isolating the wordlines. Metal layers give a decreasing contribution as their distance from the floating gate increase.
- the spectrum of particles generated by neutrons is roughly the same in all the devices.
- the threshold LET (LET_{th}) has been estimated as $\approx 2\,\mathrm{MeV} \cdot \mathrm{cm}^{-2}\mathrm{mg}^{-1}$ for 60-nm devices and $\approx 0.5\,\mathrm{MeV} \cdot \mathrm{cm}^{-2}\mathrm{mg}^{-1}$ for 48-nm devices.

Finally, Flash memory devices does possess sensitivity to atmospheric neutron irradiation. Neutron effects increas with the memories scaling trends as

i the charge used to store a bit decrease with each new generation, thus the information can be easily corrupted;

ii the density of cells increase, which means that many cells can be affected by a single ion or a grazing incident neutron.

However, other circuits are many orders of magnitude more vulnerable, thus Flash will not be the reliability-limiting factor for a typical system.

Appendix A

The Geant4 code

A.1 Physic List

In the development of a GEANT4-based application, it is the user's responsibility to decide which physics processes are required, and then to include them in the physics list [3]. The following are the fundamental categories you need to provide:

- `ConstructParticle()` choose the particles you need in your simulation, define all of them.
- `ConstructProcess()` for each particle, assign all the physics processes relevant to your simulation.
- `SetCuts()` set the range cuts for secondary production, i.e. a threshold on particle production: particle unable to travel at least the range cut value are not produced.

A description of Geant4 physics lists and their use, along with simple and detailed examples, can be found in the GEANT4Application Developer's Guide. Here, the complete code of the *Physic List* used in this work is reported, including *particles*, *electromagnetic processes* and *Hadronic processes*.

```
#include "PhysicsList.hh"
#include "G4VUserPhysicsList.hh"
#include "G4ios.hh"
```

```
#include "globals.hh"

//Radiation-matter interaction
#include "G4EmStandardPhysics.hh"
#include "G4EmLivermorePhysics.hh"
#include "G4EmExtraPhysics.hh"
#include "G4DecayPhysics.hh"
#include "G4HadronElasticPhysics.hh"

#include "HadronPhysicsLHEP.hh"
#include "HadronPhysicsQGSP_BIC.hh"
#include "HadronPhysicsQGSP.hh"
#include "HadronPhysicsQGSP_BERT.hh"
#include "HadronPhysicsQGSP_BERT_HP.hh"
#include "G4IonPhysics.hh"

PhysicsList::PhysicsList() : G4VModularPhysicsList()
{
defaultCutValue = 1*mm;

RegisterPhysics(new G4EmLivermorePhysics);
RegisterPhysics(new G4EmExtraPhysics);
RegisterPhysics(new G4DecayPhysics);
RegisterPhysics(new G4HadronElasticPhysics);
RegisterPhysics(new HadronPhysicsLHEP);
}

void PhysicsList::ConstructParticle(void)
{
G4VModularPhysicsList::ConstructParticle();
}

void PhysicsList::ConstructProcess(void)
{
G4VModularPhysicsList::ConstructProcess();
}

void PhysicsList::SetCuts(void)
{
SetCutsWithDefault();
```

```
SetCutValue(1*nm,"e-");
SetCutValue(1*nm, "alpha");
SetCutValue(1*nm, "GenericIon");
}
```

A.2 Detector

```
#include "Detector.hh"
#include "AnalysisManager.hh"

Detector::Detector(G4String name) : G4VSensitiveDetector(name)
{
collectionName.insert("hits");
evtIndex = 0;
}

Detector::~Detector()
{
}

void  Detector::Initialize (G4HCofThisEvent *HC)
{
}

void  Detector::EndOfEvent (G4HCofThisEvent *HC)
{
evtIndex ++;
}

G4bool Detector::ProcessHits(G4Step* aStep,G4TouchableHistory* ROhist)
{
//Create a hit
G4double eDep = aStep->GetTotalEnergyDeposit();
if ( eDep > 0.0 )
{

G4Track* tr = aStep->GetTrack();
G4ParticleDefinition* pd = tr->GetDefinition();
```

```
G4StepPoint* preSP = aStep->GetPreStepPoint();
G4ThreeVector prePos = preSP->GetPosition();

G4StepPoint* postSP = aStep->GetPostStepPoint();
G4ThreeVector postPos = postSP->GetPosition();

const G4ThreeVector vp = tr->GetVertexPosition();
const G4LogicalVolume* lv = tr->GetLogicalVolumeAtVertex();

LOG << evtIndex << " " <<
pd->GetParticleName() << " " <<
lv->GetName() << " " <<
tr->GetKineticEnergy()/MeV  << " " <<
eDep/MeV << " " <<
aStep->GetNonIonizingEnergyDeposit()/MeV << " " <<
prePos/nm << " " <<
postPos/nm << "\n";
}

return true;
}
```

A.3 Particles Source definition

The G4GeneralParticleSource class allows specification of particle source. The class defines the particle type, position, direction (or angular) and energy distributions. In the following, a neutron beam which cover the 80% of chip area, with ISIS and Terrestrial spectrum energy distribution, is described.

```
/gps/ene/type User
/gps/hist/type energy
```

ISIS spectrum...(see A.4)
Terrestrial spectrum...(see A.5)

```
/gps/pos/type Plane
/gps/pos/shape Square

/gps/direction 1 0 0
/gps/pos/centre -0.5 0 0 mm
/gps/pos/rot1 0 0 1

/gps/pos/halfx 1008e-9 m
/gps/pos/halfy 1008e-9 m
/gps/pos/halfz 0 m

/gps/particle neutron
```

A.4 ISIS Spectrum definition

```
Low

/gps/hist/point 1.0000 746566.8312
/gps/hist/point 1.2115 481716.129
/gps/hist/point 1.4678 343889.7435
/gps/hist/point 1.7783 256271.0754
/gps/hist/point 2.1544 169278.6708
/gps/hist/point 2.6102 110521.6644
/gps/hist/point 3.1623 79586.21704
/gps/hist/point 3.8312 54268.43843
/gps/hist/point 4.6416 36722.0222
/gps/hist/point 5.6234 22039.47922
/gps/hist/point 6.8129 14539.0495
/gps/hist/point 8.2540 7979.816185

Medium

/gps/hist/point 10.000 5339.837402
/gps/hist/point 12.115 3397.787021
/gps/hist/point 14.678 2443.30841
/gps/hist/point 17.783 1613.747993
/gps/hist/point 21.544 1233.634243
```

```
/gps/hist/point 26.102 932.062902
/gps/hist/point 31.623 746.9470371
/gps/hist/point 38.312 577.6782719
/gps/hist/point 46.416 432.8985176
/gps/hist/point 56.234 310.7714692
/gps/hist/point 68.129 235.5816353
/gps/hist/point 82.540 177.8584062
/gps/hist/point 100.00 122.6497742
```

High

```
/gps/hist/point 121.15 85.43524311
/gps/hist/point 146.78 45.84041813
/gps/hist/point 177.83 35.0
/gps/hist/point 215.44 26.01051104
/gps/hist/point 261.02 14.26494826
/gps/hist/point 316.23 7.085080785
/gps/hist/point 383.12 3.606001887
/gps/hist/point 464.16 1.670416239
/gps/hist/point 562.34 0.849310867
/gps/hist/point 681.29 0.34455877
/gps/hist/point 800.00 0.312385683
```

A.5 Terrestrial Spectrum definition

Low

```
/gps/hist/point 1.054 8300
/gps/hist/point 1.165 81900
/gps/hist/point 1.287 76100
/gps/hist/point 1.423 70200
/gps/hist/point 1.572 60000
/gps/hist/point 1.738 57200
/gps/hist/point 1.920 50600
/gps/hist/point 2.122 50200
/gps/hist/point 2.346 54400
/gps/hist/point 2.592 43000
```

```
/gps/hist/point 2.865 33400
/gps/hist/point 3.166 26500
/gps/hist/point 3.499 18600
/gps/hist/point 3.867 16400
/gps/hist/point 4.274 17300
/gps/hist/point 4.724 18800
/gps/hist/point 5.220 15300
/gps/hist/point 5.769 12500
/gps/hist/point 6.376 11600
/gps/hist/point 7.047 8900
/gps/hist/point 7.788 7160
/gps/hist/point 8.607 6730
```

Medium

```
/gps/hist/point 9.512 5530
/gps/hist/point 10.51 4580
/gps/hist/point 11.62 4090
/gps/hist/point 12.84 3800
/gps/hist/point 14.19 3440
/gps/hist/point 16.16  3020
/gps/hist/point 18.52 3220
/gps/hist/point 25.70 2590
/gps/hist/point 44.19 2090
/gps/hist/point 75.98 1530
/gps/hist/point 130.7 964
```

High

```
/gps/hist/point 224.6 430
/gps/hist/point 386.3 133
/gps/hist/point 664.2 39.9
/gps/hist/point 1140 10.2
/gps/hist/point 1960 2.24
/gps/hist/point 3380 0.336
/gps/hist/point 5810 0.0471
/gps/hist/point 9980 0.00987
/gps/hist/point 17200 0.00383
/gps/hist/point 29500 0.00086
/gps/hist/point 50700 0.000217
```

```
/gps/hist/point 87300 0.0000697
/gps/hist/point 150000 0.0000188
```

Bibliography

[1] Inside the ISIS linear particle accelerator. [Online]. http://www.isis.stfc.ac.uk/learning/inside-the-isis-linear-particle-accelerator11761.html (2010).

[2] Agostinelli S., *et al.* *Geant4—a simulation toolkit.* Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **506**(3), 250–303 (2003).

[3] Apostolakis J., Folger G., Grichine V., Howard A., Ivanchenko V., Kosov M., Ribon A., Uzhinsky V., Wright D. H. *GEANT4 Physics Lists for HEP.* pp. 833–836 (2008).

[4] Autran J., Roche P., Sauze S., Gasiot G., Munteanu D., Loaiza P., Zampaolo M., Borel J. *Real-time neutron and alpha soft-error rate testing of CMOS 130nm SRAM: Altitude versus underground measurements.* pp. 233–236 (2008).

[5] Baggio J., Lambert D., Ferlet-Cavrois V., Paillet P., Marcandella C., Duhamel O. *Single Event Upsets Induced by 1–10 MeV Neutrons in Static-RAMs Using Mono-Energetic Neutron Sources.* Nuclear Science, IEEE Transactions on, **54**(6), 2149–2155 (2007).

[6] Barth J. L., Dyer C. S., Stassinopoulos E. G. *Space, Atmospheric, and Terrestrial Radiation Environments.* IEEE Transactions On Nuclear Science, **50 (3)**, 466–482 (2003).

[7] Benedetto J. M., Boesch H. E. *The Relationship between 60Co and 10-keV X-Ray Damage in MOS Devices.* Nuclear Science, IEEE Transactions on, **33**(6), 1317–1323 (1986).

[8] Brewer J., Gill M., (Edited by). *Nonvolatile Memory Technologies with Emphasis on Flash: A Comprehensive Guide to Understanding and Using Flash Memory Devices.* Wiley-IEEE Press (2008).

[9] Butt N., Alam M. *Modeling single event upsets in Floating Gate memory cells.* pp. 547–555 (2008).

[10] Cai X. X., Platt S., Chen W. *Modelling Neutron Interactions in the Imaging SEE Monitor.* Nuclear Science, IEEE Transactions on, **56**(4), 2035–2041 (2009).

[11] Cappelletti P., Modelli A. *Flash memory reliability* in "Flash Memories". Kluwer, norwel edition (1999).

[12] Casagrande G. *Flash memory testing* in "Flash Memories". Kluwer, norwell edition (1999).

[13] Cellere G., Pellati P., Chimenton A., Wyss J., Modelli A., Larcher L., Paccagnella A. *Radiation effects on floating-gate memory cells.* Nuclear Science, IEEE Transactions on, **48**(6), 2222–2228 (2001).

[14] Cellere G., Paccagnella A., Larcher L., Chimenton A., Wyss J., Candelori A., Modelli A. *Anomalous charge loss from floating-gate memory cells due to heavy ions irradiation.* Nuclear Science, IEEE Transactions on, **49**(6), 3051–3058 (2002).

[15] Cellere G., Paccagnella A., Visconti A., Bonanomi M. *Ionizing radiation effects on floating gates.* Appl. Phys. Lett., **85**, 485–487 (2004a).

[16] Cellere G., Paccagnella A., Visconti A., Bonanomi M., Caprara P., Lora S. *A model for TID effects on floating Gate Memory cells.* Nuclear Science, IEEE Transactions on, **51**(6), 3753–3758 (2004b).

[17] Cellere G., Paccagnella A., Visconti A., Bonanomi M., Candelori A. *Transient conductive path induced by a Single ion in 10 nm SiO_2 Layers.* Nuclear Science, IEEE Transactions on, **51(6)**, 3304–3311 (2004c).

[18] Cellere G., Paccagnella A., Visconti A., Bonanomi M., Candelori A., Lora S. *Effect of different total ionizing dose sources on charge loss from programmed floating gate cells.* Nuclear Science, IEEE Transactions on, **52**(6), 2372–2377 (2005a).

[19] Cellere G., Larcher L., Paccagnella A., Visconti A., Bonanomi M. *Radiation induced leakage current in floating gate memory cells.* Nuclear Science, IEEE Transactions on, **52**(6), 2144–2152 (2005b).

[20] Cellere G., Paccagnella A., Visconti A., Bonanomi M. *Secondary Effects of Single Ions on Floating Gate Memory Cells.* Nuclear Science, IEEE Transactions on, **53**(6), 3291–3297 (2006a).

[21] Cellere G., Paccagnella A., Visconti A., Bonanomi M. *Subpicosecond conduction through thin SiO_2 layers triggered by heavy ions.* J. Appl. Phys., **99**, 074101–074111 (2006b).

[22] Cellere G., Paccagnella A., Visconti A., Bonanomi M. *Variability in FG Memories Performance After Irradiation.* Nuclear Science, IEEE Transactions on, **53**(6), 3349–3355 (2006c).

[23] Cellere G., Paccagnella A., Visconti A., Bonanomi M., Harboe-Sørensen R., Virtanen A. *Angular Dependence of Heavy Ion Effects in Floating Gate Memory Arrays.* Nuclear Science, IEEE Transactions on, **54**(6), 2371–2378 (2007a).

[24] Cellere G., Paccagnella A., Visconti A., Bonanomi M., Beltrami S., Schwank J., Shaneyfelt M., Paillet P. *Total Ionizing Dose Effects in NOR and NAND Flash Memories.* Nuclear Science, IEEE Transactions on, **54**(4), 1066–1070 (2007b).

[25] Cellere G., Paccagnella A., Visconti A., Bonanomi M., Beltrami S., Harboe-Sørensen R., Virtanen A. *Effect of Ion Energy on Charge Loss From Floating Gate Memories.* Nuclear Science, IEEE Transactions on, **55**(4), 2042–2047 (2008).

[26] Cellere G., Gerardin S., Bagatin M., Paccagnella A., Visconti A., Bonanomi M., Beltrami S., Harboe-Sørensen R., Virtanen A., Roche P. *Can Atmospheric Neutrons Induce Soft Errors in NAND Floating Gate Memories?* Electron Device Letters, IEEE, **30, 2**, 178–180 (2009).

[27] Compagnoni C., Spinelli A., Gusmeroli R., Lacaita A., Beltrami S., Ghetti A., Visconti A. *First evidence for injection statistics accuracy limitations in NAND Flash constant-current Fowler-Nordheim programming.* pp. 165–168 (2007).

[28] Crisenza G., Ghidini G., Manzini S., Modelli A., Tosi M. *Charge loss in EPROM due to ion generation and transport in interlevel dielectric.* Electron Devices Meeting, 1990. IEDM '90. Technical Digest., International, pp. 107 –110 (1990).

[29] Dodd P., Schwank J., Shaneyfelt M., Felix J., Paillet P., Ferlet-Cavrois V., Baggio J., Reed R., Warren K., Weller R., Schrimpf R., Hash G., Dalton S., Hirose K., Saito H. *Impact of Heavy Ion Energy and Nuclear Interactions on Single-Event Upset and Latchup in Integrated Circuits.* Nuclear Science, IEEE Transactions on, **54**(6), 2303–2311 (2007).

[30] Dodds N., Reed R., Mendenhall M., Weller R., Clemens M., Dodd P., Shaneyfelt M., Vizkelethy G., Schwank J., Ferlet-Cavrois V., Adams J., Schrimpf R., King M. *Charge Generation by Secondary Particles From Nuclear Reactions in BEOL Materials.* Nuclear Science, IEEE Transactions on, **56**(6), 3172–3179 (2009).

[31] Dozier C. M., Brown D. B. *Effect of Photon Energy on the Response of MOS Devices.* Nuclear Science, IEEE Transactions on, **28**(6), 4137–4141 (1981).

[32] Dunn C., Kaya C., Lewis T., Strauss T., Schreck J., Hefley P., Middendorf M., San T. *Flash EPROM disturb mechanisms.* Reliability Physics Symposium, 1994. 32nd Annual Proceedings., IEEE International, pp. 299–308 (1994).

[33] Ginami G., Canali D., Fattori D., Girardi G., Scintu P., Tarchini L., Tricarico D. *Survey on flash technology with specific attention to the critical process parameters related to manufacturing.* Proceedings of the IEEE, **91**(4), 503–522 (2003).

[34] Goodman J., Ellsworth R., Murdin P. *Cosmic Rays: Extensive Air Showers* (2001).

[35] Gosling J. T. *The Solar Flare Myth.* Journal Of Geophysical Research, **98 (11)**, 937–949 (1993).

[36] Guertin S. M., Nguyen D. M., Patterson J. D. *Microdose Induced Data Loss on Floating Gate Memories.* Nuclear Science, IEEE Transactions on, **53**(6), 3518–3524 (2006).

[37] Hess W. N., Patterson H. W., Wallace R. *Cosmic-Ray Neutron Energy Spectrum.* Physical review, **116**, 445–457 (1959).

[38] Huang C. Method and system for correcting soft errors in memory circuit. United States Patent, 7644341 (2010).

[39] Ibe E., Taniguchi H., Yahagi Y., Shimbo K.-i., Toba T. *Impact of Scaling on Neutron-Induced Soft Error in SRAMs From a 250 nm to a 22 nm Design Rule.* Electron Devices, IEEE Transactions on, **57**(7), 1527–1538 (2010).

[40] JESD89A. Measurement and Reporting of Alpha Particle and Terrestrial Cosmic Ray-Induced Soft Errors in Semiconductor Devices. Technical report, JEDEC Solid State Technology Association (2006).

[41] Johnston A. *Reliability and radiation effects on compound semiconductors.* World Scientific (2010).

[42] Katznelson R., Frohman-Bentchkowsky D. *An erase mode for FAMOS EPROM devices.* IEEE Trans. on El. Dev, **27 (9)**, 1744–1752 (1980).

[43] Klein D. Method and system for dynamically operating memory in a power-saving error correcting mode. United States Patent, 6838331 (2005).

[44] Kobayashi S., Nakai H., Kunori Y., Nakayama T., Miyawaki Y., Terada Y., Onoda H., Ajika N., Hatanaka M., Miyoshi H., Yoshihara T. *Memory array architecture and decoding scheme for 3 V only sector erasable DINOR flash memory.* VLSI Circuits, 1993. Digest of Technical Papers. 1993 Symposium on, pp. 97–98 (1993).

[45] Krishnaswamy S. Design, analysis and test of logic circuits under uncertainty. PhD Thesis, University of Michigan (2008).

[46] Lambert D., Baggio J., Ferlet-Cavrois V., Flament O., Saigne F., Sagnes B., Buard N., Carriere T. *Neutron-induced SEU in bulk SRAMs in terrestrial environment: Simulations and experiments.* Nuclear Science, IEEE Transactions on, **51**(6), 3435–3441 (2004).

[47] Larcher L., Cellere G., Paccagnella A., Chimenton A., Candelori A., Modelli A. *Data retention after heavy ion exposure of floating gate memories: analysis and simulation.* Nuclear Science, IEEE Transactions on, **50**(6), 2176–2183 (2003).

[48] Lee J., Im H.-S., Byeon D.-S., Lee K.-H., Chae D.-H., Lee K.-H., Hwang S. W., Lee S.-S., Lim Y.-H., Lee J.-D., Choi J.-D., Seo Y.-I., Lee J.-S., Suh K.-D. *High-performance 1-Gb NAND flash memory with 0.12 μ m technology.* Solid-State Circuits, IEEE Journal of, **37**(11), 1502–1509 (2002).

[49] Lenahan P. M., Dressendorfer P. V. *Hole traps and trivalent silicon centers in metal/oxide/silicon devices.* Journal of Applied Physics, **55**(10), 3495 –3499 (1984).

[50] Liaw J. SRAM cell design for soft error rate immunity. United States Patent, 664945 (2003).

[51] Massengill L. *Cosmic and terrestrial single-event radiation effects in dynamic random access memories.* Nuclear Science, IEEE Transactions on, **43**(2), 576–593 (1996).

[52] Mielke N., Marquart T., Wu N., Kessenich J., Belgal H., Schares E., Trivedi F., Goodness E., Nevill L. *Bit error rate in NAND Flash memories.* pp. 9–19 (2008).

[53] Modelli A., Bez R., Visconti A. *Multi Level Flash Memory Technology.* Int. Conf. Solid State Devices and Materials, pp. 516–517 (2001).

[54] Moore G. E. *Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff.* Solid-State Circuits Newsletter, IEEE, **20**(3), 33–35 (2006).

[55] Murray S. Propagation of 1–10 MeV solar flare protons in interplanetary space. PhD Thesis, California Institute of Technology (1971).

[56] Nguyen D., Scheick L. *TID, SEE and radiation induced failures in advanced flash memories.* pp. 18–23 (2003).

[57] Nguyen D., Lee C., Johnston A. *Total ionizing dose effects on flash memories.* Radiation Effects Data Workshop, 1998. IEEE, pp. 100–103 (1998).

[58] Nguyen D., Guertin S., Swift G., Johnston A. *Radiation effects on advanced flash memories.* Nuclear Science, IEEE Transactions on, **46**(6), 1744–1750 (1999).

[59] Oldham T., Suhail M., Friendlich M., Carts M., Ladbury R., Kim H., Berg M., Poivey C., Buchner S., Sanders A., Seidleck C., LaBel K. *TID and SEE Response of Advanced 4G NAND Flash Memories.* pp. 31–37 (2008).

[60] Oldham T. R., Lelis A. J., McLean F. B. *Spatial Dependence of Trapped Holes Determined from Tunneling Analysis and Measured Annealing.* Nuclear Science, IEEE Transactions on, **33**(6), 1203–1209 (1986).

[61] Oldham T. R., Ladbury R. L., Friendlich M., Kim H. S., Berg M. D., Irwin T. L., Seidleck C., LaBel K. A. *SEE and TID Characterization of an Advanced Commercial 2Gbit NAND Flash Nonvolatile Memory.* Nuclear Science, IEEE Transactions on, **53**(6), 3217–3222 (2006).

[62] Peurrung A. *Recent developments in neutron detection.* Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **443**(2-3), 400–415 (2000).

[63] Pia M., Basaglia T., Bell Z., Dressendorfer P. *Geant4 in scientific literature.* pp. 189–194 (2009).

[64] Platt S., Török Z., Frost C., Ansell S. *Charge-Collection and Single-Event Upset Measurements at the ISIS Neutron Source.* Nuclear Science, IEEE Transactions on, **55**(4), 2126–2132 (2008).

[65] Platt S. P., Török Z. *Analysis of SEE-Inducing Charge Generation in the Neutron Beam at The Svedberg Laboratory.* IEEE Transactions On Nuclear Science, **54(4)**, 1163–1169 (2007).

[66] Platt S. P., Cassels B., Török Z. *Development and application of a neutron sensor for single event effects analysis.* Journal of Physics: Conference Series, **15**(1), 172 (2005).

[67] Raparla V., Lee S., Schrimpf R., Fleetwood D., Galloway K. *A model of radiation effects in nitride-oxide films for power MOSFET applications.* Solid-State Electronics, **47**(5), 775–783 (2003).

[68] Ricco B., Torelli G., Lanzoni M., Manstretta A., Maes H., Montanari D., Modelli A. *Nonvolatile multilevel memories for digital applications.* Proceedings of the IEEE, **86**(12), 2399–2423 (1998).

[69] Rodbell K., Heidel D., Tang H., Gordon M., Oldiges P., Murray C. *Low-Energy Proton-Induced Single-Event-Upsets in 65 nm Node, Silicon-on-Insulator, Latches and Memory Cells.* Nuclear Science, IEEE Transactions on, **54**(6), 2474–2479 (2007).

[70] Roth D., Kinnison J., Carkhuff B., Lander J., Bognaski G., Chao K., Swift G. *SEU and TID testing of the Samsung 128 Mbit and the Toshiba 256 Mbit flash memory.* pp. 96–99 (2000).

[71] Saks N., Dozier C., Brown D. *Time dependence of interface trap formation in MOSFETs following pulsed irradiation.* Nuclear Science, IEEE Transactions on, **35**(6), 1168–1177 (1988).

[72] Salinas C. J. S., Bellido J., Wahl D., Saavedra O. *COSMIC RAYS AND ASTROPHYSICS.* Proceedings of the 3rd School on Cosmic Rays and Astrophysics, volume 1123 of *American Institute of Physics Conference Series.* American Institute of Physics Conference Series (2009).

[73] Scheick L., McNulty P., Roth D. *Dosimetry based on the erasure of floating gates in the natural radiation environments in space.* Nuclear Science, IEEE Transactions on, **45**(6), 2681–2688 (1998).

[74] Schwank J. R., Winokur P. S., McWhorter P. J., Sexton F. W., Dressendorfer P. V., Turpin D. C. *Physical Mechanisms Contributing to Device "Rebound".* Nuclear Science, IEEE Transactions on, **31**(6), 1434–1438 (1984).

[75] Schwartz H., Nichols D., Johnston A. *Single-event upset in flash memories.* Nuclear Science, IEEE Transactions on, **44**(6), 2315–2324 (1997).

[76] Seifert N., Gill B., Zia V., Zhang M., Ambrose V. *On the Scalability of Redundancy based SER Mitigation Schemes.* pp. 1–9 (2007).

[77] Selmi L., Fiegna C. *Physical aspects of cell operation and reliability* in "Flash Memories". Kluwer, norwel edition (1999).

[78] Shimomura Y., Spears W. *Review of the ITER Project.* Applied Superconductivity, IEEE Transactions on, **14**(2), 1369–1375 (2004).

[79] Siefert N., Gill B., Foley K., Relangi P. *Multi-cell upset probabilities of 45 nm high-k + metal gate SRAM devices in terrestrial and space environments.* Proc. of IEEE Reliability Physics Symp. (IRPS), pp. 181–186 (2008).

[80] Silburt A., Evans A., Burghelea A., Wen S.-J., Ward D., Norrish R., Hogle D. *Building a reliable internet core using soft error prone electronics.* pp. 227–232 (2008).

[81] Snyder E., McWhorter P., Dellin T., Sweetman J. *Radiation response of floating gate EEPROM memory cells.* Nuclear Science, IEEE Transactions on, **36**(6), 2131–2139 (1989).

[82] Stassinopoulos E. G. *Microelectronics for the natural radiation environments of Space, chapter I: Radiation environments of Space.* Reno, NV, Proc. 1990 IEEE NSREC Short Course (1990).

[83] Suh K.-D., Suh B.-H., Um Y.-H., Kim J.-K., Choi Y.-J., Koh Y.-N., Lee S.-S., Kwon S.-C., Choi B.-S., Yum J.-S., Choi J.-H., Kim J.-R., Lim H.-K. *A 3.3 V 32 Mb NAND flash memory with incremental step pulse programming scheme.* Solid-State Circuits Conference, 1995. Digest of Technical Papers. 42nd ISSCC, 1995 IEEE International, pp. 128–129 (1995).

[84] Tanzawa T., Umezawa A., Taura T., Shiga H., Hara T., Takano Y., Miyaba T., Tokiwa N., Watanabe K., Watanabe H., Masuda K., Naruke K., Kato H., Atsumi S. *A 44-*mm^2 *four-bank eight-word page-read 64-Mb flash memory with flexible block redundancy and fast accurate word-line voltage controller.* Solid-State Circuits, IEEE Journal of, **37**(11), 1485–1492 (2002).

[85] Tipton A., Pellish J., Hutson J., Baumann R., Deng X., Marshall A., Xapsos M., Kim H., Friendlich M., Campola M., Seidleck C., LaBel K., Mendenhall M., Reed R., Schrimpf R., Weller R., Black J. *Device-Orientation Effects on Multiple-Bit Upset in 65 nm SRAMs.* Nuclear Science, IEEE Transactions on, **55**(6), 2880–2885 (2008).

[86] Truscott P., Lei F., Dyer C. S., Frydland A., Clucas S., Trousse B., Hunter K., Comber C., Chugg A., Moutrie M. *Assessment of Neutron- and Proton-Induced Nuclear Interaction and Ionization Models in Geant4 for Simulating Single Event Effects.* IEEE TRANSACTIONS ON NUCLEAR SCIENCE, **51(6)**, 3369–3374 (2004).

[87] Warren K. M., Sierawski B. D., Weller R. A., Reed R. A., Mendenhall M. H., Pellish J. A., Schrimpf R. D., Massengill L. W., Porter M. E., Wilkinson J. D. *Predicting Thermal Neutron-Induced Soft Errors in Static Memories Using TCAD and Physics-Based Monte Carlo Simulation Tools.* Electron Device Letters, IEEE, **28**(2), 180–182 (2007).

[88] Watanabe N. *Neutronics of pulsed spallation neutron sources.* Reports on Progress in Physics, **66**(3), 339 (2003).

[89] Weller R., Mendenhall M., Reed R., Schrimpf R., Warren K., Sierawski B., Massengill L. *Monte Carlo Simulation of Single Event Effects.* Nuclear Science, IEEE Transactions on, **57**(4), 1726–1746 (2010).

[90] Yoshikawa K., Yamada S., Miyamoto J., Suzuki T., Oshikiri M., Obi E., Hiura Y., Yamada K., Ohshima Y., Atsumi S. *Comparison of current flash EEPROM erasing methods: stability and how to control.* Electron Devices Meeting, 1992. IEDM '92. Technical Digest., International, pp. 595–598 (1992).

[91] Ziegler J. F. *Terrestrial cosmic rays.* IBM J. Res. Developmen, **40 (1)**, 19–39 (1996).