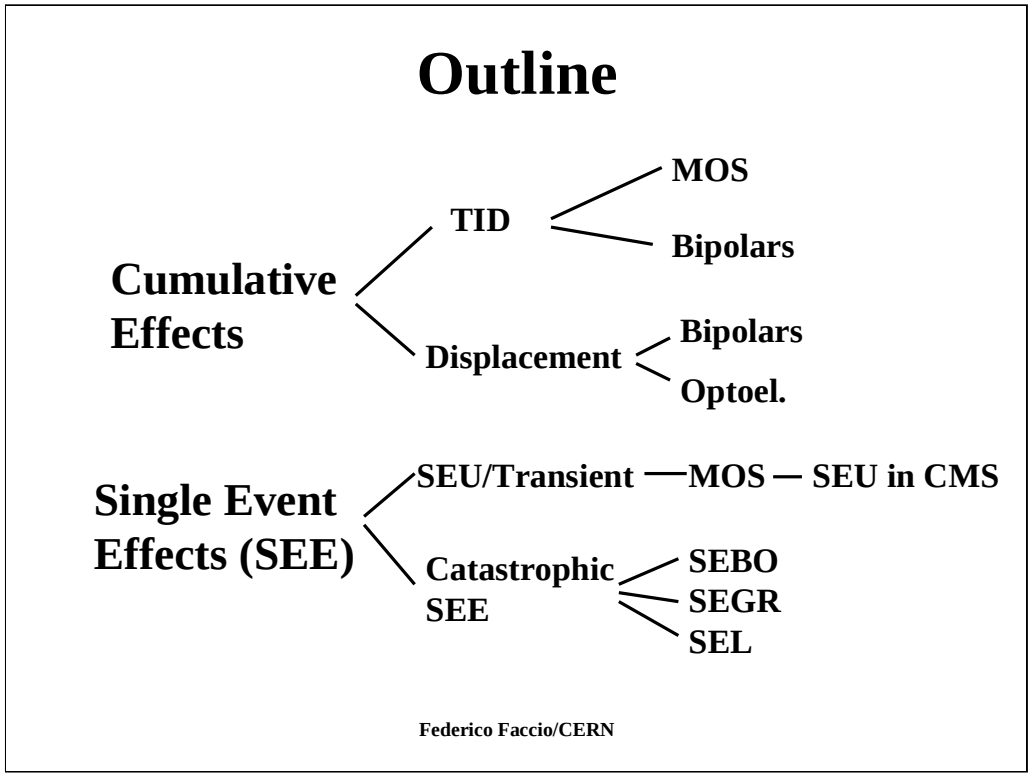
Radiation effects in the electronics for CMS

F. Faccio
CERN



Cumulative effects are gradual effects taking place during the whole lifetime of the electronics exposed in a radiation environment. A device sensitive to TID or displacement damage will exhibit failure in a radiation environment when the accumulated TID (or particle fluence) has reached its tolerance limits. It is therefore in principle possible to foresee when the failure will happen for a given, well known and characterized component.

On the contrary, Single Event Effects are due to the energy deposited by one single particle in the electronic device. Therefore, they can happen in any moment, and their probability is expressed in terms of cross-section. A device sensitive to SEE can exhibit failure at any moment since the beginning of its operation in a radiation environment.

Cumulative Effects

Radiation effects accumulating during the whole LHC lifetime, due to the energy deposited by radiation in the electronics

Ionization ⟹ **TID (dose)**

Non-ionization ⟹ **Displacement (fluence)**

Federico Faccio/CERN

Total Ionizing Dose (TID) effects are a typical case of cumulative effects. The ionization dose is deposited by particles passing through the materials constituting the electronic devices.

This happens during the whole time the device is exposed to radiation.

The same is true for displacement damage.

TID is the measurement of the dose, that is the energy, deposited in the material of interest by radiation in the form of ionization energy. The unit to measure it in the International System (SI) is the Gray, but the radiation effects community still uses most often the old unit, the rad. One should get used to both, because the dosimetry people speak about Gray, whilst electronic engineers working on the effects speak about rad. Luckily, the equivalence between the two is easy to remember:

1 Gray (Gy) = 100 rad

Displacement damage is not measured in any unit, just in its effects on the devices. The displacement damage is expressed in terms of the particle fluence, in particles/cm^2.

Total Ionizing Dose (TID)

Ionization in SiO_2
In CMS: (charged hadrons, electrons, gammas, neutrons)

⇓

Creation of electron-hole pairs

⇓

Buildup of charge/defects

⇓

Device degradation

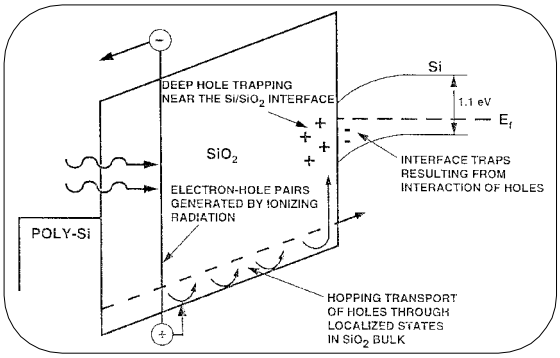
Federico Faccio/CERN

In the CMS environment, ionization effects will be induced by the ionization energy deposited by charged hadrons, electrons, gammas and neutrons (even though the last two are not directly ionizing, they can induce ionizing energy depositions).

The heart of TID effects is the energy deposition in silicon dioxide, because the electron-hole pairs created in this material do not completely recombine in a very short time. In the presence of an electric field in the oxide, a great amount of the pairs does not recombine, and both electrons and holes start to drift in the electric field. Electrons, with a much higher mobility, can easily leave the oxide. Holes instead can be trapped in defect centers in the oxide. Additionally, this process can create (or better activate) defects at the silicon-oxide interface.

The charge buildup and the activation of defects are the two reasons for device degradation induced by TID.

TID in CMOS devices



F.B.McLean et al., HDL-TR-2129 internal report, 1987

Trapped holes ⟹ Vt shift, noise, leakage fast formation, annealing

Interface states ⟹ Vt shift, mobility, transcond. Slow formation, no anneal. below 400°C

Federico Faccio/CERN

Charge buildup in the oxide is due to the trapping of holes in the oxide. This happens in the bulk of the oxide.

If we look at the gate oxide of MOS transistors, these charges will screen or enhance (depending on the polarity of the transistor) the gate electric field. This will lead to a threshold voltage shift. In the lateral oxide instead (as in the LOCOS or STI oxide to isolate transistors from each other), they might attract an image charge in the semiconductor which can invert the interface and open leakage paths. This happens only in NMOS transistors.

The defects formed at the interface between silicon and silicon dioxide (this is the region where the conductive channel forms in a MOS transistor) are called interface states. They trap charge from the channel, which leads to both a threshold voltage shift and also affects the mobility of carriers in the channel.

The two types of effects, the trapping of holes and the creation of interface states, have a very different dynamic. Holes are trapped very quickly, and can be detrapped by thermal energy (this is called annealing). Therefore, increasing the temperature is a good method to anneal the trapped charge. Interface states instead exhibit a slow formation, and they do not anneal at temperature below about 400°C.

These two different dynamics of the defects and trapped holes have to be taken into account in the testing of the devices and ICs. For MOS transistors and ICs, it exists a test procedure to evaluate the possible failure modes induced by both effects.

Bias dependence

Worst case bias most commonly used:

*** Transistors:**

NMOS	PMOS
$V_{gs}=V_{dd}$	All grounded
$V_{ds}= 0$ V	

*** Circuits:**

Powered as in operational condition

Federico Faccio/CERN

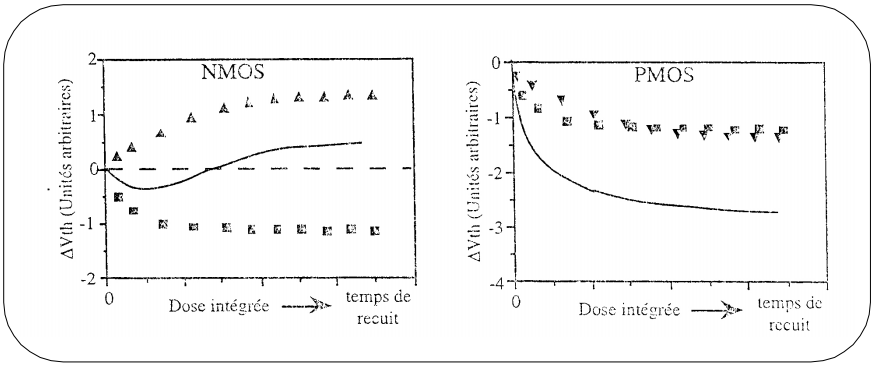
The two TID induced phenomena in the oxide are very sensitive to the applied bias. As said, radiation-created electron-hole pairs have a probability to recombine that is lowered by an applied electric field.

In the test of MOS transistors, the worst case bias condition is most often used. This condition maximizes the TID effects, hence it gives the worst possible picture for the device degradation (conservative test).

In the case of ICs, the worst condition is determined by a complex combination of individual transistors bias, and the only possible way of testing is to apply the bias such that the circuit is as close as possible to the operational condition. In some cases this would require a complex series of input signals (clocks), and a compromise solution is simply to apply the power to the circuit (no dynamic signal). The applied power supply should be the highest foreseen for the circuit use.

In all known cases, the CMOS circuits exposed with no bias (all terminal grounded or floating) exhibit a considerably lower degradation than their biased counterparts. Therefore, all TID tests on CMOS circuits have to be performed under bias.

Threshold voltage shift



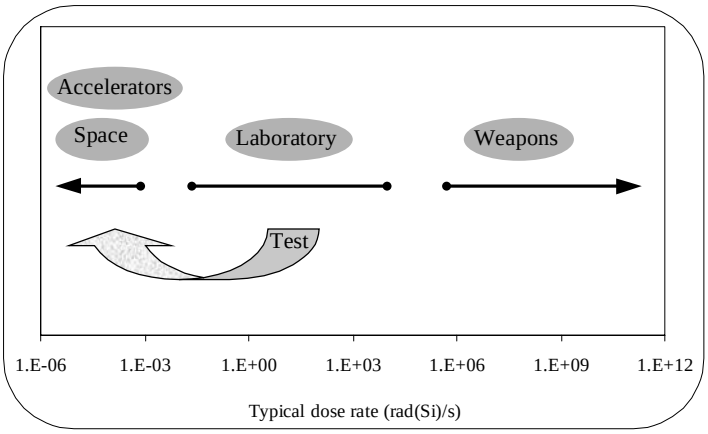
Federico Faccio/CERN

In the overhead, the typical displacement of the threshold voltage shift with irradiation is shown for NMOS and PMOS transistors.

For NMOS transistors, trapped holes tend to decrease the threshold whilst interface states tend to increase it. As the trapped holes have a faster dynamic, they prevail at the beginning of the irradiation. Then, they gradually anneal due to the temperature (this depends on the technology and on the temperature). At the same time, the interface state accumulates and, as they do not anneal, they finally dominate the threshold voltage shift. This is a typical case, but in some technologies this “rebound” is not shown (sometimes the trapped holes anneal so fast that the threshold voltage shift is always positive, whilst in the case of thicker oxides the trapped holes always dominate and the threshold voltage constantly decrease).

For PMOS transistors, both the trapped holes and the interface states tend to increase (in absolute value) the threshold voltage, and no rebound is observed.

Implications for testing



Federico Faccio/CERN

As said, these different effects and trapping dynamics have serious consequences on the testing. As these effects are time-dependent, it is important to apply a methodology that allows one to have reliable estimates for the device behaviour in the real radiation environment.

The problem in that case is that, in the real environment, the dose rate is normally very low. The TID is in fact accumulated over a long period of time (in LHC, over the expected 10 years of operation). In the laboratory, where the test is run, we cannot wait for ten years, and we need to accelerate by a considerable factor the dose deposition. Typically, the LHC-foreseen TID is deposited in a few hours, sometimes in a few days.

A reliable methodology should allow the experimenter to qualify the components in the laboratory for the dose rate in the real application. For CMOS technologies, there exist several methodologies, slightly differing from each other, to do so. Therefore, for CMOS technologies it is possible to qualify the components in the laboratory. Unfortunately, the present methodologies are very conservative, and can lead to the rejection of components that might well survive in the real environment.

Example: qualification for space

Parameter	ESA/SCC Basic Spec. No. 22900	MIL-STD-883, Method 1019.4
Scope	Test method for steady-state irradiation testing of ICs and disctretes during technology evaluation & qualification or procurement for space application	Test method for steady-state irradiation testing of packaged semiconductor ICs
Radiation Source	^{60}Co gammas (ionizing); electron accelerator (ionizing and displacement); alternate sources permitted	^{60}Co gammas (ionizing)
Dosimetry	Intensity ±5%, field uniformity ±10%	Intensity ±5%, field uniformity ±10%
Pb/Al container	Minimum 1.5mm Pb and 0.7mm Al unless no demonstrated dose enhancement	Minimum 1.5mm Pb and 0.7mm Al unless no demonstrated dose enhancement
Dose	±10% of specification	±10% of specification; an additional 0.5x overtest for "rebound"
Dose Rate	Exposure time ≤ 96 h; Window 1, Standard Rate is 1 to 10 rad(Si)/s; Window 2, Low Rate is 0.01 to 0.1 rad(Si)/s; or lower rate if agreed to by parties to test	50 to 300 rad(Si)/s or lower dose rate (≥ dose rate of intended application) if agreed to by parties to test
Anneals: *Room temperature* *Elevated temperature*	 For 24 h At 100°C for 168 h	 None "Rebound" at 100°C±5°C for 168±12 h
Temperature: *Irradiation* *Test*	 20°C±10°C 25°C±3°C	 24°C±6°C 25°C±5°C
Bias: *During irradiation & anneals* *Between irradiation & test*	 ±10%; Worst-case bias Device leads shorted (e.g., in conductive foam)	 ±10%; Worst-case bias Device leads shorted (e.g., in conductive foam)
Test sequence: *Time between irradiation & test* *Time between multiple irradiations*	 Begin within 1 h, end within 2 h 2 h maximum	 Begin within 1 h, end within 2 h 2 h maximum

Federico Faccio/CERN

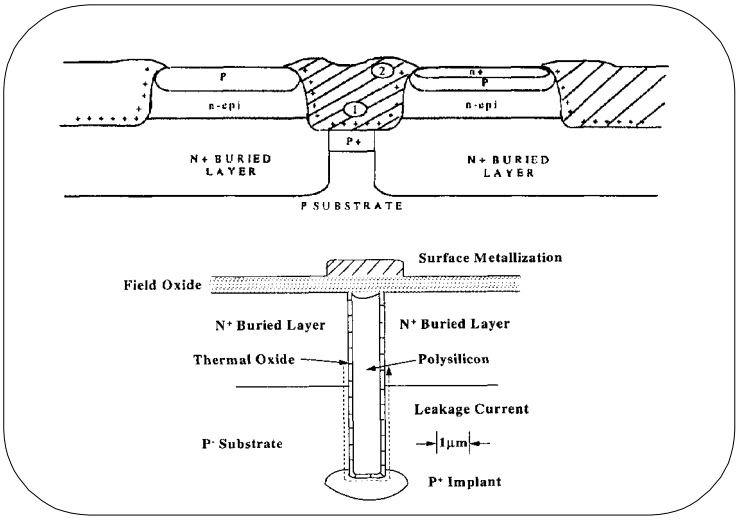
Two of the test procedures that codify the test of components for space (and military) applications are summarized in the overhead.

The principle is always the same: the measurement immediately following a fast irradiation gives the worst case picture for the effects induced by the trapped holes. Then, an annealing at high temperature accelerates the annealing rate of the holes and the formation of interface states. The measurement after the high temperature cycle gives therefore a worst case picture for effects induced by the interface states.

In reality, the real performance of the CMOS circuit will be somewhere between these two extremes, but the approach for space (and especially for military applications) is to reject the components failing at any of the two steps.

TID in bipolar devices

Substrate, sidewall and surface inversion (in oxide-isolated processes)



R.L.Pease et al., IEEE Trans. Nucl. Science. Vol.32, N.6, 1985

E.W.Enlow et al., IEEE Trans. Nucl. Science. Vol.36, N.6, 1989

Federico Faccio/CERN

The TID effects in bipolar devices are also due to charge trapping in the oxide and creation of interface states. The effects can be shared in two categories:

1) inversion of the silicon under a thick oxide, opening a conductive channel

2) effects decreasing the gain of the transistor

The inversion channel can be formed in several places, depending on the technological characteristics:

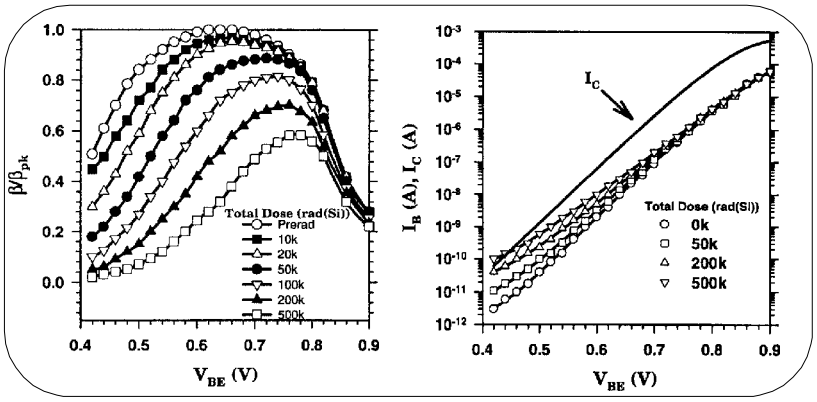
a) Substrate: opening of the channel between two buried layers

b) Sidewall: inversion near the sidewall oxide, shorting collector and emitter of npn transistors

c) Surface: inversion of the surface. Even though the surface oxide is generally thinner than the isolation recess oxide, the effect might be important when the transistors are working at low current levels.

TID in bipolar devices

Gain degradation: Increase of the surface component of the base current



R.N.Nowlin et al., IEEE Trans. Nucl. Science. Vol.39, N.6, 1992

TID acts on the gain by increasing the surface component of the base current (the bulk component being mainly sensitive to displacement damage).

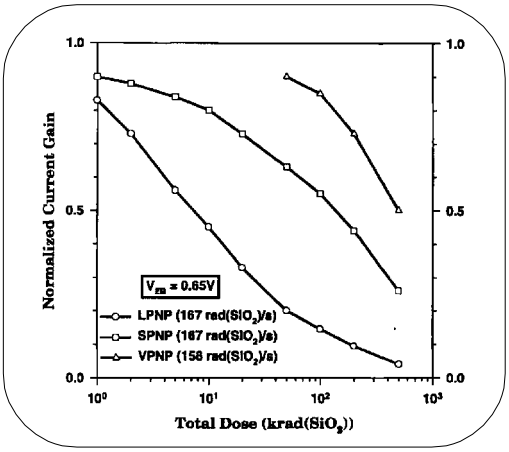
The increase in this surface current component is mainly due to an increase of interface states at the surface of the base and a positive charge buildup near the emitter-base junction (both increasing the minority carrier recombination rate).

Excess base current is in general the dominant effect, with the collector current being constant.

The sensitivity is higher at lower injection levels, as in this case there is more sensitivity to surface phenomena.

PNP transistors

Vertical PNP transistors are the less sensitive



D.W.Emily, IEEE NSREC short course, 1996

Federico Faccio/CERN

In conventional bipolar processes, the lateral PNP transistors are very sensitive to TID effects.

Vertical PNP are generally less sensitive than all other devices, including vertical NPN.

Factors affecting TID response of bipolar transistors

- **Transistor polarity**
- **Oxide thickness over base-emitter region**
- **Oxide trap efficiency**
- **Vertical and fringing electric field**
- **Base and Emitter surface concentration**
- **Emitter perimeter-to-area ratio**
- **Transistor geometry (ratio of lateral to vertical current flow)**
- **Injection level**
- **Dose rate**
- **Temperature**

Federico Faccio/CERN

Oxide thickness: the thicker the oxide above emitter-base junction area, the grater the TID effects

Oxide trap efficiency: the more degraded the oxide (during manufacturing), the grater the TID effects

Electric field: difficult to have a clear picture of the electric fields, and to generalize to several technologies

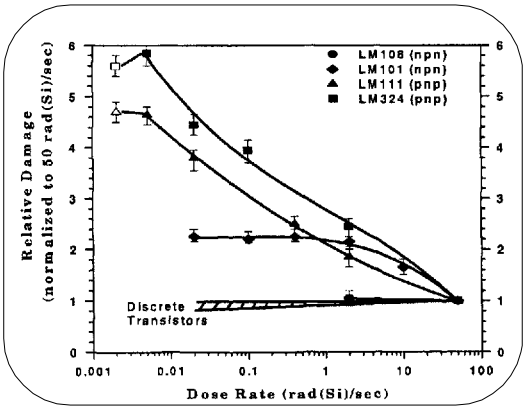
Surface doping concentration: the more heavily doped the base or emitter surface, the lower the TID effects

Emitter perimeter-to-area ratio: the grater the ratio, the grater the TID effects

Transistor geometry: vertical structures have lower sensitivity to TID effects than surface lateral structures (or substrate PNP, where 20% of the current is lateral)

Injection levels: in almost all cases, degradation is higher at low injection

Low dose rate (LDR) effect



A.H.Johnston et al., IEEE Trans. Nucl. Science. Vol.41, N.6, 1994

Summary: LDR appears to be consistently inconsistent

Federico Faccio/CERN

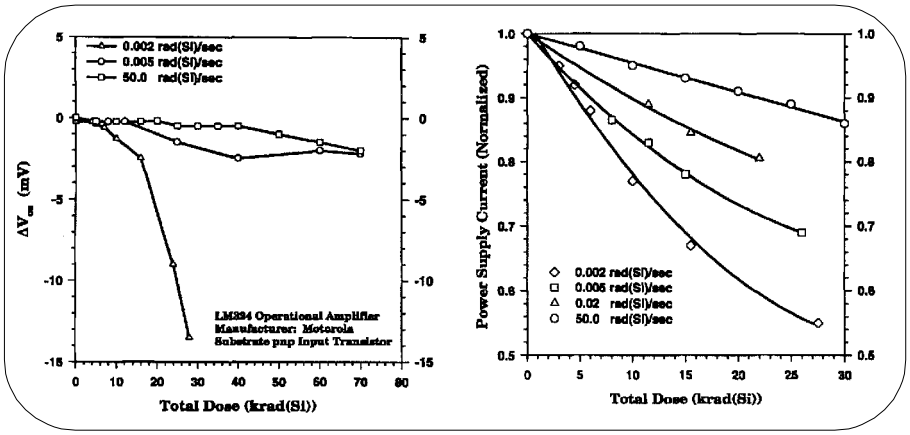
LDR effect: extreme variability in different processes (nonexistent in some, severe in others)

"True" rate effect, not time dependent effect as in MOS

This effects seem to happen for a near zero electric field in the oxide during irradiation, and the net trapped-hole density is higher at low rate.

In addition to this "true" rate effect, there are time dependent effects following irradiation, especially after high dose rate irradiation

Examples of LDR effects

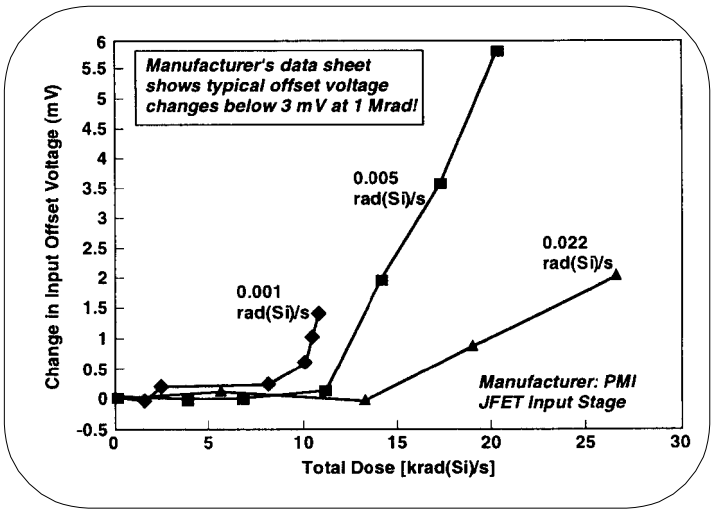


A.H.Johnston et al., IEEE Trans. Nucl. Science. Vol.42, N.6, 1995

For transistors, both NPN and different PNP transistor show an enhanced degradation at low dose rate (when they show it): their excess base current increases by a factor typically 10 to 20 more at low dose rate (0.1rad/s) compared to high rate (1000rad/s). This effect does not seem to saturate at 0.1rad/s.

Recent results indeed seem to show that the effect does not saturate even down to 0.001rad/s.

Examples of LDR effects



A.H.Johnston et al., JPL internal report, 1999

Effects on TID response

Bias: highly process dependent

Temperature: higher degradation at high T

Annealing: different behaviour of NPN and lateral PNP

Federico Faccio/CERN

Bias sensitivity:

- can vary greatly between processes

- several types of ICs at 50rad/s have shown no difference between biased and unbiased irradiation

- other ICs show an enhanced degradation under bias

- Therefore, the first time a part type is tested, one should evaluate its bias sensitivity without relying on extrapolating results from similar parts but different manufacturer

Temperature effect:

- in general, irradiation at high temperature enhances the damage. One should be careful not to exceed a temperature of about 90°C, above which trapped holes anneal!

Post-irradiation Effects:

- again it depends on the process

- transistors: NPN recovers (more at high T), lateral PNP continue to degrade (both at room T and high T)

- Circuits respond therefore differently whether the dominant degradation mechanism is related to NPN or PNP transistors. Since in most cases the dominant mechanism is not known, a 25°C anneal may be the optimal T to accelerate the PNP degradation without too much recovery in NPN.

LDR: possible test procedures

- **High temperature test (also advised by JPL, but for TID above 30krad)**

- **JPL advice below 30krad:**
  test at 0.005-0.01 rad/s (1.5x max TID)
  test at 50 rad/s (2x max TID)
  do not use the device if failure in any test
  do not use devices with no LDR data!

Federico Faccio/CERN

For high TID levels (above 30krad), there is a need to accelerate the damage mechanism. Several laboratories have proposed to use high temperature during irradiation at high dose rate (50 rad/s or more). The temperature maximizing the damage varies with the technology, between about 60 and 150°C. Recent results have shown that, even when this temperature is found, the high dose rate damage might still be a factor of 6 below that measured for a low dose rate test.

JPL advice:

1) above 30krad: use a dose rate of 0.5-2 rad/s
heat the device at 90°C
compare with high dose rate test at room T

in that case, add a wider security factor of 2 to the radiation design margin.

2) below 30krad: test at both high dose rate (50rad/s) and low dose rate (preferably 0.005 rad/s) and compare the results. In this case, the test at low dose rate is manageable in a few weeks, and can give a direct answer regarding the device sensitivity to LDR effects. If the part fails at 1.5 times the foreseen TID in any of the two tests, then do not use it.

It is noticeable that JPL discourages strictly the use of any bipolar linear device without any data supporting its behaviour at low dose rate.

Displacement damage: sensitive devices

- **Bipolar linear ICs**
- **Optocouplers**
- **Some type of optical sources**
- **Optical detectors**

Displacement in bipolar devices

Gain degradation due to increased recombination of minority carriers in the base

Displacement damage equation:

$$1/h_{FE} - 1/h_{FE0} = \Phi / [K(2\pi f_T)]$$

NB: The majority of linear ICs are still manufactured in old junction-isolated processes, BUT using less conservative approaches (more PNP transistors used in critical places)

Federico Faccio/CERN

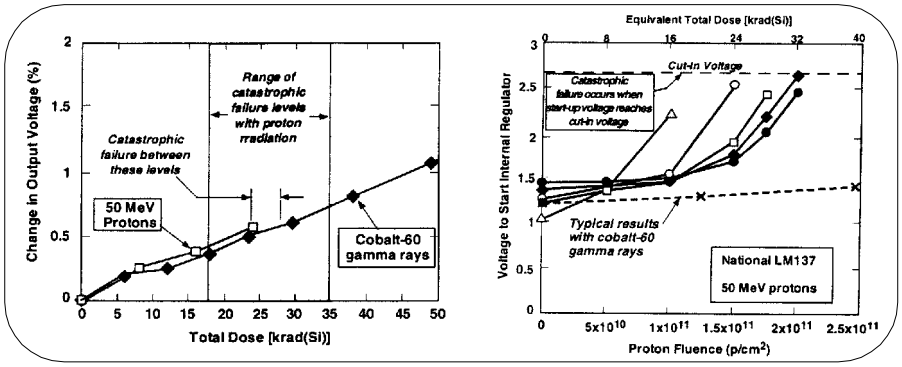
The sensitivity of bipolar transistors to displacement damage is due to the radiation-induced increase of the bulk component of the base current. Such increase is in turn due to the increased recombination of minority carriers in the base. This effect is particularly important in bipolar devices with great base thickness (lateral and substrate PNP transistors).

Also, the effect is more important when the devices is operated at a low injection level (one order of magnitude more damage than in the case of high injection, close to the gain peak in the Gummel plot).

Though new processes with higher bandwidth and thinner base region are available, most of the linear ICs are still manufactured in junction-isolated processes that have changed very little over the past 25 years. But older circuits were designed using a very conservative approach, which has been partially abandoned nowadays. The compromise PNP devices now available in such processes have better reproducibility, hence are now commonly used in critical positions in the circuits (as for input stages).

Displacement in bipolar devices: examples

LM137 negative voltage regulator



B.G.Rax et al., to be published in IEEE Trans. Nucl. Science, Vol.46, n.6, December 1999

Federico Faccio/CERN

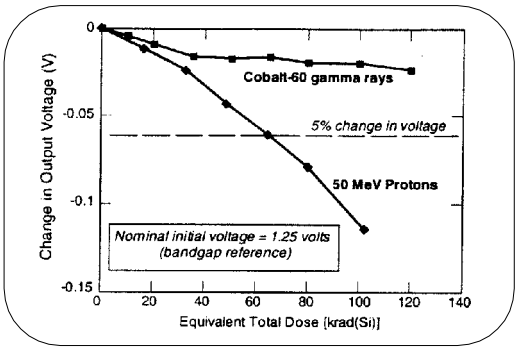
The proton test (50MeV protons) of the LM137 negative voltage regulator from National Semiconductors has shown catastrophic failure at an equivalent dose varying in a wide range (18 to 35 krad). In some cases, the failure occurred at fluences close to 10^{11} p/cm^2. When irradiated with gamma rays, no failure was observed up to a much higher TID, proving that the mechanism responsible for failure is displacement damage.

The failure was due to the increase of the minimum input voltage (for the device to operate) above the cut-in voltage. Therefore, even an increase of the input voltage to 40V could not bring the device into an operation condition.

It is important to notice that another LM137, manufactured by Linear Technology, did not show this failure mode and could operate up to fluences of 10^{12} p/cm^2. Hence, there is a clear difference between vendors.

Displacement in bipolar devices: examples

LM117 positive voltage regulator



B.G.Rax et al., to be published in IEEE Trans. Nucl. Science, Vol.46, n.6, December 1999

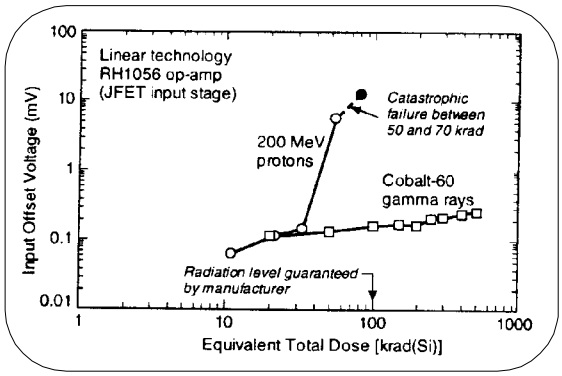
Federico Faccio/CERN

The LM117 from National showed a behaviour similar to that of the LM137 from the same manufacturer (start-up failure mechanism). This time, the effect occurred at higher radiation levels.

Nevertheless, the LM117 showed a large change in output voltage, a factor 5 higher than for the negative regulator and independent on the bias during irradiation.

Displacement in bipolar devices: examples

RH1056 radiation-hard op-amp



B.G.Rax et al., to be published in IEEE Trans. Nucl. Science, Vol.46, n.6, December 1999

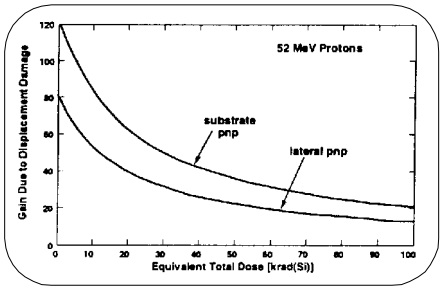
Federico Faccio/CERN

The RH1056 is guaranteed by the manufacturer to a total dose of 100krad, and works satisfactorily up to 1Mrad. Nevertheless, it fails catastrophically when irradiated with 200MeV protons, even though the total dose deposited is well below 100krad. This is clearly due to displacement damage.

Other non-rad-hard operational amplifiers, as the OP27 (Analog Devices) and the LT27 (Linear Technology), show a high increase in the input bias current when irradiated with protons. This increase, for the LT27, is significantly higher than for a pure gamma irradiation. Despite this increase in the input bias current, due to a malfunction of a lateral PNP compensation stage, both the OP27 and the LT27 continue to operate to equivalent TID above 100krad (in contrast with the RH1056 case).

Displacement in bipolar devices

Effects for lateral and substrate PNP



B.G.Rax et al., to be published in IEEE Trans. Nucl. Science, Vol.46, n.6, December 1999

B.G.Rax et al., to be published in IEEE Trans. Nucl. Science, Vol.46, n.6, December 1999

Displacement damage effects are generally negligible below $3 \cdot 10^{10}$ p/cm^2 (50MeV) also for PNP transistors

At levels above about $3 \cdot 10^{11}$ p/cm^2 , they start to become significant also for NPN transistors

Federico Faccio/CERN

Using data in literature, Rax et al. have calculate the nominal gain and resulting effects from displacement for substrate and lateral PNP transistors in a junction-isolation process from National Semiconductor.

Below a level of about $3 \cdot 10^{10}$ p/cm^2, the effect is generally negligible also for PNP transistors (for 50MeV protons).

Above $3 \cdot 10^{11}$ p/cm^2, the displacement effects start to be noticeable also for the NPN transistors, which have normally a higher f_T.

Displacement for bipolar devices

- **Results on biased and unbiased devices are almost identical**
- **Data in literature for NIEL show that 50MeV protons are 1.75 times more damaging than neutrons**
- **An exam of the circuit design might indicate the sensitivity to displacement damage**
- **Devices with very demanding electrical specifications are potentially more sensitive**

Federico Faccio/CERN

Experimental results have shown no evident bias dependence of the displacement damage effects.

Data available in the literature indicate that 50MeV protons are about 1.75 times more damaging than neutrons (1MeV equivalent). These data refer to NIEL (Non Ionizing Energy Loss), and are in agreement with results on a circuit (LM111 comparator from National Semiconductor).

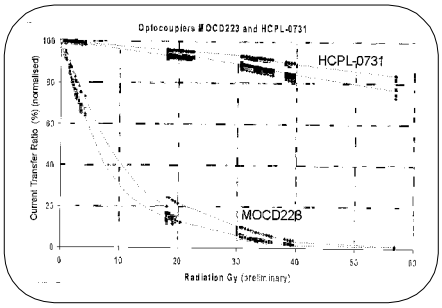
An exam of the circuit design, whenever this is accessible, might already reveal whether the circuit is very sensitive to displacement damage effects. This can help in deciding whether the displacement damage test is necessary. Key factors are the use of lateral PNP transistors in current mirrors or input stages, and the output stage design.

Devices with very high demands on electrical specifications are also potentially more sensitive (to all radiation effects). For instance, requirements on very low input offset voltage and/or input offset and bias current, or very low noise, … Any small modification of the transistor behaviour might have dramatic consequences on such ICs, even though the same modification would be negligible in circuits with wider design margins.

Displacement for optocouplers: example

Radiation tests of the ATLAS DCS front-end electronics at the CERN TCC2 area for the CAN Fieldbus - B. Hallgren

MOCD223 from Motorola => normalized CTR 0.65%
HCPL-0731 from HP => normalized CTR 77%
ILD206A from Siemens => normalized CTR 3.5%



B.Hallgren, CERN, 1999

Federico Faccio/CERN

Optocouplers are known to have a particularly high sensitivity to displacement damage. As an example, the results obtained at the TCC2 experimental area at CERN indicate a very variable degradation in the Current Transfer Ratio (CTTR) of 3 different types of optocouplers. This work has been done by B.Hallgren for the ATLAS collaboration. The radiation environment of the TCC2 experimental area is not exactly known, but is composed by gamma rays and neutrons (the energy spectrum of which has not been measured). The degradation observed on optocouplers is due to the neutrons, inducing displacement damage. Similar effects have in fact been observed by the ATLAS working group on power supplies, during irradiation performed at a pure neutron source (Prospero).

In some cases, it is possible to improve the situation by increasing the bias current of the optocoupler (method often used by the European Space Agency, ESA).

Single Event Effects (SEE)

Very localized event induced by a single particle (whilst TID and displacement are gradual cumulative effects).

They can be:

Transient	=>	**spurious signals propagating in the circuit**
Static	=>	**errors overwriting information stored by the circuit**
Permanent	=>	**or "Hard Errors", they are destructive events**

Federico Faccio/CERN

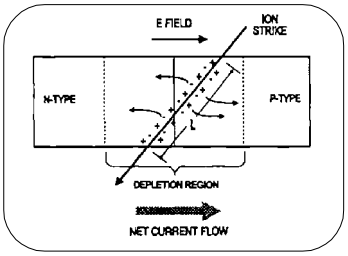
Transient errors are frequent in analog circuits, or in combinational logic. The generated signals are asynchronous, they can propagate through the circuit during one clock cycle and also sometimes propagate to a latch and become static.

Static errors can be corrected by outside control. They overwrite information stored in the circuit, but a rewrite or power cycle can correct the error with no permanent damage.

Permanent or hard errors are those leading to a permanent error, which can be the failure of the whole circuit. They cannot be recovered unless detected at their very beginning in some cases (as for Latchup). In that case, it is possible to interrupt the destructive mechanism and bring back the circuit to functionality.

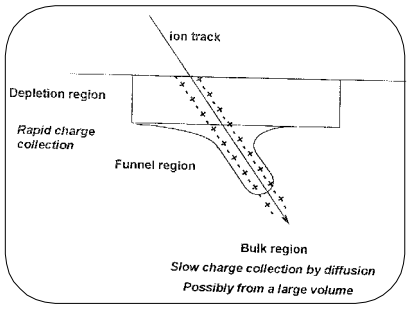
Single Event Upset (SEU)

Along the ion track, e-h pairs are created. In presence of an electric field (depleted junction), the charge will flow and a current spike might be observed.



L.Massengill, IEEE NSREC short course, 1993

Charge collection has a prompt and a slow component, and might extend far from the depleted junction (funneling)



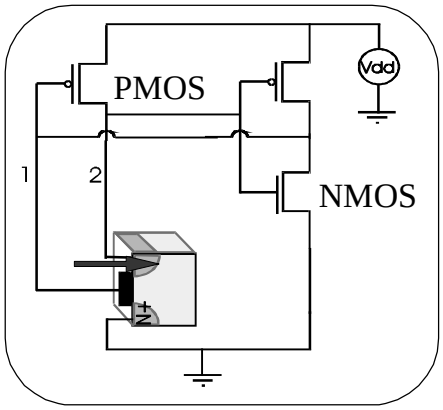
Federico Faccio/CERN

E.L.Petersen, IEEE NSREC short course, 1997

The incoming ionization particle loses energy in the semiconductor through Rutherford scattering (Coulomb interaction) with the lattice structure. The energy is transferred to the lattice as an ionization tail of free electron-hole pairs. In the bulk of the semiconductor, these will recombine with no effect. In a p-n junction or in its proximity, the pairs will be separated and collected, giving rise to a current spike.

The charge collection will have a fast (of the order of hundreds of ps or less) and a slow component (of the order of ns). The mechanism of charge collection are multiple, and the collection region might extend also relatively far from the junction through a phenomenon called “funneling”.

SEU: example



Where? ⟹ **Sensitive Volume SV**

Enough charge? ⟹ **Critical Energy E_{crit}**

Federico Faccio/CERN

The collection of charge at a circuit node might give origin to a transient (for analog circuits or combinational logic).

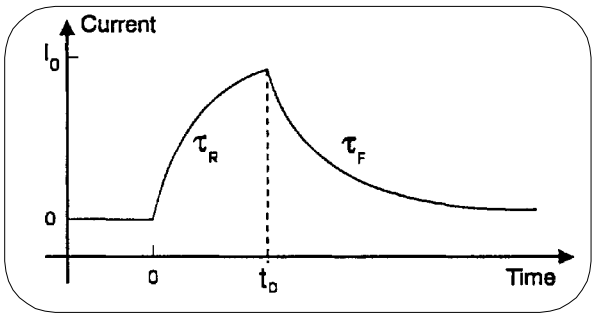
In other cases, the charge collection might upset the content of a memory cell, and in that case there will be a static error (Single Event Upset). An example is in the case of the SRAM cell shown in the figure. The charge collection at the drain of the NMOS transistor will temporarily change the state of node 2. Before the deposited charge might be evacuated to the power supply through the open transistor of this inverter, the second inverter (whose input is node 2) switches. This changes the state of node 1, which in turn enforces the wrong state at node 2. In this way, the error is latched into the memory cell.

Not all particles deposing energy in the semiconductor will induce SEU. Only the energy deposited in a range sufficiently close to the sensitive node can be collected and eventually lead to upset. Therefore, we can define a Sensitive Volume, which corresponds roughly with the volume where the charge deposited can be collected and actively participate to SEU. This is normally assumed to be a rectangular parallelepiped (RPP).

Also, not all particles deposit enough energy in the SV as to provoke a SEU. Very small energy depositions will lead to a minor change in the state of the node, which will not be transmitted or latched as a change of state (error). A critical charge must be exceeded for that. As in average 3.6eV are necessary to produce an e-h pair in silicon, the critical charge can easily be translated into a critical energy.

Critical charge from SPICE simulation

For ASICs, it is possible to localize the sensitive nodes and to estimate the critical charge from SPICE simulations. One can therefore also estimate the sensitive area from the layout of the circuit.



L.Massengill, IEEE NSREC short course, 1993

Federico Faccio/CERN

The current spike induced by an ionizing particle can be simulated in SPICE. One common way is with a double-exponential spike, but most often a triangular shape for the current spike gives sufficiently precise results. This procedure can be very helpful in the design of ASICs, where it allows to study the possible SEU mechanism and act to reduce the total sensitivity of the circuit. Wherever absolute hardness against SEU needs to be achieved, special “hardened” architectures can be used.

The amplitude of the current is increased until the upset is observed in simulation, at which point the integral of the current can reveal the critical charge. This process, repeated on the different nodes of the circuit that seem to be sensitive, will help understand the SEU performance of the circuit, and improve it.

The knowledge of the sensitive nodes of the circuit also allows for the estimate of the sensitive area of the circuit. This can easily be done by looking at the layout of the circuit.

SEU

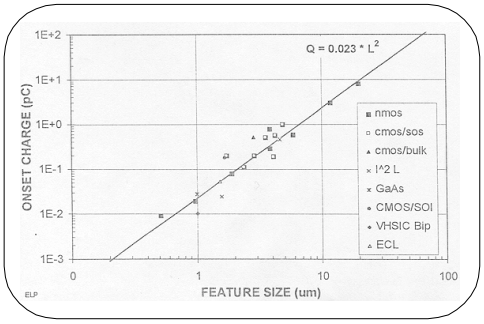
Not a destructive event => a given rate can be tolerated (depending on the system)

How often does it happen?

Environment

Which particles?
Which energy?
Which fluence?

Device sensitivity



Federico Faccio/CERN

E.L.Petersen, IEEE NSREC short course, 1997

A given rate of SEU can be tolerated, this rate depends of course on the system. Therefore, often the best approach to reduce the impact of SEU is to increase the tolerance of the system! (acting with redundancy, encoding of information, introducing error detection and protection schemes, using multiple voting,…).

But to understand whether the rate the system can stand will be exceeded in the application, there is a need to estimate how often SEU will happen. To do so, one needs to know:

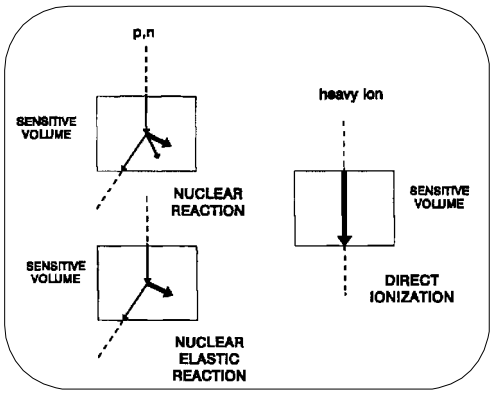
1) The radiation environment. Not only the kind of particles, but also their energy distribution and their fluence. For instance, it is impossible to get to an estimate of the upset rate starting from an environment description in terms of TID and equivalent 1MeV neutron fluence.

2) The specific sensitivity of the device. This might be argued by the technology used, but in general there is such a variability that one needs to test the device to really know. In that case, it is important to well target the irradiation source used in the test. This source has to be representative of the real environment (for example, using 1MeV neutrons for testing when the environment is represented by neutrons with energy up to 400MeV is NOT representative and will lead to completely meaningless results).

SEU: which particles?

Heavy ions (space) => high dE/dx (LET, in MeV•cm^2/mg)

Hadrons (LHC) => low dE/dx, but nuclear interactions



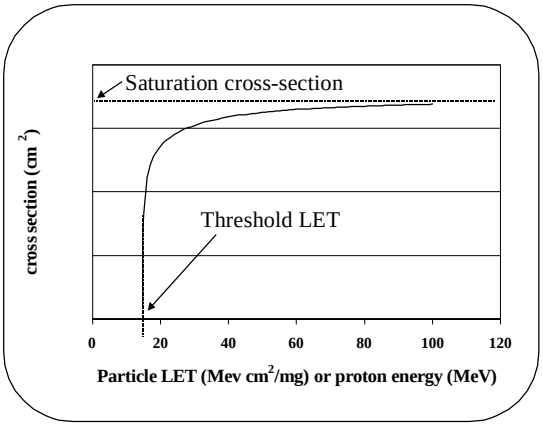
Federico Faccio/CERN

In space, heavy ions are naturally abundant and represent one of the biggest sources of SEU. They have enough stopping power to deposit a significant amount of energy (hence charge) in a small volume of semiconductor. In the Sensitive Volume, this might easily lead to SEU. Each ion with sufficient LET crossing the sensitive volume will finally induce SEU (if the track inside the SV is long enough).

Instead, hadrons as protons, pions and neutrons, cannot deposit by direct ionization enough charge to upset state-of-the-art devices. This case is particularly important for us, as in LHC the radiation environment will mainly be dominated by charged hadrons and neutrons. These particles can nevertheless induce SEU. They can in fact interact (elastic or, more important, inelastic nuclear interaction) with nuclei in the SV or in its close surroundings. The recoils from the interaction can in turn have a dE/dx high enough to induce SEU.

Characterizing the SEU sensitivity of a device

Device Cross-section (σ)



Proton beams:

$$\sigma = \frac{N_{events}}{\Phi} \quad (cm^2)$$

Heavy ion beams:

$$\sigma = \frac{N_{events}}{\Phi \cos\theta} \quad (cm^2)$$

Federico Faccio/CERN

The device SEU sensitivity is characterized by its cross-section curve. This curve is normally measured irradiating the device with a particle beam, most often with protons or heavy ions.

In the case of a proton beam irradiation, the cross-section is calculated as the ratio of the number of SEU over the proton fluence. This result is irrespective of the incident angle during the test, and in this case cross-section is plotted as a function of the proton energy.

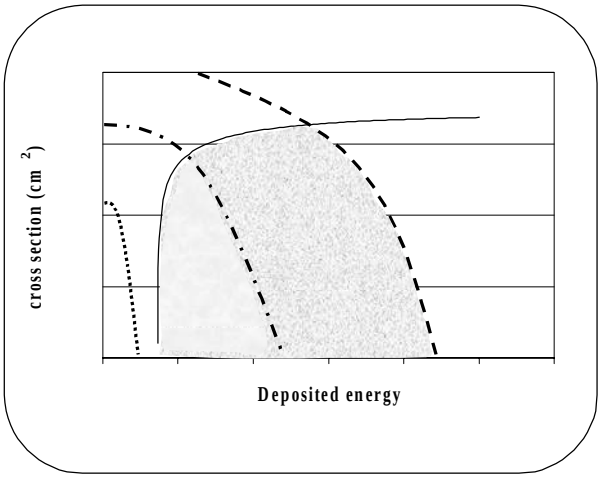
In the case of a heavy ion beam, the cross-section is plotted as a function of the particle LET. Since the change in LET during test is sometimes done by tilting the device of an angle θ (normally below 60°), both the effective LET and the fluence need to be corrected to take this tilting into account:

$LET_{effective} = LET/\cos\theta$

for the cross-section, the formula taking into account the tilt angle is given in the overhead.

From the heavy ion cross-section curve, one can get the Threshold LET (below which there is no SEU) and the saturation cross-section (the total area of the device which is sensitive to SEU). Both the heavy ion and the proton cross-section curves indicate the sensitive area of the device at a given LET/proton energy. This area can be divided by the total number of sensitive nodes (for example, for an SRAM memory, the total number of bits stored) to obtain the sensitive area of each individual cell. This corresponds to the surface of the Sensitive Volume.

SEU rate estimate



Federico Faccio/CERN

The cross-section curve can be used to estimate the SEU rate the device would experience in a given radiation environment. To do so, there is the need to have both the environment and the cross-section curves plotted on the same scale, as visually done in the overhead.

One possibility is to translate the heavy ion cross-section curve as a function of the energy deposited by the ion in the sensitive volume. To do so, there is "only" the need to know the thickness of the sensitive volume, which is not easy. Nevertheless, a reasonable value might be guessed, leading to a meaningful estimate.

Of course, to complete the process, one needs to have also a description of the environment in terms of the probability to have the given energy deposited in the Sensitive Volume.

This approach has been followed by several authors, recently it has been shown to lead to results in good agreement with experiments by Normand (Boeing Space and Avionics). We have used a very similar approach as Normand, and also obtained a good agreement with experimental benchmarks (Huhtinen and Faccio, "Computational methods to estimate Single Event Upset rates in an accelerator environment"). In such work, we also show that an irradiation with protons directly give an estimate of the SEU rate in LHC: the measured cross-section at an energy of about 60MeV or more, multiplied by the total hadron flux (above 20MeV) foreseen in the position of interest in CMS, gives the expected upset rate. See the paper for the detailed description of the calculations leading to this conclusion.

Destructive SEEs (Hard errors)

- **SEBO => Single Event Burnout occurring in power MOSFET, BJT (IGBT) and power diodes**
- **SEGR => Single Event Gate Rupture occurring in power MOSFET & (less often) in CMOS ICs**
- **SEL => Single Event Latchup occurring in CMOS and bipolar ICs**
- **They can be triggered by the nuclear interaction of charged hadrons and neutrons**

Federico Faccio/CERN

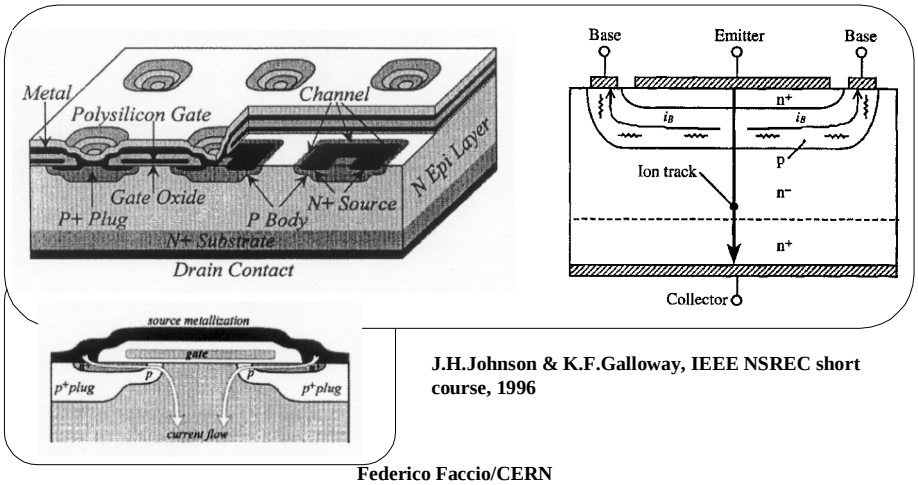
These destructive events are most often triggered by heavy ions, as the energy needing to be deposited to initiate the event is in general higher than for SEU. Nevertheless, the threshold of some devices for one of these destructive effect can be sufficiently low to endanger its survival in a “more benign” radiation environment where heavy ions are absent. Such cases have been observed already.

For instance, SEL threshold well below 10MeVcm^2/mg have been measured on ICs manufactured on some technologies. These devices will definitely be exposed to latch-up in an environment composed of high energy neutrons, as the CMS environment.

Another typical literature case is the failure observed by both European and Japanese train manufacturers because of SEBO or SEGR of a power MOSFET in the train engine. Such destructive event was induced by atmospheric neutrons, and could be reproduced in the laboratory with energetic protons or neutrons.

SEBO (SEB)

Double-diffused MOS (DMOS) power transistor and power BJT transistors are vulnerable



J.H.Johnson & K.F.Galloway, IEEE NSREC short course, 1996

Federico Faccio/CERN

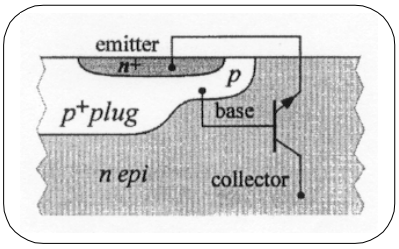
Power transistors (both MOSFET and BJT) are vulnerable to SEB.

The cross-section view of a DMOS power transistor is shown in the overhead. The thick epitaxial drain region is required to drop the large drain to source voltages that the transistor must block when operating in the OFF state. Typically, thousands of cells are connected in parallel to effectively create a very wide channel to achieve the large currents required in the ON state.

The vertical structure of a power BJT transistor, also shown in the overhead, is very similar to that of the DMOS. This is why they are both susceptible to SEB.

SEBO (SEB)

Mechanism: passage of the ion in the OFF state, generating a transient current. A regenerative feedback occurs until second breakdown sets in and permanently destroys the device (short source-drain or emitter-collector).



J.H.Johnson & K.F.Galloway, IEEE NSREC short course, 1996

Important mechanism in the regenerative feedback: avalanche-generated hole current in the collector region of the parasitic (or main) bipolar transistor.

Federico Faccio/CERN

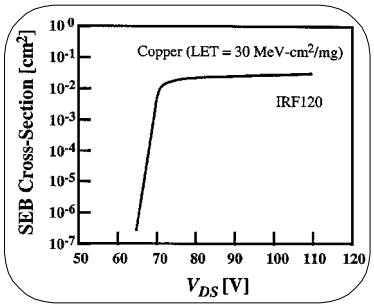
The sensitivity to SEB of the power devices is in the OFF state. In that case, the device is blocking a high drain-source (collector-emitter) voltage. The passage of the ion induces a current transient, turning on the parasitic bipolar structure in the MOSFET or the main transistor in the BJT. At that point, a regenerative feedback mechanism might set in, and the current increases until second breakdown and finally permanent device destruction.

A key component of the regenerative feedback is the avalanche-generated current in the collector region of the parasitic (or main) BJT.

For this reason, power P-channel MOSFETs are much less sensitive to SEB than their N-channel counterpart (impact ionization rate for holes is much less than for electrons).

SEBO: experiments

A cross-section curve as a function of Vds is often used



J.Titus et al., IEEE Trans Nucl. Science, Vol.43, 1996

Trends: **susceptibility decreases with incident angle**
susceptibility decreases with temperature

Noticeable amount of experiments on neutron and proton-induced SEBO (and probably observed even at TCC2-CERN)

Federico Faccio/CERN

Aside from the conventional cross-section curve as a function of the heavy ion LET, it is usual practice to plot the cross-section as a function of the applied voltage Vds. This shows that, below a given Vds, no SEBO is observed. Therefore, derating the device is a possible solution, though the necessary derating might be quite important (more than 50%).

Experiments with heavy ions have shown that the SEB susceptibility decreases with the angle of incidence of the ion and with the temperature of the test.

Recent results have been published on SEB tests of power MOSFETs exposed to neutron and proton beams. In general, 400 and 500V MOSFETS exhibited SEBO when operated at voltages above 300V. Even 200V n-channel MOSFETs underwent SEBO by neutrons at voltages above 190V (IRF250) and by protons when operated above 170V (2N6798). Measured cross-sections varied with the applied voltage, ranging from 10^{-6} (for 400V parts operated at 400V) to 10^{-10} cm^2 (for 200V parts operated at 200V or 400V parts operated at 300V). SEBO could also be induced by 14MeV neutrons from a D-T generator in 400V and 500V MOSFETs.

There has been an evidence of probable SEBO also in the radiation tests the LHC machine team is running at CERN at the TCC2 facility. In this environment, dominated by gammas and neutrons, the standard VME power supplies (WES V422B) failed three consecutive times, after a very variable TID (and operation time). The range is so wide (failures after an operation time between 6 and 263 hours in the same conditions) that it is very reasonable to think that the origin of the failure is a SEE. The component that was traced back as responsible for failure was a power MOSFET BZU357 (rated for 1000V, 7A). Note that the derating in the application was important, since the device operates normally at 300V!

SEBO: decreasing sensitivity of power MOSFETs

- **Acting on the technology**
- **derating (reduce source-drain bias)**
- **use p-channel MOSFET instead of n-channel**
- **increase the temperature (but problems for long-term reliability)**

Possible methods to decrease the susceptibility to SEBO in power MOSFETs:

1) the most effective method is to extend the length of the p+ plug as far as possible, without interfering with the channel region. This is a technological solution, reducing the resistance of the base region of the parasitic bipolar junction transistor inherent to the power MOSFET. This in turn increases the current necessary to get regenerative feedback.

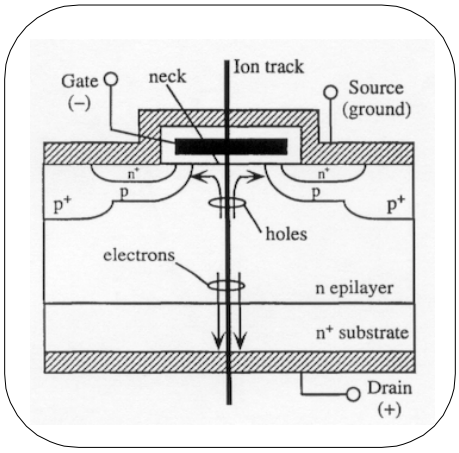
2) Decrease the source-drain bias (derating). This reduces the electric field in the base-collector depletion region, reducing the impact ionization.

3) p-channel MOSFETs are less sensitive to burnout than n-channel MOSFETs. For instance, no SEBO has been observed on p-channel MOSFETs irradiated with neutrons, protons or heavy ions.

4) As experimentally shown, the SEBO sensitivity decreases with the temperature. Nevertheless, increasing the temperature might not be possible, and also it rises doubts concerning the long-term reliability of the parts.

SEGR in power MOSFETs

SEGR is caused by heavy-ion-induced localized dielectric breakdown of the gate oxide



J.H.Johnson & K.F.Galloway, IEEE NSREC short course, 1996

Federico Faccio/CERN

SEGR is caused by a heavy ion strike in the neck region of the power MOSFET, as represented in the figure in the overhead.

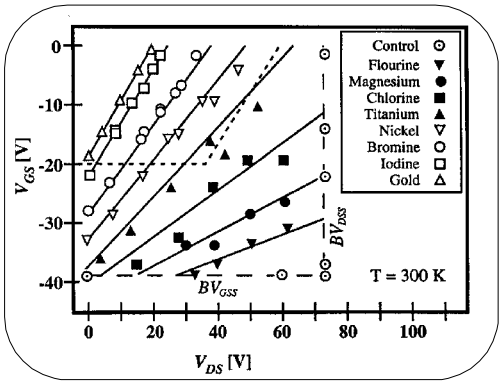
In the presence of a negative gate potential (which is the case for OFF power MOSFET devices), the field across the gate oxide can be large, but not as large as to cause the isolator failure. However, the charge deposited by the ions might change this situation, especially when a high Vds is applied.

The charges deposited by the ion are separated, the holes migrating towards the Si-SiO2 interface (the neck region) and the electrons towards the n-doped substrate (drain). The movement of all these excess carriers produces voltage drops that locally weaken the space-charge region between the n-epitaxial layer and the p-diffusions. This permits a dangerously large electric field in the oxide, which can exceed the critical field and lead to a localized gate rupture. Once the rupture is initiated, current flow through the gate oxide to the poly results in a thermal runaway condition, locally melting the silicon, dielectric and poly.

Both n-channel and p-channel power MOSFETs are sensitive to SEGR.

SEGR: experiments

To date, there exist no non-destructive SEGR test method!



C.F.Wheatley et al., IEEE Trans. Nucl. Science, Vol.41, N.6, Dec.1994

SEGR is sensitive to changes in both Vgs and Vds

Federico Faccio/CERN

The test for SEGR is quite difficult and expensive, as there is no experimental way to stop the destruction mechanism and restore the device back in operational condition. Therefore, a wide set of devices is needed to a meaningful test.

The chart in the overhead shows clearly, for one particular device that the sensitivity to SEGR is strongly influenced by both the gate and the drain potentials. The quantitative behaviour of other device is different, still the qualitative behaviour is the same. The dotted line in the figure represents the maximum permissible operating region as specified by the manufacturer. In this region, only ions with LET close to 18 MeVcm^2/mg can induce gate rupture.

To date, there is no evidence that SEGR can be induced by either proton or neutron irradiation. Recent experiments on 500V p-channel power MOSFETs (Motorola MTP2P50E) has shown no SEGR during an irradiation with high energy neutrons.

SEGR: decreasing sensitivity of power MOSFETs

- **acting on the technology**
- **derating (reduce source-drain bias)**

Federico Faccio/CERN

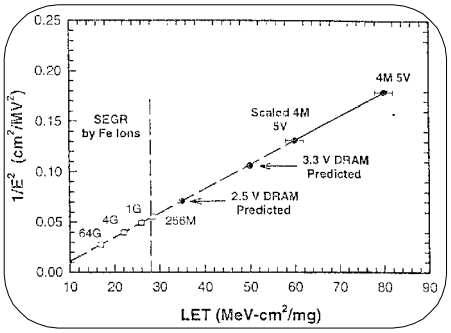
Possible solutions to reduce the power MOSFETs sensitivity to SEG:

1) Introduce technology modifications to lower the sensitivity. This might include the increase of the gate oxide thickness or the removal of the polysilicon gate that lies above the neck region of the device.

2) As in the case of SEBO, the sensitivity can be reduced by derating the device operation.

SEGR in CMOS technologies

Recent concerns in possible trend of SEGR in modern technologies



F. Sexton et al., IEEE Trans. Nucl. Science, Vol.44, N.6, Dec.1997

The most complete study so far has shown that SEGR should not become a significant concern if the gate oxide electric field is limited below 5 MV/cm

Federico Faccio/CERN

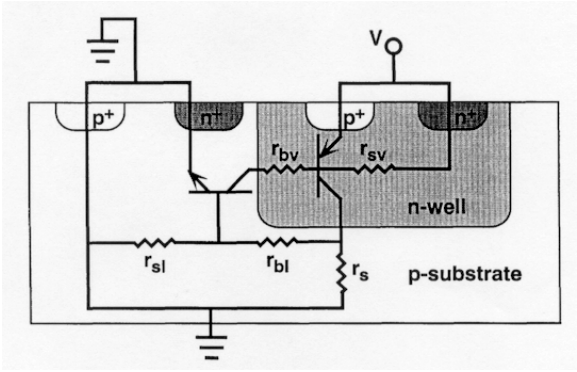
Recent extrapolations from available data have raised a concern on the possible increased sensitivity to SEGR of modern CMOS processes. This extrapolation predicted a threshold for SEGR below 29MeVcm^2/mg for an operating voltage of 2.5V.

The most complete study on that issue, performed by Sexton and co-workers at Sandia Nat. Lab. (New Mexico), has instead concluded that modern technologies will be more SEGR resistant at a given electric field. This is because, as the oxide gets thinner, the breakdown field increases due to the reduced defect creation by hot carriers in the oxide. However, there is a great deal of uncertainty in how voltage may be scaled with decreasing oxide thickness. This work shows that SEGR should not get a significant concern for devices that operate with gate oxide electric field below 5MV/cm.

Single Event Latchup (SEL)

Electrical latchup might be initiated by electrical transients on input/output lines, elevated T or improper sequencing of power supply biases. These modes are normally addressed by the manufacturer.

Latchup can be initiated by ionizing particles (SEL)



A.H.Johnston et al., IEEE Trans. Nucl. Science, Vol.43, N.6, Dec.1996

Federico Faccio/CERN

Semiconductor manufacturer are aware of possible electrical latchup initiated by transients on input/output lines, or improper power supply sequencing. Circuits are often protected against these failure modes.

Nevertheless, circuits operating in a radiation environment might be subject to an ionizing particle-induced latchup, SEL. The simplest model for SEL is the two transistor model shown in the overhead for a CMOS technology. The two parasitic bipolar transistors are interconnected such that the collector current of each BJT feeds the base current of the other. In such structure, an increase in pnp collector current gives an increase in the npn base current. This in turn increases the collector current of the npn, which gives an increase in the pnp base current. This positive feedback is such that, if the overall gain of the thyristor pnpn is high enough, any perturbation (for instance, an ionizing particle strike) turning on one of the parasitic BJT structures can trigger latchup.

After the latchup is initiated, it can be interrupted by promptly cutting the power supply to the circuit. In that case, the circuit can be saved from destruction and can be returned in the operational condition.

The value of the resistors shown in the picture in the overhead is very important to determine whether a structure has a high sensitivity to SEL. The charge deposited by an ionizing particle and not recombined flows to the power supplies through the well and substrate contacts. If the resistance along this path is high, the consequent voltage drop is high. Therefore, the local voltage might be quite different from Vdd or Vss, which means that the emitter-base junction is forward bias. This can start the injection process in the positive feedback structure.

Reducing SEL sensitivity

The best solution is to decrease the gain of the parasitic pnpn structure. Technological and layout solution can help in that respect:

Technological	**=> use of epitaxial substrates and retrograde wells** **=> use of trench instead of junction isolation**
Layout	**=> increase the distance between complementary devices** **=> use guardrings** **=> use lots of substrate and well contacts**

Federico Faccio/CERN

Several solutions are possible to decrease the SEL sensitivity of CMOS ICs. Some of them are technological solutions, other simply layout practices that can be quite effective. They all aim at reducing the gain of the parasitic pnpn thyristor, decreasing the gain of the two BJT transistors and decreasing the resistance of all circuit points to Vdd and Vss. The reduction of such resistance allows the flow of the ion-deposited charge to the power supplies without the forward bias of the emitter-base junction of the parasitic BJTs.

On the technological solutions, the use of epitaxial substrates and retrograde wells are well known. Epitaxial substrates limit the depth of the collection region after an ion strike, therefore limiting the amplitude of the current flow. The use of retrograde wells ensures a lower resistance path from every point in the circuit to Vdd. Trench isolation is also effective to reduce the sensitivity to SEL, as it dramatically decrease the gain of the lateral BJT parasitic structure. An extreme case is for the SOI technology, where the oxide isolation is complete, and no parasitic pnpn structure exists.

Also on the layout side, it is possible to reduce the resistance along the current paths by an abundant use of substrate and well contacts, well distributed all over the circuit. The extreme of such approach is the systematic use of guardrings at the edge of the wells and in the substrate close to the wells. This technique has been shown several times to be very effective, though area-hungry. Another approach is to decrease the gain of the lateral parasitic transistor. To do so, one can increase the distance between complementary devices.

SEL: experiments

Experiments aim at tracing the cross-section curve

SEL sensitivity is enhanced by temperature, hence the test should be done at the maximum foreseen T

Though in general modern technologies should be less sensitive to SEL, there are exceptions!

SEL can be induced by high energy protons and neutrons, and this case is quite frequent

Federico Faccio/CERN

The typical SEL experiment, either run with heavy ions or protons, aims at tracing the cross-section for the device. When SEL is detected, the power line is cut for a short time, then the operating conditions can be restored.

As SEL sensitivity is highly temperature dependent (it increases with T), it is recommended that the test is run at the maximum temperature foreseen in the aimed application.

Modern technologies are in principle less sensitive to SEL because they most often have thin epitaxial layers, retrograde doping profiles and isolation is achieved with shallow trench (STI). Nevertheless, some circuits manufactured in advanced processes with relatively thin epitaxial layers have been shown to have a high sensitivity to latchup. Even though these are exceptions which might be due to obscure design choices, it is important to consider that it can happen!

The threshold for SEL can be low enough as to represent a problem in a proton or neutron environment. Measured proton cross-section on several components show a very wide variability (up to a factor of 300), and the same experiments have pointed out that it is difficult to directly correlate proton and heavy ion SEL results. The differences between the two types of experiments can be explained with the difference in the charge collection process of long-range heavy ions and short-range proton recoils in the different device types. Therefore, it is not always simple to extrapolate the SEL sensitivity of a device in a proton or neutron environment from the available heavy ion data, especially for devices with a low LET threshold for latchup.