POLITECNICO DI TORINO

SCUOLA DI DOTTORATO
Dottorato in Elettronica e Comunicazioni - XXII ciclo

Tesi di Dottorato

Project solutions for low-cost space missions

Stefano Speretta
Matricola 143631

Tutore
Prof. Leonardo M. Reyneri

Coordinatore del corso di dottorato
Prof. Ivo Montrosset

Marzo 2010

Acknowledgements

During these three years, while working on this thesis, I had the opportunity to work together with many people and I have now the possibility to acknowledge them.

First of all, I should thank my advisor, Prof. Leonardo M. Reyneri, for introducing me in this project and for the opportunity I had to learn many new things. I should thank him for the freedom I had during this work and for the patience he had in helping me. I should also thank the whole AraMiS team at Politecnico di Torino for their help during these three years, in particular Prof. Claudio Sansoè, Dante Del Corso and Claudio Passerone. A special thank should go to Nikos Andrikos, Marco Borri, Alberto Hurtado, Danilo Roascio and Maurizio Tranchero for the continuous support I got for them and because they made these three years funnier.

I should also thank the guys at ISISpace (The Netherlands) where I had the opportunity to work for a short period, in particular Wouter Jan Ubbels, Wouter Weggelaar and Gerard Albers for their help and for the fun on Friday afternoons.

Last but not least, I should thank my family and Elisa for standing me during this period, for being patient with me and for their continuous support, in particular in the bad moments.

Torino, May 18^{th}, 2010

Stefano Speretta

Summary

With the name "space mission", developers usually reference to artificial satellites orbiting around the Earth or around other celestial bodies. Satellites are complex machines made by several sub-systems which integrate state-of-the-art electronic, optical and mechanical devices. Beside the harsh space environment, the satellite should also stand launch accelerations an vibrations which usually last for only few minutes but can cause severe damages to it, if proper solutions have not been employed. Satellite development and launch are quite expensive, as compared to every-day electronic systems, because the environment poses strong constraints on it and because no servicing can be performed to repair damaged systems. Only few organizations world-wide could afford these high costs which were mainly due to fact that satellites were always seen as an ad-hoc system, developed and produced in just few items for a specific mission. Costs were extremely high (in the range of multi hundred millions or billion dollars) and only military and governative organizations could afford it. In the last years, this concept changed dramatically, such that many companies or even universities started developing their own satellites: cost reduction strategies allowed to shift mission cost down to less than one million for small missions and this triggered a widespread interest in space systems, also from the didactic point of view.

Many organizations around the world started researching about cost effective solutions for space systems and this thesis is going in the same direction. Many approaches can be followed to save money in the development, production, testing and operational phases of the project. In literature many approaches are suggested, but the most well-known one is the CubeSat: a cube-shaped 10 cm wide satellite weighting maximum 1 kg. This is a basic unit, which can be composed to create bigger structures: unfortunately this standard is limited to a mechanical modularity in the design. Taking modularity to a high level, involving the whole satellite could help in further reduce system costs and increase performances: this is the idea that lies behind the AraMiS architecture, which will be presented in this work.

This thesis will cope with several problems related to space systems development, and show some solutions that can help in both keeping system development and production cost low while still achieving good performances. In Chapter 2 the space environment in which satellites will have to operate will be presented, showing how to numerically evaluate satellite environmental constraints, with focus on low Earth orbit (LEO). Chapter 3 deals with particle interactions with matter and will be in particular about radiation effects on electronic components. Particle transport in matter will be addressed to evaluate the

shielding effects that a thin layer can have on particle fluxes, to better understand the radiation environment inside the satellite. A novel technique will be presented to compute protons transport in matter which speeds-up computation by many orders of magnitude.

Space systems development costs will be addresses in Chapter 4: a cost model developed by NASA will be presented, and based on it, cost reduction solutions will be presented. Modularity and cost-sharing between multiple missions will appear as optimal solutions for reducing development costs, while the use of commercial components (COTS) will be presented as a way to simplify procurement and further lower system cost. In Chapter 5, an overview of many low-cost design techniques will be presented, with a focus on those employed in the development of AraMiS.

In Chapter 6, the AraMiS architecture will be analyzed, focusing on the different modules this architecture is composed by and on the advantages that this novel architecture has to achieve high performances and fault tolerance with a low development and production cost. Chapter 7 deeper analyzes three AraMiS sub-systems, which were developed during these three years: a latch-up protection system used to protect commercial components from latch-up effects, a wireless data communication bus, developed for reducing harness mass and routing problems in a small satellite, a power management sub-system and a power distribution bus, used to route power to all the satellite sub-systems and to supply them. These sub-systems are the satellite backbone and their modularity and scalability gives great flexibility to the AraMiS architecture.

Chapter 8 addresses instead some of the tests that were performed at different levels on the system, to qualify it for space operations. Radiation tolerance tests were also preformed on some of the components that are used in the satellite to ensure their endurance.

Contents

List of Acronyms

AraMiS	Architettura Modulare per Satelliti (Modular architecture for Satellites)
ASIC	Application-Specific Integrated Circuit
AMCM	Advanced Missions Cost Model
BER	Bit Error Rate
CERN	Centre Européenne pour la Recherche Nucléaire (European Organization for Nuclear Research)
CMRR	Common-Mode Rejection Ratio
CMOS	Complementary Metal Oxide Semiconductor
COTS	Commercial Off-The-Shelf
CPU	Central Processing Unit
CREME	Cosmic Ray Effects on Micro-Electronics
DRAM	Dynamic RAM
EDAC	Error Detection And Correction
ELDRS	Enhanced Low Dose Rate Sensitivity
EMI	Electromagnetic Interference
ESA	European Space Agency
ESD	Electro Static Discharge
EVA	Extra Vehicular Activities
FAI	Fédération Aéronautique Internationale (International Aeronautic Federation)
FIFO	First In First Out
FIR	Fast Infrared

FPGA	Field Programmable Gate Array
GEANT4	GEometry ANd Tracking v.4
GENSO	Global Educational Network for Satellite Operations
GEO	Geostationary Earth Orbit
GPS	Global Positioning System
GSFC	Goddard Space Flight Center
IC	Integrated Circuit
IEEE	Institute of Electrical and Electronics Engineers
INFN	Istituto Nazionale di Fisica Nucleare (National Institute for Nuclear Physics)
IrDA	Infra-Red Direct Access
ISS	International Space Station
ITAR	International Traffic in Arms Regulation
ITU	International Telecommunication Union
JAXA	Japan Aerospace eXploration Agency
JPL	Jet Propulsion Laboratory
LED	Light Emitting Diode
LEO	Low Earth Orbit
LET	Linear Energy Transfer
LHC	Large Hadron Collider
LNA	Low Noise Amplifier
MBU	Multiple Bit Upset
MCU	Micro Controller Unit
MEO	Medium Earth Orbit
MIPS	Million Instructions Per Second
MIR	Medium Infrared
MOS	Metal Oxide Semiconductor

MPPT	Maximum Power Point Tracker
MULASSIS	MUlti-LAyered Shielding SImulation Software
NASA	National Aeronautics and Space Administration
NIEL	Non-Ionizing Energy Loss
N-MOS	N-type Metal Oxide Semiconductor
NTC	Negative Temperature Coefficient
OBC	On-Board Computer
PA	Power Amplifier
PCB	Printed Circuit Board
PDA	Personal Digital Assistant
PiCPoT	Piccolo Cubo del Politecnico di Torino (Small Cube from Politecnico di Torino)
P-MOS	P-type Metal Oxide Semiconductor
PPM	Pulse-Position modulation
PWM	Pulse-Width modulation
rad	Radiation Absorbed Dose
RAID	Redundant Array of Inexpensive Disks
RAM	Random Access Memory
RF	Radio Frequency
RH	Rad-Hard
RZI	Return to Zero Inverted
SAA	South Atlantic Anomaly
SCR	Silicon Controlled Rectifier
SEE	Single Event Effects
SEFI	Single Event Functional Interruption
SEL	Single Event Latch-up
SET	Single Event Transient

SEU	Single Event Upset
SIR	Serial Infrared
SHF	Super High Frequency
SMD	Surface Mount Device
SNR	Signal to Noise Ratio
SOC	System On Chip
SOI	Silicon On Insulator
SOS	Silicon On Sapphire
SPENVIS	Space Environment Information System
SPF	Single Point of Failure
SRAM	Static RAM
SRIM	Stopping and Range In Matter
SSTL	Surrey Satellite Technology Limited
STI	Shallow Trench Isolation
TID	Total Ionizing Dose
TMR	Triple Modular Redundancy
TTL	Transistor-Transistor Logic
UART	Universal Asynchronous Receiver and Transmitter
UFIR	Ultra Fast Infrared
UHF	Ultra High Frequency
USB	Universal Serial Bus
VFIR	Very Fast Infrared

Chapter 1

Introduction

Space systems are relatively young since they appeared for the first time on October 4^{th}, 1957, when the Russian Sputnik was launched. This was the first artificial satellite orbiting around the Earth but it opened the so called "space race" between Russia (URSS) and USA. Forty years later, artificial satellites became part of the every-day life: real-time weather forecast satellite images are usually published in news or GPS signals are used every day for navigation.

Satellites have always been looked as a highly expensive business, which requires deep knowledge and experience to achieve success: this idea was primarily linked to the high initial costs required for their development and launch. Furthermore, the impossibility to repair and substitute parts (this was true up to the first Hubble Space Telescope servicing mission in 1993) makes design even harder because it requires advanced fault tolerance solutions and extreme reliability. But with the evolution of satellites and the birth of commercial space companies, this market had a gradual growth up to now, when several private companies are operating, also providing launch services. Many more launch opportunities appeared then, thus lowering also the price for a satellite launch and this allowed an even higher business increase. Low cost design techniques played an important role in the aerospace market growth in the past years, but they can still play an important role in future developments. Many research institutions and commercial companies are in fact trying to further reduce satellite costs, and the latest results of this process is the CubeSat concept: a really small satellite, built using commercial components, that everybody could buy and assemble from a kit. This work goes in the same direction: the novel AraMiS (Italian acronym for Modular Architecture for Satellites) architecture will be described: its goals are to go beyond the CubeSat concept and create a true modular architecture. The main idea is the development of distributed and intercommunicating on-board units, that can be assembled together to fit mission specific requirements, thus allowing an effective cost sharing between different missions. This solution wants both to create cheaper systems and make design time faster.

The proposed architecture is intended for different satellite missions, from small systems weighting about 5 kg to bigger ones. Modularity is intended in many ways: from the mechanical point of view, like in the CubeSat concept, to allow the composition of

bigger structures in a simple way. But it is also intended from the electronic point of view: fitting such a wide range of applications, requires that most of the internal sub-systems are developed in a modular and scalable way.

1.1 Design Goals and proposed solutions

The main goal of this thesis is the definition and implementation of innovative solutions for reducing satellite cost, without reducing system safe margins with respect to traditional solutions. These techniques should be applied to the development of a satellite architecture that can fit many target missions: scalability and modularity are the two most important characteristics this novel architecture should have. The target environment for this new satellite is the Low Earth Orbit with an operating life of 5 years, even if the system should be flexible enough to be employed also in different conditions, with reduced performances.

The most effective way to reduce the cost of a nano- and micro-satellite missions is to reduce design and non-recurrent fabrication costs as much as possible, which usually account for more than 60 – 70% of the overall budget. Reducing them can be achieved by sharing the design among a large number of missions. Design reuse is the rationale behind the AraMiS project, that is, to have a modular architecture based on a small number of flexible and powerful modules which can be reused as much as possible in different missions.

Commercial components can also help in reducing mission cost bu they required proper safety margins to be considered during design to allow safe operations in the harsh space environment. Fault tolerance is a key topic throughout the whole work because this is the only way to guarantee safe operations in space for a long period.

1.2 Main contributions

After the initial documentation phase about the state-of-the-art low cost systems and design techniques, the first efforts were spent in defining the space environment constraints: this is extremely important because it allows to better select components and solutions for the mission.

The following step has been the definition of the satellite architecture, to better estimate satellite size, weight and internal organization. After this phase the design and development phase could start, which was centered around several satellite sub-systems: a latch-up protection system, a wireless data communication bus and a power management and distribution system.

The latch-up protection system is intended to protect components or systems against latch-up and in particular to prevent system damages due to high energy particle strikes. The wireless communication bus was developed to address the problem of harness inside the satellite: in small satellites usually there is not much free space for cabling and their integration is also quite complex; a wireless solution can become interesting also from the mass point of view. The power management system is usually quite mission-dependent and requires ad-hoc development and tailoring for the specific needs, which can increase

system cost and testing time. The basic idea that lead the development of this system was to make it modular and scalable, such that it could be adapted to many different situations and missions, thus lowering development and testing costs. Modularity was also exploited for power distribution: every power management module can be connected in parallel or series to the others to better to meet mission specific requirements.

1.3 Contents organization

This thesis will cope with several problems related to space systems development, and show some solutions that can help in both keeping system development and production cost low while still achieving good performances. In Chapter 2 the space environment in which satellites will have to operate will be presented, showing how to numerically evaluate satellite environmental constraints, with focus on low Earth orbit (LEO). Chapter 3 deals with particle interactions with matter and will be in particular about radiation effects on electronic components. Particle transport in matter will be addressed to evaluate the shielding effects that a thin layer can have on particle fluxes, to better understand the radiation environment inside the satellite. A novel technique will be presented to compute protons transport in matter which speeds-up computation by many orders of magnitude.

Space systems development costs will be addresses in Chapter 4: a cost model developed by NASA will be presented, and based on it, cost reduction solutions will be presented. Modularity and cost-sharing between multiple missions will appear as optimal solutions for reducing development costs, while the use of Commercial Off-The-Shelf (COTS) components can also help to simplify procurement and further lower system cost. In Chapter 5, an overview of many low-cost design techniques will be presented, with a focus on those employed in the development of AraMiS.

In Chapter 6, the AraMiS architecture will be analyzed, focusing on the different modules this architecture is composed by and on the advantages that this novel architecture has to achieve high performances and fault tolerance with a low development and production cost. Chapter 7 deeper analyzes three AraMiS sub-systems, which were developed during these three years: a latch-up protection system used to protect commercial components from latch-up effects, a wireless data communication bus, developed for reducing harness mass and routing problems in a small satellite, a power management sub-system and a power distribution bus, used to route power to all the satellite sub-systems and to supply them. These sub-systems are the satellite backbone and their modularity and scalability gives great flexibility to the AraMiS architecture.

Chapter 2

The space environment

This chapter is focused on describing the environment around the Earth, with special interest to low orbits, where low cost satellites are usually operating. Higher orbits or interplanetary ones put strong constraints on space systems which makes harder to implement cost reduction solutions. The space environment will be described to better understand which constraint will be mainly influencing satellite design. The first section is focused on the atmosphere top layers and on the issues related to them. The second section is dealing with ionizing particles sources and distribution around the Earth. The last section is not strictly related to the space environment and will be addressing launch vehicles constraints that should be addressed to ensure that the satellite will survive orbit injection.

2.1 The solar activity

The Sun is the source of energy for our planetary system and its activity has great influence on many aspects, from atmosphere density to particle flux. Most of solar parameters, like emitted energy or particle flux, radio emissions or Sun spot numbers are directly linked to its internal activity: the Sun follows a regular 11 years cycle which can be seen in Figure 2.1. Long term planning is extremely complex and subject to errors, while short term forecasts are more accurate: real time observation of the Sun can provide useful warnings of solar flare activity, as large proton events are usually associated with the strong emission of electromagnetic radiation, such as visible light, radio waves and soft X-rays during a flare.

Solar peak activity events can be harmful for satellites, but they are even worse for astronauts: sun spot prediction is thus vital in case Extra Vehicular Activities (EVA) have to be planned. Statistical forecasting based on past observations was the most used before the space age: now real-time particle flux measurement are available, allowing for example, astronauts to enter the safe areas in the Space Shuttle or the International Space Station (ISS) in case of strong radiation peaks. Before satellites came into operations, particles could only be directly measured with sounding rocket or balloons experiments.

Solar radiation spans over an extremely wide range of frequencies, starting from short

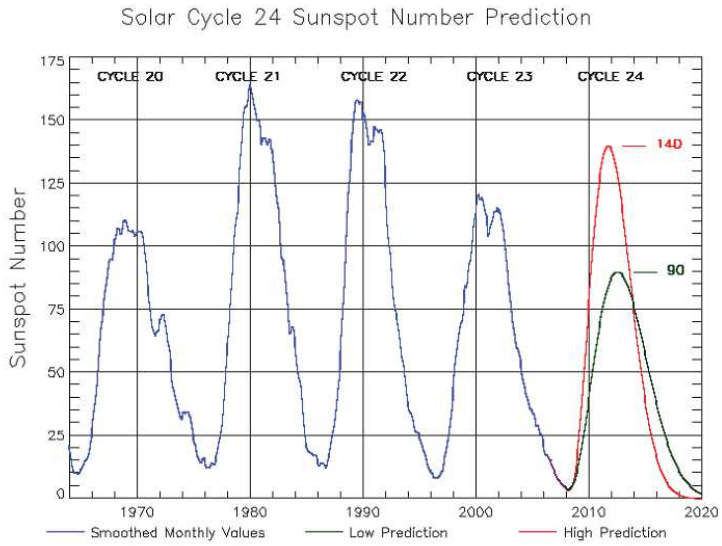


Figure 2.1: The Solar cycle (showing also a prediction for the 24^{th} cycle) [1].

waves (around 10 MHz) and going up to X- and γ-rays. The solar spectrum is depicted in Figure 2.2, showing similarities with the black-body radiation spectrum with an equivalent temperature of 5800 K. The integral over frequency of that curve is also called solar constant, equal to the total radiated power by the Sun which is equal to 1344 $\frac{\mathrm{W}}{\mathrm{m}^2}$ at the Earth distance.

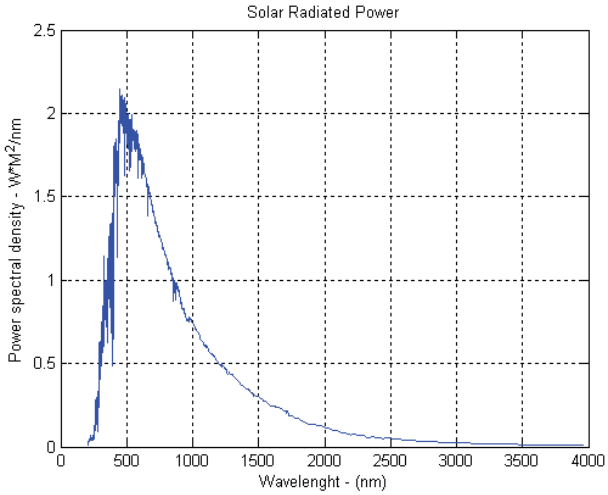


Figure 2.2: Solar radiated spectrum [2].

2.2 The top atmosphere layers

2.2.1 Residual air drag

The atmosphere top layers are conventionally separated by the lower layers by the Kármán line: it is an imaginary line set at an altitude of 100 km above the Earth's sea level, and is commonly used to define the outer space boundary. 100 km is the altitude where the required flying speed for an aircraft (speed required for its wings to generate enough lift to fly, which is dependent on air density) is equal to orbital velocity. The precise altitude is not exactly 100 km, but the Fédération Aéronautique Internationale (FAI) decided to use the round number because it was easier to remember.

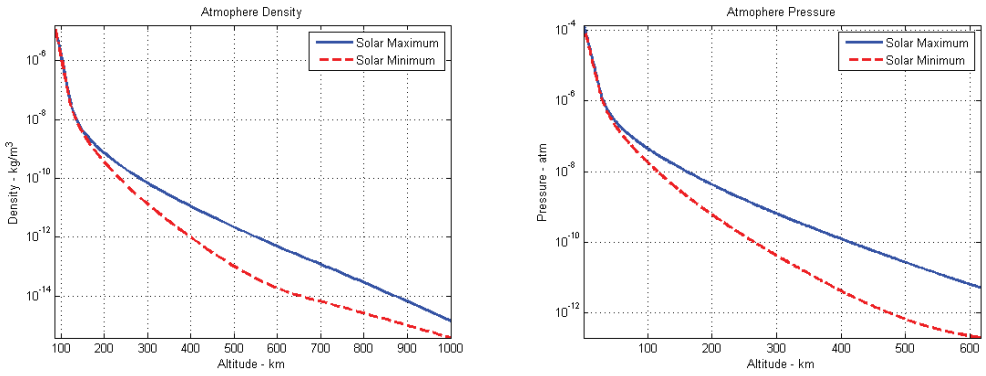
Solar activity mainly influences atmosphere temperature, which then makes gas layers expand showing a higher density and pressure at the top. Satellites orbiting in these top layers, also called Low Earth Orbit (LEO), experience a strong air-drag that can act as a de-orbiting force, actually slowing them down and making them decay down to the Earth. Drag is dependent on several factors: satellite speed (which is actually around 7 km/s in LEO orbit), atmosphere density (see Figure 2.3) and satellite ballistic coefficient, which measures its ability to overcome air resistance in flight and is define as follows:

$$Bc = \frac{m}{C_d \; A} \qquad \left[\frac{kg}{m^2}\right] \tag{2.1}$$

where m is satellite mass, C_d is the drag coefficient, equal to 2.5 for LEO satellites, and A is satellite frontal area.

Considering circular orbits, the period reduction rate $\frac{dP}{dT}$ can be computed from Newton's and Kepler's laws and is equal to:

$$\frac{dP}{dT} = \frac{3 \; \pi \; a \; \rho(a)}{Bc} \tag{2.2}$$



(a) (b)

Figure 2.3: Atmosphere model (computed with MSIS-86 [3]).

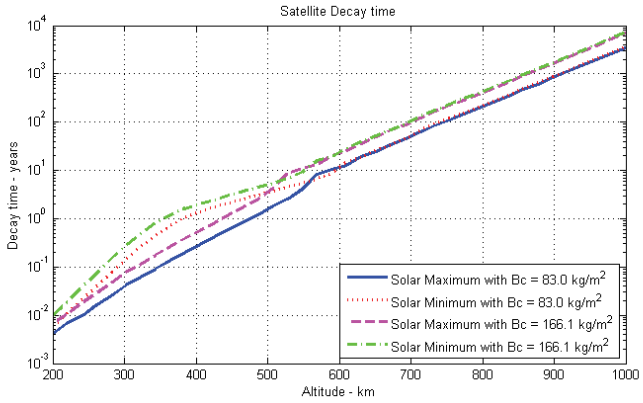


Figure 2.4: Satellite decay time. Satellite characteristics: **mass** 6 kg, **area** 29 cm^2, **Bc** 83 $\frac{kg}{m^2}$ and **mass** 48 kg, **area** 115 cm^2, **Bc** 166.1 $\frac{kg}{m^2}$ (computed using MSIS-86 [3] and equation 2.2).

where a is orbital radius, ρ is atmosphere density as a function of orbital radius and Bc is the ballistic coefficient. By integrating this law until satellite reaches 180 km altitude (this number was selected because at that altitude decay time is only few hours) decay time can be computed.

In the dense atmosphere layers air-drag can make satellites decay in just few days to few months while in higher orbits air-drag is greatly reduced thus allowing satellites to last up to hundred years. It is also easy to see that satellites with high ballistic coefficient will last longer in orbit, even in lower ones. There can be huge differences in orbit duration if the satellite is launched during solar maximum or minimum due to the different density of top atmosphere layers, but the two curves converge together for really low orbits (below 200 km) or for higher ones (above 600 km).

Objects with high ballistic coefficient and flying above 800 km can keep orbiting for hundreds of years if no alternative de-orbiting system is provided: this leads to the accumulation of space junk (or debris) in higher orbits that made collisions risk between satellites quite high.

2.2.2 Temperature range

Atmosphere pressure above 100 km is too low to allow convection to play a role in thermal exchange, so the only means to exchange heat in orbit are conduction and radiation. Thermal exchange with the environment is obtained through radiation, while inside the satellite both components can be important.

Sun direct light is usually the biggest external contribute to satellite heating (this is not the case for satellites always behind Earth shadow in Sun-Synchronous orbit) and other contribution is from the Earth reflected radiation. Earth radiation can be furthermore divided into two components: albedo and thermal emission. The albedo is the amount of

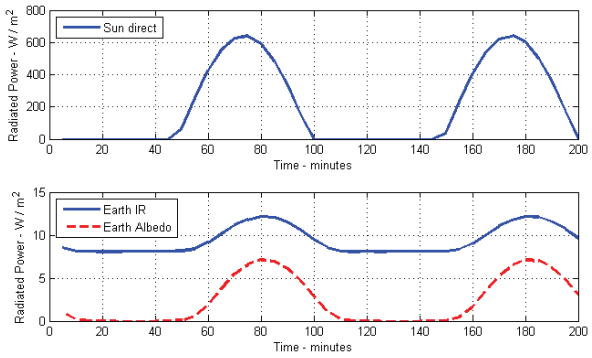


Figure 2.5: Solar radiation computed on a 800 km 87° orbit, on a surface 45° off from velocity vector (computed using SPENVIS [5]).

visible light the Earth is actually reflecting back to space: this parameter is widely varying due to cloud coverage or soil characteristics, but the planet average value is 0.3 [4]. The Earth thermal emission is instead the emission of a black body having the same temperature as the Earth (different on the day and night side): night side average temperature is constant over the position on the globe and is around -25°C, while day side temperature varies with the looking angle cosine and has usually a peak value of 33°C [5].

Since thermal exchange is varying across the orbit depending on Sun position, satellite temperature is varying from a minimum value (at the end of the dark part of the orbit) and a maximum value just before sunset. Under direct Sun radiation the temperature rises but the increase rate is limited by satellite thermal capacity. Thanks to the low orbit period (in LEO it is around 90 minutes) and to satellite thermal capacity, temperature changes are not severe and the equilibrium point is around 20°C. Thermal simulation is a quite complex task due to the complexity of satellite shape and material selection and can be precisely carried out only with numerical analysis. Figure 2.6 shows the temperature trend the Piccolo Cubo del Politecnico di Torino (Small Cube from Politecnico di Torino) (PiCPoT) [6] satellite was expected to experience in orbit.

Beside extreme temperature values, satellite should be able to stand temperature transients (some Celsius degrees per minute, usually, but in some cases can be much more severe). Furthermore, thermal cycles are continuous for the whole satellite lifetime: in LEO usually a satellite experiences 15 cycles per day, for an average mission duration of several years. This makes more than 5000 cycles per year which can be destructive for many electronic and mechanical components.

2.2.3 Plasma interactions

The topmost layer of the atmosphere has extremely low density and it is in direct contact with ionized particles coming from the Sun and ultraviolet radiation: all these contributions can make this layer ionized and thus highly reactive with orbiting objects. Particles

generated or attracted by the satellite are usually flying next to it, creating a long wake in the gas. While ions and electrons will be constrained by the magnetic field of the Earth, neutral particles generated by the spacecraft will be free to travel with the vehicle until disturbed by collisions. This will be especially of concern for large structures carrying out active emissions of neutral particles, e.g. water and waste dumps, thruster firings, atmospheric venting, etc. It is therefore important to clean every surface of the satellite and keep the satellite in a clean environment to reduce grease and dust contamination. Material selection is also important because vacuum can degrade materials breaking atomic links. Siliconic glue molecules, for example, is made by long oxygen chains kept together by weak links: when put in a low pressure atmosphere, these links can break, freeing oxygen atoms and making the glue ineffective.

Neutral particles can influence performances of satellite sub-systems such as optical transducer: dust clouds can stand in front of telescope apertures or solar cells thus reducing the effective light received and degrading experimental results. Charged particles instead can help electric current conduction and arcing: big solar arrays usually have exposed solar cell interconnects, prone to current leaking or arcing with other cell strings or satellite parts. Ionized particles can also be harmful for non conducting materials: since electric charge is not able to move freely through dielectric materials, charges will be collected, making potential build-up, until a discharge path is found. High voltage arcs can originate between two dielectrics, actually making holes in them (and sputtering atoms in clouds around the satellite). High potentials can also damage electronic components on the outside or the inside of the satellite if proper Electro Static Discharge (ESD) precautions are not taken. Most of the times dielectric materials are coated with low conductance films or paint, which helps removing collected charges and reducing potential build-up.

Ionized elements can also be much more chemically reactive with satellite surfaces: the most troublesome of these elements is atomic oxygen (already reactive when bi-atomic), but it is even worse when in mono-atomic state. It can react with thin organic films, advanced composites and metallic surfaces actually eroding them. Proper material selection is fundamental to ensure survivability to corrosion: gold plating for example is an effective

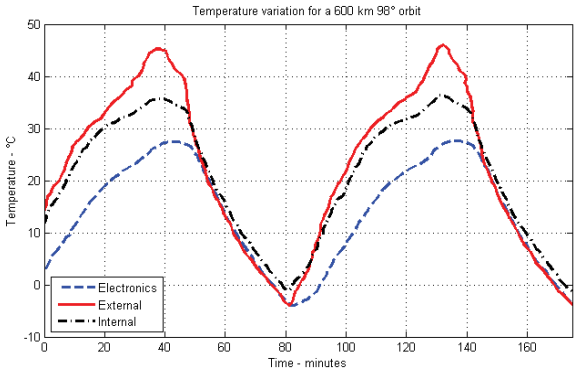


Figure 2.6: PiCPoT thermal simulation [7].

way of reducing corrosion in exposed metallic surfaces and Kapton films can be used to cover satellite external surfaces.

Material	**Corrosion rate per year**		
	500 km	**800** km	**1000** km
Kapton	1 um	294 nm	3.54 nm
Teflon	28 nm	0.8 nm	98 pm
Polycarbonate	3.4 um	98 nm	11.8 nm
Epoxy resin	0.96 um	27.8 nm	3.35 nm
Gold	0	0	0
Silver	5.94 um	172 nm	20.7 nm

Table 2.1: Atomic oxygen corrosion rate for 87° inclination orbits during solar maximum with different altitude (computed using SPENVIS [5]; data from the previous mentioned model and from literature suggest that Gold is experiencing no corrosion due to atomic Oxygen).

2.3 The radiation environment

The space environment is characterized by the presence of high energy particles: they were generated mainly by nuclear reactions in the universe and due to the highly energetic nature of those reactions, they were accelerated to extremely high speed and energy (up to 99.9999% of light speed and 57 EeV or 57×10^{18} eV in GRB 080916C126 [8]).

The electronVolt (eV) is, by definition, the amount of kinetic energy gained by a single electron, unbound from any atomic nucleus, when it accelerates through an electric potential difference of one volt. Thus it is 1 volt (1 joule per coulomb) multiplied by the electron charge (1.602×10^{-19} C). Therefore, one electron volt is equal to 1.602×10^{-19} J.

Before going on, it is useful to define two terms that will be used quite often:

- **Flux (ϕ)**: number of particles per unit area and per unit time
  ϕ = Particles/(Area×Time)
  Measurement unit: $Particles/(cm^2 \times s)$

- **Fluence (Φ)**: number of particles per unit area (time integral of the flux)
  $\Phi = \int \phi \, \mathrm{dt}$ = Particles/Area
  Measurement unit: $Particles/cm^2$

Radiation can be divided into 4 main components, which are:

- **Atomic nuclei**, from protons, α particles up to heavy ions
- β **particles**
- **Photons**
- **Neutrons**

Atomic nuclei are made by protons (having a positive charge) and neutrons (without an electrical charge), so they have a total positive charge. They have a short range (several centimeters) in air and cannot penetrate the outer layer of skin or a thin sheet of paper. These particles lose energy with every interaction with matter they experience, until they acquire one or more electrons and stabilize as neutral atoms. α particles consists of two protons and two neutrons and are identical to the nucleus of a Helium atom.

β particles can be either negative (electrons) or positive (positrons): They originate in the nucleus of an atom that undergoes radioactive decay: an electron is generated by a reaction changing a neutron into a proton, while a positron is generated by changing a proton into a neutron. β particles are smaller and more penetrating than alpha particles, but their range in tissue is still quite limited. When their energy is spent after interactions with matter, an electron attaches to an atom, while a positron collides with an ambient electron and the two particles annihilate each other, producing two gamma rays.

γ and X-rays are electromagnetic radiation given off by an atom as a means of releasing excess energy: they are quanta of energy that have no charge or mass and can travel long distances through air (up to several hundred meters), body tissue, and other materials. A gamma ray can pass through a body without hitting anything, or it may hit an atom and give that atom all or part of its energy. γ-rays are made by photons with energy higher than 100 keV, while X-rays are made by photons with lower energies.

Neutrons are particles with rest mass equal to proton mass but without an electrical charge: they cannot directly generate ionization, but generate it due to secondary effects. When neutrons collide with protons, β particles can be generated, which then act as ionizing radiations.

The main particles that can be found in space are electrons, protons, ions nuclei, photons and neutrons, each of them with a different flux as a function of energy. Particles can be divided depending on their sources:

- **Van Allen belts**, trapping electrons and protons,
- **Solar wind and flares**, photons, protons, neutrons and light ions,
- **Galactic sources**, generating photons and heavy ions.

2.3.1 Van Allen belts

The Van Allen belts are areas around th Earth were ionized particles are kept in place by the Earth magnetic field lines: the two belts, as can be seen in Figure 2.7, are not completely symmetric because of the solar wind which compresses the sun ward side and elongates the opposite one. Sometimes, due to solar activity and magnetic field fluctuations, particles spilled from the inner belt can cause the polar auroras phenomenon.

The outer belt is filled with electrons, while the inner one is filled both by electrons and protons as can be seen in Figure 2.8.

Sometimes a third belt can appears for a limited period of time in case of highly energetic events such as solar flares or high altitude nuclear explosions. On March 24, 1991, for example, an external belt originated due to a particularly strong solar flare

while on July 9, 1962, a new belt at 400 km appeared during the Starfish Prime nuclear experiment [9] and lasted for around 5 years.

Radiation belts are composed by trapped particles coming from the Sun, thus the solar cycle has great influence on them. The variation of solar radiance with the 11 years cycle induces a periodicity of the low altitude trapped proton and electron fluxes: during solar maximum the Earth's neutral atmosphere expands compared to solar minimum conditions, so that the low altitude edges of the radiation belts are eroded due to increased interactions with the residual atmosphere upper layers.

Particles in the radiation belts distribute according to magnetic field lines in different magnetic shells. These shells are surfaces generated by rotating a magnetic field line around the Earth magnetic dipole axis, and thus satisfying the following equation:

$$R \;=\; L\; cos^2\; \lambda \tag{2.3}$$

where R is the distance in Earth radii from the idealized point dipole near the Earth's center and λ is the magnetic latitude; L is the magnetic shell radius measured in Earth radii. This model was first introduced by McIlwain [10], taking also into account higher harmonic to model charged particles motion due to magnetic field perturbations.

The low altitude trapped particle population is also influenced by secular changes in the geomagnetic field: the location of the center of the geomagnetic dipole field drifts away from the center of the Earth at a rate of about 2.5 km/year (the separation currently exceeds 500 km), and the magnetic moment decreases with time. The combined effect is a slow inward drift of the innermost regions of the radiation belts.

The separation of the dipole center from the Earth's center and the inclination of the magnetic axis with respect to the rotation axis produce a local depression in the

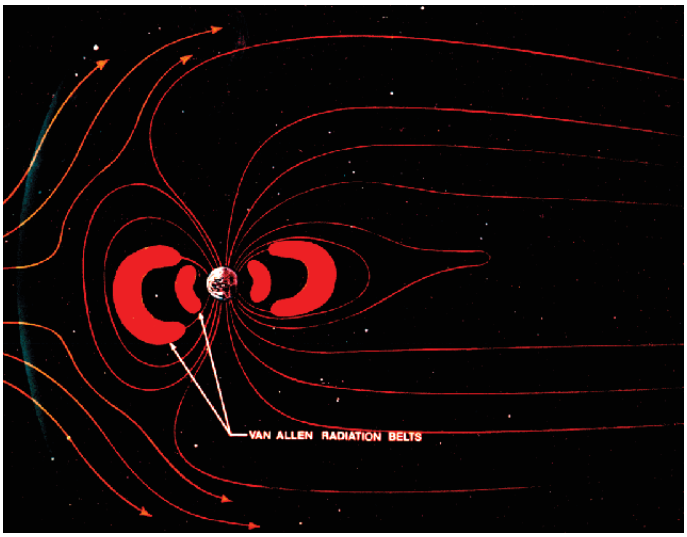
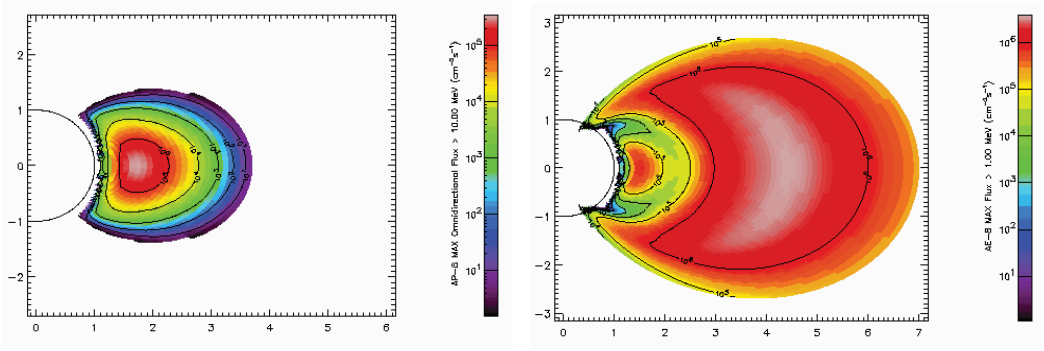


Figure 2.7: The Van Allen Belts.



(a) Trapped proton belt. From AP-8 MAX (b) Trapped electron belt. From AE-8 MAX

Figure 2.8: Van Allen belts particle distribution at solar maximum generated with AP-8 MAX (for protons) and AE-8 MAX (plots unit is the Earth radius) [5].

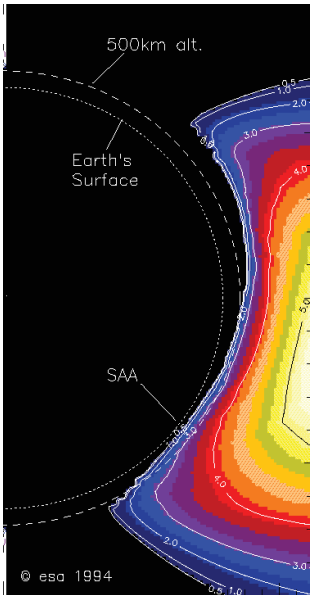
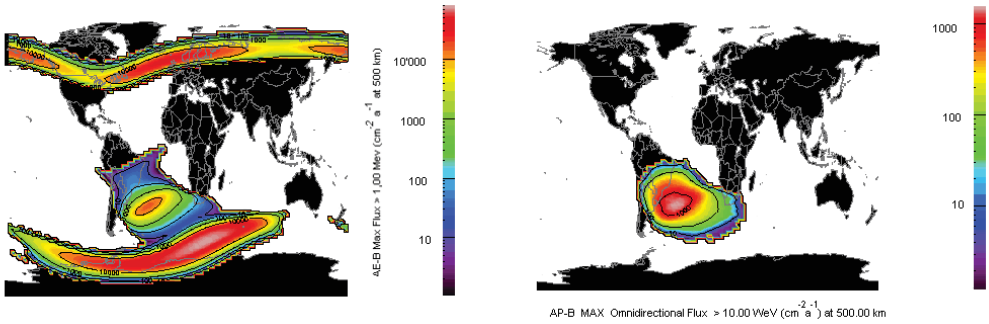


Figure 2.9: The South Atlantic Anomaly (SAA).

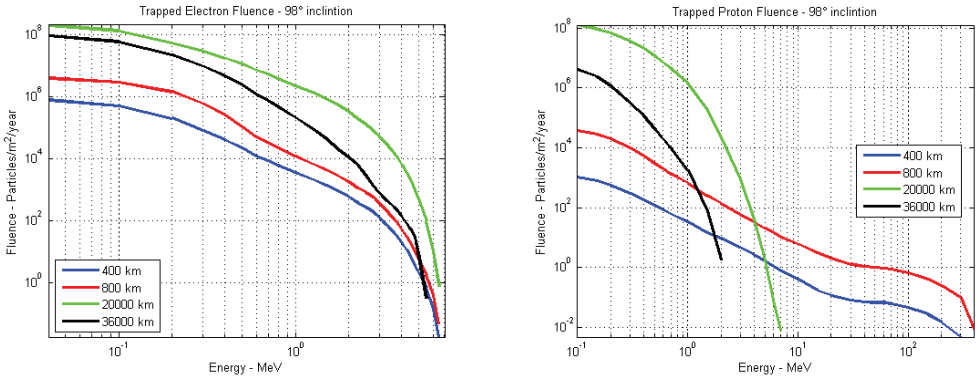
low altitude magnetic field distribution at constant altitude. As the trapped particle population is tied to the magnetic field, the lowest altitude radiation environment (below about 1,000 km) peaks in the region where the magnetic field is depressed. This region is located in the South East of Brazil, and is called the SAA.

Based on the geomagnetic coordinate system, standard models were developed to describe particle motion in the belts: the most used ones are AE-8 [11] for electrons and AP-8 [12] for protons, each of them able to represent particle flux during solar maximum and minimum. AP-8 represents mainly the inner radiation belt while AE-8 represents the



(a) Trapped electrons. From AE-8 MAX (b) Trapped protons. From AP-8 MAX

Figure 2.10: The South Atlantic Anomaly (SAA) projected on a world map. See [5] for further details.



(a) Trapped electrons (b) Trapped protons

Figure 2.11: Trapped proton and electron spectrum (computed using AP-8 and AE-8 during solar maximum).

whole area around the Earth up to around 7 Earth radii.

From the output of these models, the total amount of particles that a satellite can see while passing in the Van Allen belts can be computed: integration is performed on every orbit and this can give the energy spectrum of the incoming particles, as can be seen in Figure 2.11, where results were computed for several orbits with 98° inclination.

Particle flux in high orbits (above 2000 km) is isotropic but in lower orbits particles behavior is dominated by local magnetic field lines and by the upper layers of the atmosphere: particles trajectory (usually an helix when they are moving in a magnetic field) is deflected by the different density of the atmosphere layers generating the so called East-West effect [13]. Flux has two peaks on the East and West side of the satellite with respect to the velocity vector: differences can be as high as three or four times when compared to the front or back side.

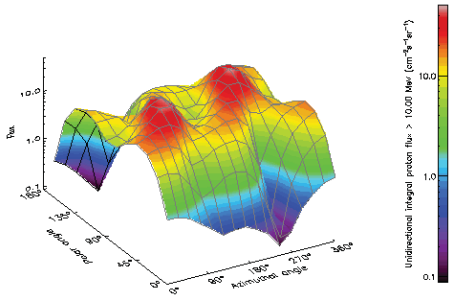


Figure 2.12: Proton flux anisotropy (represented on the plane perpendicular to magnetic field lines; polar angle = 90° and azimuth = 180° is coincident with satellite velocity vector) [5]

The gas giant planets Jupiter, Saturn, Uranus and Neptune, all have intense magnetic fields with radiation belts similar to the Earth's outer belt. Jupiter's belt is the strongest, first detected via its radio emissions in 1955 though not understood at the time. Jupiter's belt is strongly affected by its large moon Io, which loads it with many ions of sulfur and sodium from the moon's volcanoes. Saturn seems to have an "inner belt" similar to the Earth's, observed by Pioneer 11 during its 1979 fly-by and probably produced by cosmic rays which eject neutrons from Saturn's planetary rings.

2.3.2 Solar wind and flares

Solar wind and flares are the second most important source of radiation in Earth orbit and they are heavily dependent on the solar cycle. Protons and ion nuclei are the main particles ejected by the Sun and they are emitted in particular during solar flares, which can be seen as giant explosions on the Sun surface, but are also emitted (with a much lower flux) during the whole Sun life.

As shown in Figure 2.13, a solar flare can last for several days, but its effects can be seen just few minutes after the actual flare appears on Sun surface. Relativistic energy particles and photon can be seen first, while lower energy particles can take up to hours to travel due to the interaction with the solar corona and interplanetary medium. The frequency of such events can be seen in Figure 2.1.

Due to the interplanetary magnetic field, the most direct propagation line from the Sun to the Earth originate from the heliographic longitude of 60° West: solar flares originated in other regions can likely result in a negligible increase in particle count on the Earth. Early detection systems pay attention to this to avoid generating false alarms.

The ability of ionized particles to penetrate the Earth magnetic field is dependent on their rigidity, which is defined as follows:

$$R = \frac{mv_{\perp}}{B} \qquad [eV] \tag{2.4}$$

where m is particle mass, $v_{\perp}$ is particle speed component perpendicular to magnetic

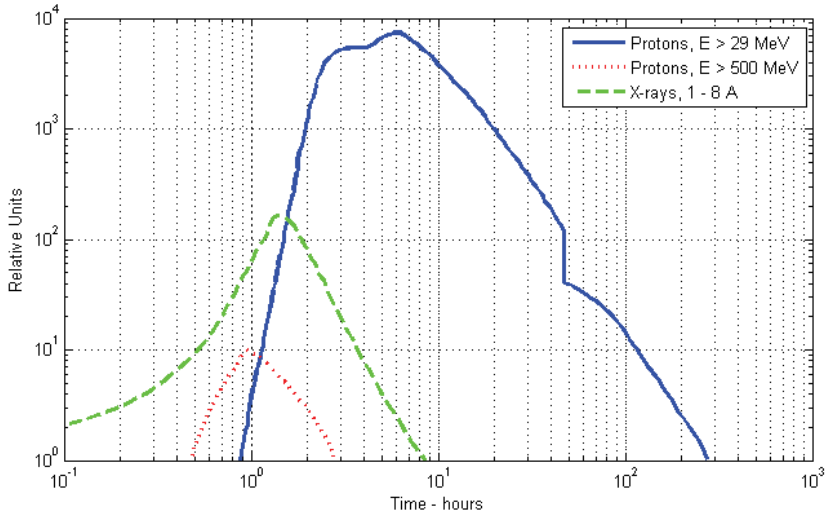


Figure 2.13: Solar flare time evolution [14].

field lines and B is the magnetic field strength. For every point and direction of approach in there is a rigidity threshold value (called geomagnetic cut-off) preventing particles with lower rigidity coming from a certain direction to reach the specified point (see Figure 2.14); particle flux is forced to zero when the Sun is shadowed by the Earth, so no straight propagation is possible. Particle spectrum along the orbit is computed by integrating the particle distribution multiplied by the geomagnetic cut-off and the Earth shadowing over the orbit.

The de-facto standard for proton spectrum modeling is the JPL-91 model [15][16]: it

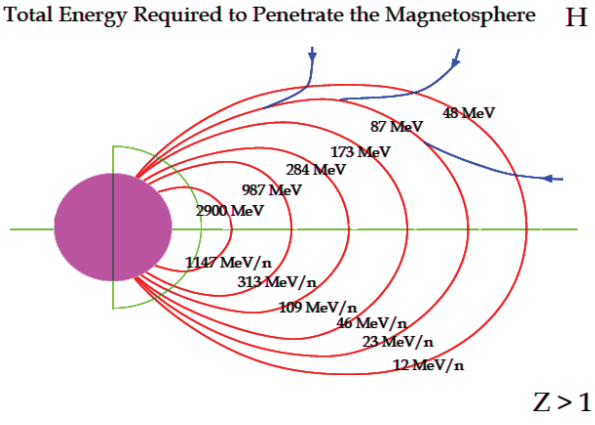


Figure 2.14: Geomagnetic cut-off for protons (below) and heavy ions (above).

is mainly based on data from solar cycle 19, 20 and 21 and it is the first model to properly model the asymmetric behavior of the Sun during the cycle (7 active years and 4 non-active years). The spectrum is accurately modeled for the active period, while in the non-active period no particles are ejected from the Sun. In the year 2002 the model was renamed to JPL, after verifying that it was still consistent over solar cycle 22 and 23 and not requiring anymore to be linked to the original dataset used to generate the model [17]. The proton spectrum along different orbits with 98° is plotted in Figure 2.15.

Beside protons, the Sun is also emitting several kind of nuclei generated by the internal nuclear reactions: this radiation contribution is modeled together with galactic particles and thus will be discussed in section 2.3.3.

From Sun nuclear reactions neutrons are also generated and accelerated to extremely high energies but quite few of them are actually able to reach the Earth: a free traveling neutron has a beta-decay half-life of 886 seconds. This is comparable to the light travel time from the Sun (499 sec), so only the fastest (relativistic) neutrons can make it to the Earth before decaying, while the others experience a β^- decay, where a neutron (n) decays generating a proton (p), an electron (e_-) and an anti neutrino($\bar{v_e}$):

$$n \rightarrow p + e_- + \bar{v_e} \tag{2.5}$$

The directly arriving neutrons contain essential information about the original particle acceleration, because they are not affected by the magnetic field. They are difficult to measure on the Earth because they trigger nuclear reactions when colliding with top atmosphere layers and thus decay to secondary particles: the only way of measuring them is thus by using satellite based detectors.

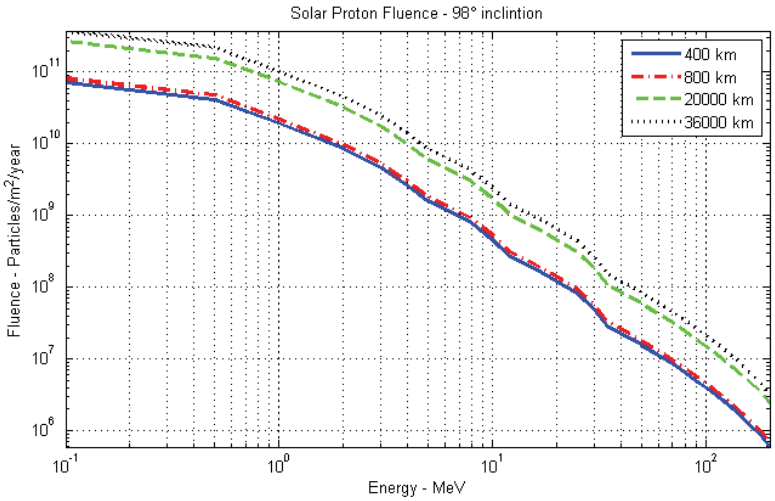


Figure 2.15: The Solar proton spectrum (computed using JPL model at solar maximum).

2.3.3 Galactic sources

Particles generated from galactic sources are also called cosmic rays due to the fact that they traveled extremely high distances before arriving to us (most of the times particles traveled from thousands to billion light years).

Cosmic rays are composed mainly by positively charged atoms which had their electrons stripped by the extremely high energy processes that accelerated them: protons are by far the most common, but other nuclei are also present (see Figure 2.16 for further details). Beside heavy nuclei, high energy photons are common (mainly γ-rays but also X-rays) but their influence on electronic systems is limited because their flux is low and their penetration in metals is limited.

Particles traveling to the Earth are also experiencing a deflection due to the magnetic field (geomagnetic cut-off) and Earth shadowing, as was happening with solar protons (see section 2.3.2 for further details).

The cosmic ray background is almost constant in all directions since all active galactic objects are producing cosmic rays (stars, supernovae, black holes), anyway a stronger flux is emitted by our galaxy, the Milky Way, because distance is lower. Cosmic rays measured by satellites (in particular measured before they interact with the atmosphere) are called primary cosmic rays because they were actually emitted by their cosmic source and traveled up to the Earth. After the top atmosphere layers, cosmic rays start to interact with residual gas molecules thus generating other particles through nuclear reactions: these are called secondary rays and are the ones usually measured from ground. Modeling cosmic rays is important to take into account their influence on the satellites: different models are available and they are distinguished by the metric used to model particle fluxes. For physical research purposes, it is important to model particle flux as a function of energy

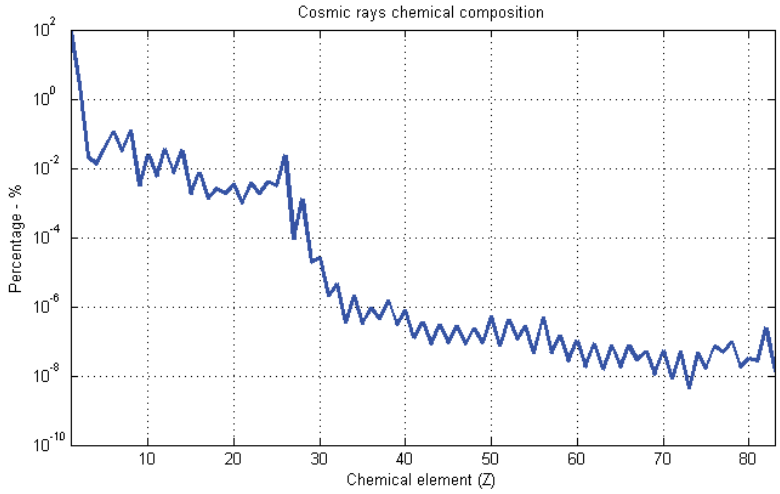


Figure 2.16: Cosmic rays flux as a function of chemical composition (computed using CREME-86 [18]).

or their atomic mass, while our interest is more related to the effects they generate in electronic components and in particular on their ability to transfer energy to electronic devices. The most used model for this purpose is called CREME [18] and it allows to compute particles flux as a function of their Linear Energy Transfer (LET) (see section 3.4 for further details). Furthermore, CREME allows to integrate cosmic rays contribution in different orbits, as can be seen in Figure 2.17, where four different orbits were analyzed.

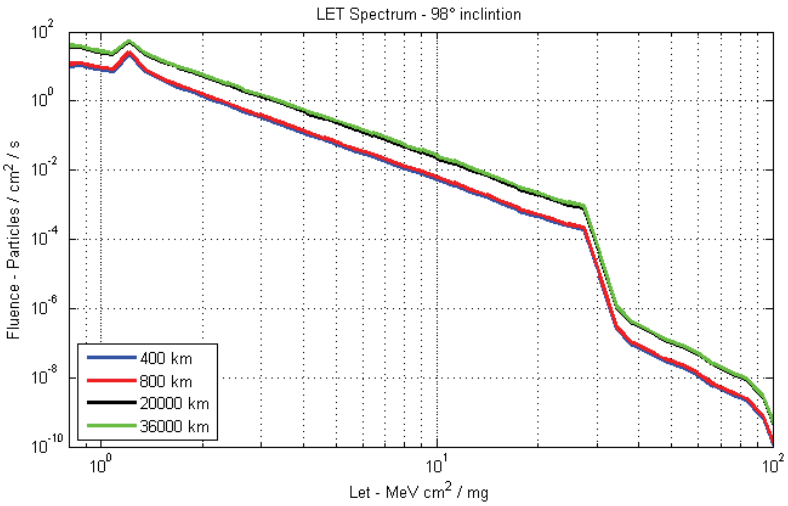


Figure 2.17: LET spectrum in different orbits (computed using CREME-86 [18]) with 2 mm Al shielding.

2.4 Satellite launch

This section is not completely related to the space environment, but is focused on the environment before and during satellite launch: failure to evaluate this variable can easily lead space missions to failure.

Before launch satellite is bolted on a release mechanism on the launch vehicle: this release mechanism can be an explosive bolt, or a spring based system. Explosive mechanisms are usually employed in bigger satellites, where a strong grip on the launch vehicle is important. Spring based mechanisms usually are used for smaller satellites (like CubeSats) because they allow multiple satellite to be released from the same device (the deployer acts like a dispenser releasing all satellites in sequence).

Satellites are launched by means of rockets, which impose strict requirement on satellite mechanical structure because it has to stand violent accelerations and vibrations. Specifications vary widely from rocket to rocket and compliance with these specification should be tested before launch to avoid satellite failure before injection into orbit.

Different kind of vibration test specifications were created to ensure survivability and the main ones are: sine-wave sweep, random and shock tests. In sine-wave sweep test a

sinusoidal vibration is applied to the satellite and then swept over the frequency range of interest to verify if vibrations are amplified by the mechanical structure: this amplified level may be the result of a vehicle anomaly, a primary structural resonance or a locally-induced perturbation. The frequency sweep rate is usually prescribed by every testing standard, but most of the times a value of one octave per minute sweep is used.

In random vibration test the stimulus is containing all frequencies at once (whose instantaneous magnitude cannot be explicitly defined) to better reproduce launch conditions. Shock tests are used instead to simulate separation mechanism firing and are performed by making the satellite undergo a sudden acceleration to verify if mechanical and electronic systems are still fully operational. For further details on test specifications see section 8.1.

Beside vibration problems, the satellite is also going to experience a sudden pressure decrease during rocket ascent and fairing separation: proper holes should be provided in satellite mechanical structures to allow the air inside the satellite to exit without damaging the system (see Figure 2.19 for an example pressure variation trend). Usually before rocket launch, the fairing is filled with gas up to a pressure slightly higher than ambient pressure to avoid contaminating gas and dust particles entering before launch.

(a) CalPoly single barrel P-POD[19]

(b) Planetary Systems Corporation separation ring

Figure 2.18: Separation mechanisms.

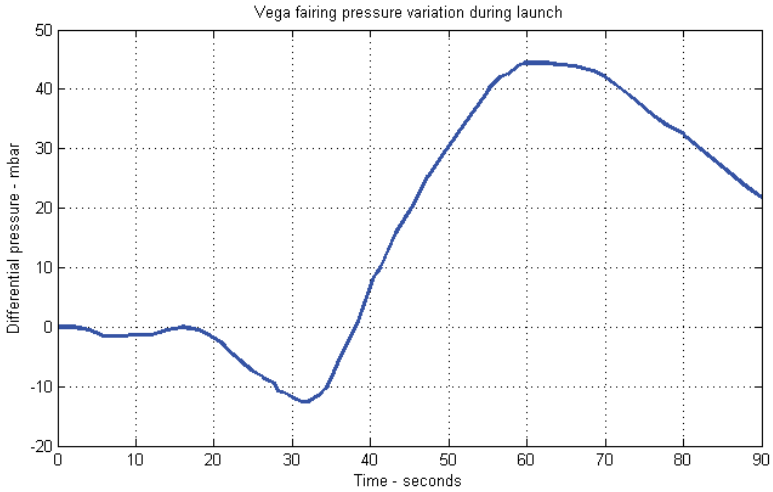


Figure 2.19: Vega fairing internal pressure during launch [20].

Chapter 3

Radiation effects on electronic components

Ionizing particles, penetrating through a material, lose energy in interactions with the atomic lattice and leave a charged wake behind them. These interactions can originate several effects in the substrate, thus influencing its physical properties. When electronic components are exposed to ionizing radiations, their electrical characteristics can change even dramatically, from the original, non irradiated ones.

This chapter will focus on the three main effects that can be noticed on electronic components: total ionizing dose, single event and displacement damages effects. For better understanding these effects and how to shield components from them, section 3.1 and 3.2 deal with particle interaction in matter and how to numerically compute particle transport: this can be used for evaluating the radiation spectrum that can be found behind a particle shield or inside the satellite and will give a better understanding of the next sections that describe the effects produced by high energy particles.

3.1 Particles interaction in matter

Charged particles passing through matter lose their kinetic energy by electromagnetic and nuclear interactions with the atoms of the lattice and this results in two main effects: collision energy loss and atomic displacement. Particles interactions which result in the excitation or emission of atomic electrons (for example ionizing the atom) are referred to as energy loss by ionization or energy-loss by collisions while the Non-Ionizing Energy Loss (NIEL) processes are interactions in which the energy transmitted by the incoming particle results in atomic displacements or in collisions where the knock-on atom does not move from its lattice location and the energy is dissipated in lattice vibrations (phonons). Depending on the particle, different effects and mechanisms arise and all of them contribute to the stopping power, that is defined as the derivative of energy over the penetration distance: it is in general measured in $\frac{MeV}{cm}$ on a specified material or, more in general, in $\frac{MeV\ cm^2}{g}$ which can be reduced to the first unit by multiplying it by material density.

Heavy particles (this way are called particles heavier than electrons, so protons and

heavy ions), when accelerated to high energy (higher than few tens of keV), have their electrons stripped out and, as bare nuclei, they travel through the lattice and lose energy due to interactions with the lattice atoms electrons and only rarely interactions take place with the nuclei of the medium. When the interaction is so strong to eject one electron from its orbital, the stripped electron is called δ-ray. The interaction that generated the stripped electron is called primary, while further interactions between this electron and lattice atoms are called secondary.

The energy lost by positively charged particles traveling through matter is described by Bethe-Block equation, which holds for the region between 100 keV/amu up to about 100 MeV/amu (this notation is used to be independent from ion size), while for higher or lower energies other effects start to be significant and require correction coefficients to properly match particle behavior. The Bethe-Block equation states that:

$$\left(\frac{dE}{dx}\right) = \frac{4\pi k_0^2 z^2 e^2 n}{mc^2\beta^2}\left[\ln\frac{2mc^2\beta^2}{I(1-\beta^2)} - \beta^2\right] \qquad \left[\frac{eV}{m}\right] \tag{3.1}$$

with:

k_0 = the Boltzmann constant ($8.99 \times 10^9\,\mathrm{Nm^2C^{-2}}$)

z = atomic number of the heavy particle (1 for Protons),

e = electron charge ($1.602 \times 10^{-19}\,\mathrm{C}$),

n = number of electrons per unit volume in the medium (6.9902×10^{29} in Silicon),

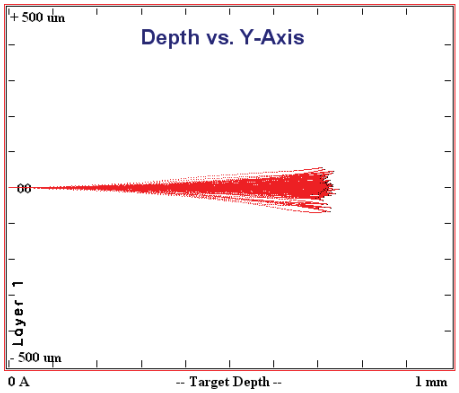
m = electron rest mass ($9.109 \times 10^{-31}\,\mathrm{kg}$),

c = speed of light in vacuum ($2.99 \times 10^8\,\frac{\mathrm{m}}{\mathrm{s}}$),

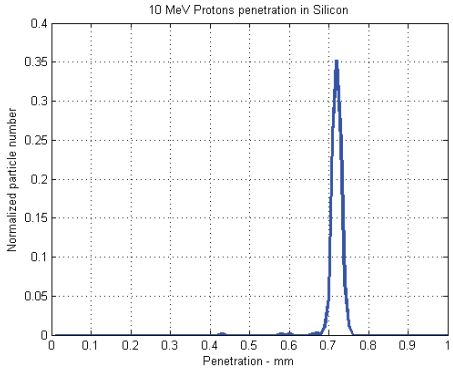
$\beta = \frac{V}{c}$ = speed of the particle relative to c,

I = mean excitation energy of the medium (169 eV for Silicon).

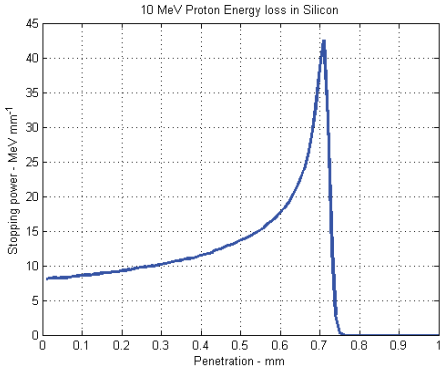
By solving this equation for protons (as incoming particle) and for Silicon (as lattice) the curve depicted in Figure 3.2c can be computed: as it can be clearly seen, the stopping power has a maximum at 0.7 mm and then drops sharply to zero. This can be easily explained by looking at Bethe-Block equation: the slower the particle is and the lower the β will be and so the higher the $\frac{dE}{dx}$ will be. This can be explained in an easier way by thinking that the particle exchanges more energy with the lattice when it is slower (or with lower kinetic energy). The steep decrease in stopping power is corresponding to the depth at which most of the protons stop (see Figure 3.1b). Not all the protons stop exactly at the same depth because, as can be seen from Figure 3.1a, protons trajectory is slightly divergent (the variance in the penetration depth is also called straggling) but almost linear and with small deviations due to the interaction with matter. Particles range can be successfully approximated with a gaussian function, with average value equal to particle range and standard deviation equal to the straggling. By comparing Figure 3.1c



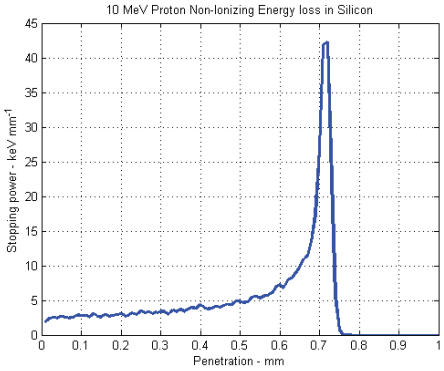
(a) Protons trajectory



(b) Protons penetration



(c) Protons ionizing energy loss



(d) Protons non-ionizing energy loss

Figure 3.1: 10 MeV Protons penetration in Silicon (computed with SRIM [21]).

and Figure 3.1d it can also be seen easily that the energy loss by ionization is many orders of magnitude higher than the non-ionizing component.

Electrons lose energy by collisions while traversing an absorber, just as massive charged particles do. In addition, because of their small mass and depending on their kinetic energy, they will undergo a significant energy-loss by radiative emission, called Bremsstrahlung.

The treatment of the energy loss by collisions for incoming electrons follows the same lines as for massive charged particles, and the equation is quite similar to Bethe-Block one [22]:

$$\left(\frac{dE}{dx}\right) = \frac{1}{(4\pi\epsilon_0)^2}\frac{2\pi e^4 nZ}{mv^2}\left[\ln\frac{mv^2E}{2I^2\,(1\,-\,\beta^2)}\;+\;(\ln 2\;+\;1)\,\beta^2\right] \qquad \left[\frac{eV}{m}\right] \tag{3.2}$$

with:

Z = lattice atomic number (14 for Silicon),

e = electron charge (1.602×10^{-19} C),

m = electron rest mass (9.109×10^{-31} kg),

n = number of electrons per unit volume in the medium (6.9902×10^{29} in Silicon),

v = incoming particle speed,

c = speed of light in vacuum (2.99×10^{8} $\frac{\text{m}}{\text{s}}$),

$\beta = \frac{v}{c}$ = speed of the particle relative to c,

E = incoming particle energy.

Since electrons are hitting lattice electrons and their mass is the same, both particles will experience strong deviations. When a charged particle experience a decelerations (thus changes its kinetic energy) a photon is emitted whose energy is the same as the energy lost by the electron. This mechanism is called synchrotron emission when it happens in vacuum and Bremsstrahlung when it happens in a different medium. Radiative energy loss is expressed by the following equation:

$$\left(\frac{dE}{dx}\right) = \frac{nEZ\,(Z\ +\ 1)\,e^4}{137m^2c^4}\left(4\ln\frac{2E}{mc^2}\ -\ \frac{4}{3}\right) \qquad \left[\frac{eV}{m}\right] \tag{3.3}$$

with:

Z = lattice atomic number (14 for Silicon),

e = electron charge (1.602×10^{-19} C),

m = electron rest mass (9.109×10^{-31} kg),

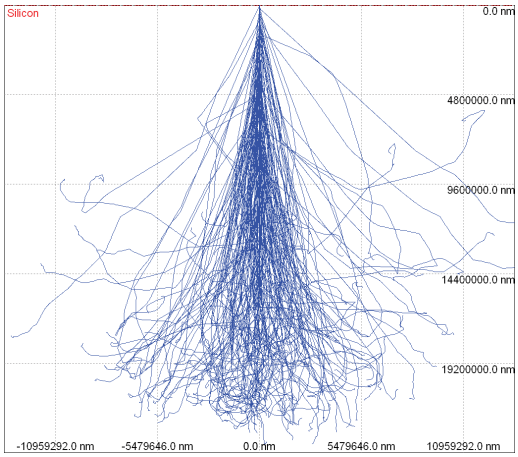
n = number of electrons per unit volume in the medium (6.9902×10^{29} in Silicon),

c = speed of light in vacuum (2.99×10^{8} $\frac{\text{m}}{\text{s}}$),

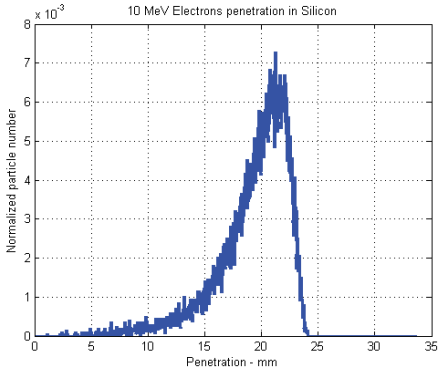
E = incoming particle energy.

The total energy loss experienced by electrons is thus the sum of the previous two components, and the results of the previous two formulas are depicted in Figure 3.2c. As it was discussed before, electron trajectory is not a straight line, but it is influenced by the knocks with other electrons, as can be seen in Figure 3.2a: the same result can be seen by looking at Figure 3.2b where the position where incoming electrons stop after penetration is plotted. Since the path is not linear, a penetration depth is quite difficult to define, as can be seen by comparing Figure 3.2b with Figure 3.1b. Protons in fact stop almost at the same depth (excluding a small variance, the straggling) while this cannot be said for electrons.

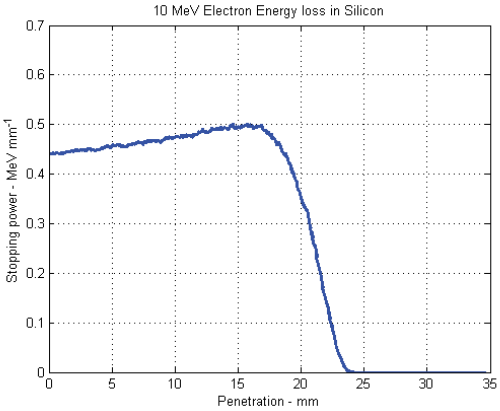
As it was computed for protons, also electrons lose part of their energy by a non-ionizing interactions, even if this mechanism is weaker than the ionizing ones, as an be seen in Figure 3.2d. Furthermore, a comparison between non-ionizing energy loss between electrons and protons can be seen in Figure 3.23.



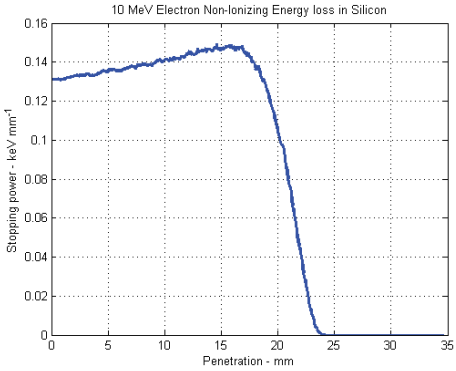
(a) Electrons trajectory



(b) Electrons penetration



(c) Electrons stopping power



(d) Electrons stopping power

Figure 3.2: 10 MeV Electrons penetration in Silicon (computed using Casino [23]).

To properly compute particle transport in matter, the simple equations presented above are not enough to take into account all the effects that may arise, so these computations are in general performed with specialized softwares. Particle transport codes are in general used for high energy physics experiments and most of the software were developed for this reason. GEANT4 [24] is the most used due to its accuracy and speed: it was developed by CERN for particle accelerator research and it is the basic tool used to develop the LHC accelerator [25]. GEANT4 performs 3D Monte Carlo simulations for determining the particle trajectory inside the medium and achieves extremely good results; anyway, to have reliable simulation, an extremely high number of particles is needed, thus requiring a long time. When particles propagation need to be calculated through a planar structure, MULASSIS [26] can also be used: it is a sub-set of the full GEANT4 package limited to planar geometries. Just to make an example, the plot in Figure 3.3 took about 12 hours to be computed with MULASSIS using one billion particles to achieve the result. With a

lower number of particles, due to the probabilistic approach in the Monte Carlo method, the result is influenced by a huge number of spurious peaks that got instead averaged with such a long simulation. This simulation took as input the solar proton fluence in one year in LEO and the result is the proton spectrum inside the shield. This is a quite important value since it allows to properly select electronic components by defining the radiation environment they have to stand. The simulation requires also such a high number of particles because of the wide range of energy and fluence the input spectrum spans.

These tools furthermore allow to compute matter interaction with a wide rage of particles, ranging from protons and electrons up to neutrons, muons, quarks and neutrinos.

Other tools can be used to compute particle propagation through simpler geometries and for single particles, like SRIM [21] or Casino [23]. The former is useful for computing positively charged particles, such as protons and heavy nuclei, interaction with matter; the software is in reality two fold: a Monte Carlo solver is used to compute interactions with matter, but many pre-computed tables are also used to speed-up the computation. These tables allow to compute in few seconds particle range and energy loss for a wide variety or materials and compounds. The latter software, Casino, is used instead to compute electrons interaction with matter by Monte Carlo simulation. Performances are quite similar to SRIM but there are no pre-computed tables to simplify simulations. While protons propagation can be easily modeled using SRIM tables, range for electrons has a too wide variability range, thus making this approach more complex. Based on pre-computed tables, a faster and approximate particle propagation code was developed to evaluate proton fluence behind the AraMiS mechanical structure: this new approach is approximate but allows to compute output spectrum in few seconds and it could be really useful for defining the best shielding geometry. A precise simulation can be then performed

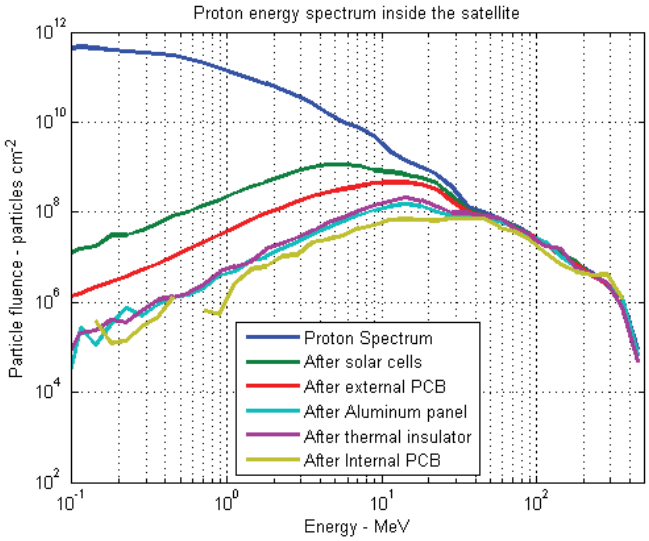


Figure 3.3: Protons through a multi-layer shield (computed using MULASSIS).

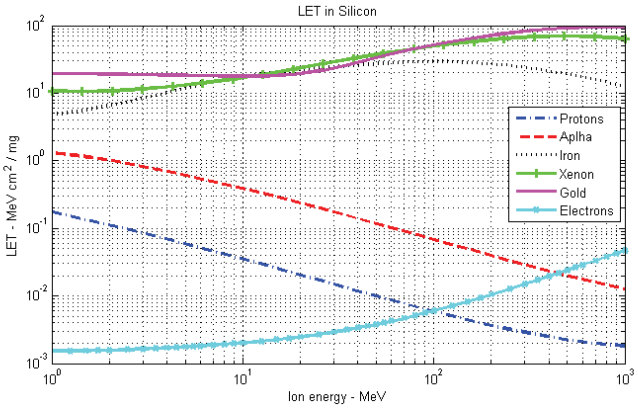


Figure 3.4: LET in Silicon (computed with SRIM [21] for protons and heavy ions, with ESTAR [27] for electrons).

to verify the result, without needing a full (and 12 hours long) simulation at every design step. Further details about it can be found in section 3.2.

The part of incoming particle energy that induces ionization on an electronic device can generate two main effects, depending if charge is injected in an insulating or in a conducting area. If the injected charge is located in an insulating area, it will not be able to recombine in a short time and it will be thus trapped: this will generate problems due to the accumulation of charge, leading to total ionizing dose problems (see section 3.3). If the insulating area is inside a capacitor, charge can be injected there, thus changing the stored value (further details about this effect can be found in section 3.4). If charge is injected in a conducting area the electron-hole couples can experience different effects whether or not an electric field is present. If no electric field is present, recombination will annihilate them in a short time, thus leading to no effects. In case an electric field is instead present (such as for example on a high resistive region or inside a p-n junction) holes and electrons will drift in opposite directions thus giving birth to a current that may generate upsets in the circuit (see section 3.4). These effects linearly depend on incoming particle LET, which expresses the amount of energy lost in ionizing interactions with matter (see Figure 3.4). When instead, the prominent effect generated in the device is the displacement of lattice atoms, electrical and optical characteristics of the device are degraded: further details about it can be found in section 3.5.

3.2 Fast ion transport in matter

Particle transport codes are used to compute the amount of energy lost and the trajectory a charged particle follows during its interaction with matter. Those codes are in general used, for space applications, to compute whether a charged particle is able or not to pass through a thin layer of matter that acts like a shield. This shield can be able to completely stop this particle or just attenuate its energy. Current computer codes for this application

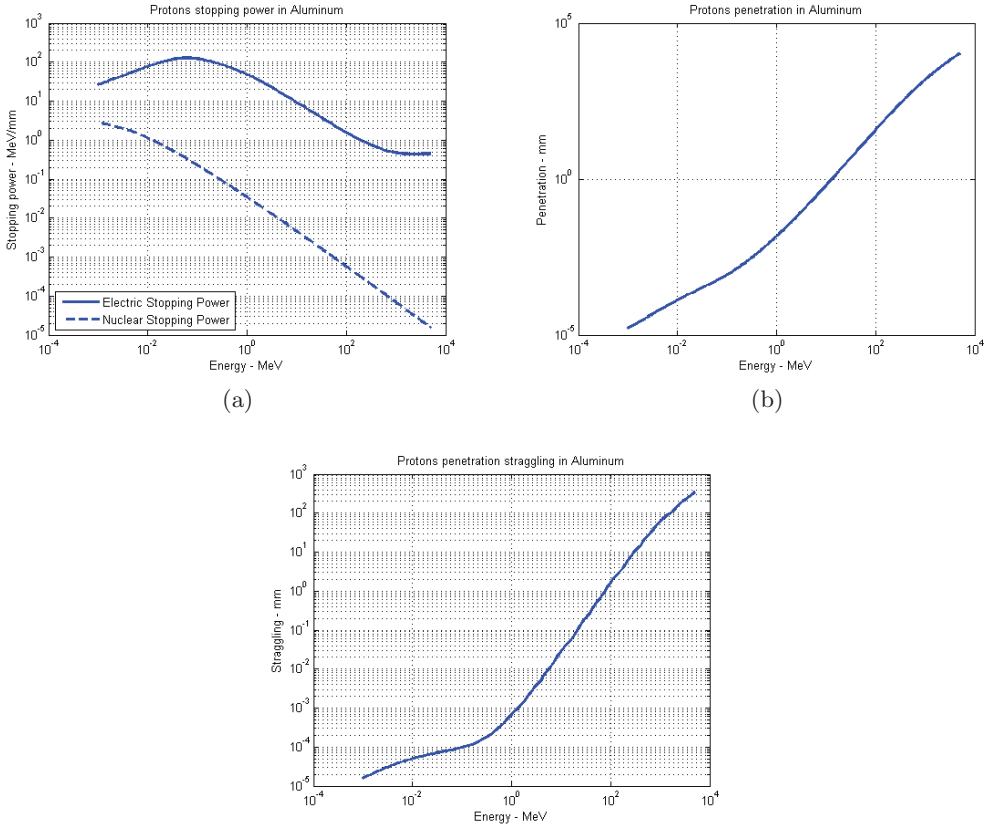


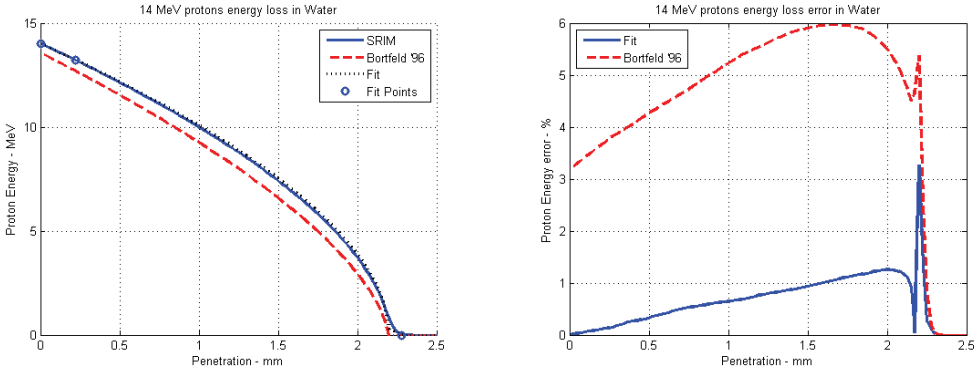
Figure 3.5: Protons penetration in Aluminum (computed using SRIM [21]).

use a statistical approach that requires the computation of all the interactions generated by a huge number of particles while passing through matter. Many equations are available for computing particle transport but they do not take into account all the contributions and thus, can lead to errors but the computation time is extremely lower.

As it was discussed at the end of section 3.1, protons and heavy ions transport in matter can easily be approximated with pre-computed tables as can be seen in Figure 3.5: these tables are computed using SRIM [21] in just few seconds. But data in these tables were computed using Monte Carlo simulations, so they could not be expressed in a closed form equation.

The approach suggested here to speed-up this computation is to take advantage of a closed-form equation for approximating particle transport, but fitting this equation with the results found in SRIM tables such that the total error associated with an analytical solution is minimized.

The equation that needs to be approximated is the Bethe-Block equation, that is expressed as a differential equation: as it can be seen in Figure 3.1c,this is a quite complex



(a) (b)

Figure 3.6: 14 MeV Protons energy loss in water (computed using SRIM [21]).

equation to approximate. But what is actually needed is the integral of this equation, that is represented in Figure 3.6. This second equation expresses the actual amount of energy the particle has during the interaction with a layer of matter. As it is suggested in [28], this function can be approximated with the following equation:

$$E(x) = \left(\frac{R - x}{\alpha}\right)^{\frac{1}{p}} \quad [MeV] \tag{3.4}$$

where R is proton penetration and the two constants α and p equal to $3.38 \times 10^{-5} \frac{\mathrm{m}}{\mathrm{MeV}}$ and 1.6. These the two coefficients were computed from Bethe-Block equation for water and they only apply for that material. But, as the authors say, this does not take into account particle straggling, which increases the error in the region near to the maximum penetration. This error can be easily solved by adding a further parameter:

$$E(x) = \left(\frac{R - x}{\alpha}\right)^{\frac{1}{p}} + k \quad [MeV] \tag{3.5}$$

Using this third parameter, as could be seen in Figure 3.6, the error is greatly reduced. The three constants needed, α, p and k can be computed by fitting the previous equation using SRIM tables for water. For fitting three parameters, at least three points are needed, as can be seen in the previous figure; they have coordinates $(0; E_0)$, $\left(\frac{R}{10}; E_0 - \left(\frac{R}{10}\frac{dE}{dx}\right)\right)$ and $(R_s; 0)$ where E_0 is the incoming particle energy, R is particle Bragg peak position, R_s is the Bragg peak position plus the straggling and $\frac{dE}{dx}$ is incoming particle stopping power. The first and the last point are quite straightforward to understand, since they describe the curve before the first interaction (the first point) and when all the energy has been deposed (the last). The coordinates of the third point are computed by approximating Bragg curve in the first part with a line with tangent equal to the stopping power. This approximation has been verified over a wide range of energies and it showed good consistency. By performing a numerical fit of equation 3.5 with the previous three points,

the curve represented in Figure 3.6 can be computed. As can be seen from Figure 3.7 the error between the approximated curve and the energy loss computed with SRIM is around 3% at 10 MeV, which is below SRIM results error bar (about 8%). The plotted error is the maximum error over the full penetration range, which usually shows a peak near maximum penetration depth (see Figure 3.6), while it is quite lower in the other points.

Given the positive results just pointed out, this idea will be applied to proton transport through a planar shield to compute the particle spectrum behind it. Using equation 3.5, energy loss through the shield can be computed: the same procedure can be repeated to compute energy loss for a continuous particle spectrum as can be seen in Figure 3.8 (see attenuated proton spectrum line). One assumption has been made for computing this result: all the protons pass through the shield, even if this is not true, since, instead, most of the low energy ones will not pass. This further problem can be solved by computing protons range in the shield material and imposing that, if the shield is thicker than the penetration depth, no proton will pass through. But this is still incomplete since straggling has been neglected. Protons penetration depth in the shielding can be approximated by a gaussian curve (see Figure 3.1b) with average equal to particle range and variance equal to particle straggling. In this way, the probability, for a proton with energy E, passing through a shield with thickness equal to x is:

$$p(x) = \frac{1}{2}\left[Erf\left(\frac{R - x}{\sqrt{2}\,S}\right) + 1\right] \tag{3.6}$$

with R equal to particle penetration in the shielding material and S equal to particle straggling. This probability can be computed for all the points in the input spectrum and the result can be simply multiplied by the attenuated energy spectrum, as depicted in Figure 3.1b: the result can be seen in Figure 3.9, where the result is compared to MULASSIS output. This method is not extremely precise (the error is around 10%) but it

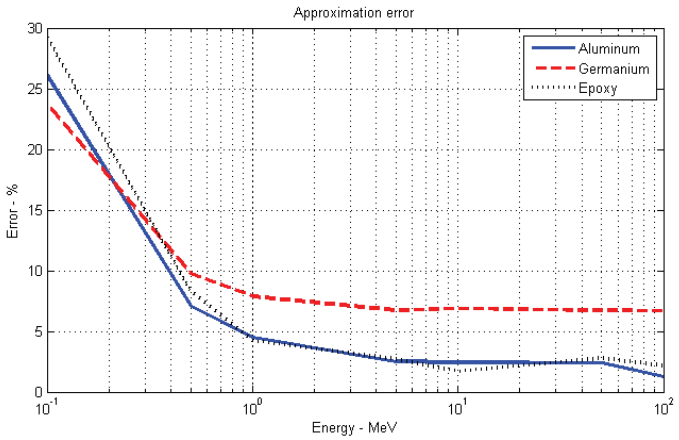


Figure 3.7: Maximum approximation error as a function of energy (computed using SRIM and the proposed algorithm).

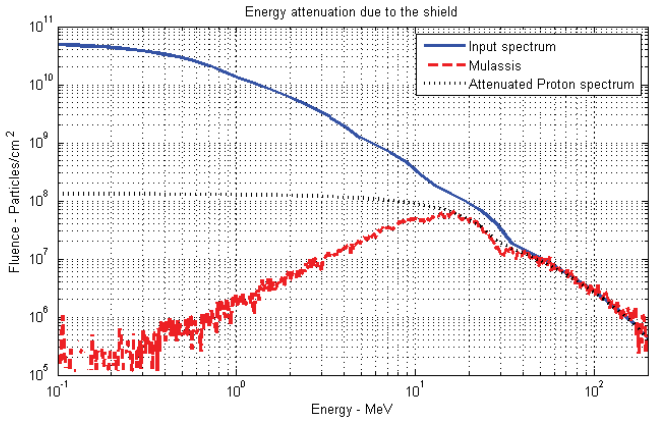
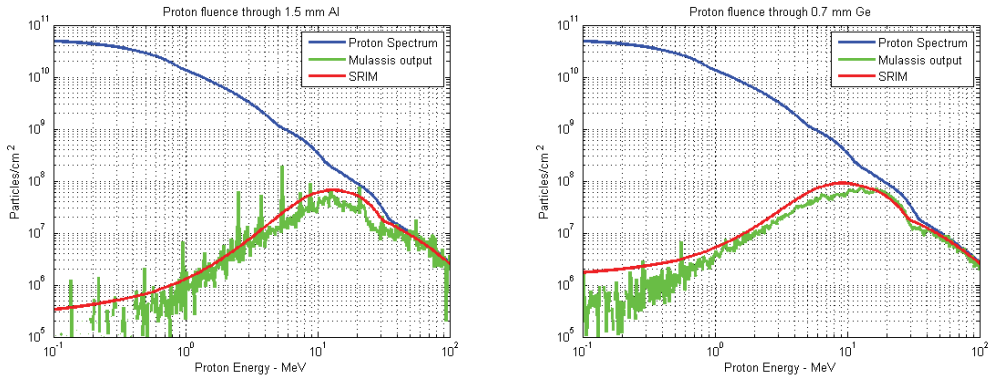


Figure 3.8: Proton energy after passing through a 1.5 mm thick Al shield (computed using MULASSIS and the proposed algorithm).

is really fast: 12 seconds are necessary for computing the result, while using MULASSIS at least 12 hours are needed. The same approach can be used if incoing particles are heavy ions since SRIM is also able to generate those stopping power plots but it cannot be applied to electron transport since, as can be seen from Figure 3.2, the average trajectory is not a straight line, thus penetration is widely changing from particle to particle.

The previous considerations enlight the advantages of using this approximated method: it can be used during the first evaluation of a particle shield efficiency and for selecting proper materials to shield incoming particles. This method allows to evaluate in a short time a great variety of different materials and, when the optimal compromise has been selected, a full Monte Carlo simulation can be performed.



(a) (b)

Figure 3.9: Protons penetration in Aluminum and Germanium (computed using SRIM [21] and the proposed method).

3.3 Total Ionizing Dose

A charged particle penetrating through a semiconductor creates, due to Coulomb interaction, a charged wake of holes-electrons pairs: the number of generated carriers is directly dependent on particle LET and on electron-holes generation energy (in 3.6 eV in Si, 17 eV in SiO_2). When these couples are generated in a bulk semiconductor, recombination acts to remove them without generating other issues. When instead, an electric field is present, these free charges start moving (generating a photo-current) through the material and their effect is quite different if the material is an insulator or a conductor. In a conductor this current will flow to other parts of the circuit modifying circuit behavior (SEE) without permanent effects. The holes and electrons mobility is quite different in dielectric ($20\,\mathrm{cm}^2/\mathrm{Vs}$ for electrons and $10^{-4} \div 10^{-4}\,\mathrm{cm}^2/\mathrm{Vs}$ for holes) so electrons will drift away from dielectrics much faster, while holes will be trapped: this trapped charge quantity in insulating layers can degrade electronic devices and also stop operations.

The unit of measurement of deposed charge is the gray, however, the most used unit is the rad (Radiation Absorbed Dose (rad)) which is equal to 0.01 gray, $10\,\frac{\mathrm{mJ}}{\mathrm{kg}}$ or $100\,\frac{\mathrm{erg}}{\mathrm{g}}$. Absorbed dose can be converted also to electron-holes number by means of the following equation:

$$\frac{e-h\ pairs}{rad\ cm^3} = \frac{Dose \times Density}{Electron-Holes\ creation\ energy} = \\ = \frac{\left[\left(\frac{100\ erg}{g}\right)\left(\frac{10^{-7}\ J}{erg}\right)\left(\frac{eV}{1.6\times10^{-19}\ J}\right)\right]\ \rho}{E_{e-h}}$$

which is equal to $8.1 \times 10^{12}\,\frac{\mathrm{holes}}{\mathrm{rad\ cm}^3}$ in SiO_2. It should be noted that total ionizing dose should always be expressed together with the material in which is was calculated (for example 100 rad(Si)). Total Ionizing Dose (TID) can also be expressed as a function of incoming particle energy, according to the following equation:

$$TID = \int LET(E)\Phi(E)dE \tag{3.7}$$

where E is incoming particle energy, *LET(E)* is incoming particle linear energy transfer as a function of energy and Φ is particle fluence.

Total absorbed dose is a cumulative effect that originated by trapped charges is an insulating layer; after particle penetration several mechanism should be taken into account to understand all the effects (see also Figure 3.10 for further details):

- generation of e-/h pairs (17 ± 1 eV/pair in $Si0_2$);
- prompt recombination of trapped electrons (due to their higher mobility). In $Si0_2$, electrons are much more mobile than the holes and they are swept out of the oxide, typically few picoseconds. However, in this first amount of time, a fraction of the electrons and holes will recombine while holes;

- free carriers trapped in the oxide start to migrate due to the electric field or diffusion. This second process causes the short-term recovery of radiation degradation and it is quite slower to fully complete (about 1 second) but it is very sensitive to the applied electric field, temperature, oxide thickness, and oxide processing history;
- formation of trap states in the dielectric. The effect generated by these trap states can last from several hours even up to years, but anyway they undergo a gradual annealing;
- formation of interface traps by means of reactions with free protons: these hydrogen nuclei are impurities in the Silicon that were trapped during manufacturing processes. These atoms do not interfere with device properties until they get activated (by for example, a trapped hole in the dielectric).

These physical mechanisms influence an electronic device in different ways, depending on its technology: the two most important device technologies are the bipolar and the MOS one.

In MOS transistors a thin oxide layer is used to insulate the gate from the channel and its characteristic are important in determining transistor performances. As it was specified before, defects can accumulate in the dielectric or near the interface with the Silicon, as depicted in Figure 3.11 and they can degrade performances, as described in this equation:

$$V_{th} = V'_{th} - \frac{Q_f}{C_{ox}} - \frac{N_{it}e\,(2\Phi_f)}{C_{ox}} - \frac{N_{ot}e}{C_{ox}} \tag{3.8}$$

where V'_{th} is the transistor threshold voltage, Q_f is the total trapped charge due to impurities, C_{ox} is the gate capacitance, N_{it} it the total number of interface trapped charges,

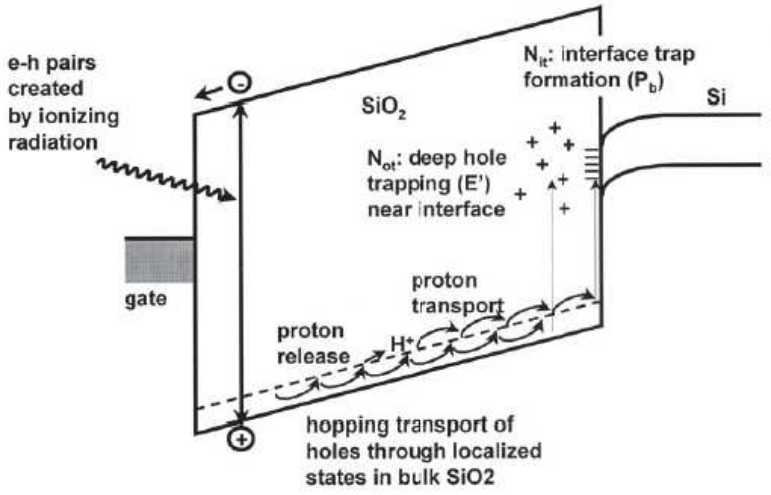


Figure 3.10: Trapped holes in Silicon dioxide (Marc Poizat, ESA).

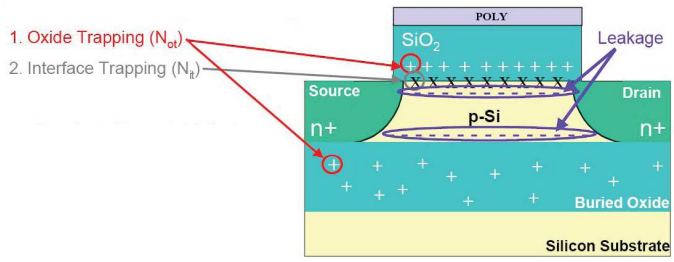


Figure 3.11: TID in MOS transistor (Marc Poizat, ESA).

N_{ot} is the total number of oxide trapped charges and e is the electron charge. From this equation it is easy to see which effects can be generated by trapped charges: the most important effect is the shift in threshold voltage. The last component in equation 3.8 depends on oxide trapped charges and this contribution is always positive in N-MOS and P-MOS because trapped charges are always holes. The second contribution in the same equation is instead dependent on interface trapped charges and oxide-Silicon potential, which can be positive or negative according to transistor type, which makes this component negative for N-MOS and positive in P-MOS. The first parasitic component is instead due to lattice imperfections, an it is not strictly linked to radiation issues. It is easy to understand that, P-MOS transistors will see their threshold voltage increase with radiation (in absolute value) while N-MOS transistors can see the threshold increase or decrease according to

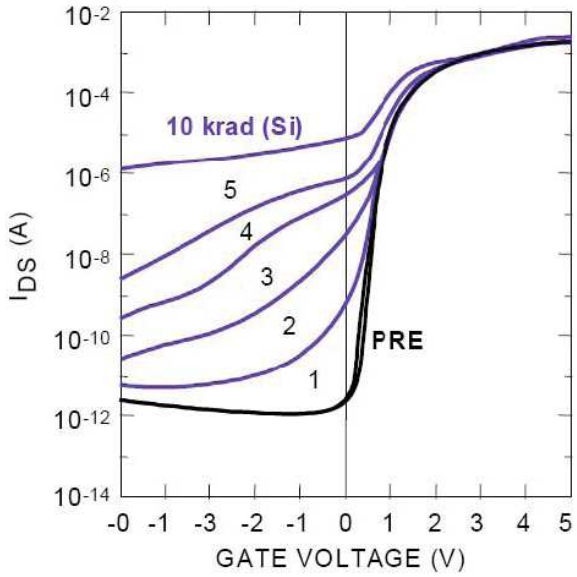


Figure 3.12: Leakage current in irradiated MOS transistor (Marc Poizat, ESA).

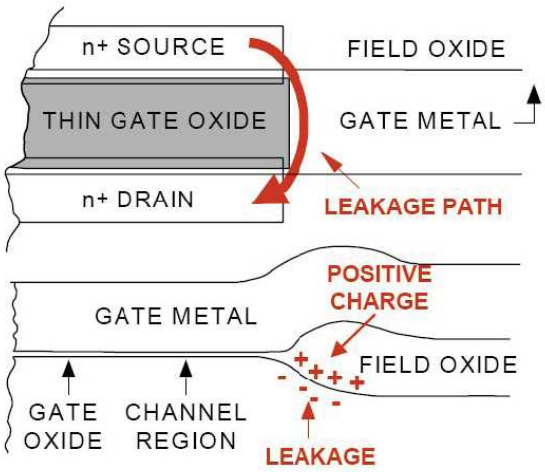


Figure 3.13: Parasitic MOS transistor (Marc Poizat, ESA).

the magnitude of the two total dose shift components.

This problem is anyway going to be less severe with new technology nodes since gate thickness is reducing, thus making the electric field increase in that area: this has a good effect on trapped charges that can then drift out of the dielectric by themselves due to tunnel effect (see section 5.1 for further details).

Together with threshold voltage shift, MOS transistors will also see a big increase in sub-threshold current, as can be seen in Figure 3.12, which is again mainly due to degradation in the gate oxide. This effect can also explain quite well the increased current consumption that can be measured on components after irradiation. Even if this problem is usually not dramatic due to the high quality in gate dielectric, this is not the case for the buried oxide or for the shallow-trench one, which is in general of poor quality. Lower quality oxide presents a higher number of trap levels in the energy gap that can collect holes after radiation. This can cause a threshold shift in all the parasitic transistors, as the one that can be seen in Figure 3.13 which is on the side of the real transistor, below the gate end border, above the channel. Here the trench oxide, which is quite thick, creates a parasitic transistor with a high threshold voltage (higher than maximum supply voltage in the non-irradiated device) which can be turned on if the threshold down shifts due to radiation. This is a common problem in commercial electronic components that can solved by modifying transistor layout (see section 5.1 for further details).

The last component that can degrade transistor performances under radiation is direct gate current: with technology scaling, in fact, gate thickness reduced dramatically down to few nanometers. Radiation can thus enhance, by creating more trap levels in the oxide energy gap, this current, which is mainly due to tunnel effect through the potential barrier.

Circuit bias can also play an important role in radiation tolerance, especially in MOS transistor which show higher tolerance if they are left powered off: this makes them particularly suited for use in cold redundant systems, where the cold spare will degrade

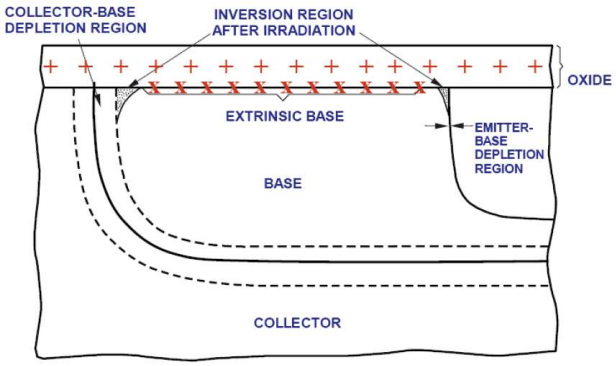


Figure 3.14: Base region depletion in Bipolar transistor (Marc Poizat, ESA).

less due to radiation (for further details, see section 5.4).

Bipolar transistors do not use directly an oxide layer to operate, but dielectrics are anyway present for technological reasons: a passivation layer is used on top of the transistor to protect it, but under radiation this structure can become a parasitic MOS device. On top of the base region, this passivation layer can trap charges (see Figure 3.14) and this can act as a transistor shorting the base region, which reduces base current actually going towards the emitter, thus reducing the gain. A leakage current can also originate from base region to the collector one (see Figure 3.15). These two effects contribute to reduce transistor gain and increase leakage components and their effects is more severe if the transistor is left un-biased because the electric fields can act to increase trapped charge drift and recombination.

Furthermore, bipolar devices show an important dependence on radiation dose rate: the lower the dose rate is and the higher the defect creation and degradation will be. This contribution, called Enhanced Low Dose Rate Sensitivity (ELDRS), can be explained by considering proton migration through the device [30]. Radiation releases H^+ ions in the oxide from defects, which, under positive bias, are driven to the interface, where they

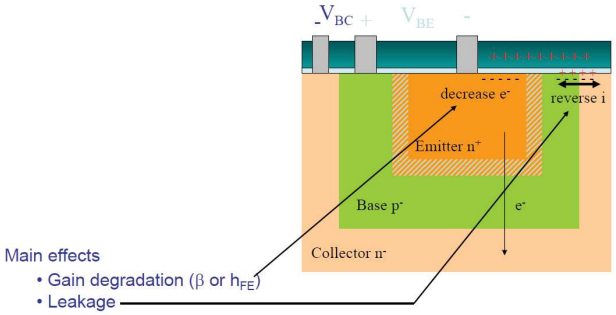


Figure 3.15: Leakage currents in Bipolar transistors (Marc Poizat, ESA).

directly interacts with oxide defects: the difference between low dose rate and high dose rate can be explained by examining the relative roles of holes and H^+, both of which are driven towards the interface. Holes have higher mobility so that they get to the interface region first. Under low dose rate conditions, the positive charge in the oxide does not interfere seriously with H^+ migrating toward the interface because the holes have time to migrate close to the interface and/or get annihilated via electron tunneling. As a result, their ability to inhibit the later motion of H^+ is reduced. Thus, depassivation of dangling bonds occurs relatively efficiently at low dose rates. Under high dose rate conditions the positive charge in the oxide builds rapidly to a high value and acts as an electrostatic fence during irradiation, shielding the interface from H^+ and resulting in a smaller level of defects creation.

When electric fields are present, even at the relatively low field magnitudes typically found in bipolar field oxides, the ELDRS effect is dramatically reduced because the field acts to sweep the holes out of the bulk of the $Si0_2$, reducing interface trap formation.

3.4 Single Event Effects

High energy particles lose their energy mainly through ionization and, if this energy loss took place in a region with an electric field, the injected charge drifts according to field lines and induces a very short duration current pulse in the circuit (usually less than 10 ns). After this initial current peak, the pulse can propagate through the circuit, inducing an anomalous behavior that depends on circuit layout and technology. Injected current

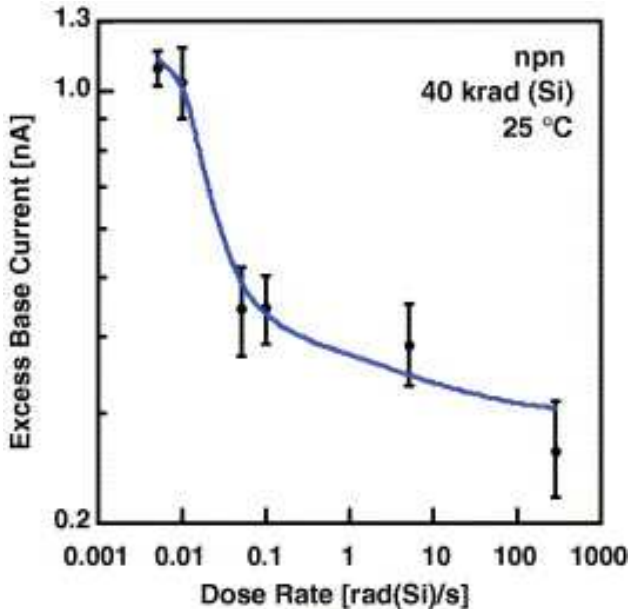


Figure 3.16: ELDRS in Bipolar transistor [29].

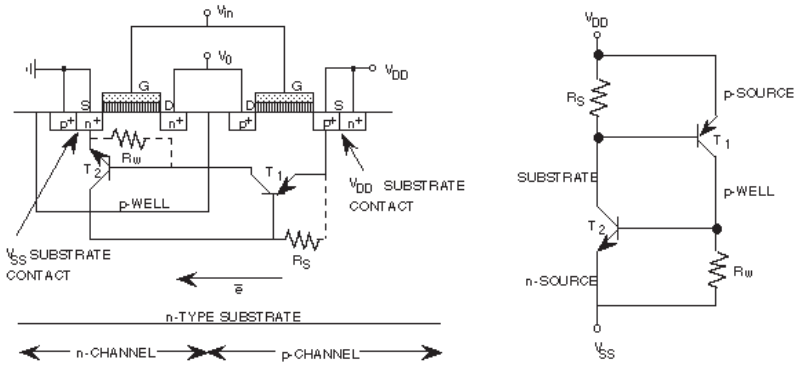


Figure 3.17: Source of latch-up in CMOS [31].

magnitude linearly depends on particle LET, since this parameter describes the amount of energy lost in ionizing interaction with matter (see Figure 3.4).

The effect can manifest in a transient way (Single Event Transient (SET)), or this can become permanent if particular circuit area are hit: if the hit area is a memory element, the stored value can be modified (Single Event Upset (SEU)) or, if a parasitic SCR is triggered, an anomalous current consumption can be generated in the circuit (Single Event Latch-up (SEL)). When in general, the effect of an ionizing particle causes a device malfunctioning, this is called Single Event Functional Interruption (SEFI) but it is not strictly related to a particular effect but to the interruption of correct operations in the system.

Bulk CMOS designs contain two parasitic bipolar transistor structures that form a four-layer structure, similar to a Silicon Controlled Rectifier (SCR), shown in Figure 3.17. This SCR is not involved in normal operations but transient signals at the input or output terminals can inadvertently trigger it on, when it starts drawing very large currents that may cause catastrophic failure, and can only be turned off by temporarily removing

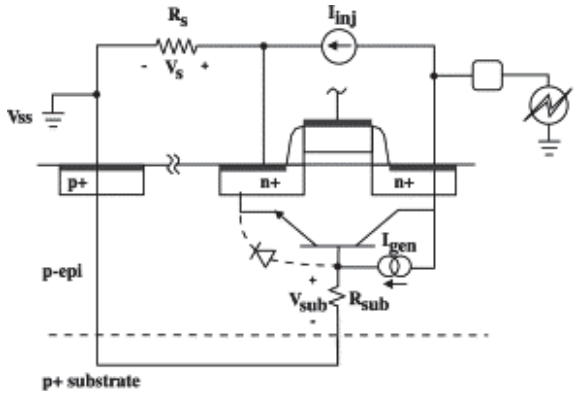


Figure 3.18: Snapback equivalent model in a CMOS circuit.

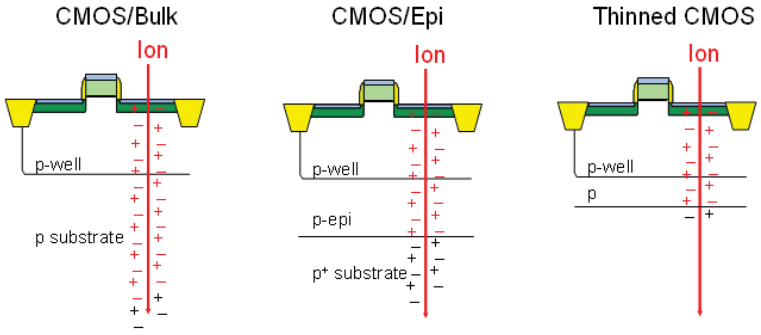


Figure 3.19: Charge injection in different types of CMOS processes.

power. All CMOS designs use special guard bands and clamp circuits at I/O terminals to prevent this from happening in standard circuit applications, however, in a radiation environment transient signals are no longer confined to I/O terminals, and current pulses from heavy ions can trigger a latch-up in internal regions of a CMOS device as well as at I/O circuitry. In most circuits, currents of several hundred milliamperes or more will flow in the localized region where latch-up is triggered, rapidly heating that section to extremely high temperatures and usually generate localized damages to the silicon and metalization, and maybe to other regions.

Furthermore, a micro-latch-up is defined as a latch-up initiated by a single heavy ion in which the current is limited by the device internal circuitry. Since the current is limited, the micro latch-up is not destructive, but the effect on device functionality can be significant (SEFI).

Snapback has many of the characteristics of latch-up, but can take place within a single MOS transistor structure: it can occur thus in Silicon On Sapphire (SOS) and Silicon On Insulator (SOI) technologies that do not contain four-layer parasitic structures or in discrete devices. A single high-energy particle may trigger snapback when the parasitic bipolar transistor that exists between the drain and source of an MOS transistor amplifies avalanche current that results from the heavy ion. This results in a very high current between the drain and source region of the transistor, with subsequent localized heating.

Many variables affect latch-up, including the bias conditions applied during testing. Latch-up tests should be made under conditions of maximum power supply voltage. Another important variable is temperature. At 125° C the threshold LET for latch-up decreases by about a factor of three when compared to its room temperature value. Therefore, latch-up testing must be done at the highest temperature required in the application. A null latch-up result at room temperature cannot provide any direct information about the likelihood of latch-up at higher temperatures.

SET are generated when a cosmic particle strikes a sensitive node within combinational logic or an analog high impedance node. A voltage disturbance is produced at that node which may propagate through the circuit generating short pulses (from few tenths of picoseconds up to few tenth of microsecond) with varying amplitude depending on circuit

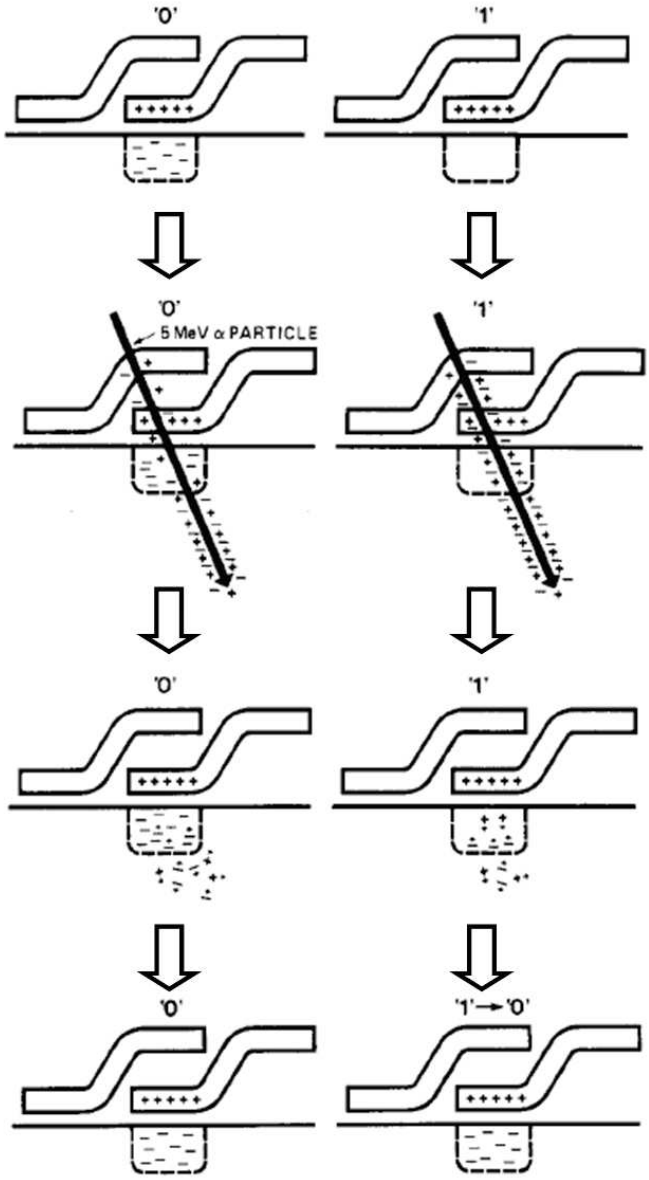


Figure 3.20: Dynamic RAM Single Event Upset (SEU) [32].

topology and injected charge. As can be seen in Figure 3.19, different IC technologies can show different charge collection characteristics: better performances can be seen for thin substrates, that reduce the area where charge collection can take place. For further details on SET propagation in electronic circuits, please see section 7.1. These transient effects can become permanent if they get latched by a memory element: this can happen if the transient is taking place next to a clock transition causing a memory element to store the altered value. Transients are usually quite short in time, but with increasing

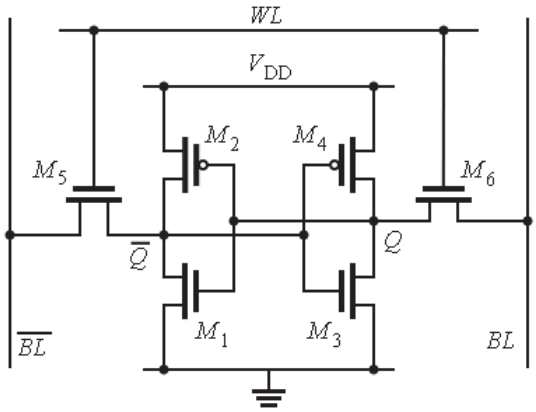


Figure 3.21: Static RAM Single Event Upset (SEU).

clock frequencies the probability is increasing.

But beside combinational logic, also memory elements can suffer SEU when an high energy particle hits directly the storage element. The memory element can be of three main types: a flip-flop, so made with combinational logic, a static Random Access Memory (RAM) cell, a dedicated cell in general made using six transistors, or a dynamic RAM cell, made using one transistor and one capacitor.

The behavior of a dynamic RAM cell is depicted in Figure 3.20: on the left the cell is initialized to logic 0, while on the right it is initialized to logic 1. When the cell is hit by an ionizing particle, the charged wake left can be clearly seen: after charge is injected electrons and holes are swept in opposite directions by the electric field that pushes electrons back into the well and the hole outside of it. In this way, a stored 0 is not modified by the ionizing particle, while a stored 1 is upset because electrons generated by the charged particle are injected into the cell memory, thus showing a stored 0.

In a static RAM instead, the memory element is made by two inverters with a feedback loop (see Figure 3.21): if the ionizing particle is able to inject a current in the Q or $\overline{Q}$ node in the previous figure, this parasitic component gets integrated by MOS input capacitance and, if total injected charge is higher than a threshold value, the upset is triggered.

These upset mechanisms can be prevented with particular layout solutions, but this usually requires an ad-hoc production process which makes these kind of devices expensive. The usual solution employed is to use instead redundancy in memory cell elements and logic circuits and comparing the results to notice eventual upsets. Data stored in memory are also protected with error detection and correction codes.

3.5 Displacement damages

Displacements occur when the primary interaction between an incoming particle and the lattice results in the displacement of one or more atoms from their lattice position. A vacancy is the absence of an atom from its normal lattice position while the atom that

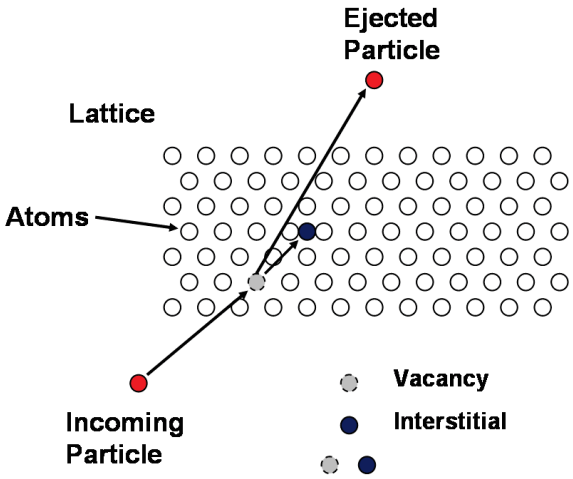


Figure 3.22: Lattice displacement.

moved into a non lattice position is called an interstitial. The combination of a vacancy and an adjacent interstitial is known as a Frenkel pair. This simple configuration is in general the result on a low energy particle hitting the lattice and creating a single Frenkel pair. Most of the times the knocked-out atom still has enough energy to hit and displace a neighbor atom, thus creating a more complex structure called cluster.

Once defects are formed by incident radiation, they will reorder to form more stable configurations: vacancies are in fact unstable and tend to be filled by near atoms, thus creating a sort of migrating vacancy that tends to link to impurities and other interstitial, creating a Frenkel pair. This effect is also called defect reordering and it depends on temperature and excess carrier concentration. Defect reordering is usually called annealing, which typically implies that the amount of damage and its effectiveness are reduces with time: in general, the reordering of defects with time or increased temperature to more stable configurations can also result in more effective defects. This process is often referred to in the literature as reverse annealing in contrast to the more typical process of forward annealing.

In general, any disturbance of lattice periodicity may give rise to energy levels in the band gap and they can have a major impact on the electrical and optical behavior of semiconductor materials and devices. Displacement damages can induce disturbance in carrier generation, recombination, trapping and tunneling. In principle, any combination, or all, of these processes can occur and the role a particular level plays depends on variables such as carrier concentration, temperature, and the device region in which it resides (e.g., in a depletion region) [33].

Given the complexity of analysis of this topic and its dependence on device technological and layout features, a complete modeling of the effects is extremely hard to achieve. For this reason a high level approach is in general adopted: instead of modeling the exact atoms behavior, the amount of energy that actually leads to these effects is computed.

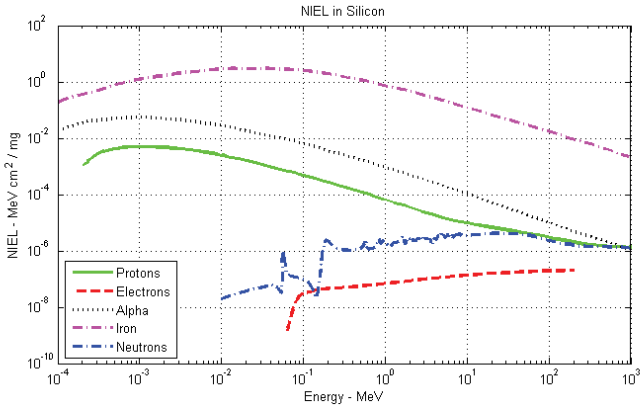


Figure 3.23: NIEL in Silicon [22].

This approach allows also to compute the contributions due to different radiation environments and different particles (see Figure 3.23). The actual amount of energy that an incoming particle loses in lattice interactions and that generates displacement damages is called NIEL and it is in general computed by Monte Carlo simulations considering the kinematic interactions between the incoming particle and the lattice and its measurement unit is $MeV\ cm^2\ /\ mg$, as for the LET that describe energy transfer by ionization and excitation per unit length. The stopping power (see section 3.1) can then be seen as the sum of ionizing and non-ionizing energy loss. Displacement damages in space are generated by higher mass particles (so protons and heavy ions) more than from electrons: protons then are the major component due to their relative abundance with respect to heavy ions (see Figure 2.16 for further details).

NIEL is thus a function of incoming particle, varying for every kind of particle and substrate, that expresses the actual portion of energy lost in the non-ionizing interaction. To compute the total amount of energy deposed by a mono-energetic particle beam, it is only necessary to multiply particle fluence times the NIEL. This is also valid for a continuous energy spectrum and the only difference is that the single product is exchanged with the integral of the NIEL function times the particle energy spectrum as in the following formula:

$$NIEL\ =\ \int S(E)\ NIEL(E)\ dE \tag{3.9}$$

where *S(E)* is the particle fluence as a function of energy. As this equation was used to compute the total non-ionizing deposed energy for an energy spectrum, we can also calculate the equivalent damage produced by a mono-energetic spectrum with this formula:

$$Fluence\ =\ \frac{\int S(E)\ NIEL(E)\ dE}{NIEL(E_{beam})} \tag{3.10}$$

where E_{beam} is the particle energy of the mono-energetic beam. This equivalence holds

for all the incoming particles variety and allows to compute the equivalent particle fluence that generates the same amount of displacement damages as the input spectrum. This is really useful for radiation tolerance testing since it is impossible to generate on ground a particle spectrum equivalent to the one that a space component will interact with in orbit.

The great advantage that this approach offers is also that, as it was demonstrated by many experiments [34][33], device degradation from particle-induced displacement damage shows a linear relationship with NIEL for a variety of electrical parameters, incident particles, and device materials. This simplifies a lot all the computations since the NIEL calculations describes the energy deposited into the formation of Frenkel pairs and do not consider the processes by which stable electrically active defects are formed.

The devices that experience the most displacement damages are the ones that employ a p-n junction, so from bipolar transistors to optical sensors and transducers. In particular, for diodes, an increase in leakage current and voltage drop is in general experienced which translates in an h_{FE} reduction in bipolar transistors. Photo-diodes and photo-transistors experience reduced photo-currents and increased dark currents, while for the latter a reduction of current gain is also present. LEDs and laser diodes will see an output power reduction as a function of the dose and this will also influence opto-couplers. Solar cells usually are the components experiencing the worst degradation because they are exposed on the external side of the satellite, and thus face direct particle flux: they in general experience a reduction in short-circuit current an open-circuit voltage, which then generates a reduction in power output. Since displacement damage tend to modify the material energy gap by introducing further energy levels, they also change optical properties of the materials, in particular the highest degradation can be seen in the light transmission coefficient. For further details about displacement damage, see sections 5.2 and 8.2.1.

Chapter 4

Low-cost vs. traditional space missions

Spacecraft structures have evolved in the past forty years, passing from the first ones (the Russian Sputnik and the American Explorer-1) which were quite simple and low-size missions to present satellites, like the Hubble Space Telescope, the communication or the weather monitoring satellites that are now part of everyday life.

Satellites have always been looked as a highly expensive business, which requires deep knowledge and experience to achieve success: this idea was primarily linked to the high initial costs required for their development and launch. In the first years only governative entities could afford a satellite launch, also because it was not provided as a commercial service but it was a military affair. But with the evolution of satellites and the birth of commercial space companies, this market had a gradual growth up to now, when several private companies are operating, also providing launch services. Many more launch opportunities appeared then, thus lowering also the price for a satellite launch and this allowed an even higher business increase.

This chapter is devoted to the analysis of satellite costs to show that low-cost solutions can be an alternative: a brief analysis of satellite sub-systems is performed, to better show how a satellite is actually composed. Beside the usual approach to space systems development, a low-cost approach will be also presented, highlighting its advantages and disadvantages.

4.1 Traditional satellite architecture and components

An unmanned spacecraft is composed by at least these three components: the payload, the spacecraft bus and the launcher adapter. The payload is that particular equipment that needs to be put in orbit; the spacecraft bus is the rest of the satellite, which holds the payload an provides communication an power supply to the payload, while the launcher adapter is the shaft used to attach the satellite to the rocket and release it during orbit injection.

The payload is the sub-system generally dictating mission constraints: bus size, available power and communication data rate are heavily dependent on the particular payload carried and this influences also orbit selection. Bus is also designed considering the selected orbit since it dictates available power, thermal conditions, link budget and attitude control requirements. The bus is usually divided into many sub-systems, that could be present or not according to specific needs.

The power sub-system is used to supply power to the payload and to the rest of the satellite: it should have a power source (usually solar cells, but in some cases nuclear power generators are also used); a power storage unit, used to supply energy in case of eclipse (when solar panels are used) or to supply power consumption peaks; a power conversion and distribution system that is needed to route power to all satellite sub-systems. Power sub-system is highly dependent on the specific mission the satellite should perform, so its usually an ad-hoc system: communication satellites usually have low eclipse time but have a constant high power consumption due to their radio equipments, while radar satellites usually have a low average power consumption but with extremely high and short peaks. Power storage or generation modules experience a performance degradation in orbit due to components aging thus requiring a safety margin to allow to meet mission requirements even at the end of operating life.

The communication sub-system links the spacecraft to the ground operators which control the on-board payload. Information is usually flowing bi-directionally: from ground commands and eventually software updates are sent to the satellite (this is called uplink) while data about satellite status of health (housekeeping data) and payload data are transmitted to ground (this is called instead downlink). The two links (up and down) have different requirements and most of the times are also operated in two different frequency bands. Uplink usually does not have many requirements on channel data rate since the amount of data is limited; for security reasons usually this link is also encrypted to avoid that external operators could take control of the satellite. The downlink channel is used to transfer both housekeeping and payload data, which may have completely different requirements: usually the housekeeping channel does not require a high data rate, while payload data are usually more demanding on this side. It is not uncommon then to employ two separate channels, thus allowing also contemporary use. Housekeeping channel usually does not depend on the particular mission and can be seen as a "standard" component, while the other one is usually an ad-hoc system.

The command and data handling sub-system is responsible of managing the satellite: it is responsible of distributing commands to the other sub-systems and to coordinate their operations. Commands from ground are usually directed to this system (sometimes also called On-Board Computer (OBC)) that elaborates them before delivering them to the rest of the satellite. The OBC is usually a high performances computer with a data storage system that is used to temporarily store data from the payload or telemetry to relay to ground when connection is established. Usually other tasks are also running on the OBC, like the attitude control code or a scheduler used to perform automatic operations.

The attitude determination and control (sometimes even orbit control can be used) system is responsible of monitoring satellite position and orientation with respect to several reference points and control satellite actuators. The simplest spacecrafts rely on magnetic

field measurement and Sun sensors to estimate satellite orientation, while in bigger and more complex systems star sensors or GPS receivers are employed. Attitude control is performed by means of actuators which can be magnetic torquers and reaction wheels for simple attitude control up to thrusters for orbit control. Liquid or gas thrusters are generally used due to their well known technology: propulsion can be achieved with a chemical propulsion system (generally employing hydrazine as fuel), a cold gas one (using pressurized gas as fuel) or a ionic thruster, using a combination of pressurized gas and electrostatic / magnetic acceleration on gas ions.

The mechanical sub-system is tightly linked to the payload, since it should provide mechanical support and alignment between the payload and the rest of the satellite. This system is usually ad-hoc developed to better adapt to the specific mission. This sub-system is also connected to the launcher adapter to fix the satellite during launch.

4.2 Satellite cost analysis

Satellite cost analysis is important to estimate economical needs at the beginning of the project: it is also important to roughly estimate satellite size according to the available budget. Two main approaches are employed for satellite development: the traditional one relies on space-born components since they ensure the highest success probability to the mission, without budget considerations, while the second approach relies more on the available budget and requires the selection of the best compromises between cost and performances. The main difference between the two approaches can be identified in the most important mission constraint, which is in the first case a performances constraint, while in the latter it is an economical constraint.

4.2.1 Satellite development and production

The first approach is usually pursued by big governative entities, like national space agencies (ESA, NASA, JAXA, ...) where projects are usually extremely big and require a long development time (the cost is usually measured in thousands millions Euro and ten year for development): with this in mind it is clear that every solution that can increase success probability by few percent is highly preferred, even if total cost can increase by millions. Budget in these projects is usually not rigidly limited so a limited incease can be tolerated. In commercial applications, instead, the most important parameter is the income, so every solution that can reduce the income by few percent is generally avoided. Performances requirements in this second case are not the driving constraint, but are not anyway discarded: cost effectiveness is the most important parameter to take into account.

The low cost approach started to be considered in the latest years, when commercial satellite operators started to enter the market. Probably the first company that tried to pursue this latter approach is SSTL [39], a spin-off from Surrey University (UK): their approach was mainly to look for a cost effective high performance satellite for commercial applications, like Earth monitoring.

Furthermore, in the last few years a new trend in space development appeared: the CubeSat concept [40]. The main idea that lies behind the CubeSat is the development of an

extremely small satellite (only 1 kg) with an extremely small form factor ($10 \times 10 \times 10$ cm) which can then be developed and built in a short time and with a low cost. These three different approaches lead to three different cost models that can be used to roughly estimate mission cost, as can be seen in Figure 4.1.

NASA Advanced Missions Cost Model (AMCM) [35] has been used to model satellite cost as a function of satellite weight: this tool has been developed by NASA Johnson Space Center taking into account 122 NASA missions (from the first missions up to the year 2002) and modeling their cost taking into account different parameters. The results plotted in Figure 4.1 were computed for the Earth observation satellites class: two different lines were actually plotted: the first one represents satellite cost for developing one satellite from the beginning, by designing everything for the first time: this was the usual approach for past space missions where everything was ad-hoc produced, thus giving a high mission cost. The second line represents the cost for developing a satellite, taking a previous mission as a reference and upgrading the design. In particular, the third revision has been considered, assuming that afterwards the satellite can be considered as a "mass production" item. As can be clearly seen, the cost dropped by factor of about 1.6 for the whole weight range: this is not a high relative change, but can give a good saving by considering that total mission cost could approach several hundred million Euro.

A low cost approach in developing satellites could prove quite effective in reducing mission cost: as can be seen again from the previous figure a factor of about 8 in cost can be saved. Data plotted were taken from SSTL product overview [36] to highlight this saving margin: the saving is appreciable with low satellite weight, while for bigger sizes the cost approaches NASA cost model, partially confirming those data. The CubeSat approach lead to a further cost saving but this is limited to really small size missions: CubeSats became in fact a de-facto standard among small satellites. Many sub-systems were developed for this class of satellites such that they became standard commercial components. These components can be considered as off-the-shelf, so produced in a bigger scale than regular satellite components, allowing a further cost saving. The CubeSat standard is anyway limited to small sizes (6 kg is considered the bigger satellite size). Costs plotted in figure

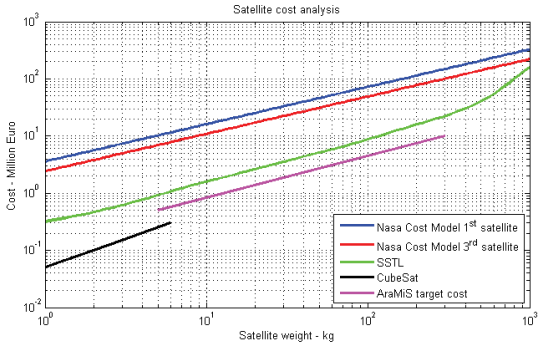


Figure 4.1: Satellite development and production cost as a function of weight (NASA model from [35], SSTL data coming from [36], CubeSat data were derived from [37][38]).

were derived from [37][38] which are essentially on-line shops where CubeSat components can be ordered, thus showing that space components can be like "consumer" products.

To bridge the gap between extremely low cost missions (Cubesats) and bigger size missions, an innovative architecture is proposed in this work and it will be better described in the next chapters: this architecture target cost can be seen in the previous figure last line. Further details about it can be found in section 4.3 and 6.

4.2.2 Launch cost

Beside mission development cost, satellite launch is also an important part which could account for as much as 50% of total mission cost: Table 4.1 sumerizes general cost trends for small to big size satellites for different launch vehicles, making a distinction between European and American rockets and the others.

Satellite Class	**Launch cost per kilo (Euro)**	
	EU or USA	**Non EU or USA**
1 kg - 10 kg	60000	40000
10 kg - 100 kg	30000	20000
$\geq$ 100 kg	25000	15000

Table 4.1: Approximate Satellite launch cost per kilo with different launch vectors [41] (this chart does not consider interplanetary missions).

It should also be noted that the previous mentioned costs can be greatly reduced if several satellites share the same launch cost: most of the times, in fact, launch cost is mostly paid by the primary satellite, which is the one which actually dictates the injection orbit, but other secondary satellites can be also launched, if their orbit constraints match with the primary one. In this case, secondary satellites will pay a lower cost and this launch strategy is usually the most employed for small satellites, which could not afford a dedicated launch.

4.2.3 Recurrent and Non-Recurrent costs

The total cost of a project can be divided into two main components: recurrent and non-recurrent costs. Recurrent costs are related to system production and represent the amount of money spent for producing every single item and verify its functionalities. Non-recurrent costs instead are related to researching, developing, designing, and testing a new product and these costs are paid only once. When the device enters production, its cost will be computed by dividing the non-recurrent cost by the number of produced devices and adding to that the production cost. From the previous simple formula, it can be clearly understood that final device cost heavily depends on the number of produced items: by sharing development cost over three items (as most of the times happens in the space market, were each system is developed ad-hoc) or over three hundred pieces makes a great difference. Furthermore, space systems development requires a huge effort due to the performances required: no fault should stop system operations and the space environment

is extremely harsh. Development and qualification cost is then the major component in a space project: beside reducing production cost, then, it is also extremely important to reduce non-recurrent costs.

This can be achieved by sharing the same costs over a higher number of missions: as can be seen from Figure 4.2 the cost of a satellite greatly reduces if it is not developed from the beginning, but its design is based on previous revisions: data represented in figure were computed using NASA AMCM model [35] to show how the cost of a 5 kg mission can reduce if it is based on previous versions. Total mission cost can be reduced by more that a factor of 2 because the actual non-recurrent costs are now shared over a large number of revisions, instead over just one.

4.3 The low-cost approach

Different strategies can be exploited to reduce space missions cost: as it was identified in the previous section, two main costs can be defined, recurrent and non-recurrent ones. The proposed solutions to solve these two problems are presented in the next sections: non-recurrent costs can be lowered by sharing the cost of system development over multiple satellites, thus achieving an effective re-use, while recurring costs can only be reduced by reducing the cost of the components employed in the design: space qualified components are extremely expensive and complex to procure, so it is proposed to adopt commercial components (COTS) that can show enough tolerance to the harsh space environment.

4.3.1 Modularity

A system is defined modular if it is divided in different sub-systems that can be considered as independent between each other: with a proper interface between every module it is possible to enclose every system function in a black box. This concept allows to create a wide variety of systems by simply exchanging these modules: when a certain function is needed in the system, the corresponding module can be used. Function separation in different blocks allows also to easily upgrade every module without changing the existing

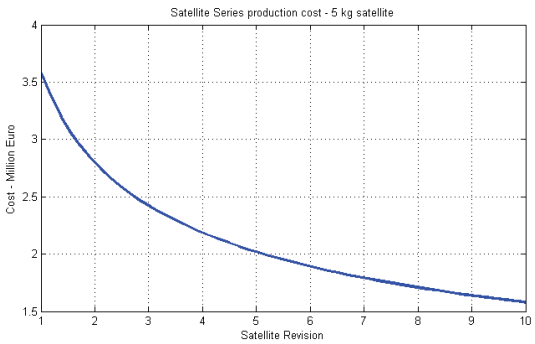


Figure 4.2: Cost per satellite revision (data from NASA AMCM model [35]).

system. In this way all the modules can be used in several different missions without a complete re-design. This strategy allows to greatly reduce non-recurrent costs in several ways: design costs are shared among several missions, qualification tests can be already performed at module-level, thus requiring smaller and cheaper test equipments, system upgrade requires limited changes to functionalities that require upgrades, and not to the whole system. This approach will be better described in section 6.

4.3.2 Commercial Component

Commercial components can be used instead for reducing satellite procurement and production costs: since they are readily available from the next-door component supplier, procurement is quite simple and their cost is much lower than the military equivalent: a Xilinx commercial top grade FPGA can cost about a hundred Euro per device (buying a small quantity), while the military equivalent can cost about one thousand times more. Apart from the pure component cost, device technology is usually different: military components vendors carefully control and qualify their technology to meet high levels of radiation tolerance while most commercial parts suppliers do not identify or control technology parameters that affect radiation hardness. Consequently, radiation hardness of COTS microelectronics is often low and highly variable and, more importantly, the manufacturer will not guarantee radiation hardness levels. For example, state-of-the-art CMOS microprocessors, like for example the Intel Pentium II chip, will fail at total-dose levels as low as a few kilorads and suffer latch-up if exposed to high-energy particles in space. However, COTS parts may exhibit higher levels of radiation hardness that result from unintentional variations in processing. In such cases, qualification by the customer is required to identify those fabrication lots or wafers that can meet system requirements. Qualification can be expensive and time consuming, and there is still the concern that process variations across a fabrication lot (or even a wafer) create uncertainty in radiation hardness. For user qualified product, the higher the radiation requirements of the system, the greater is the testing expense to identify satisfactory standard commercial electronics.

In the past, most of the components were able to stand the space radiation environment, while present technology devices sometimes cannot do it anymore due to technology advances, which usually badly tolerates particle interaction. But the fast expansion that the space market saw in the past years keeps on requiring performances increase and cost reductions that cannot be achieved with military components that show a slower development trend due to their limited market. While defense electronics was once 60 to 70% of the electronics industry in the early 1960s, it is now about 0.5% of a 150 billion dollars market.

Although COTS parts can be qualified for using in radiation environments, safety margins will be lower than with military component and system hardening techniques will be required to enhance these design margins. In addition, qualification to meet a complete set of radiation requirements is unlikely. The use of COTS presents many challenges and concerns for the system engineer. These concerns include:

- possible increased sensitivity to radiation as technology advances,

- variability in radiation hardness from lot to lot,
- changes in processes and designs without notification leading to degraded radiation hardness,
- absence of traceability to production lot,
- rapid obsolescence of components,
- hardness assurance becoming the costumer's responsibility,
- radiation data unavailable or limited.

In designing a cost-effective system, COTS should be used whenever possible. However, where a design is considered mission critical, an assessment for the use of radiation-hardened parts that provide sufficient margin in radiation environments must be made. For example, if a processor were required to control attitude, vectoring, location, or event sequencing for an entire system, it may be necessary to use a rad-hard processor. In addition, in most military systems some type of RH memory is needed to protect operating states or modes of the system, possibly positional and directional data, or other mission critical data.

Shielding provides significant protection for energetic electrons in the Earth's radiation belts, however, it will not provide complete protection to the energetic galactic cosmic particles or protons. If COTS parts sensitive to single-event upset are selected (such as a dynamic ans static RAM, or microprocessor), an Error Detection And Correction (EDAC) mechanism must be implemented. For a microprocessor, a watchdog timer that monitors the processor's health at specific time intervals is needed. If the microprocessor were to upset and not respond properly, the watchdog timer may initiate power cycling, switch the system to a redundant unit or refresh the memory registers. Voting logic or triple redundant logic designs are also methods of mitigating soft error upsets, such as in the case of FPGAs.

Another key issue associated with life-cycle costs is part obsolescence. Since COTS suppliers constantly introduce new products into the marketplace, devices chosen for a system become obsolete in a very short time. This may force certain decision about the up-front procurement of parts. Gate arrays for example, from Xilinx and Actel are introduced for about two years before another product family is introduced. Since military systems often take longer to develop, they generally are required to last between 15 to 30 years.

Chapter 5

Low-cost design techniques

A large part of satellite cost is due to mainly electronic components, since they have to stand the harsh space environment. These components can cost as much as 10 to 1000 times their commercial equivalent and their cost can be as much as 20% – 30% of the full mission cost.

Before going on, it is useful to define single devices performances under radiation in a LEO orbit: this allows to divide components in three main categories considering their radiation hardness [42]:

1. Commercial:
   - Process and Design limit the radiation hardness
   - No radiation tolerance control or lot characterization
   - Hardness levels (not specified by the producer, but evaluated by the costumer):
     Total Dose: 2 to 10 krad (typical)
     SEU Threshold LET: 5 MeV cm^2 / mg
     SEU Error Rate: 10^{-5} errors/bit-day (typical)
     Latch-up Threshold LET: 20 MeV cm^2 / mg
   - Customer assumes all risks
2. Rad Tolerant:
   - Design assures rad hardness up to a certain level
   - No radiation tolerance control and limited lot characterization
   - Hardness levels:
     Total Dose: 20 to 50 krad (typical)
     SEU Threshold LET: 20 MeV cm^2 / mg
     SEU Error Rate: 10^{-7} – 10^{-8} errors/bit-day (typical)
     Latch-up Threshold LET: 50 MeV cm^2 / mg
   - Usually tested for functional fail only

3. Rad Hard:

   - Designed and processed for particular hardness level
   - Wafer lot radiation tested
   - Hardness levels:

     Total Dose: $>$ 200 krad to $>$ 1 Mrad
     SEU Threshold LET: $80 \div 150$ MeV cm^2 / mg
     SEU Error Rate: 10^{-10} – 10^{-12} errors/bit-day
     Latch-up Threshold LET: $>$ 100 MeV cm^2 / mg

Commercial components procurement is usually quite simple since they can be purchased from almost any electronic components supplier while radiation-hardened ones were most of the times developed for military purposes, thus requiring particular authorizations just for requesting informations about them. Many of the space-rated components developed in US, for example, are protected by International Traffic in Arms Regulation (ITAR) and so they can only be purchased by authorized costumers.

From the previous points, it is clear that in order to reduce system cost it is important to start at component level and then continue throughout the design. The main goal of this section is to show that most of the times commercial components can be hard enough for space missions, if employed with proper solutions. Total ionizing dose, single event effects and displacement damages will be analyzed, in particular focusing on which kind of solutions can be adopted to guarantee components survivability in space.

5.1 Total Ionizing Dose effects evaluation

As described in section 3.3, TID effects are proportional to the integral of absorbed particle energy, making devices experience problems after a certain threshold TID has been absorbed. The main effects that can be seen in active devices is the gain reduction and threshold voltage drift in transistors, which can then influence many other parameters, like open-loop gain or transition time. These problems could then impact system behavior if proper TID mitigation techniques are not adopted.

TID hardness at device level can be achieved using proper technological solutions, like the use of circular gate structures in MOS transistor [43]: off-state leakage in N-type Metal Oxide Semiconductor (N-MOS) is commonly attributed to radiation-induced oxide trapped charge in the Shallow Trench Isolation (STI) oxide, which can invert the p-substrate in N-MOSs at the STI edge, thereby creating a shunt leakage path between source and drain, increasing off-state leakage: since oxide thickness is higher in the STI, the electric field strength is not enough to deplete the trapped charge region (see section 3.3 for further details). Anular MOS greatly reduce this effect by eliminating the area with an higher oxide thickness under the gate between source and drain (see Figure 5.1).

Sub-micron MOS technologies exhibit also a high tolerance to total dose [44] thanks to the extremely thin gate dielectric: thanks to the high electric field present there, trapped holes can be removed in just few hours [45]. This relies on the fact that as the channel

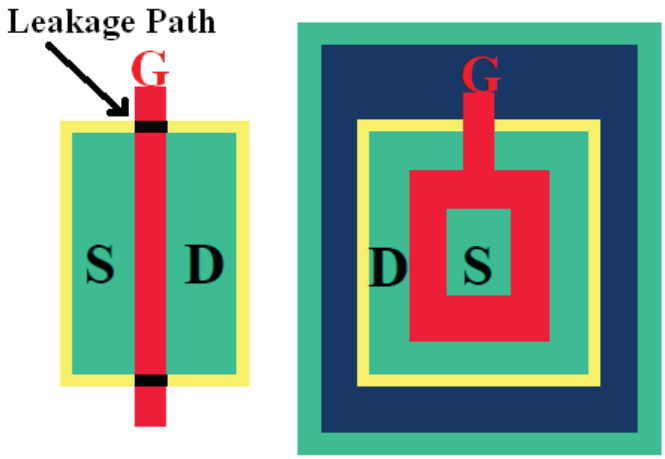


Figure 5.1: Anular MOS layout (From F. Faccio, CERN).

length of MOS transistors reduces, so does the thickness of the gate dielectric and the amount of holes trapped. This follows from the expression for the flat band voltage (the voltage at which a constant Fermi level is achieved across the device):

$$\Delta V_{FB} = \frac{q}{\epsilon_{ox}\epsilon_0} b(t_{ox} - 2h_1)\frac{t_{ox}}{2} \qquad [V] \tag{5.1}$$

q is the electron charge, ϵ_{ox} and ϵ_0 the dielectric constants of the gate oxide and vacuum and b the fraction of holes which are trapped. The parameter h_1 is the distance over which trapped holes can recombine with electrons tunneling from the substrate or from the gate. It depends on the time between irradiation and measurement but for typical times this is $\sim 3\,\text{nm}$. For thin and ultra thin oxides (less than 6nm) essentially no net hole trapping will occur [46] (see also Figure 5.2).

In bipolar transistors particular process and geometrical choices may also help to improve the device radiation tolerance. For example, the thick oxide layer which is overlaying the base-emitter junction (see Figure 5.3) is one of the main sources of degradation: the build-up of positive trapped charge in the oxide will deplete the lowly doped p-type base region causing an increase in the recombination current and thus reducing transistor gain. Reducing the thickness of this oxide layer can contribute to increase transistor radiation tolerance [48].

Furthermore, the passivation layer used on top of the IC die can contribute to device degradation: experiments have shown that the type of passivation layer is important in determining whether ELDRS does or does not occur. This was demonstrated by observing that ELDRS was eliminated by removing the passivation layer (made in silicon nitride) from various integrated circuits [49][50].

The previous solutions are not employed in consumer ICs since they limit transistor integration density and thus make device cost increase. Radiation tolerant devices have

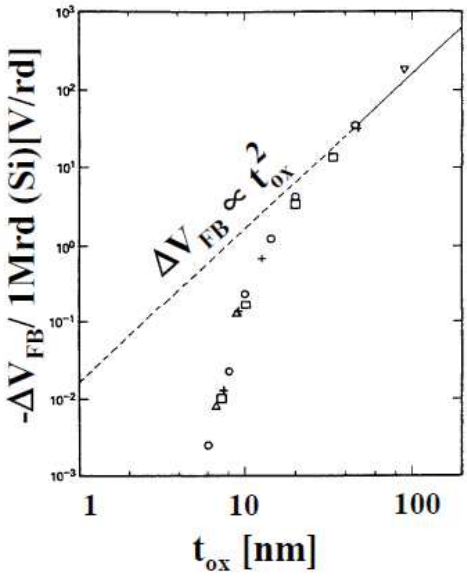


Figure 5.2: Gate oxide thickness variation as a function of threshold voltage shift in submicron MOS transistors[47].

an extremely small market: the number of produced devices is small so development costs cannot be divided over million pieces (as happens instead for consumer products). IC development was thus rejected as a technique to achieve devices tolerant to total ionizing dose. Anyway, there are some commercial technologies that show a high total dose tolerance and thus should be preferably used in space, even if they are not rad-hard, in the sense that they were not developed for that purpose: for further details about it, see section 5.5.

Derating component operative values can also help to improve reliability and radiation tolerance: in case a component is going to suffer high performance degradation due to radiation, it is useful to derate its operating parameters (like for example output current for power devices) as much as required to allow the device to still meet operational requirement

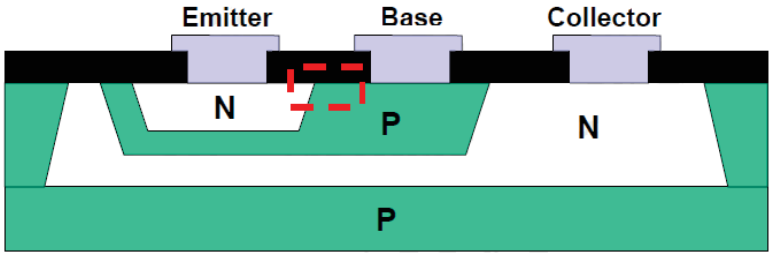


Figure 5.3: NPN transistor layout (from A. Paccagnella, Università di Padova).

after radiation degradation. General derating guidelines, specified in ESA ECSS-Q-ST-30-11C [51] (see Table 5.1) help also improving components reliability due to natural aging. Derating device parameters (like current or power dissipation) can be performed also by balancing the stress over multiple devices: this approach can also be called hot redundancy (see section 5.4) and allows to further increase the degradation limit that the system is able to tolerate. It is quite used for power devices, where dividing the drive current, for example, by a factor of 2 reduces also the power consumption by a factor of 4. It is important to point out that in case of radiation the device should only degrade in performances and not experience a functional failure: in this last case derating is not effective and only redundancy can give some help. Radiation derating margin should be taken into account after testing device performances after irradiation to avoid over or under estimating this parameter (for further details on radiation testing see section 8.2).

Typical requirements for semiconductors	
Rated voltage	$< 80\%$
Rated current	$< 75\%$
Rated power	$< 60\%$
Rated junction temperature	$< 70\%$

Table 5.1: Derating guidelines according to ESA ECSS-Q-ST-30-11C [51].

When no technology can be considered hard enough for the mission the only approach left is shielding. Although this can be expensive since it is increasing the total weight, sometimes it is the only solution. Effective materials for reducing TID can be Aluminum (usually employed for mechanical structures), epoxy resins and laminates (mainly due to their lower weight with respect to Aluminum), Tantalum (used mainly for shielding nuclear weapons radioactivity [14]) and in general molecules with an high ratio of Hydrogen (more

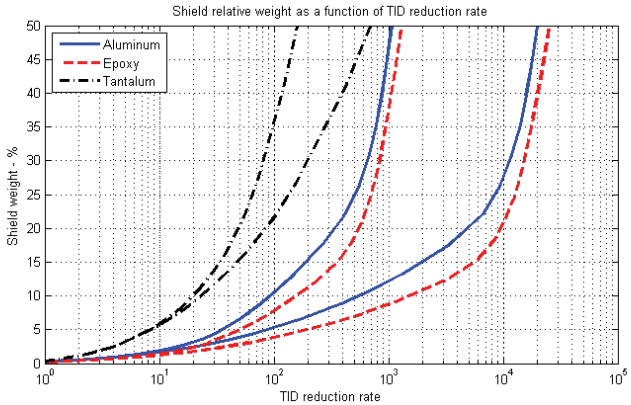


Figure 5.4: Satellite particle shield weight (maximum and minimum, considering orbits ranging from 400 to 36000 km) as a function of TID reduction rate (different materials considering a satellite $16.5 \times 16.5 \times 16.5\,\mathrm{cm}^3$ big weighting 10 kg).

protons mean a stronger Coulomb interaction and thus an higher stopping power). An example of different TID attenuation factor is depicted in Figure 5.4.

5.2 Displacement damage effects

Displacement damages generate lattice imperfections that could impact device performances: the problem is more severe in large area devices, such as solar cells or optical transducers, where lattice imperfections can limit an efficient electrons-photons coupling. Solar cells are mounted on the external part of the satellite, thus making them experience an extremely high particle flux, even in LEO. Cell shielding is the most used approach but the cover should not attenuate incoming optical radiation. The most used material for this purpose is quartz glass (SiO_2) cut in thin slices with optical properties carefully matched with solar cell conversion spectrum (see Figure 5.5). Cover glass should then be glued on top of the solar cell with a special resin to ensure no optical losses. By properly selecting the best cover glass (transmission spectrum and thickness) low losses can be achieved (2 - 5 %) with high displacement damage reduction. The actual degradation the solar cell experience is directly proportional to the equivalent particle flux, but the proportionality coefficient varies widely (see Figure 5.6). Solar cell technology can be selected to ensure better tolerance to incoming particles: commercial solar cells (developed without taking into account the space environment) achieve quite low hardness and are thus not good for long lasting missions in space.

The other components that suffer high degradation due to displacement damages are optical devices, like sensors and emitters. Their degradation in performances is linear with incoming particle NIEL (see section 3.5 for further details) and is in general impacting conversion efficiency [55] (the more the defects in the lattice, the more recombination traps can be found, thus reducing radiative recombination). Space-developed optical devices usually are also enclosed in metal/quartz packages which show low degradation under

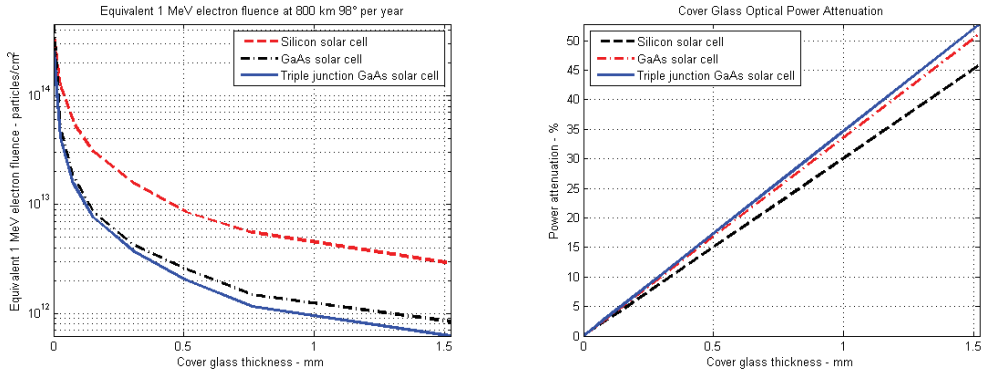


Figure 5.5: Solar cell damages and power attenuation as a function of cover glass thickness (computed taking degradation data from space qualified solar cells datasheets [52][53] and cover glasses optical attenuation from [54]).

radiation while commercial ones are enclosed in epoxy packages which show more lattice damages with higher particle dose: again degradation is almost linear with NIEL. By carefully selecting the proper technology it can be shown that the major degrading factor is actually the plastic package (double heterojunction diodes, for example show a high tolerance level to displacement damage [55]). For further details on the radiation hardness of commercial LEDs and photo-diodes refer to section 7.2.

Laser diodes show also an high radiation tolerance, as shown in [56] where a power degradation of 23% has been demonstrated for a proton fluence of 2×10^{14} 30 MeV particles per square centimeter (for a 5 years mission at an altitude of 800 km the equivalent 30 MeV proton fluence is about 3×10^9 particle, so five orders of magnitude lower than the tested value). Plastic optical components are in general more prone to displacement damages than quartz ones and this can also be seen in optical fibers. In general optical attenuation in media is linearly dependent on path length: this makes optical fiber particularly suited as dose sensors. Given the fiber losses per meter as a function of radiations, sensors with different sensitivities can be built by making the fiber longer or shorter.

Optical devices are also employed for communication [57] over a short or long distance and opto-couplers can be quite common in satellites because they allow to galvanically isolate two systems. As we said before, LEDs and photo-diodes performances degrade with radiation and the opto-coupler Signal to Noise Ratio (SNR) degrade too. In order to guarantee a minimum end of life value, a high current should be used to control the LED (higher than strictly needed at the beginning of the mission) and most of the times this increased power consumption cannot be tolerated. To avoid this, there are two solutions: using rad-hard components (that does not show performances reduction due to radiation) or employ an active gain control on the emitter driver. This solution allows to better tolerate COTS components degradation in space. Further details about the use of commercial components for developing a communication bus can be found in section 7.2.

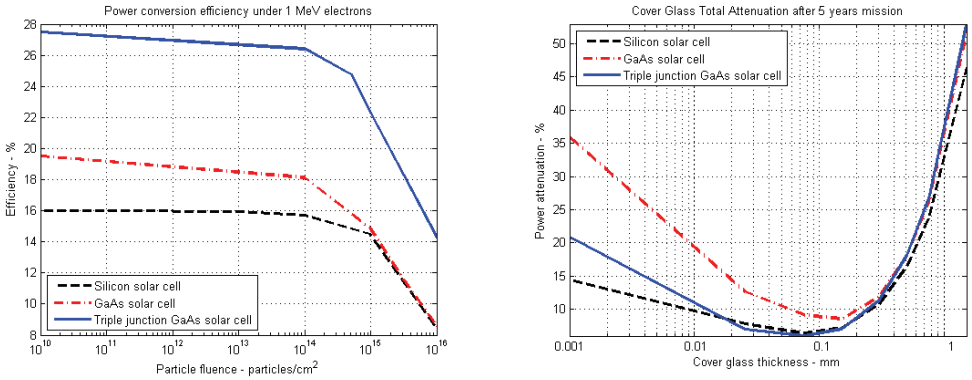


Figure 5.6: Solar cell efficiency reduction and total power loss (efficiency reduction and optical attenuation) as a function of cover glass thickness (computed taking degradation data from space qualified solar cells datasheets [52][53] and cover glasses optical attenuation from [54]).

5.3 SEE mitigation

High energy particles, when impacting with active semiconductors can deposit charge in different circuit parts, thus creating spurious currents or influencing data stored in memory devices. The effects can be divided in two main types: destructive and non-destructive. The latter can be seen as disturbances that only influence the information stored or elaborated by the circuit, while the former can lead to permanent system malfunctioning. Single event effects can be divided in many categories, but the most frequent are latch-ups (SEL), SEU and SET. CMOS devices can be fabricated in many different ways and this can have large influence on single event effects tolerance: Figure 5.7 shows two different processes made on bulk material that both use a separate n-well region to fabricate p-channel devices. This well structure has no direct function other than providing an isolated region for the p-channel devices and maintaining a reverse bias across this junction isolates the well region from the p-substrate; the problem is that this structure contains a parasitic bipolar transistors that can be turned on by high-energy particles that can short the power supply, thus generating latch-up. The highly doped p+ substrate that is used in the second process reduces bulk resistance, making latch-up less likely compared to standard bulk processes. The low-resistivity substrate also reduces the amount of charge that can be collected from the n+ drain, which improves single-event upset hardness compared to bulk processes. Figure 5.8 shows two CMOS structures that eliminate the junction-isolated well structure, thus eliminating the possibility of latch-up. The first process is Silicon On Sapphire (SOS), which results in two separate p- and n-doped islands on an insulating sapphire substrate while the second process is Silicon On Insulator (SOI), which uses special processing to grow an isolated silicon dioxide insulating layer on a bulk silicon substrate.

Another approach to reduce susceptibility can be the use of doped silicon "trenches" to greatly increase the current needed to trigger and sustain latch-up, making these types

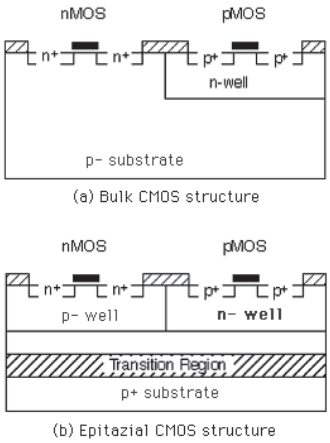


Figure 5.7: Cross-sections of bulk and epitaxial CMOS processes [31].

of events much less likely in space. All these techniques will generally raise the minimum LET threshold and can make SEL probability reduce to acceptable level but the only certain way to eliminate it is to use an SOS or SOI CMOS process that remove one of the parasitic transistors. Anyway, these solutions are not used in significant volume production and they are both highly specialized and costly.

Latch-up can be present in bipolar devices too since transistors are enclosed in wells but the doping concentration is quite higher that in CMOS devices, thus leading to extremely robust parasitic SCR which require extremely high energy particles to be triggered and, in conclusion, makes them almost latch-up free.

The high internal currents that result from latch-up can heat the local, latched region to very high local temperature within a few microseconds. Even if the latch-up is detected and clamped within about 100 us the localized heating may be sufficient to cause electro-migration, burnout of metalization or degradation of MOS contacts. Although burnout is easily identified, the other mechanisms degrade device reliability, effectively introducing latent damage. Micro latch-ups can create also other problems, since the sharp current consumption increase can be limited by the fact that only a small part of the IC is actually latched, thus limiting the effectiveness of protection strategies (see Figure 5.9).

Once latch-up occurs, a device will remain in the high current, latched condition until power is removed. Power cycling will be required each time that a latch-up occurs, which will temporarily shut down sections of the subsystem that share power supplies. Along with power cycling, circuits and subsystems affected by any component that undergoes latch-up will have to be reinitialized. Power cycling and reinitializing may be acceptable with a very low latch-up probability, but will generally be unacceptable if it is frequent; there may also be critical phases of a mission during which latch-up and power cycling cannot be accommodated, because there is insufficient time to recover within the operational window. A latch-up protection system is deeply analyzed in section 7.1.

SEU and SET are generated because an high energy ion induces a short-duration pulse of current in a p-n junction: if the charge is collected by a storage element (e.g., memory or flip-flop) and it exceeds the critical charge required to switch the circuit, it will change

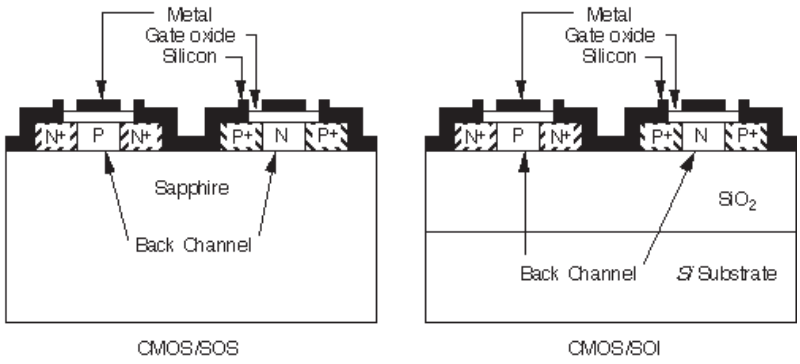


Figure 5.8: Cross-sections of CMOS/SOS and SOI processes [31].

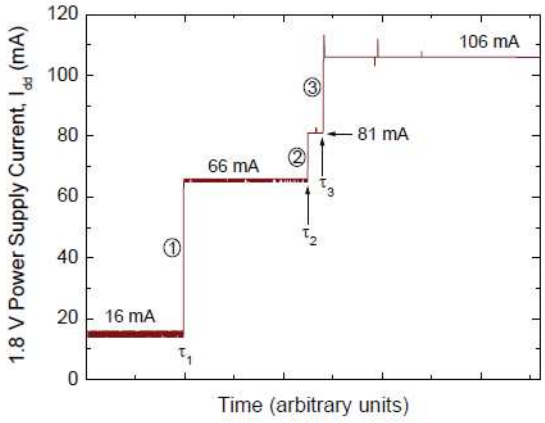


Figure 5.9: Current evolution for three successive laser pulses (simulating ion strikes) incident on an IC [58].

state, and the information that was previously stored will be lost. Even though the circuit changes state, it still functions normally, and reinitializing or rewriting it can restore its original configuration. In a complex ASIC, SEUs will appear at random locations, depending on the particular region that is struck by a high-energy particle. The term SEU describes the situation where the passage of the particle through the device produces only a single upset, while Multiple Bit Upset (MBU) can happen when the particle charged wake influences many devices at the same time [31].

Silicon-on-Insulator substrates can help reducing charge collection because the oxide layer eliminates the well/substrate junction which is one of the major sources of photo currents (currents generated by the interaction with ionizing particles). The reduced photo current collection results in increased circuit upset levels and improved recovery times [59] (the time needed by the injected charge to recombine). Another way to reduce SEU sensitivity is making the device bigger, such that the critical charge needed to flip the stored value is higher: this solution makes high scale integration more complex and thus is almost always rejected, also because it is going in the opposite direction of technology trends. The most effective technique, anyway is to duplicate (or triplicate) the storage device so that error detection of correction can be performed. It should be pointed out that, with the technology scaling trend, devices are much smaller than the charge trace left by the particle (see Figure 5.10 for details) which means that both replicas can suffer upset.

Besides the effects on storage cells, single-event interactions can produce transient output pulses in combinational logic or analog circuits that do not contain storage elements. These transients are usually of short duration (from few nanoseconds to few microseconds), but may indirectly produce errors in storage elements if they occur at critical time periods, such as during clock or data transitions.

In general SEU and SET issues cannot be completely solved at device level: the most

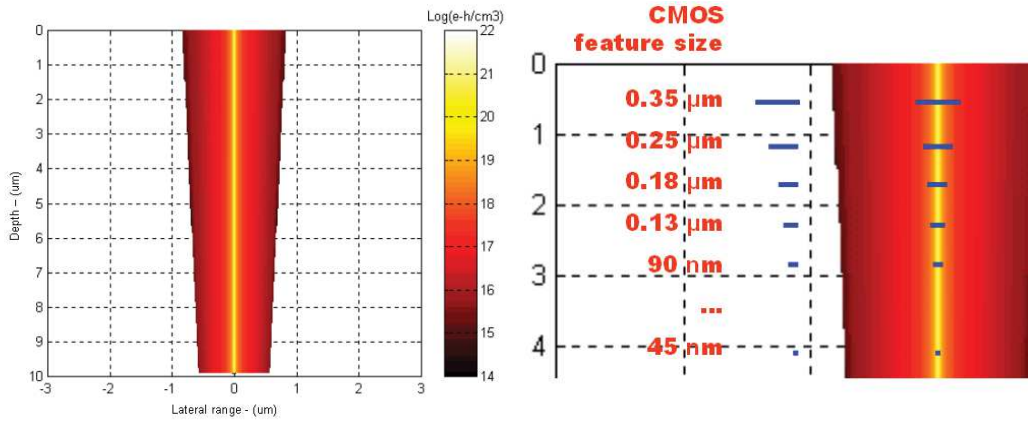


Figure 5.10: Charge trace left by a 275 MeV Fe ion (24 Mev cm^2/mg) in Silicon (P. Foulliat, EWRHE 2004). In the right picture gate sizes for different technology nodes are reported.

used approach is to employ redundancy and voting techniques (see section 5.4) which allow to completely mask single event effects. In memory elements usually error detection and correction techniques are also employed. To further reduce SEU or SEL probability in extremely sensitive devices, they can be kept switched off (not in stand-by mode) and turned on only when needed: this way the upset probability and power consumption can be further reduced.

Shielding is in general not effective against single event upset because high LET particles have a quite deep penetration, thus requiring a thick shield to reduce particle fluence of a limited factor: a 10 mm thick Aluminum shield is only able to reduce 10 MeV cm^2 / mg particles flux by a factor of 8 (see Figure 5.11). The flux is thus reducing from one particle per square centimeter every 3 minutes to one every 20 minutes: shielding can only reduce by a small amount the particle flux but cannot eliminate the problem.

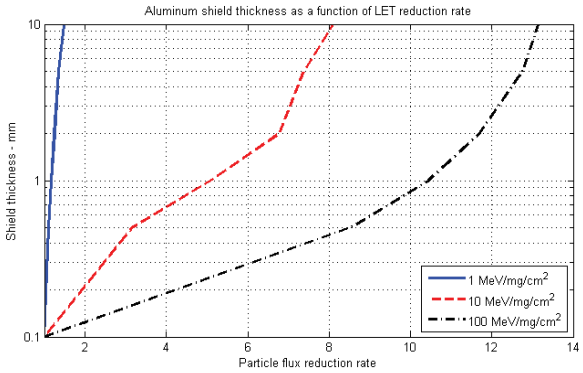


Figure 5.11: Shield thickness as a function of particle flux reduction rate.

5.4 Redundancy

Redundancy is a design approach that requires the replication of a critical component to increase its performances or fault tolerance. By just describing it with such simple words, it is clear that this approach can be applied in many different fashions and to many different levels.

In general there are four main different approaches dealing with redundancy:

- Hardware redundancy, which is basically component replication,
- Information redundancy, which deals with error detection and correction,
- Time redundancy, used to deal with transient faults,
- Software redundancy, or N-version programming.

One of the most well-known applications of hardware redundancy is the so called Redundant Array of Inexpensive Disks (RAID) which was developed to increase the fault tolerance of computer hard-disks for critical applications like servers. The basic idea was to use cheap hardware (with a low reliability) and to improve it with redundancy, by storing multiple copies on different disks, such that, in case of failure, no data is lost. The proposed approach is to apply similar techniques to space systems in order to reduce the cost.

Software redundancy or, more in general, N-version programming is used to reduce the probability of undetected bugs in a program creating a crash: complex systems cannot be 100% tested before their use and a small possibility for a bug to exist in the code is always present. What can be done is to develop two (or in the general case N) completely independent softwares, developed by different people, with different habits and conventions. In this way, it is unlikely that two independent groups could solve the same problem with the exact same solution, so bugs will be different in the two versions. This means that if one of the two versions crashes, the other one should be still operating correctly: in this way an error detection system could notice that the systems are giving different outputs and thus implement a recovery action. In general two replicas are not enough to correct an error with a majority voting strategy, so replicas are usually odd in number: an example of this is the NASA Space Shuttle on-board computer that is seven-fold and each replica was built by different developers. A fault affecting two systems together is in general called common-mode fault and it is what is avoided by using the redundant replicas. The overhead introduced by this technique is quite high, since several system replicas need to be present at the same time and they should be developed by different teams, thus multiplying the cost of one single instance times the number of replicas. This technique is thus used only in very complex and expensive systems (or when a reliable life support is needed) as happens in the Space Shuttle.

Time redundancy implies the evaluation of a certain parameter twice (or again N times) in time to ensure that the result is still the same. When this comes to electronic systems, it is in general implemented by executing twice the same sub-routine (if in software) or by feeding an electronic circuit with the same data twice and comparing the result. This

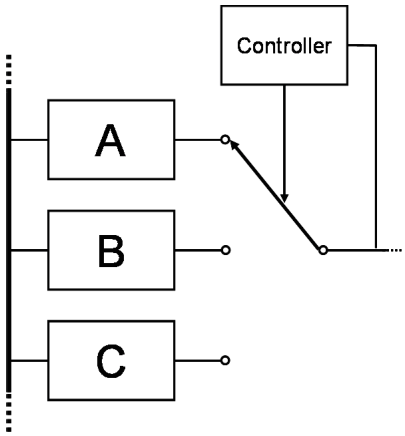


Figure 5.12: An example of cold redundancy with three replicas and a controller selecting which one to actually use.

technique allows to detect (and correct if more than two replicas are again used) a transient error that happened only in one of the N executions of the same algorithm. This technique is quite simple to implement, but may be limited by the existence of common-mode faults (the same system is used to compute N times the result) or by errors in stored values, thus making it ineffective. The main advantage is that it does not cost much in terms of hardware, but it multiplies execution time by N (the number of iterations that should be computed), which can then cost in terms of hardware if the system has critical time constraints and a faster processor or logic circuit should be employed.

Information redundancy is in general used for securing data storage or transfer by adding further information that will allow to detect or correct errors. The simplest approach requires to transmit or store twice the same data so that, if one copy is corrupt, the second can be used to detect the error: the cost is linear with the redundancy level and usually cannot be tolerated. Smarter approaches take advantage of data encryption or compression to actually reduce the amount of data to be stored or transmitted. Just to go back to the first example, RAID1 employs two disks for storing two independent copies of the same data, while other implementations like RAID2, store data with error detection and correction codes (single bit correction, double bit detection) to reduce the overhead.

The previous three approaches are mostly implemented in software (even if with some exceptions) involving complex algorithms: they were not fully analyzed in this work and developed since most of the efforts were spent on hardware.

Hardware redundancy takes advantage of components replication to increase system reliability and it can be implemented in three different ways: when all the replicas are powered on at the same time this is called hot redundancy while when only one replica is operating at once it is defined cold; the third solution, defined as hybrid or sometimes warm, employs a mix of the previous two techniques.

In cold redundancy, many powered-off spare devices are used in the system with just

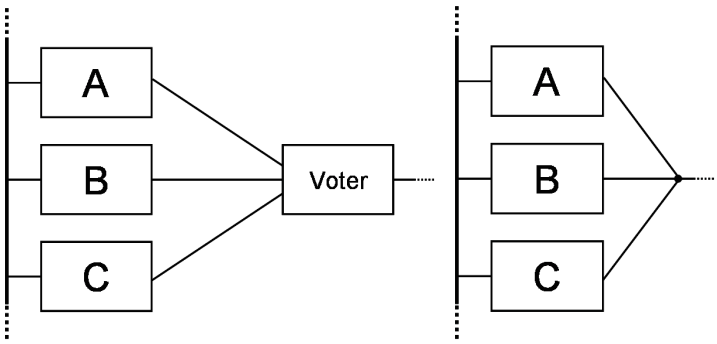
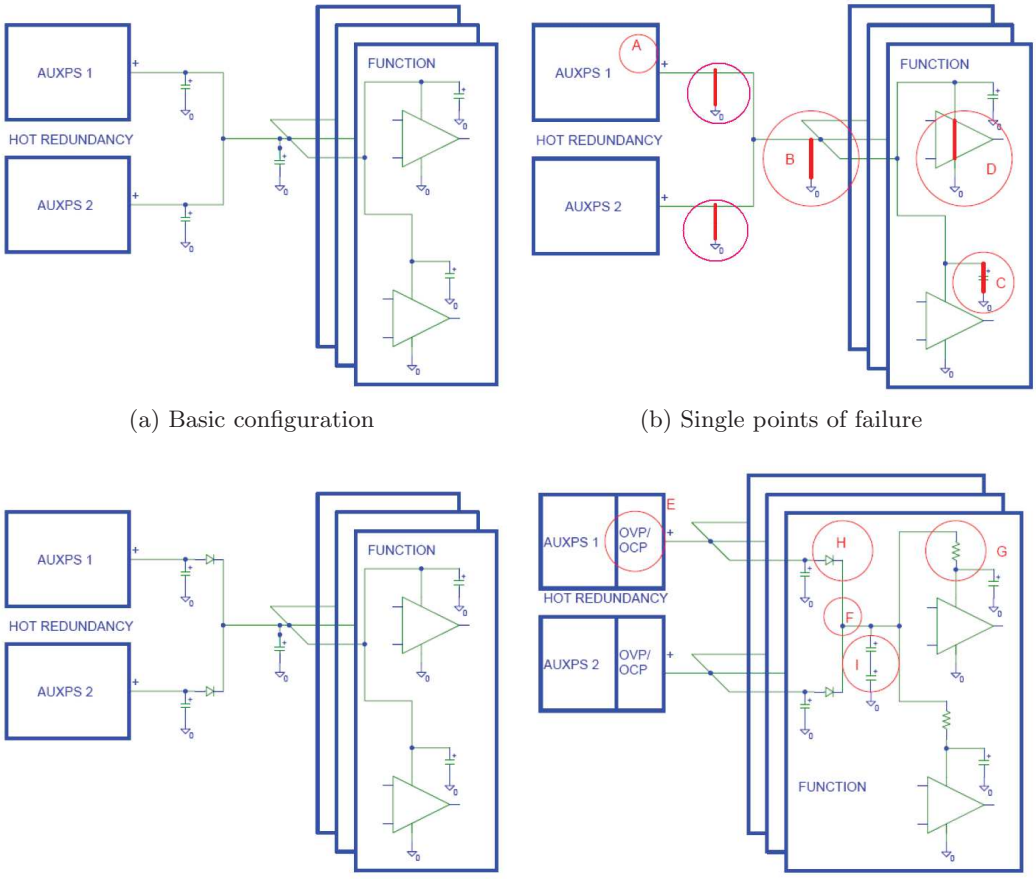


Figure 5.13: Two examples of hot redundancy: TMR system with voter and a redundant driver example, with three components sharing the output.

one on at once: error correction is not feasible since only one replica is operating, but detection is still possible. In case a fault is affecting the replica under use, it can be switched off and substituted by one of the cold spares (see Figure 5.12). Several strategies can be used for selecting which device to use, like sequentially use all the replicas: this approach makes the controller extremely simple because it does not have to detect errors, but has to just select which one to use. This can be useful for TID effects mitigation when powered down devices suffer lower damages than the others: in this way unpowered devices will survive longer and cold redundancy ensures the use of the less damaged one. This technique is not effective for example for bipolar ICs since they suffer more for TID degradation if left unpowered while MOS devices show higher degradation when powered on (see section 3.3 for further details). Cold redundancy is efficient from the power consumption point of view (only one at once is on) but can be slow in recovering from errors and furthermore, errors cannot be masked, so they will be present on output for the controller reaction time; this makes cold redundancy in general not suited for critical sub-systems.

Hot redundancy implies instead the contemporary use of all the replicas at the same time connected together in different fashions. For power circuits in general all the outputs are connected in parallel so that current can be balanced between all the devices, thus actually lowering constraints on the single device and also allowing it to tolerate higher degradation due to radiation. It should be noted that a faulty replica should not prevent the others to operate correctly (for example a short circuit at the output can stop the whole system) so before the parallel connection, usually a protection circuit is employed. The other solution for hot redundancy is to use all the components at the same time and then use a voter for selecting the correct result based on all the outputs with a majority voting (most of the time this is called modular redundancy, for example Triple Modular Redundancy (TMR), see Figure 5.13). The critical point is then moved from each replica to the voter, since a failure there would stop the whole system. Hybrid solutions can also be employed by mixing the previous two methods together.

While using redundancy, the designer should ensure that each replica is independent



(a) Basic configuration

(b) Single points of failure

(c) Power supply protection

(d) SPF-free configuration

Figure 5.14: Single Point of Failure (SPF) prevention (from Olivier Mourra, ESA).

from the others: when one is faulty, it should not influence the others. While at a first look, this seems quite easy, many problems can arise: in Figure 5.14a it can be seen that the two power supply modules (the blocks on the left) are not completely independent, since their output is connected together and a short in one of the two can short also the second. This is usually called Single Point of Failure (SPF). After identifying it, it is possible to solve it by slightly modifying the circuit (see Figure 5.14b): the output of the power supplies can be protected against shorts (Figure 5.14c) and then, by properly duplicating all the connections and protecting critical devices, all SPF can be removed (Figure 5.14d).

5.5 COTS component selection guidelines

From the system designer's perspective, the important and distinguishing feature between rad-hard and commercial components is the person that guarantees that the device is

going to stand the specified radiation level: rad-hard components are produced for that purpose and so they are guaranteed by the manufacturer, while with COTS components it is up to the designer to accept the risk.

First of all, the radiation environment should be modeled to exactly know the limits the devices should be able to tolerate, in terms of TID, displacement damage and single event rates. Given these numbers a first decision should be taken: going for military / space-qualified or commercial components. This selection should also be performed keeping budget in mind since the usual price ratio between the two families is around 100 - 1000 times.

The only way to certify that devices are able to survive in space is to do proper testing: this can be made by the manufacturer (rad-hard components) of by the designer. Since testing can be quite expensive, in many cases the designer can rely on tests performed by other developers and grouped in databases: one of the most complete and up-to-date is provided by the IEEE Radiation Effects Data Workshop [60], but others are also present, like the GSFC Radiation Data Base [61] or the ESA Radiation Effects Database [62]. These archives contain the results of radiation tests performed on many different commercial components and allow to perform a pre-selection of the devices to employ. Unfortunately most of these components are a bit outdated since in general older components (which already have a reliability record of several years) are preferred to new ones. Many times, anyway, mission requirement cannot tolerate such old components so non already tested devices have to be selected.

This small selection guide wants to enumerate some guidelines for proper selecting commercial components for space applications, taking into account in particular technology benefits from the radiation tolerance point of view.

The first problem that will be analyzed is the latch-up: commercial CMOS devices are prone to it, so they need to be protected because even if the probability of such events is low, it can cause component destruction and maybe mission failure. Since the only strategy to protect commercial devices from latch-up is power cycling them, protected devices cannot guarantee 100% availability: in case this cannot be tolerated for some components, latch-up resistant devices should be employed instead. Bipolar ICs are a viable alternative since they show a higher resistance to this phenomenon, but not all the components can be made bipolar. Micro-processors or, in general, complex digital systems, cannot be bipolar because they are usually not available on the market or their power consumption would be too high. In this case, if the availability constraint cannot be softened, rad-hard components should be employed. When the same digital function can be performed with analog circuits, it can be performed using bipolar ICs that are not prone to latch-up.

An example of this strategy could be the control of power converter stages: they should be available 100% of the time, since a power cycling there would shut down part of the satellite. This strategy has been applied to the development of a Maximum Power Point Tracker (MPPT) for solar panels and further details can be found in (section 5.4, where an hybryd analog - digital control loop has been developed.

Once a rough selection has been performed between bipolar and CMOS ICs, a more

detailed technology selection should be performed to select the best device for each application.

Technology		Hardness
Bipolar	TTL/STTL	≈ 1 Mrad
	ECL	≈ 10 Mrad
	Linear	≈ 10 krad
CMOS	NMOS	≈ 1 krad
	CMOS Bulk	≈ 10 krad
	CMOS RH	≈ 10 Mrad
	CMOS SOS/SOI	≈ 10 Mrad

Table 5.2: TID tolerance as a function of device technology [42].

In small size systems, in general processors are not employed and the preferred solution is to go for micro-controllers (from extremely low power ones to hundred MIPS): these devices are usually not listed in the previous mentioned data bases, but data about them can be found by looking at CubeSat Developers Workshop [19] proceedings, where flight data are sometimes presented. In general MSP430 from Texas Instruments has quite a long flight heritage [63] and has been tested also under radiation [64]: it should be noted that these results are for the MSP430F1xx series. Higher power micro-controllers have been employed in small satellites too, such as Atmel AT91SAM7A1 [65], based on an ARM7 core.

Memories are also employed in all digital systems and their selection is a critical task since many different technologies are becoming available both for the volatile and non-volatile ones.

Volatile memories belong to two main families: static and dynamic RAMs. Static RAM (SRAM) are faster and more power hungry when compared to dynamic ones but they are much simpler to use because no external refresh controller is needed; on the other side, Dynamic RAM (DRAM)s can be highly integrated, reaching high capacity (many gigabit per chip) while static ones are limited to few megabit per chip. Since the SRAM cell is bigger and more complex, it was always thought to be more resistant to upset, but while the technology is scaling, DRAM proved to be more resistant. With new technology nodes cell area is smaller, thus smaller charge collection occurs, while cell capacity is not scaling much, making dynamic RAMs more robust. They are thus preferred to SRAM, but the added complexity in the circuit should also be taken into account.

Among non-volatile memories, FLASH are the most used because they guarantee quite good performances in terms of radiation tolerance: the main disadvantage they have is that after a certain TID threshold has been passed the device cannot be programmed anymore. This drawback is not caused by the FLASH cell itself, but by the charge pump circuit needed to generate the higher voltage for programing or erasing it [66]: this problem is more severe in NAND devices with respect to NOR ones. Nevertheless they are employed because typical TID threshold before failure is in the order of 30 krad - 50 krad. NAND devices furthermore are becoming also bigger and bigger thanks to the spread of solid

state disks, so that devices from 4 to 64 gigabit are commercially available.

Commercial memories are anyway SEU and SEL prone: for SEL protection power cycling is in general enough and cell itself is robust to SEU so no bit flip are in general expected, even if errors in the read-out circuit can happen. SEU furthermore are in general not able to turn the programming / erasing circuit on, thus altering stored values.

Newer technologies are also emerging, like magnetic [67] and ferroelectric [68] RAM: both cells show high TID tolerance (up to around 100 krad) but suffer from heavy ion induced problems. Magnetic RAMs are prone to latch-up down to quite low LET (around 10 MeV cm^2 / mg, which can make them experience latch-up on average every hour in an 800 km orbit) while Ferroelectric ones show better performances. The main problem these memories can experience is due to the control logic, which can suffer upsets causing memory rewrite or memory erase (there is no need to turn charge pumps on as in FLASH memories, so it is easier to damage stored data).

Optical components should also be selected taking into account all the different technologies employed in commercial devices for space operations: double heterojunction LEDs showed a far higher displacement damage tolerance when compared to single heterojunctions ones [55] or regular junction emitters.

The previous guidelines can help identifying the best commercial device for space application from the technology point of view. Anyway, large variations can happen between devices developed with the same technology and even between different samples of the same component. Batch lot testing is the best approach to guarantee survivability: a big number of devices built in the same production lot (and so with exactly the same technology) should be collected and, by testing some samples, the whole lot can be declared as qualified.

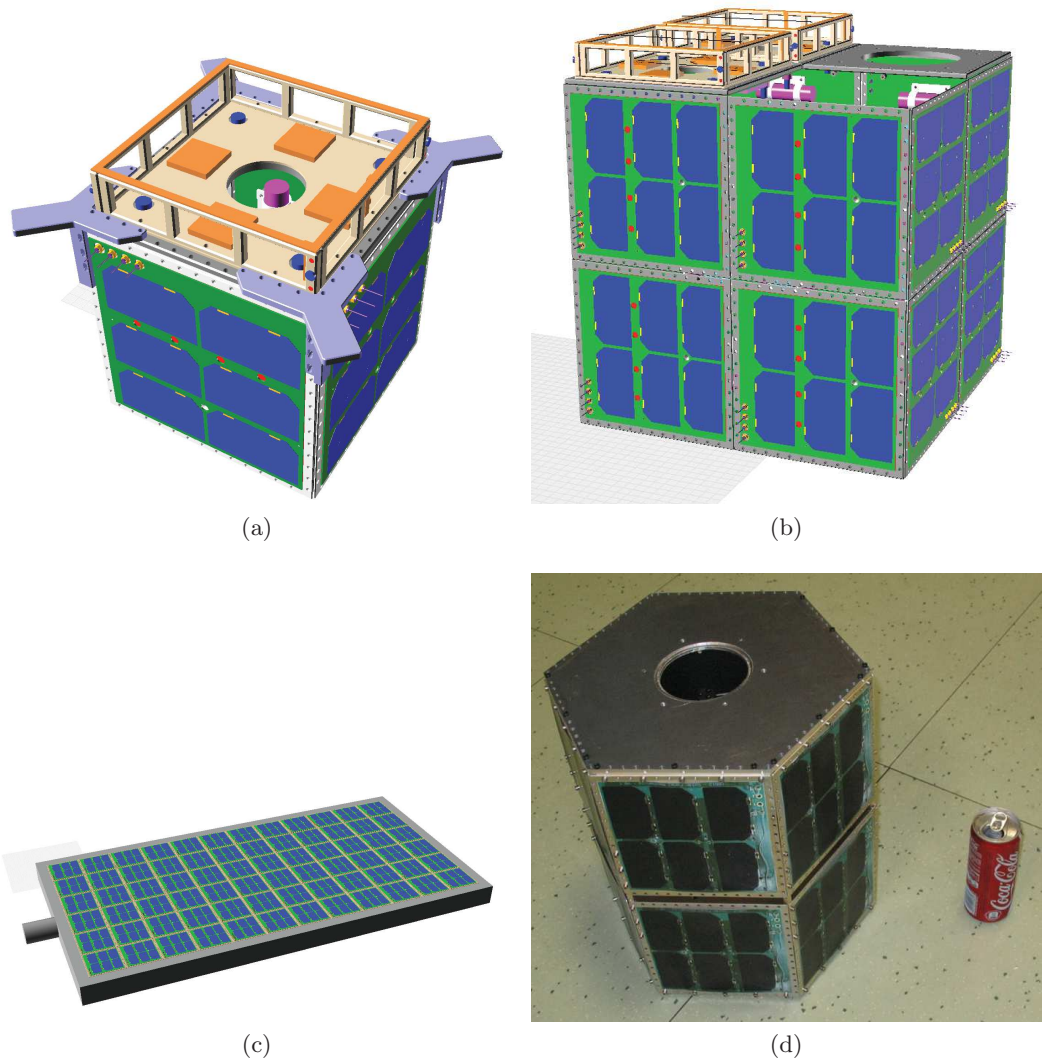
Chapter 6

The AraMiS architecture

Avionics for satellites is a market which is continuously growing in these years, especially thanks to the availability of low cost launch vectors. This cost reduction enabled many institutions (industries but also universities) to develop their own satellites. A first result of this process was the CubeSat concept: a really small satellite, built using COTS components, that everybody could buy and assemble from a kit. Architettura Modulare per Satelliti (Modular architecture for Satellites) (AraMiS) is a project that wants to go beyond this concept and create a true modular architecture. The main idea is the development of distributed and intercommunicating on-board units, built with COTS components to reduce system cost, that can be assembled together to fit mission specific requirements, thus allowing an effective cost sharing between different missions. This solution wants both to create cheaper systems and to make design time faster.

The most effective way to reduce the cost of a nano- and micro-satellite missions is to reduce design and non-recurrent fabrication costs as much as possible, which usually account for more than 90% of the overall budget. Reducing them can be achieved only by sharing the design among a large number of missions. Design reuse is the rationale behind the AraMiS project, that is, to have a modular architecture based on a small number of flexible and powerful modules which can be reused as much as possible in different missions.

This architecture is intended for different satellite missions, from small systems weighting about 5 kg to bigger missions: several configurations are depicted in Figure 6.1 to show some of the potentialities of the proposed architecture. Modularity is intended in many ways: from the mechanical point of view, like in the CubeSat concept, to allow the composition of bigger structures in a simple way. But it is also intended from the electronic point of view: fitting such a wide range of applications, requires that most of the internal subsystems are developed so that they can be composed together to increase performances. The most straightforward example is the power management sub-system: to get the maximum solar power, solar cells are mounted on all the available surfaces but their number is varying from mission to mission, thus requiring to re-design each time this system. The new approach makes use of a standard module, as can be seen from in Figure 6.1, and it replicates it as many time as needed to fit mission constraints.



(a) (b)

(c) (d)

Figure 6.1: Different configurations that can be built with the AraMiS architecture [69].

The whole satellite is anyway composed by many other sub-systems that allow to be combined together to achieve the desired flexibility level; the main sub-systems that compose this architecture are:

- mechanical,
- power management,
- telecommunication,
- attitude determination and control,

- on-board control and processing,
- payload.

In the following sections a brief description of all the sub-systems will be presented, while a deeper description can be found in chapter 7, where a detailed analysis of the developed sub-systems is presented.

6.1 Mechanical sub-system

The mechanical sub-system is the backbone of a satellite: it is used to hold in place all the different sub-systems and to protect them from the external environment conditions. The main material used for building the AraMiS structure is Aluminum, used in particular for its low weight: the backbone is made by metallic square rods while the power management and telecommunication sub-systems are mounted on thin panels that are screwed to the rods. The power management tiles (called in this way due to their shape) are used to cover the satellite: solar panels are mounted on the external face so that power generation can be easily achieved. The number of these tiles is mainly dependent on satellite size and required power and this gives high freedom to mission designers since size and generated power can be increased by simply adding more modules.

Aluminum surface should be conductive for one main reason: while in the upper atmosphere layers, the satellite will be surrounded by plasma that can electrostatically charge non-conductive parts, thus making potential build-up until a spark takes place (see section 2.2.3). There are two main conversion coatings for Aluminum: anodization and alodination. The first one is an electrochemical passivation process that grows a thick oxide layer on the metal to greatly increase corrosion resistance but its drawback is that it makes the surface dielectric, so prone to arcing in the plasma environment. Alodination is instead a coating process that uses a particular paint (Alodine® from Henkel was the original one, but many brands are now available) that creates a protective and conductive layer on the metal surface: this process is heavily used in spacecrafts (it is also regulated by the MIL-DTL-5541F [70] standard) but may be toxic: most of the available products in fact use hexavalent chromium in the paint, which leaves the surface yellow colored, but requires proper handling. Other products are also available that do not employ chromium anymore, but use instead a zirconium salt. A conducting surface is useful also from the Electromagnetic Interference (EMI) shielding point of view, since it allows to create a perfect Faraday cage to enclose the electronic systems: to ensure proper shielding all the tiles should show perfect electrical connection to the mechanical structure and between themselves up to the maximum frequency involved in the electronic circuits. In our case the highest frequency is 2.4 GHz (without considering the harmonics) and this requires a perfect electrical conductive surface with holes smaller $\frac{\lambda}{10}$. This constraint was met by imposing the maximum distance between the mounting screws to be lower than 12.5 mm (it is 11.75 mm in AraMiS).

All the tiles are expected to be connected on the external faces of the satellite and the payload is expected to be mounted inside, anyway if requested by the particular mission,

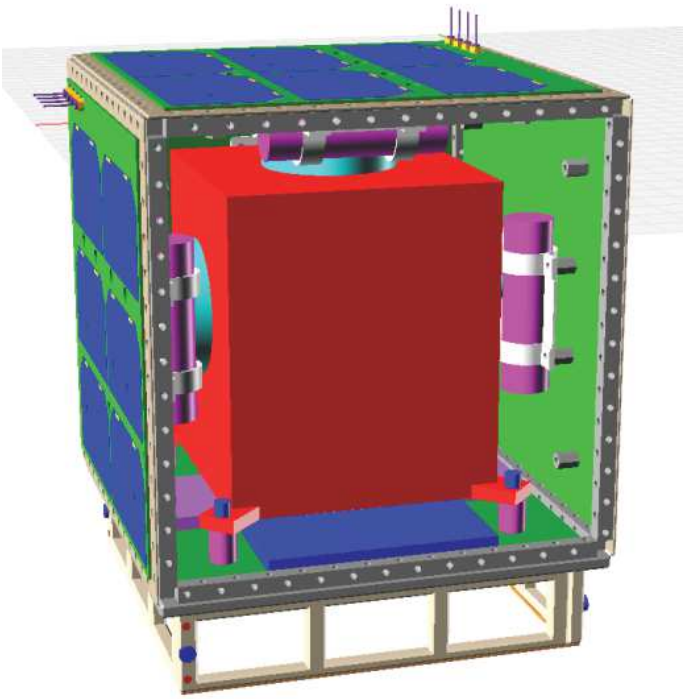
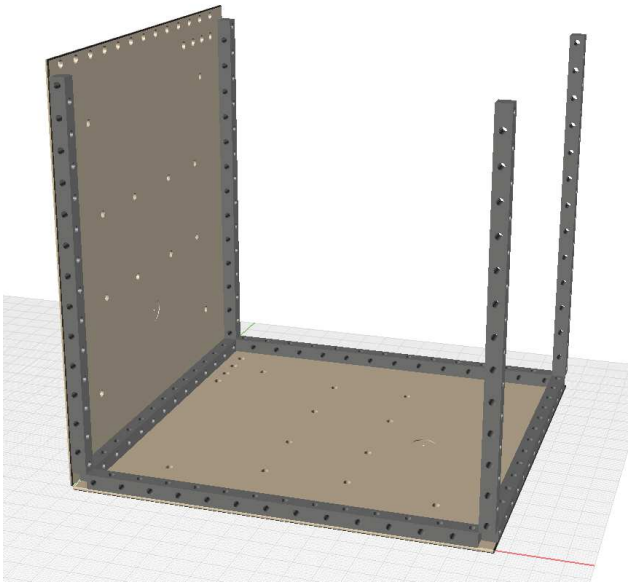


Figure 6.2: Payload bay inside AraMiS (in red) [69].

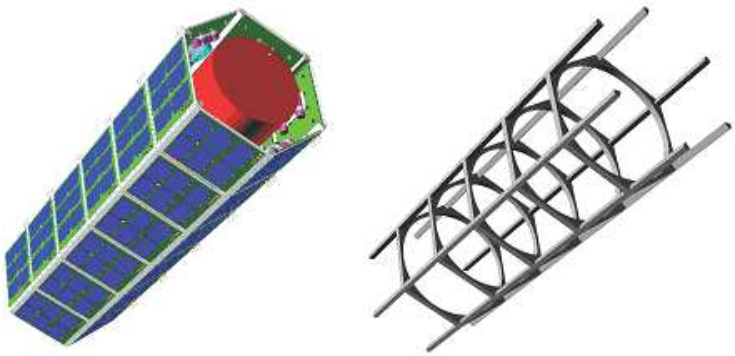
an empty panel can be mounted on the external structure to house a particular payload. When the payload is mounted on the internal part of the satellite, it is expected to be mechanically connected to the Power Tiles, as depicted in Figure 6.2, while for particular needs, it can be connected to the backbone, as in Figure 6.3b. In the basic, single cube configuration, a payload bay $10 \times 10 \times 10$ cm big is available in the center of the satellite; in case a camera needs to be employed in the payload, a hole of 5 cm in diameter is provided at the center of the telecommunication tile to fit the optics. The configuration depicted in Figure 6.1b can house a payload $23 \times 23 \times 23$ cm wide, while in telescope configuration (see Figure 6.1d), a cylindric payload with a diameter of 20 cm can be used with an hexagonal prismatic configuration or a 32 cm diameter with an octagonal one.

6.2 Power management sub-system

The power management sub-system is responsible of generating, storing and delivering power to all the other satellite sub-systems. This one is a particularly critical part since a destructive failure here can shut down everything, thus leading to mission failure. Fault tolerance is then particularly important and most of the design solutions were selected for this reason.



(a)



(b)

Figure 6.3: AraMiS mechanical structures for the single cube configuration (a) and for the telescope configuration (b) [69].

Power management is usually quite mission-dependent thus requiring ad-hoc development and tailoring for the specific needs, which can increase system cost and testing time. The basic idea that lead the development of this system was to make it modular, such that it could be adapted to many different situations and missions, thus lowering development and testing costs. Since it's primary goal was to be employed on a small satellite (similar in size to CubeSats, but with higher internal capacity, see Figure 6.1) and bigger ones, it has to be modular and the best solution was to put these modules on the external faces on the satellite, thus making it able to be adapted to many different architectures.

To allow the use of these modules to build a big system it is necessary to interconnect

them in a flexible way: in particular when big structures are assembled, it is normal to deal with high power demanding payloads that usually require a high voltage power supply. A conversion from a low voltage bus can allow this but it reduces overall efficiency due to Joule losses in the cables. The best solution is to have a modular bus, connected to the previous modules, that can be configured according to specific needs: the new approach to solve power distribution problems is to connect these modules in series, if a high voltage bus is need, or in parallel, when a high current is needed from the bus; an hybrid solution can also be employed to meet mission specific requirements.

As pointed out in the previous paragraphs, power management can be divided into two main parts: power generation and storage and power distribution. They will be better analyzed in section 7.3 and 7.4.

6.3 RF Communication sub-system

The AraMiS communications sub-system follows the modularity philosophy of the satellite. In fact, a basic communication tile is provided as standard, while dedicated tiles can be foreseen in case of special applications.

This communication module is used to receive command and control packets from the ground station and to send back telemetry and status information. The bandwidth needed to exchange this kind of information is usually low, so the RF link was designed for low speed and low power.

This module was designed using commercial components which required proper solutions to achieve a good fault tolerance level: two different channels are used in the bands allocated by ITU for satellite communications. The first channel lays in the UHF 437 MHz band, the second in the SHF 2.4 GHz band. The data content of the two links is equivalent, thus providing two interchangeable possibilities to communicate with the satellite. To reduce occupied bandwidth, both channels implement an half-duplex protocol, sharing the same frequency for downlink and uplink.

The UHF downlink was designed to be compatible with the amateur packet radio, so it uses the AX.25 protocol. The reason for this choice was to enable the reception of AraMiS telemetry by radio amateurs around the World or to take advantage of the GENSO [71] ground station network to collect data from orbits unreachable from our ground station. The S-Band link is organized in a similar way but, to avoid the computational overhead of some of the operations required by AX.25 (scrambling and bit-stuffing), the transceiver uses a modulation scheme which is not directly compatible with amateur stations.

As already mentioned, the basic RF subsystem is built on a standard AraMiS tile. The UHF link is based on a transceiver from Texas Instruments / Chipcon, model CC1020. This chip implements a complete digital UHF radio, with one output and one input channel, with good input sensitivity and output power in excess of 1 mW. To complete the UHF channel it is therefore necessary to connect the transceiver to a microprocessor on the baseband side, to a Power Amplifier (PA) on the RF transmitter side and to a Low Noise Amplifier (LNA) on the RF receiver side. A 34 dBm power amplifier is employed to guarantee a loud and clear signal on the Earth, also using an hand-held receiver: this also

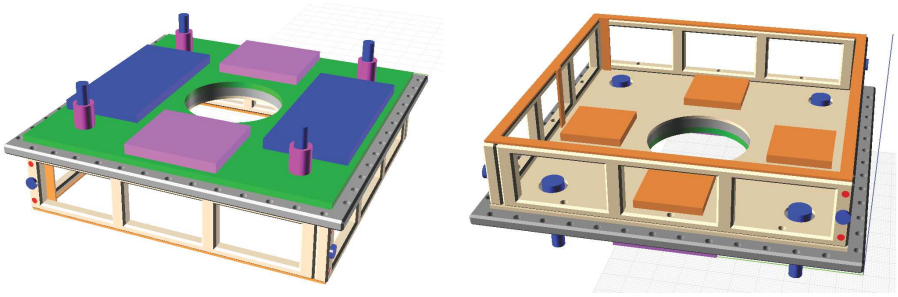


Figure 6.4: TT&C module [69].

ensures a low bit error rate up to 9600 bps.

An electromechanical switch is used to connect the single AraMiS UHF antenna to the output of the PA or the input of the LNA. This device is more robust than an active switch (no radiation-related problems), and more power efficient than many diode based circuits.

The processor supervising the UHF link is a MSP430 and its functions are essentially:

- to exchange baseband command / status / telemetry packets with OBC and Payload processors;
- to encode / decode packets, by performing scrambling / descrambling, bit-stuffing, insertion and removal of prologue and header information, so that OBC and Payload do not need to cope with communications details;
- to supervise operation of the transceiver and RF subsystem (power sequencing, antenna switching).

The S-band link is based on the Texas Instruments / Chipcon CC2510 transceiver. This device incorporates a complete radio modem and a 8051 core, thus it is not necessary to use an external processor as in the UHF channel. Again, the transceiver was not intended for high power operations, so an external power amplifier had to be employed to achieve a 35 dBm output: this allows to safely have a 500 kbps link with a 10^{-5} Bit Error Rate (BER) down to 10° above the horizon with a 3 m dish on ground.

Two different antennas are employed on the channels: a hula hop antenna is used in UHF to guarantee almost omni-directional coverage (even if with low gain) while a 2×2 array of patch antennas is used in S-band. This second antenna allows to achieve a gain of 12 dB with a -3 dB lobe of 60°.

Figure 6.4 depicts he structure of the communication module, built on an Aluminum tile, which has the two antennas on the external side and the electronics on the other. The 5 cm hole for an eventual camera can be clearly seen from that picture.

6.4 On-Board Computer sub-system

In AraMiS the OBC unit is mainly responsible of managing the system, in particular of:

- creating and transmitting (by Transceiver board) Beacon packets,
- decoding and executing commands,
- executing attitude control algorithm,
- storing housekeeping data,
- controlling Payload sub-systems.

The attitude control algorithm is the most time consuming task the OBC should perform: it takes data from the magnetic field, gyroscopes and Sun sensors on the Power Management Tiles and it should compute attitude corrections for all the actuators. The OBC should have a 3D map representing the position of every tile with respect to satellite reference frame and from that map and the position correction, the torque required for all the tiles is computed. Then the command is sent to the actuators for execution.

The OBC is also responsible of controlling Payload boards which should perform all the operations needed to accomplish the mission: OBC has also the possibility to perform some mission specific tasks, in case the payload requires a limited CPU time, even if it was not developed for this purpose. The last task the computer has is the acquisition and storage of housekeeping measurements that can be used to verify satellite status.

The unit is made by 2 MSP430 [72] micro-controllers that check every operation with each other: this system is used to prevent single even transients that influence only one of the two. Furthermore, every micro-controller is protected by an advanced watch-dog that checks software operations and memory locations against a golden value stored in memory. This supervisor is implemented on a small FPGA and allows to verify if code and data memory got corrupted or software crashed. In case of errors, the supervisor performs a processor reset so that operations can be continued: in case of more that 3 successive reboots in a short time, a complete memory reprogramming can be performed to prevent code corruption. A basic firmware version is stored in a read-only memory and it is used in case of re-programming. Processor code can be also uploaded from ground to add new features or correct bugs: in case the new software generates too many crashes a re-program is performed using the old reference code, so that even with a broken firmware, the system is able to guarantee basic operations.

6.5 Attitude determination and control sub-system

This sub-system is responsible of sensing and modifying satellite attitude to keep the on-board instruments aiming in their target and ensure proper antenna pointing for ground communications. Attitude control can be performed in passive or active way: passive attitude control is usually achieved by mounting a permanent magnet in the satellite, that will then act as a compass in the Earth magnetic field. This system is extremely simple

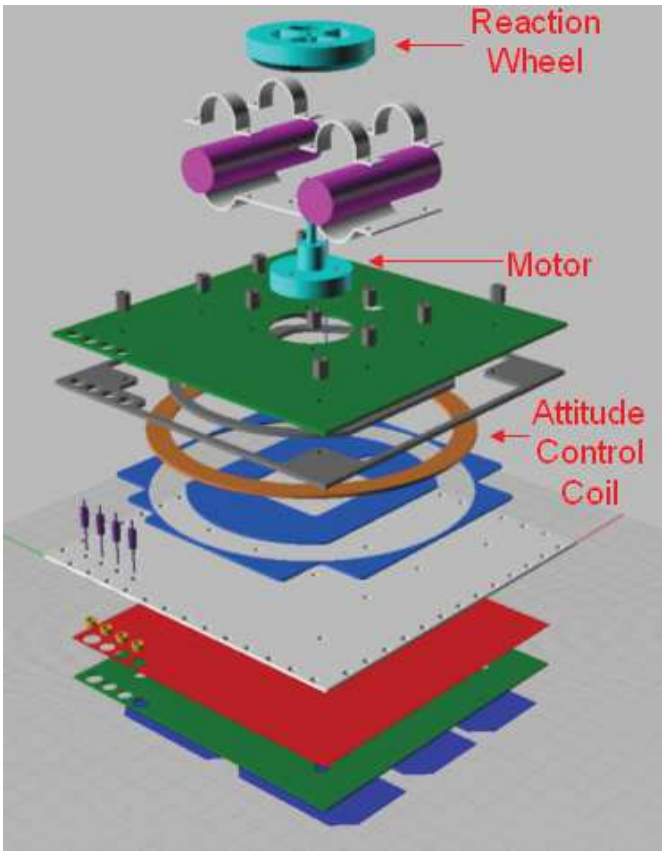


Figure 6.5: AraMiS external panel exploded view and attitude control actuators [69].

to implement and does not even require a power supply. The main drawback is that no fine control can be performed and the satellite will be spinning due to the Earth magnetic field varying from North to South along the orbit. Active control is performed instead by means of controlled actuators that modify satellite attitude on commands: control can be automatic when target direction and corrections are computed on-board while it can be manual when this is computed on ground. In AraMiS, attitude control is performed in automatic way by the satellite itself and this required computational capabilities on-board and attitude actuators and sensors to be put in a closed control loop.

Three kind of attitude sensors are used in AraMiS: magnetic, spin and Sun sensors. They were selected because they do not require much space on-board and they can be implemented with low power consumption devices. Since the main AraMiS goal it to use commercial components, the best solution to ensure fault tolerance and system survivability is to use multiple sensors, better if they can be put on all the satellite faces to allow contemporary measurements from different points. In this way many measurements can be taken from different sensors and an eventual fault can be tolerated, furthermore multiple measurements allow to improve precision. It was thus decided to install all the

sensors on each of the faces and in particular to mount them on each Power Management Tile: the tile should already have a small control unit that is taking care of power handling and adding just the sensors does not make its complexity grow much.

The Sun sensor is implemented using a small camera with a pin-hole optic that allows to identify a light spot moving onto the sensor array, thus allowing to compute Sun position from the spot position. This sensor allows to compute the Sun looking angle that allows, when combined with the sensor orientation, to compute the Sun position with respect to the satellite reference frame. Sensor precision is quite high and allows to compute the Sun direction with an error of approximately one tenth of degree. The spin sensor is an integrated gyroscope that can be used to measure satellite spin over all the axes. Since every sensor is single axis, a sensor on each tile is used so that all the possible spin axis can be measured. Precision in this case is quite low, in particular when the spin rate is low (one turn every few seconds) but this information can be easily computed from other sources (like solar cells available power or from the Sun sensor) while performances are better when spin rate is higher and cannot thus be measured with other sensors. Magnetic sensors are able to measure the 2D magnetic field and the 3D value can be computed by using multiple sensors. The output of this sensor represents the 2D angle towards magnetic North pole and precision is limited to about one degree, but power wires in the satellite can create spurious signals when they start conducting high currents.

Attitude actuators used in general on small satellites are magnetic torquers (coils used to apply a momentum to the satellite thanks to the Earth magnetic field) and reaction wheels (electric motors used to generate a torque). Magnetic actuators are in general quite slow in reacting because the magnetic field strength is quite low in orbit, thus limiting the actual force that can be generated, while reaction wheels have a quite fast response time since they only require to spin-up a motor. The two previous actuators have completely different characteristics, so they can be both used to achieve an effective control. Magnetic actuators in fact can only generate momentum acting on planes that are not parallel to the Earth magnetic field one: this implies that only two satellite axis can be effectively controlled with coils (no momentum around the nadir axis could be generated). Reaction wheels instead generate a torque around their rotation axis which means that, by properly installing many reaction wheels in the satellite, torques acting on any axis can be applied to the satellite.

The attitude control system has two main modes of operations: de-tumbling and attitude control. The first mode is used just after satellite separation from the launch vector: since separation is performed by mechanical or explosive clamps, a random tumbling movement can be induced due to the difference in the release time of the clamps. These oscillations are stronger for small satellites due to their low mass. The attitude control system should then act to stop the oscillations: in this situation reaction wheels cannot be properly used since the movements are rather random, so magnetic torquers are the best solution because they can impose a fixed orientation with respect to the Earth magnetic field and the tumbling energy can be slowly dumped. When satellite is stable, the real attitude control system can be turned on and reaction wheels can be used to induce torques on the various axis to reach proper orientation on every axis.

6.6 Payload

The payload is heavily mission dependent and the architecture was developed to allow high flexibility on it: the main requirements that the AraMiS architecture poses on the payload is its compatibility with the on-board power distribution and data bus. Different payloads can be fitted in the various configurations but mechanical fixtures should be developed to connect them to the mechanical structure.

The payload is housed in the inner part of the satellite and in general it does not have access to the external panels. In case of particular needs a Power Management Tile can be removed and an ad-hoc panel can be used to expose the internal instruments to the environment. Furthermore, to house a small camera and optics, a 5 cm diameter hole is provided in the telecommunication tile, so that, also with the smallest configuration, a camera can be fitted easily. In bigger configurations this is not a problem, like in the telescope one (see Figure 6.1d), because some Power Management Tiles can be removed and substituted with ad-hoc panels without reducing much the total solar generated power.

6.7 Environmental constraints

Taking into account what has been said in section 2 and 3, the environmental conditions in which AraMiS is going to operate can be easily computed. AraMiS was designed for an operating life of 5 years in low Earth orbit with an altitude of approximately 700 km - 900 km, even if the system could operate as well in lower orbits. In higher orbits there would be a stronger radiation environment, thus reducing operative life.

According to solar cycle, the expected orbital lifetime can vary from about 60 years at 700 km for the single cube configuration down to about 30 years for the telescope configuration (see Table 6.1 and section 2.2.1 for further details). This means that satellite decay will not be a limiting factor for mission lifetime and probably in the future a de-orbiting solution should be adopted to limit the space debris problem. Solar power is about 1344 $\frac{\mathrm{W}}{\mathrm{m}^2}$ and this is useful to compute actual incoming power to better simulate satellite behavior.

The temperature range the internal components will experience was described in section 2.2.2: components should stand the industrial temperature range, from -20° C to 80° C. This is going to be a complex problem for batteries where the usual operating

Configuration	**Ballistic coefficient -** $\frac{kg}{m^2}$	**Lifetime - years**	
		700 km	**900** km
$1 \times 1 \times 1$	83	59	850
$2 \times 2 \times 2$	166	113	1688
hexagonal telescope	48	29	553
octagonal telescope	59	35	674

Table 6.1: Orbital lifetime for different AraMiS configurations at different altitudes (computed with MSIS-86, see section 2.2.1 for further details).

range is 0° C to 60° C. When the temperature is below 0° C their internal capacity is greatly reduced while the equivalent series resistance will be higher, thus limiting maximum discharge current. Temperature should in no case be higher than 60° C because this can permanently damage them, but this is not likely to happen since satellite internal temperature does not reach that critical value.

The external satellite surface will experience strong interactions with plasma and this requires that the external surfaces are properly protected: first of all, to prevent electrostatic discharges, the external surface should ensure a good surface conductance and this is achieved with Alodine coating. Furthermore all electric conductors should be protected against the atomic oxygen corrosion: a thick insulator or gold plating are the most used solutions for this problem.

The radiation environment has been described in section 2.3, but what really matters for satellite design is the environment inside the satellite, where most of the electronic components are. This can be performed by using radiation transport codes such as MULASSIS [26] so that the number of particles actually reaching the internal part of the satellite can be computed.

The most important thing to consider now is the shielding geometry: starting from AraMiS mechanical structure, as can be seen from Figure 6.5, there are many layers and they can be summarized as follows (from the outside to the inside):

1. solar cells: 160 um thick layer made of Germanium,
2. solar cell PCB: 0.3 mm FR4 layer (modeled as epoxy resin),
3. Aluminum panel: mechanical structure, 1.5 mm layer,
4. gap for thermal insulation, 2 mm thick,
5. internal PCB, FR4, 1.6 mm thick.

Total ionizing dose is mainly related to trapped electrons in the Van Allen Belts and their contribution can be modeled using AE-8: afterwards MULASSIS should be run using the electron spectrum previously computed to get particle fluence inside the mechanical structure. From the previous simulations is was determined that the TID level inside the satellite is equal to 1.7 krad per year during solar maximum (so this is a worst case simulation). Equivalent Aluminum shielding thickness was computed using SHIELDOSE (in SPENVIS [5]) and it is equal to 4.6 mm.

The same procedure can be performed for protons and the major source of protons in LEO is the Sun. By using JPL model, the solar proton fluence was computed and, using again MULASSIS, internal proton spectrum has been computed. Protons induce manly displacement damages: to make damage comparison (between the real case, so with a wide proton spectrum, and the test case, measured with a mono-energetic beam) easier, the equivalent 10 MeV proton flux inside the satellite has been computed and it is equal to 10^9 particles per year.

Using the previous computed equivalent Aluminum thickness, the LET spectrum can also be computed (using SPENVIS [5]) and the result is potted in Figure 6.6. The previous

computed values will be used in the next chapter to select the proper components to use in the different sub-systems.

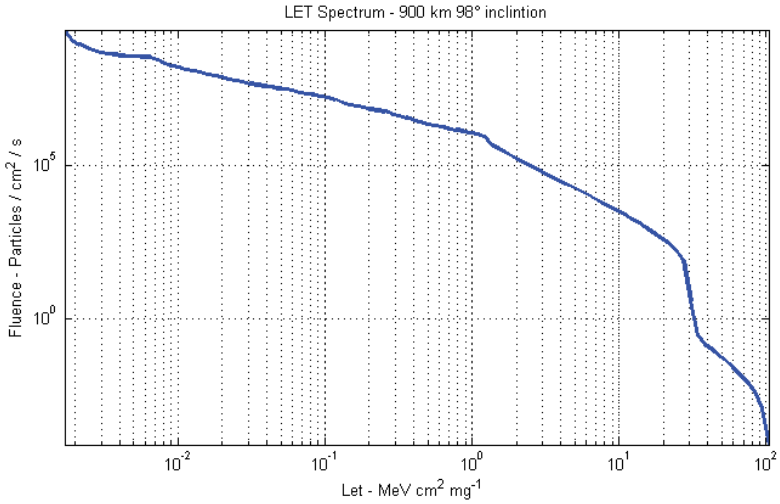
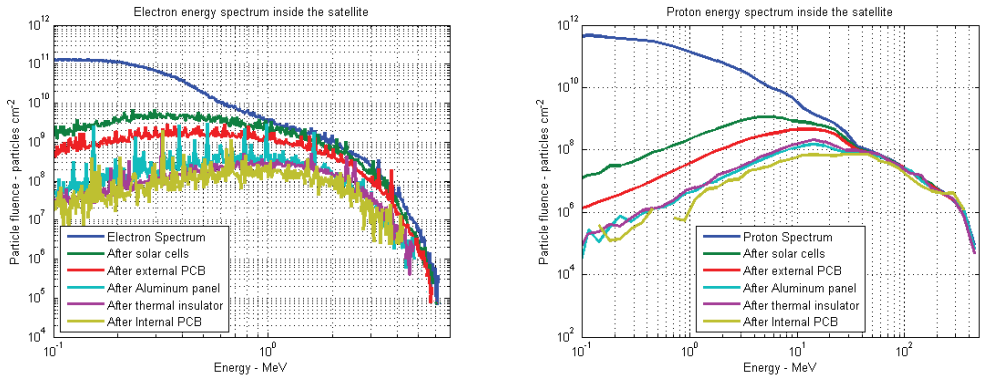


Figure 6.6: LET spectrum (computed using SPENVIS [5]).



(a) (b)

Figure 6.7: AraMiS electrons and protons particles fluxes inside the satellite (computed using MULASSIS with input energy spectra computed using SPENVIS).

Chapter 7

AraMiS sub-systems

In this section a deeper description of some of the AraMiS sub-systems will be presented: different sub-systems were analyzed and their implementation is presented.

The first section of this chapter will analyze a latch-up protection system for commercial devices: many COTS components are in fact prone to latch-up but this problem cannot be solved at device level because it would require the use of ad-hoc developed components, thus making cost increase. This protection system is suited for sensitive device protection, ranging from the single ICs to a whole system. This is key component for ensuring successful COTS device usage in space and is heavily used in all the AraMiS sub-systems.

In the next sections three satellite sub-systems will be then analyzed: a data communication bus, a power management sub-system and a power distribution bus. These three components are part of the AraMiS architecture and they are suited for small satellites ranging from few kilos up to several hundreds kilos.

7.1 A latch-up protection system

This device is intended to protect components or systems against latch-up and in particular to prevent system damages due to high energy particle strikes. The occurrence of a SEL can be discovered by monitoring the device supply current and by noticing a sharp increase in it: the only solution is then to turn power off as fast as possible. Since this system can be employed many times in a satellite built with commercial components, size and power consumption become critical issues. Beside system architecture, also radiation tolerance issues will be analyzed to show that proper design techniques allow to safely use commercial components in space with a good safety margin.

Latch-up is a catastrophic phenomenon which affects CMOS devices caused by high energy particles which trigger the parasitic transistors, mostly in CMOS devices that are not radiation-hardened. This system was developed to prevent latch-up damages on ICs by turning them off before they can become critical: the particle strike is detected by a sudden increase in current consumption which exceeds a programmed threshold. Power supply is then cut off in a short time and it is turned on again after an off time has elapsed

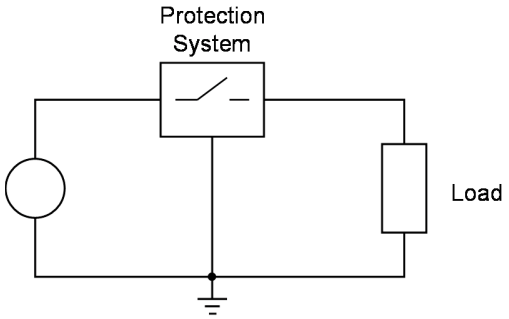


Figure 7.1: Latch-up protection circuit.

to ensure that the device will be correctly operating afterwards.

The main requirement this system had to satisfy were:

- latch-up free operations
- TID tolerance: at least 30 krad
- use of COTS components
- wide operating voltage range: 2.7 V ÷ 36 V
- low power consumption
- switch current: 2 A
- enable signal
- controlled turn-on slew rate
- fast turn-off time: < 100 us
- additional load shunt transistor
- wide temperature range: at least $-20°$ C - $80°$ C
- small footprint: less than $15 \times 15\,\text{mm}^2$

A basic block diagram of the protection circuit is shown in Figure 7.1: given the need of measuring the protected device current, the anti latch-up should be put in series with the load and it should be able to cut the power supply off. It could be mounted low-side or high-side with respect to the load, but in order not to cut the ground connection, it was decided to mount it on the high side. This makes the current sense circuit slightly more complex, but this can simplify its final use, because no modification is made to the ground track.

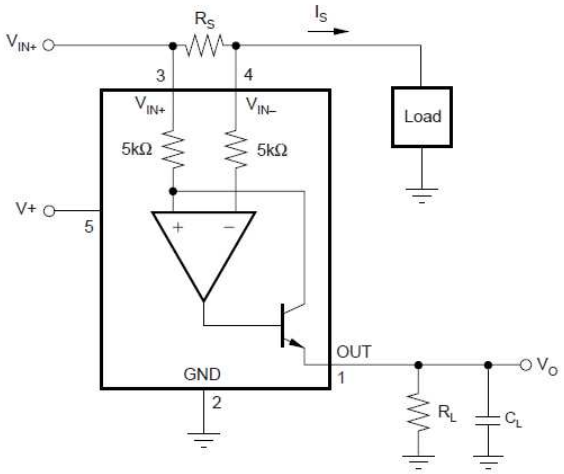


Figure 7.2: Current measurement circuit with the INA138.

7.1.1 Design description

The main building blocks that can be identified from the requirements before mentioned are:

- Current measurement circuit;
- voltage reference, identifying the maximum current value allowed;
- threshold comparator, for identifying if maximum has been passed;
- monostable multivibrator, to keep the load off after latch-up occurred;
- slew-rate control circuit, for achieving a fast turn-off and slow turn-on time;
- load shorting circuit, to fully discharge load capacitance.

Current measurement is performed by sensing the voltage drop across a high side sense resistor, which should be low in value not to waste much power, but measuring a small voltage drop (tens of millivolt) with a common mode voltage one thousand times higher requires a properly designed circuit. A differential amplifier could be use for the purpose, but it requires an extremely high CMRR (higher than 60 dB) which is complex to achieve with discrete resistors due to values mismatch. The solution was to employ an integrated hi-side current sensor, which can guarantee such high performances: component selection is really important since this device should tolerate 30 krad of TID and be latch-up free. Among all the available components on the market, the selection was the INA138 from Texas Instruments, which is manufactured on a completely bipolar process, so latch-up free, and it should ensure also good tolerance to TID.

From Figure 7.2 the output voltage of this current monitor can be computed as follows:

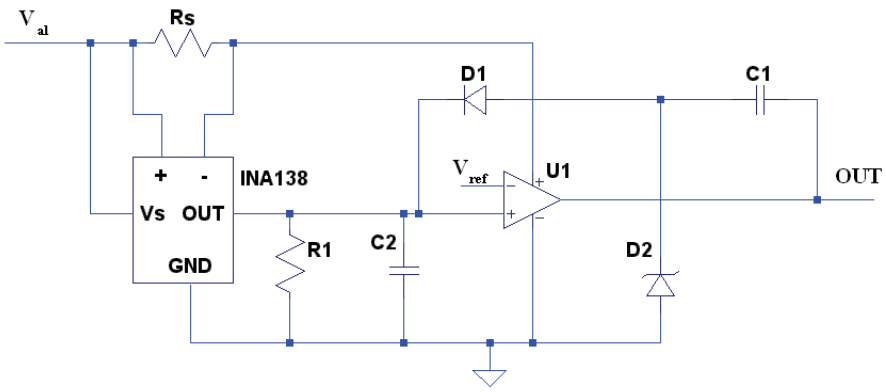


Figure 7.3: Monostable circuit with capacitive feedback.

$$V_O = \frac{I_s\ R_s\ R_L}{5\,\mathrm{k\Omega}} = I_s G \tag{7.1}$$

$$G = \frac{R_s\ R_L}{5\,\mathrm{k\Omega}} \tag{7.2}$$

The voltage reference should be able to operate with a supply voltage ranging from 3 V up to 36 V and it should also ensure low power consumption: this makes the use of series regulators almost impossible because no COTS reference was found with a low voltage output (preferably in the range of 1 – 2 V) and an input of 36 V. The selection was then made among parallel regulators because they were able to stand such wide input range: the selected device was the Analog Devices AD589. It is a 1.2 V precision zener diode temperature compensated for the military range (−20° C – 120° C). This device was also found on the NSREC components database [60] where test results are reported showing TID tolerance up to 30 krad.

A threshold comparator should be used to compare load current with the reference voltage and then detect a latch-up: to reduce losses across the sense resistor, the threshold voltage drop across the sense resistor should have been of 30 mV. From this value, the sense resistor value can be computed and also the gain programming resistor (R_L in Figure 7.2) could have been computed and its value is 200 kΩ.

To guarantee a short intervention time (in case of latch-up) and a long delay time after latch-up a monostable circuit is necessary. To reduce the number of components, the threshold comparator and the monostable multivibrator were implemented using a single operational amplifier: the circuit configuration is a non-inverting threshold comparator which is using a capacitive feedback network. The great advantage the capacitive feedback has is that it can be used, together with few diodes, to make the system change between the two states (output high and low, or load on and off) with two different transition times. In this way a fast turn-off time can be achieved together with a delayed turn-on: this will make the protection circuit react quickly to a latch-up, thus turning the load

off, and wait before turning the load on again that all the capacitance have discharged properly. The schematic is depicted in Figure 7.3, showing also the connection with the current sensor circuit.

$$\alpha = \frac{C_2}{C_1 + C_2} \tag{7.3}$$

$$C_2 = \frac{R_s\ Q_{max}}{5k\ \Omega V_{ref}} = \frac{G\ Q_{max}}{R_1\ V_{ref}} \tag{7.4}$$

$$\alpha = \frac{R_1\ C_2}{T_{off}} \ln\left(1 + \frac{V_{CC}}{V_{ref}}\right) \tag{7.5}$$

Given a target off-time after latch-up (T_{off}) and a maximum inrush charge that should flow in the load before triggering a turn-off (Q_{max}) and the reference voltage (V_{ref} equal to 1.2 V), the value of the two capacitors can be computed.

The C_1 capacitor acts also filtering spurious current peaks present on the input or single event effects on the amplifier output (since the output impedance is in the order of $10^5\ \Omega$).

The operational amplifier selected for this function is the National Semiconductors LM4250 because it was the one giving the lowest power consumption over the full supply range: this can be achieved thanks to the external bias current control, that allows, by making the device a bit slower to decrease power consumption. This device has a limited common mode input voltage range (from 0.6 V to V_{cc} − 0.6 V) but it is not a problem since the threshold voltage has been selected to be 1.2 V. Furthermore, the device does not have a rail-to-rail output, thus requiring external components to properly control the power switch. Since there is only the need for a fast turn-off time, two separate drive circuits were used: turn-off is accomplished by a transistor shorting the gate-source terminals of

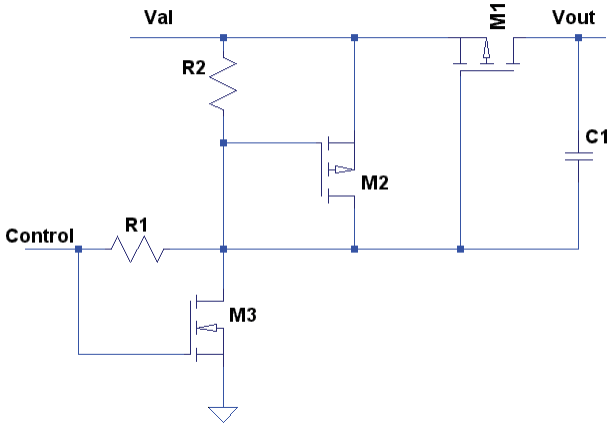


Figure 7.4: Slew-rate control circuit.

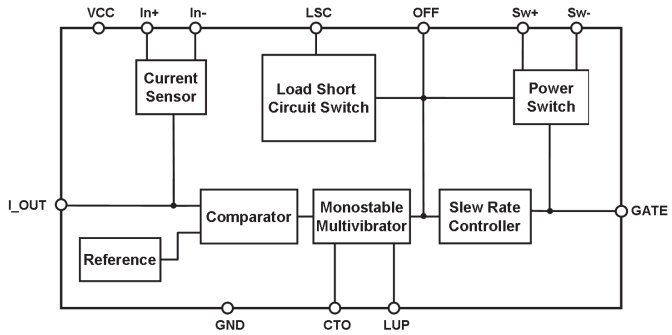


Figure 7.5: Latch-up protection circuit internal block diagram.

the power switch, while a high value resistor is used to control device turn-on time (see Figure 7.4).

The M_1 transistor is the power MOS used to turn the load on and off: this transistor is turned on by the R_1 resistor, while it is turned off by M_2 and M_3. By employing a high value resistor, turn on can be quite slow (up to some millisecond) while turn-off if faster (about 60 us). The capacitor C_1 is used to slow down turn-on time (C_{sr} in Figure 7.7, 7.8, 7.9 or 7.10) and it should be mounted externally. Since it was not possible to select a single power transistor to operate in the full voltage range range (power mosfets with threshold voltage below 3 V have a maximum drain - source voltage of 20 V), it was selected to divide this range in two parts: from 2.7 V to 18 V the internal transistor is used, while for higher voltages an external transistor will be required. The control pin for the transistor can be also used to insert a higher capacitor to further reduce turn-on slew-rate. The actual capacitance value gets multiplied by Miller effect since this capacitor is connected between the circuit output and the transistor input (gate terminal). The derivative of output voltage across time (which is actually the slew-rate) can be computed as follows:

$$\frac{dV_{OUT}}{dt} = \frac{1}{C_1}\frac{V_{al}}{R_1} \tag{7.6}$$

All the internal components that make the latch-up protection circuit are shown in Figure 7.5. A basic circuit using the latch-up protection system is depicted in Figure 7.6 where only one external component is needed: the sense resistor. This allows to adapt the device to a broad range of applications with only one external device. The latch-up threshold current can be programmed by means of R_s according to the equation:

$$R_s = \frac{V_{th}}{I} \tag{7.7}$$

$$P_{R_s} = \frac{V_{th}^2}{R_s} \tag{7.8}$$

where V_{th} has a typical value of 30 mV [73] and I is the threshold current. This resistor should also stand the power dissipation due to Joule effect.

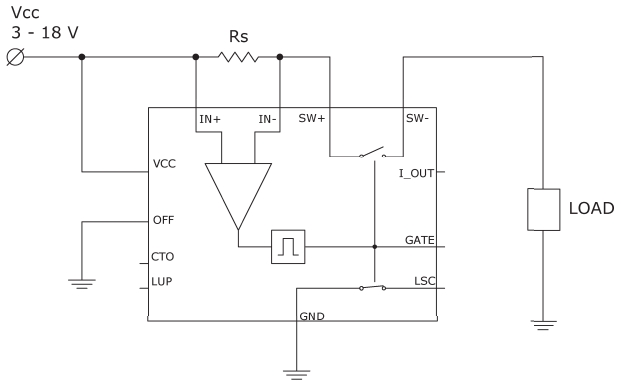


Figure 7.6: Basic latch-up protection circuit.

7.1.2 Advanced circuit configurations

Other circuit configurations can be used to take advantage of the flexibility of this system, like in Figure 7.7, 7.8, 7.9 or 7.10.

When the end-user needs to increase turn-off time, this can be accomplished by adding an external capacitor (C_{tc} in Figure 7.7) computed in this way:

$$C_{tc} = \frac{T_{OFF}}{200\,\mathrm{k\Omega}} - 100\,\mathrm{nF} \tag{7.9}$$

where T_{OFF} is the desired recovery time. Turn-on rise time can be increased by adding the C_{sr} capacitor to the circuit:

$$C_{sr} = \frac{T_{rise} - 165\,\mathrm{us}}{33\,\mathrm{k\Omega}} \tag{7.10}$$

where T_{rise} is the selected rise time. For particular needs, the threshold voltage across the sense resistor can also be increase from the default value of 30 mV by using a resistor connected to the I_{OUT} pin. The equivalent resistance (the external and the internal resistance combined) when using a resistor R_{th} is:

$$R_G = \frac{R_{th}\ R_{out}}{R_{th} + R_{out}} \tag{7.11}$$

where R_{out} is the internal gain programming resistance (200 kΩ). This equivalent resistance modifies threshold voltage according to the following equation:

$$V_{th} = \frac{V_{ref}\ 5k\Omega}{R_G} \tag{7.12}$$

with V_{ref} equal to 1.2 V. The use of the C_{ir} capacitor helps low-pas filtering input spikes, thus achieving a cut frequency given by:

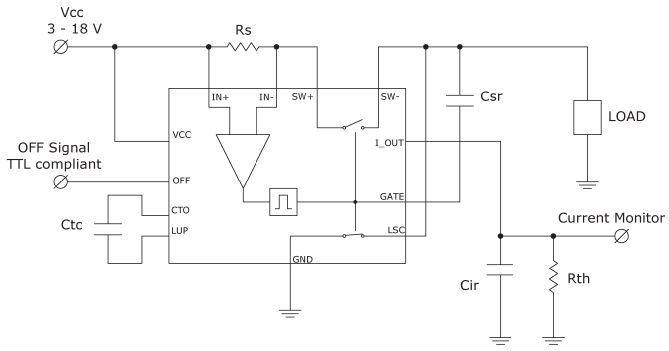


Figure 7.7: Advanced latch-up protection circuit configuration, with slew-rate, off-time and inrush current control capacitor and threshold voltage control resistor.

$$f_T = \frac{1}{2\pi\ (C_{IR} + 100\,\text{pF})\ R_G} \tag{7.13}$$

Given the added delay due to the increased capacitance, an inrush current is tolerated before triggering a turn-off: this inrush current, multiplied by the reaction time can give a charge that is actually allowed to flow into the protected device before a turn-off

$$Q = \frac{2\pi\ (C_{IR} + 100\,\text{pF})\ 3\,\text{kV}\Omega}{R_S} \tag{7.14}$$

where R_S is the sense resistor.

From the I_{OUT} pin, the load current can be measured during normal operations (it is important to not that the current value should not be measured when the power switch is turned off because the value will not be reliable):

$$V_{I_{OUT}} = \frac{I\ R_s\ 5k\Omega}{R_G} \tag{7.15}$$

where R_G has a nominal value of 200 kΩ or the value specified by equation 7.11 when an external resistor is used to program the gain.

The OFF signal can be used to turn the load off: this signal is TTL compliant and can make this system act like a load switch with current protection.

The 18 V limit of the basic configuration circuit is due to the maximum voltage across the power transistor: transistors rated 30 V - 60 V have in general a high threshold voltage (around 3 V) which makes them not suited for operations down to 3 V. Thus, a low voltage transistor (maximum 20 V) with a low threshold voltage (1 V) has been used for low voltage operations and an external one is needed in the other case. The external transistor can be controlled by the GATE pin. To proper control the slew-rate during turn-on C_{sr} should be computed as follows:

$$C_{sr} = \frac{T_{rise}}{33\,\text{k}\Omega} - C_G \tag{7.16}$$

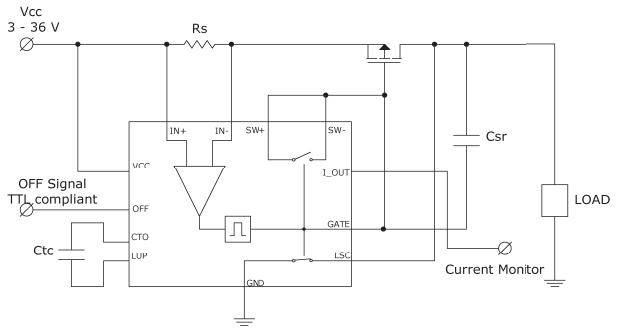


Figure 7.8: Advanced latch-up protection circuit configuration, with an external power transistor to extend operating range up to 36 V.

where C_G is the external transistor gate - source capacitance.

The configuration shown in Figure 7.9 makes use of two latch-up protection systems wired in parallel with two redundant series transistors to achieve fault tolerance: a fault in only one of the two devices cannot prevent the load to be powered, no matter which fault it is. In case of damages to one of the transistors (either in short or open circuit) the other branch can ensure proper operations, while a fault in one of the protection systems will stop only one of the two branches.

In Figure 7.10 a different application for this protection system is also shown: the use of the external diode allows the system to turn off the load after a latch-up occurred but will not make the load be turned automatically on afterwards. To turn it on again an external transistor is needed to reset the internal monostable.

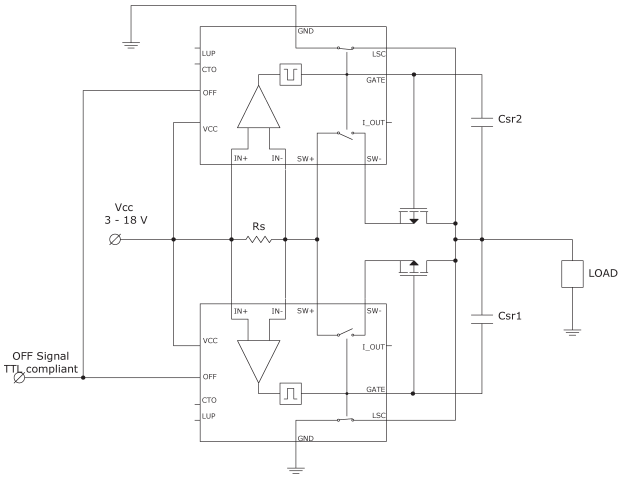


Figure 7.9: Advanced latch-up protection circuit configuration with external and redundant power switch.

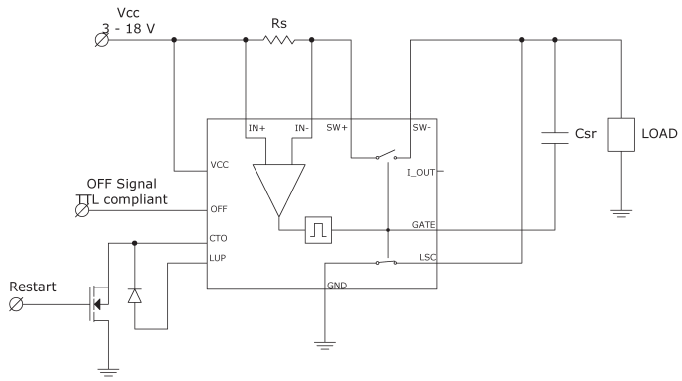


Figure 7.10: Advanced latch-up protection circuit configuration.

7.1.3 System implementation

The system was first built on a small PCB to debug and test all the components, then to reduce device size and increase commutation speed by reducing parasitic capacitances, it has been build as an hybrid circuit by Neohm Componenti (see section A) and it can be seen in Figure 7.11 and 7.12. Most of the devices were used as bare dies and where connected to the circuit by means of bonding wires. The whole system was then packaged in a metallic case mainly for reliability reasons because it has to stand multiple reflow soldering processes (with temperatures as high as 220° C) and a metallic package will suffer lower thermal stresses than plastic ones. The device will have to stand four soldering processes: the first to solder the SMD components onto the device, the second for soldering the metallic cover, then the third time the device will be soldered onto the final PCB and, in case of problems, a fourth rework process can take place. Plastic packages can stand all these thermal cycles, but a proper selection of the best resin to use would have taken much more time, thus making the device also more expensive. Once the

Figure 7.11: Latch-up protection circuit, external view (courtesy of Neohm Componenti [73]).

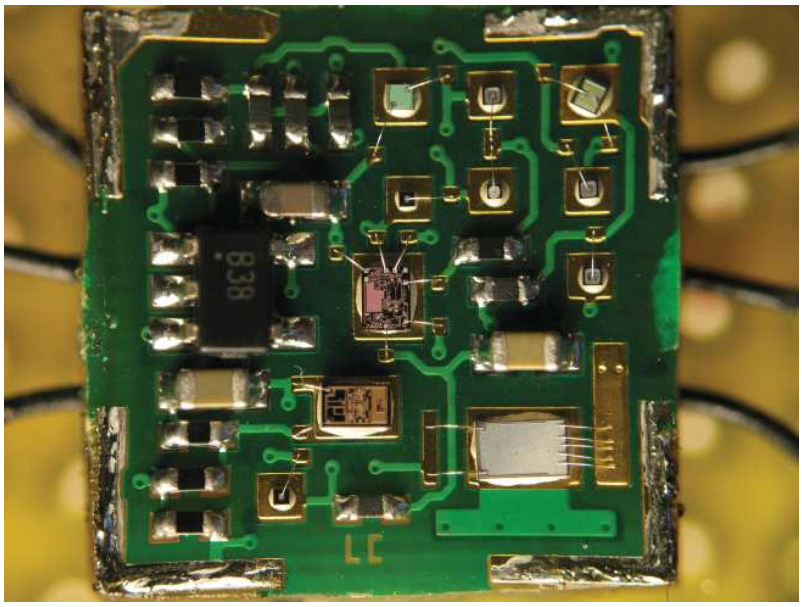


Figure 7.12: Latch-up protection circuit, internal view (courtesy of Neohm Componenti [73]).

device has been manufactured, tests were made to verify the proper device behavior, as shown in Figure 7.13: the latch-up protection main features as can be seen from the plots, like the fast turn-off time or the controlled turn-on time.

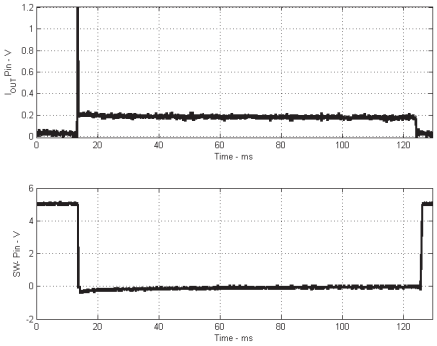
7.1.4 Radiation effects evaluation

Three main problems should be addressed to ensure the device radiation tolerance: total ionizing dose, latch-up and single event transient. The TID problem can be solved with a proper component selection and this was achieved by choosing bipolar linear devices, that show a good tolerance and by also selecting as many devices as possible from radiation effects databases.

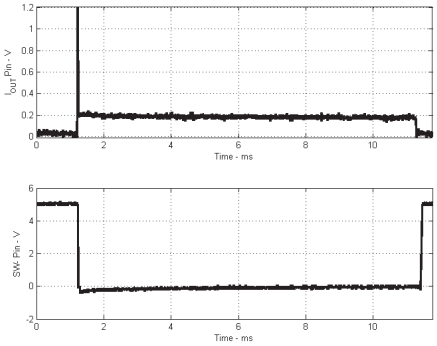
The latch-up problem was solved again using bipolar devices, that are immune to latch-up for construction, since they require an extremely high energy to trigger this event.

The last problem is the most complex to address since all components can suffer single event transient issues that can in this case trigger false latch-up detections. This problem is hard to address because there is no data available for the selected components and testing is expensive. Given the fact that the device anyway will need to be tested under a particle beam for simulate single event effects, a theoretical evaluation of the problem has been also performed.

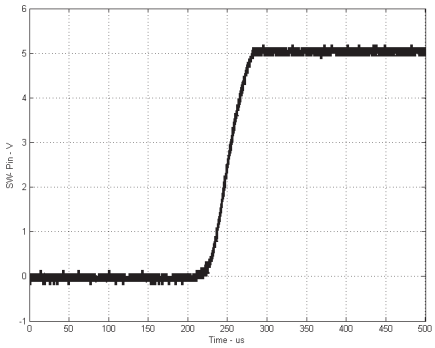
The first analyzed device is the INA138, manufactured on a traditional bipolar process (this information is not written on the device datasheet, but was provided by the Texas Instrument technical support team and, even if not completely reliable, it can help a bit



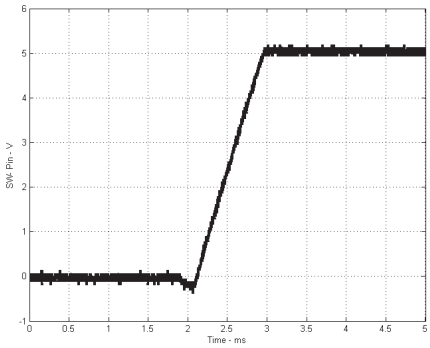
(a) Latch-up event with $V_{cc} = 5\,\mathrm{V}$.



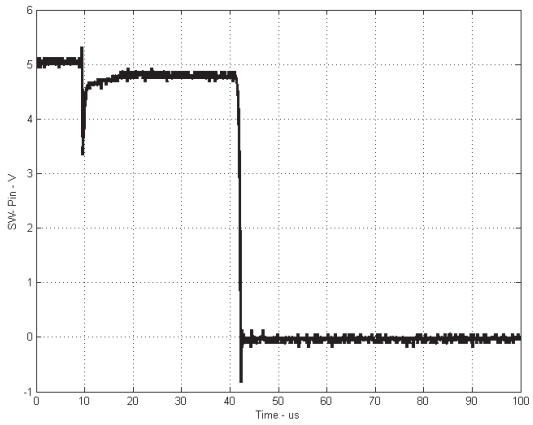
(b) Latch-up event with $V_{cc} = 5\,\mathrm{V}$ and $C_{tc} = 1\,\mathrm{uF}$.



(c) Turn-on time with $V_{cc} = 5\,\mathrm{V}$.

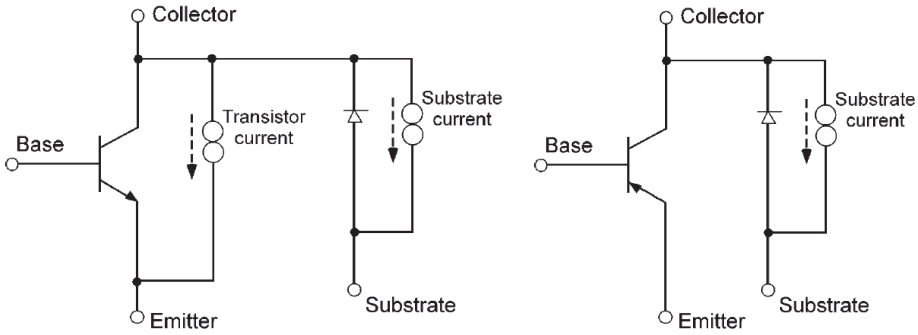


(d) Turn-on time with $V_{cc} = 5\,\mathrm{V}$ and $C_{sr} = 2.7\,\mathrm{nF}$.



(e) Turn-off time with $V_{cc} = 5\,\mathrm{V}$.

Figure 7.13: Latch-up protection circuit typical characteristics at 25 °C [73]).



(a) Vertical NPN (b) Lateral PNP

Figure 7.14: Bipolar transistors single event transient models [74].

to find a solution) using vertical NPN and lateral PNP transistors. The layout of these devices can be modeled (as can be seen in [74]) together with particle interactions in order to understand which effects can take place. The NPN transistor can be represented by two current sources (as shown in Figure 7.14) that represent the particle injected charge when the emitter / base or the collector / substrate junction is hit. Th PNP transistor is simpler and can be represented by only one current source modeling a charge injection in the collector / substrate junction. These two current sources should inject a known charge for a limited amount of time, like a port function, where the time length can vary between 100 ns and 100 us. These two are considered the average time limits of the transient signal generated by an ionizing particle according to [75][76]: since a unique value cannot be identified, different simulations were run to find the worst case in the previous range.

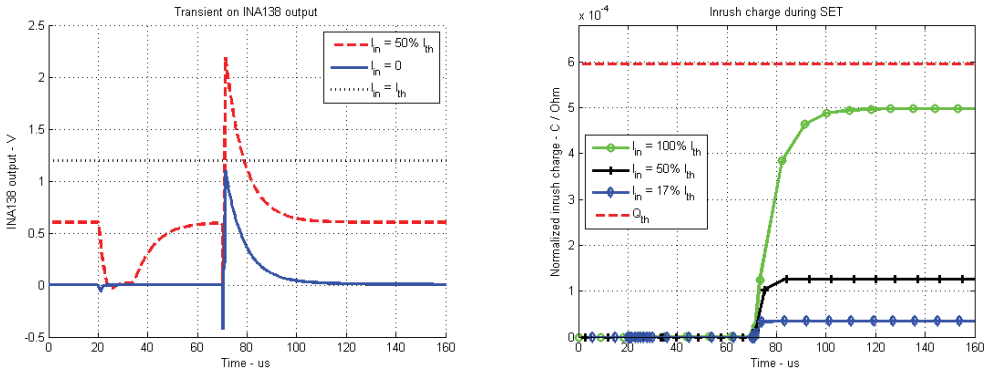
The peak current value is linked to total injected energy which then depends on particle LET: a theoretical calculation of the energy can be quite complex since there is no precise knowledge of transistor layout. It was decided then to take an empirical approach to link particle LET to injected energy that takes advantage of many experiments that were carried out to find a correspondence between heavy ion effects and picoseconds laser pulses. According to [77], an experimental equation was calculated to convert LET (in $MeV\ cm^2\ /\ mg$) to energy (in pJ) for laser pulses with a wavelength of 605 nm:

$$E\ =\ 0.328\ LET \pm\ 0.326 \qquad [pJ] \tag{7.17}$$

Making the hypothesis that every photon generated by the laser is able to generate a hole-electron pair in the semiconductor (given that the photon energy is 2.04 eV, which is higher than Silicon energy gap), the total injected charge is equal to:

$$Q\ = \frac{E\ \lambda}{h\ c}\ =\ 0.4893\ E \qquad [C] \tag{7.18}$$

where E is pulse energy, λ is laser wavelength, h is Plank constant ($4.135\ 10^{-15}\ eV\ s$) and c the speed of light in vacuum ($2.997\ 10^8\ \frac{m}{s}$). This result can be compared to some literature results, like [78], where 12 MeV Nickel ions where used to inject charge in a



(a) (b)

Figure 7.15: Single event transient on INA138: in (a) the output of the IC during a SET generated by a particle with 120 MeV cm^2 / mg LET. (b) shows the inrush charge normalized to sense resistor during a SET.

Silicon Carbide diode: ions LET (computed with SRIM) is equal to 18.58 MeV cm^2 / mg which gives an injected charge of 2.98 pC which is about ten times higher than the experimental measured value (280 fC) demonstrating the effectiveness of this technique to find an approximate upper bound to charge injection. From laser pulse charge, current pulse current can be computed this way:

$$I = \frac{Q}{T} \qquad [A] \tag{7.19}$$

where Q is pulse charge and T is pulse width. The laser beamwidth used was 1.2 um that is quite smaller than transistor size (for the INA138 it was estimated in approximately 12 um^2): pulse energy is thus completely transferred to the transistor (this is again a worst case approximation used to compute a sensitivity upper bound).

When pulse amplitude has been computed, a SPICE simulation can be performed using the modified device model to include the previous mentioned current sources: in this way device behavior can be evaluated to estimate single event effects.

These effects were evaluated for the INA138, identifying the critical nodes that can trigger an output pulse, as can be seen in Figure 7.15: the voltage pulse generated on the INA138 output can be seen as a current pulse on the load that can trigger a latch-up in case the pulse amplitude is able to exceeds the threshold. Even if the peak pulse value is above threshold, an inrush charge, as was computed in equation 7.14 should anyway flow before triggering. Expressing the INA138 output as V(t), sensitivity to SET can be computed as follows:

$$V_{ref}\ R\ C = \int V(t)dt \tag{7.20}$$

where V_{ref} is the internal reference voltage, R is the resistor used to program the INA138 gain, C is the capacitor in parallel to R, used to filter the inrush current. Using

the previous equation the filtering capacitor C can be computed in order to filter out SET pulses. Figure 7.15b was plotted assuming R = 200 kΩ and C = 62 pF (40 pF INA138 parasitic capacitance plus a 22 pF capacitor) and assuming the worst case SET pulse generated by a particle having a LET of 120 MeV cm^2 / mg; these particles have a probability of hitting the device approximately once every 10^{-10} seconds.

The same procedure can be applied to the LM4250: from simulation it can be seen that pulse duration is in the order of 30 us, which is approximately one half of turn-off time. This means that such a short pulse will not be seen by the power switch gate, thus not triggering false latch-ups.

From what can be seen in Figure 7.15, it is clear that SET are more likely to generate false latch-up detections when the load current is near to the threshold value because most of the injected charge pulse will be seen as a inrush charge. When the load current is instead quite lower than the threshold value, only the part of the charge pulse actually passing the threshold value will be considered as inrush charge: it is clear that, to better tolerate SET a threshold value at least twice as big as the normal current consumption is needed. This can be see also from Figure 5.9, where the device under test shown normal supply current on 16 mA and after the first partial latch-up, the current consumption rises up to 66 mA. A proper threshold current value for this circuit should be around 30 mA – 40 mA thus allowing a good safety margin for false latch-up detections and allows also to clearly distinguish SEL because current consumption is then quite higher that the threshold. Current threshold should also be selected while taking into account the protected device degradation under radiation: it is quite normal that the supply current rises with higher doses but this increase should not reduce the latch-up detection margin of the protection circuit.

Qualification of the latch-up protection system could not be completed due to the unavailability of radiation sources, but this process will be anyway completed before actual device use in space.

7.2 Optical wireless communication bus

A wireless communication bus [79][57] has been developed to address the problem of harness inside the satellite: in small satellites usually there is not much free space for cabling and their integration is also quite complex; a wireless solution can become interesting also from the mass point of view. The development of such a system poses anyway many issues, in particular for radiation tolerance: different techniques were employed to achieve this goal, from the radiation environment evaluation, to component selection.

This section focuses on a crucial field of application, the on-board data communication sub-system and in particular on the application of the before mentioned techniques to a practical case. Many different solutions were used to deal with the different issues that arised during the development.

The main requirement this system had to satisfy were:

- fault tolerance in communication

- single point of failure free
- TID tolerance: at least 30 krad
- use of COTS components
- 1 Mbps data rate
- extended industrial temperature range (-20° C – 80° C)
- low power consumption
- low mass increase due to harness

The first issue that had to be solved was the selection of the bus topology and communication media: fault tolerance, system mass and complexity were the most important parameters that were used to select the proper implementation. Several fault-tolerant bus systems have been developed, realized and used in space [80] and usual approaches consist in serial redundant links shared by multiple users, with point-to-point or bus (multi-point) topology. Point-to-point communication structures are very robust: a single fault or a node which starts generating data on the communication medium does not cause problems to the rest of the system [81], but it results in high costs, volume and weight due to the number of wires. Buses are cheaper, but also more sensitive to faults.

Wireless communication systems have still not been widely used in space, but they can bring several advantages to on-board communication. First of all, it is quite simple to connect modules, since no wiring is needed, then it is simple to create a broadcast channel among modules. Commercial devices can also reduce overall power consumption making their usage feasible on battery-powered system.

Given the aim of using low-cost components, it was also selected to employ commercial communication standards since they allow to re-use well known and documented solutions, speeding-up the development process. Low-power and low-cost wireless communication links are mainly IrDA, Blue-Tooth and Zig-Bee. To avoid EMI problems inside the satellite, radio based connections were discarded because most of these standards are operating at 2.4 GHz or 430 MHz and these frequencies should have been avoided for ground communications. Unfortunately both bands have a small fraction of the spectrum allocated for satellite operations, so excluding two of the most used bands was not an option and these standards had to be rejected. As an alternative, these standards could have been implemented on different frequency band, but then everything would have needed to be re-implemented from scratch, eliminating the advantages of commercial protocols.

The final selection was the IrDA standard which provides also low power consumption and acceptable data rates (in the range 1.2 kbps up to 16 Mbps). In comparison with other wireless solutions (e.g. RF based), the optical channel is fully insensitive to EMI: no electric noise can be radiated or coupled since there is no long wire connecting all the devices.

The benefits of IrDA vs wired bus communication are both in the intrinsic galvanic isolation and in the lack of wiring. Furthermore, the use of a single communication link can

lead to a single point of failure, thus a redundant communication bus should be employed. The IrDA has the great advantage that frequency multiplexing can be performed quite easily: the use of two different light wavelength can create two completely independent channels that do not influence with each other. As will be shown in section 7.2.4, the optical emitters and transducers are believed to be the most critical components from the radiation tolerance point of view and using two different devices for the two channels helps also removing common mode faults (which is what happens when two replicas of the same sub-system stop operating for the same reason).

7.2.1 Infrared Data Association

The Infrared Data Association (IrDA [82]) is a standard defining the specifications for a short-range communication protocol to exchange data over an infrared channel between small devices. The communication medium is free space and his technique is used for connecting personal devices such as mobile phones, PDA and laptops.

The IrDA Infrared Physical Layer Specification is the lowest layer of the specifications. The most important specifications are:

- Range (Standard: 1 m),
- Infrared ($[800 - 900]$ nm),
- Speed (2.4 kbps to 16 Mbps),
- Point-to-Point protocol,
- BER (lower than 10^{-8}),
- Communication mode (half-duplex),
- Modulation (Base band, no carrier),
- Angle (minimum cone $\pm 15°$),
- Maximum Irradiation Intensity (500 mW/sr).

Figure 7.16: A commercial IrDA-USB transceiver.

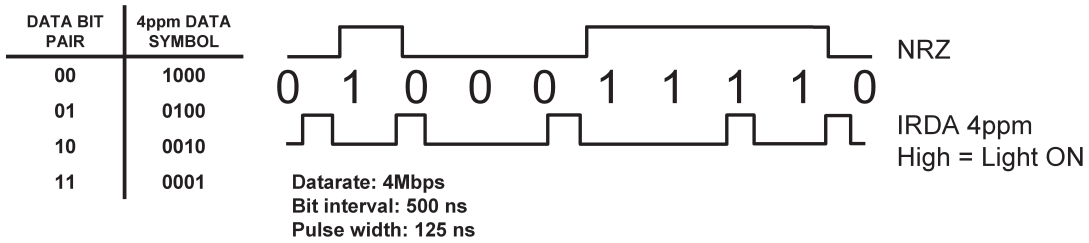


Figure 7.17: 4PPM modulation scheme.

Standard range is 1 m in direct line of sight but many devices are able to communicate up to 10 m (like TV remote controls) and the maximum radiation constraint limits peak power for eye safety reasons and for preventing receiver stage saturation. Communication mode is half-duplex because during transmission the receiver is blinded by the light of its own transmitter.

Standard transmission rates fall into three main categories: SIR, MIR, and FIR. Serial Infrared (SIR) data rates are normally used when interfacing with an RS-232 device (2.4 kbps, 4.8 kbps, 9.6 kbps, 19.2 kbps, 38.4 kbps, 57.6 kbps, 115.2 kbps). Medium Infrared (MIR) refers to rates of 0.576 Mbps and 1.152 Mbps while with Fast Infrared (FIR) a data rate of 4 Mbps is denoted. There are also two high-speed standard, the Very Fast Infrared (VFIR) ranging 16 Mbps and Ultra Fast Infrared (UFIR) ranging 100 Mbps, but the latter is still under development.

For data rates up to and including 1.152 Mbps, Return to Zero Inverted (RZI) modulation scheme is used, and a "0" is represented by a light pulse. For rates up to and including 115.2 kbps, the optical pulse duration is nominally 3/16 of a bit duration. For 0.576 Mbps and 1.152 Mbps, the optical pulse duration is nominally 1/4 of a bit duration.

At 4 Mbps the modulation scheme used is 4PPM where a pair of bits is taken together and called a data symbol: the transmission time for that symbol is then divided into four parts and only one optical pulse is transmitted in one of those parts (see Figure 7.17). The nominal pulse duration (chip duration) is 125 ns.

For 16 Mbps transmission, the HHH(1, 13) [83] code (a low duty cycle, rate 2/3, (d, k) = (1, 13) run-length limited code) is used as the modulation code to achieve the specified data rate. The HHH(1, 13) code guarantees for at least one empty chip and at most 13 empty chips between chips containing pulses in the transmitted IR signal. The 16 Mbps rate packet frame structure is based on the current IrDA-FIR (4 Mbps) frame format with modifications introduced where necessary to accommodate the requirements that are specific to the new modulation code. Furthermore, the HHH(1, 13) code is enhanced with a simple scrambling / de-scrambling scheme to further optimize the duty cycle.

7.2.2 Commercial Devices

There are many commercially available integrated solutions for creating an IrDA system but they should be analyzed to verify the possibility of their usage in the space environment. Commercial chips can rely onto two main categories: Integrated Transceiver Controllers (external optical devices) and Integrated Transceiver Modules, which includes also the LED and the photo-diode.

The easiest solution for creating a transceiver is to use an integrated module with LED and photo-diode: these modules are sold by many companies [84][85] and are also available in extended temperature range. These devices can be controlled directly with a microcontroller UART (Universal Asynchronous Receiver and Transmitter), greatly reducing the complexity of the circuit (see Figure 7.18). Power consumption is quite low (usually few milliWatts) and devices are available with a data rate up to 16 Mbps.

Most of these devices use CMOS technology, mainly for reducing the power consumption in portable applications: latch-up could then be a problem in space, so they should be protected. Moreover their behavior under radiation should be carefully evaluated also for total dose induced problems. The most critical components usually in this case are LED, photo-diode and the epoxy resin used to cover them. Another drawback of these components is the lack of flexibility: since modulation and demodulation circuits are integrated in the same package the only free parameter is usually the transmitting power, while channel coding or carrier frequency (the optical link wavelength) could not be selected.

An integrated transceiver controller can be used to solve part of the previous problems: it should be used in fact with external photo-diode and LED, allowing the designer to select the desired carrier frequency and also select the proper components able to withstand the mission radiation level. Power consumption is also very low since the integration of most of the components helps in reducing parasitic parameters. Devices are available with a data rate up to 512 kbps.

Figure 7.18: An integrated IrDA transceiver from Vishay [86].

These devices are again not developed for the space environment: most of them are latch-up prone (since based on CMOS technology), and could also show problems related to total ionizing dose.

To sum up, the use of commercial devices provides many advantages: first of all the development time and cost are by far lower because most of the devices described before require only a single chip with few external components. Size could be also reduced due to the low number of devices. Test time and complexity indeed is increased by the use of completely integrated solutions: different type of circuits are present on the same device and so multiple tests should be performed and if even a single test fails the device should not be used. On the other hand, the "discrete" (the development of the full receiver circuit) solution allows maximum flexibility: we can select each component, for example the LED and photo-diode to change carrier frequency instead of being forced to change the whole transceiver because of embedded optical devices. Given the above pro and cons we decided to go for a discrete solution.

7.2.3 Implementation

The implementation of the IrDA transceiver makes no use of transceiver ICs since this solution gives no flexibility: it was decided to develop the photo-diode amplifier and threshold comparator for the receiver and the transmitter driver. Using discrete LED and photo-diodes gives more freedom in wavelength selection and diode technology (fixed in integrated transceiver modules) to find the best from the radiation tolerance point of view (this issue will be further analyzed in section 7.2.4).

The IrDA standard specifies a single master bus and it was mainly intended for communication between a controller device (for example a computer) and a slave (for example a printer or a mobile phone): it is a point to point connection, even if in the higher software layers, point-multipoint connections could be supported too.

A single master bus has many advantages over a multi-master one, first of all complexity, because there is only one node talking at once and the master is ensuring no collision (removing also the collision detection mechanism on slaves). This makes the whole structure simpler but creates a single point of failure, stopping the whole bus operations in case of master failure. This problem is solved using a second, redundant, bus that can be still operating when the first one gets damaged.

If the slaves have to request attention to the master, they have to wait until the master polls them: this can lead to high latency in case slaves have to signal problems because the master has to keep polling all the devices. Neither any interrupt nor "attention request" mechanism from slaves is included in the IrDA standard, therefore a second low speed channel was used to signal problems or high priority error conditions. This channel has a low data rate ($\approx$1 kbps) and can be used just to raise an attention requests to the master (in theory it can be used also to transfer a small amount of data in case the high speed one is faulty). The two channels use widely different bit rates and therefore they can be separated by frequency domain filters (can be seen as a sort of frequency multiplexing). The carrier frequencies of the two channels are separated in frequency by three decades, which allows to achieve around 120 dB attenuation with a simple 2-stages

RC filter. The cost of the secondary channel is low, since it requires - besides the filters - only one additional operational amplifier in the receiver and one LED driver transistor in the transmitter because the optical devices (LED and photo-diode) are shared.

The system architecture is shown in Figure 7.19: the main channel is capable of MIR data rates (high speed, up to 1.152 Mbps) and operates as a shared bus (point-multipoint link) for data transfer between the different boards in the satellite. The two channels solutions allows to still be compliant with the IrDA standard and save the complexity needed for a multi master bus but increasing performances with the new interrupt mechanism.

The transmitter is made by the LED and by two drivers for the two channels, connected in parallel: this solution allows to better control supply current frequency spectrum to separate the two channels. The low speed channel uses a slew-rate limited driver not to generate high frequency components while the other one employs an high speed control circuit generating short pulses with a maximum frequency of 1 MHz and a duty-cycle of 10%.

For the receiver, different photo-diodes have been evaluated and their conversion spectrum has been compared to the solar emitted spectrum (see Figure 7.20) to evaluate if it is going to interfere with the received signal. Since this communication system should be used inside a satellite, there is the concrete possibility that sunlight can enter from some small holes thus blinding the receiver. The influence the Sun can have on the receiver stage can be seen from Table 7.1: a strong interfering component can be clearly seen with an amplitude at least 50 dB higher than the signal will be received by the photo-diode. The sunlight component has been considered constant since no reference has been found

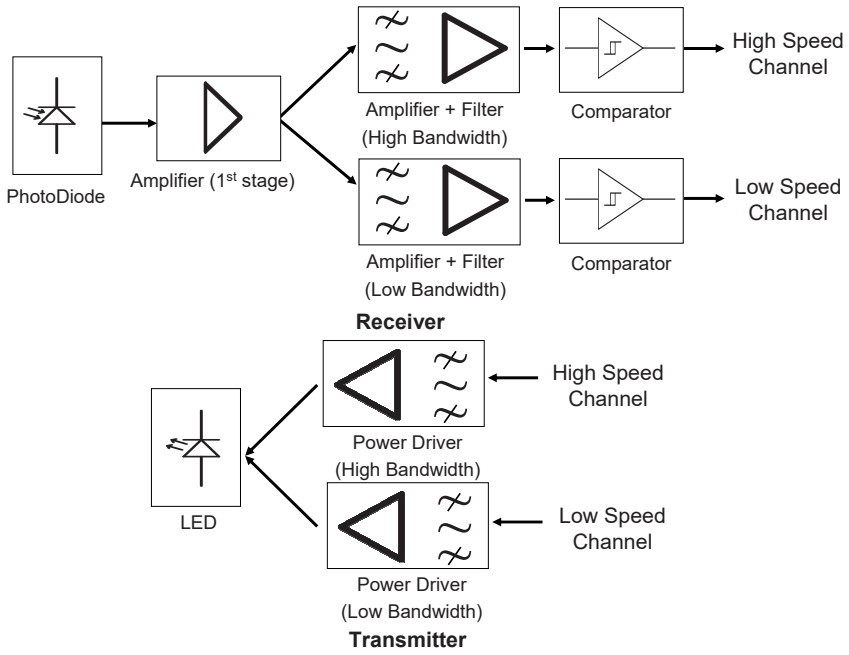


Figure 7.19: Block diagram of the Transceiver.

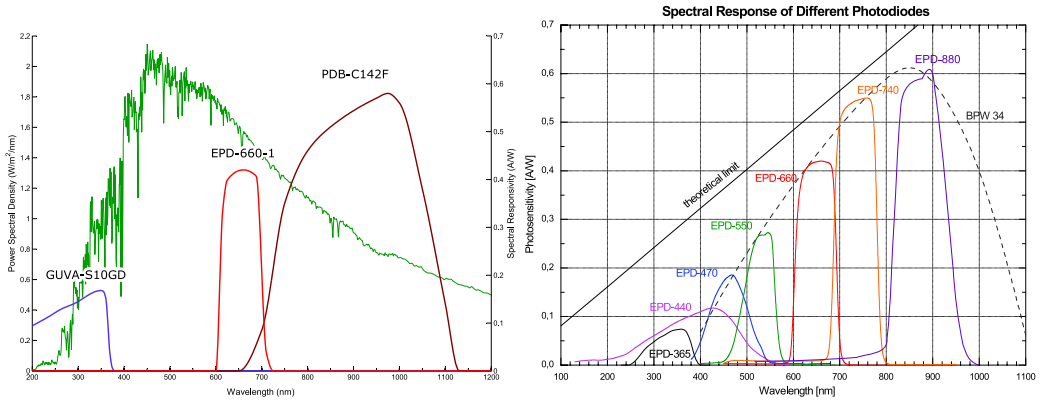
dealing with light noise with a frequency higher than 1 kHz and so two solutions have been employed to avoid receiver blinding: selecting a photo-diode with an operating band far from Sun peak power emission and employing a common mode rejection circuit to attenuate constant illumination.

Photo-diode	**Signal current** - μA	**Solar generated current** - μA	**Ratio** dB
PDB-C142F	0.92	680.00	−57
EPD-660-1	0.18	54.00	−50
GUVA-S10GD	$0.54 \cdot 10^{-6}$	0.84	−63

Table 7.1: Signal and solar current generated by three different photo-diodes (see Figure 7.20 for diodes spectrum) [79].

The photo-diode has been used in the circuit with a reverse bias (photo-conducting mode) because this can reduce diode capacitance an increase generated photo-current: the reduced parasitic capacitance allows an increase in communication speed by increasing the receiver gain at higher frequencies.

The receiver is a two stages amplifier with the first one being a transconductance amplifier with a lower gain to reduce noise pick-up while two different second stages are present for the two channels. The power supply for the channels is separated such that in case only one channel is needed, the other can be switched off. A threshold comparator is used to generate the digital output stream: since signal power can widely vary (due to transmitter distance or components degradation a fixed threshold would be hard to determine. Thus a variable threshold has been employed which is computed from the signal AC peak-to-peak amplitude, making the receiver less sensitive to these variations. The output digital signal is then fed into an MSP430 micro controller that performs the



(a) (b)

Figure 7.20: Different diodes spectral response [87] (b) compared with solar spectrum [79] (a).

decoding (from RZI decoding to protocol management).

To prevent software faults leaving the transmitter on for a long period, an analog bus guardian (used as a watch-dog for the micro-controller) has also been implemented for fault tolerance: when the LED is left on for more than 50% of the time, the guardian forces a reset on the MSP430 to avoid the Babbling-Idiot Failure [81] (this is the name used to identify a node in a network which is stuck talking on the bus, thus locking it). This can be simply achieved using an integrator connected to the LED power supply: the more the LED is on, the higher will be the integrated value that could, when over-passing a programmed threshold, trigger an MCU reset.

The protocol stack is implemented in software on the MCU: received bytes are read from the Universal Asynchronous Receiver and Transmitter (UART) FIFO and then processed. The communication link is single master, thus saving the complexity of bus arbitration and collision detection: slaves are allowed to use the bus only under request of the master.

Component selection is important to ensure system survivability in space, thus it should be performed with particular care to select the most suited devices for the purpose. LED and photo diode should be selected in order to stand the particle flux in orbit (for further details see section 7.2.4).

Once components have been selected, the system was built on a PCB as shown in Figure 7.21 where only the receiver without the photo-diode and the EMI shielding is shown. Even if the bus itself is EMI-free, the photo-diode is quite sensitive to electromagnetic interference since the generated photo-current is in the order of some nanoAmperes (more or less the same magnitude of EMI coupled currents). A metallic shielding is thus needed: without this Faraday cage, receiver gain cannot be increased too much to avoid instability (basically every disturbance could be interpreted as a signal while with the shielding this effect can be easily kept under control and sensitivity can be increased).

Transmitter and receiver should be in line of sight for a successful communication, but optical connection can also be achieved by reflecting / diffusing surfaces that distribute the light among all transceivers. Total visibility through reflections and diffusion has been verified by simulation and experimental tests for the most simple configurations of the architecture. For some configurations, small “mirrors” can be used to put all the internal satellite boards in “optical connection”. This solution is not new in small satellites [88], and the use of multiple mirrors can create a light path, achieving a wireless guided propagation.

Equivalent propagation distance (i.e. the free-space distance that gives the same attenuation found between transmitter and receiver inside the cube) has been calculated using light propagation simulations inside the satellite, as can be seen in Figure 7.22 where three cases are analyzed. Figure 7.22a depicts the light power distribution inside the satellite (it is a cube with 16.5 cm side with another 10 cm wide cube inside that represents the payload): the internal surface of the satellite is made by PCBs and their reflection coefficient has been evaluated in laboratory measurements. In this figure there are no interfering objects between the transmitter (bottom left corner) and two receivers (bottom center and top right corner).

Figure 7.22b shows a feasible way to greatly reduce propagation distance by using a simple mirror: equivalent distance is reduced by a factor of approximately 3, allowing a

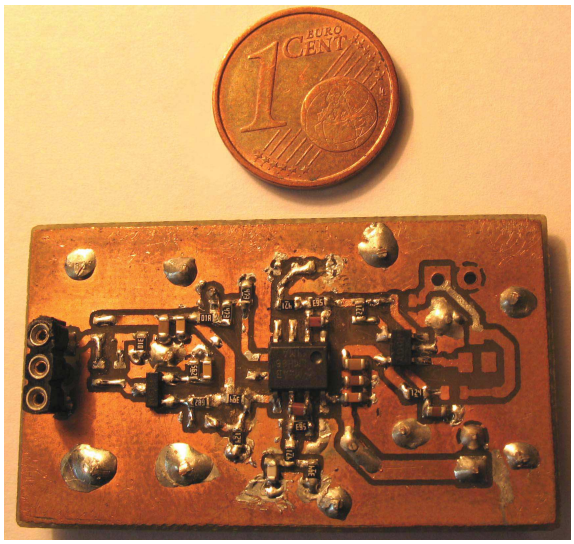
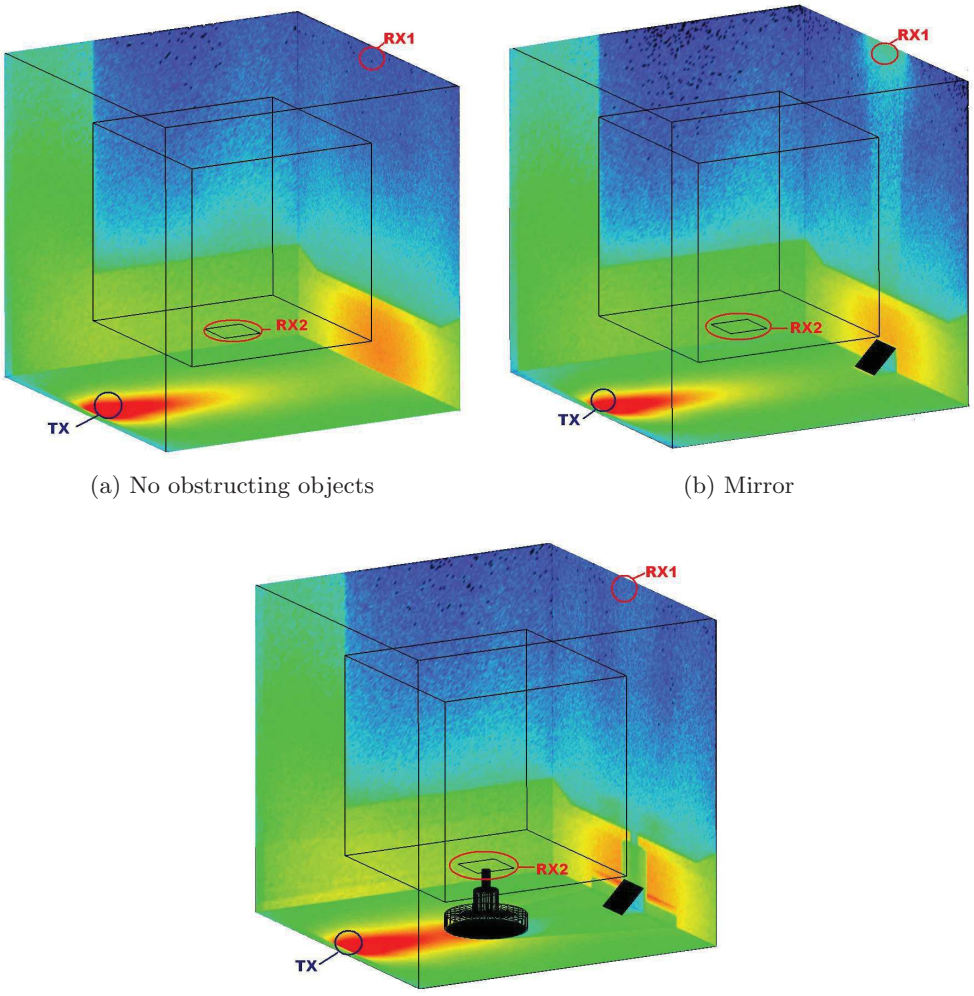


Figure 7.21: The proposed implementation without EMI shielding compared with an Euro Cent coin.

low BER communication between two opposite points of the satellite. In Figure 7.22c we can see the contribution of an obstructing object to light propagation: a motor (used for attitude control) has been inserted in the center of the faces causing an increase of 2.5 times in distance. The presence of objects inside the satellite can make distance rise up to around 2.5 m. Object placement is really important when there is no direct line of sight communication. It should be noted that most of the objects inside the satellite are made of metal which helps a lot in reflecting light by generating multiple scattering. Quantitative results are shown in Table 7.2.

The bus will be used to communicate with different sub-systems inside the satellite, each of them having different requirements regarding the data link: communication will take place between the satellite OBC and attitude determination and housekeeping sensors, attitude control actuators and the satellite payload. The link connecting the payload requires a low BER for transferring data (due to the large amount of data to be transferred): as in IrDA specifications we require a BER better than 10^{-8} and a throughput efficiency higher than 90 % for packets at least 256 bytes long. The link with sensors and actuators could tolerate a higher bit error rate since re-transmission can be employed in case of errors and the average packet size is quite low (20 - 40 bytes): given a maximum communication time of 4 ms (the maximum time needed to send one packet without errors to the receiver, considering also retransmission) we can tolerate three retransmission, being able to tolerate a BER as high as 2×10^{-3} still with a good safety margin. The increased channel usage will not be an issue since sensor read-out and actuators control should not be performed very often, while the communication with the payload is far more frequent and thus critical.



(a) No obstructing objects

(b) Mirror

(c) Mirror and obstructing objects

Figure 7.22: Optical power distribution inside the satellite (bigger cube side = 16.5 cm, smaller cube side = 10 cm).

Figure	TX - RX1	BER	TX - RX2	BER
7.22a	1.83 m	6×10^{-3}	0.15 m	$< 1 \times 10^{-12}$
7.22b	0.64 m	1×10^{-12}	0.15 m	$< 1 \times 10^{-12}$
7.22c	1.54 m	1.2×10^{-3}	0.16 m	$< 1 \times 10^{-12}$

Table 7.2: Equivalent distance between transmitter and receiver and BER (see Figure 7.22 for details).

7.2.4 Radiation issues

This system is used on the AraMiS satellite (see section 6) which is intended for low Earth orbit with an altitude below 1000 km: further details on environmental constraints can be found in section 6.7. To prevent latch-up, all the selected analog components were bipolar technology ones not to worry about it and also because in general bipolar amplifiers show better performances from the noise point of view. For the high speed channel the National Semiconductor LMH6646 operational amplifier was selected because it was developed with a bipolar on Silicon-On-Insulator technology featuring low power consumption, high bandwidth, high total dose tolerance (up to 300 krad, see section 5.5) and good SET hardness [89]. For the low speed channel, which requires a lower bandwidth, a bipolar low power op-amp was selected (Maxim MAX4092).

The MCU selected for controlling the communication and for encoding is an MSP430: this device was evaluated under radiation in ground based tests (see [64] and in flight (it was used in many CubeSat missions as OBC and showed good performances. It should be noted that this IC is made with CMOS technology and thus requires a latch-up protection system, while the operational amplifiers do not.

The main problem however was represented by the LED and the photo-diode: since the goal was to use COTS components, radiation hardness of commercial devices had to be evaluated: they have a plastic case (mainly epoxy or siliconic resin) and the die was not developed taking into account radiation problems. Double heterojunction GaAs LEDs were selected as they guarantee an higher tolerance factor to radiation damage (see section 5.5 for further details). While displacement damage data can be found to properly select the best device technology, no information could be found about the plastic package and in particular how much this can contribute to total degradation. The only drawback about plastic packages is that they are not completely hermetic, thus they can store a small quantity of water vapor that can create problems during thermal cycles under vacuum but this was not considered an issue since the internal part of the satellite is not going to experience extremely strong thermal cycles (temperature is between -10 °C and 60 °C with long transients). Since no data about plastic package degradation was available, a radiation test campaign was set up to verify the entity of this parameter and further details can be found in section 8.2.1. The test was performed with a low energy source (2 MeV protons) that allowed to analyze plastic degradation correlation with light transmission coefficient. The following LEDs were tested: Osram SFH-4502, SFH-4258, SFH-4650 and Vishay TSHG8400 and the latter showed a really good tolerance up to a fluence of 5×10^{10} protons. Few photo-diodes were also tested, like the Advaced Photonics PDI-C172SMF, PDB-C142 and PDB-C142F up to the same fluence and showed again good performances.

It can be clearly seen that the devices were tested up to fluences two orders or magnitude higher than the actual mission limit, which demonstrates that these devices could be used in space with a reasonable safe margin.

7.2.5 Results

The implemented solution has been tested in a laboratory environment and showed good immunity to external "noise" such as solar and fluorescent light. The system keeps this performance level with a high "DC noise", such as constant light power of about 700 W/m^2, which guarantees smooth operations also under indirect Sun radiation in space.

Power consumption was a key factor in the development and we achieved 1.1 Mbps link with only 3.6 mA at 3 V in reception while current consumption during transmission was about 35 mA. In these estimation the power consumption of the micro-controller was not taken into account since it is used also for other tasks.

Within this section it was shown that optical wireless communication inside a satellite can be a good system to save harness weight and complexity while still achieving a reliable communication: furthermore optical communication can be a good way to solve EMI issues inside the satellite.

7.3 Power Management Tile

Each power management module, called also Tile due to its shape (see Figure 7.23), can be mounted on a backbone mechanical structure that gives the shape to the system and also supports the other sub-systems. Beside the power handling system, each Tile also houses the attitude control sensors (2 axis magnetometers and a miniaturized gyroscope) and actuators (a magnetic torquer and a reaction wheel) since each tile is mounted on a different plane, thus offering the best support for sensors and actuators.

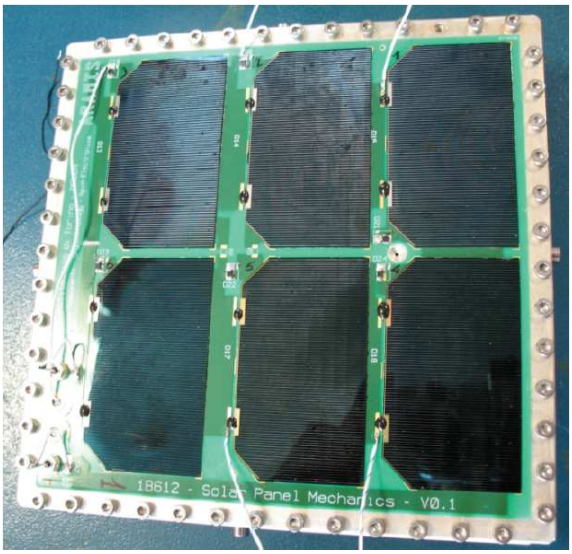


Figure 7.23: AraMiS Power Management Tile.

7.3.1 Power generation

The primary AraMiS power source are solar cells mounted on the external side of the satellite, covering almost the whole available surface, to use its power. Tile size has been selected to allow the best coverage with solar cells, so their selection is a critical point. Since solar cells are exposed to a direct particle flux, they should be strong enough to tolerate space radiation level. As it was discussed in section 5.2, the best solution to meet the operating life requirement was to employ space-rated solar cells. Among the different models on the market, the best compromise was selected by using GaAs Thin Film Triple Junction cells produced by CESI [52]. The company is based in Milano and this allowed good commercial relationship and their products quality is among the top ranked ones worldwide. Solar cell size is approximately 69×40 mm, a thickness of 160 um and a top efficiency of 27% (even if 23% efficiency cells were employed to reduce system cost and ease procurement) which allowed to fit 6 cells on a square panel 165×165 mm large for a maximum of 5 W output per tile in LEO. Cells are usually connected to each other and to the circuit by means of particular interconnecting clips which need to be weld onto the cell. Since this mechanism was considered a bit complex to integrate in our system, we selected to use bonding wires to connect one of the two terminals (the negative one on the front face of the cell) and to glue the back side one with an electrical conducting resin. The substrate on which cells were assembled is a normal FR4 PCB 0.3 mm thick which is then glued on the Aluminum panel. This solution allowed a simple building process and a strong grip between the cells and the mechanical structure. The hardest problem that had to be solved was the selection of the proper glue: most resins need to stand a curing cycle in oven up to 80 °C – 180 °C for about one to four hours but this will cause thermal expansion of all the components making bonding equilibrium temperature equal to oven temperature. This can be an issue when the thermal expansion coefficients of the materials are different, as in this case causing a residual stress to the materials when temperature is lower that the curing one (which is in our case, for the whole operating life). The lower the operating temperature will be, then the higher will be the stress, which can damage the materials, and in particular the most fragile one, solar cells. It was thus selected to use a resin with ambient temperature curing, which reduces the thermal stresses at the minimum operating temperature (around -30 °C). This solution allowed for a reliable adhesion between the materials, even after multiple thermal cycles (see section 8.1).

Solar cell configuration had to be evaluated in order to select the most reliable one in case of faults: a cell connection fault can lead to complete power loss, so proper solutions had to be adopted. A protecting diode has been put in anti-parallel to each solar cell so that in case one connection breaks, current can still flow from there. Furthermore, triple junction solar cells are quite sensitive to reverse current that can short the reverse junction and destroy the whole cell: this can be also avoided be means of the parallel protection diode. The best configuration is made by all the six cells in series, with their parallel diodes, that shows better performances when compared to other configurations, like the parallel of two string made by three cells in series: in case of fault the total power lost by the series is only one sixth of the total power, while when sub-strings are used, the power lost is around one third, so twice as much (see Figure 7.24).

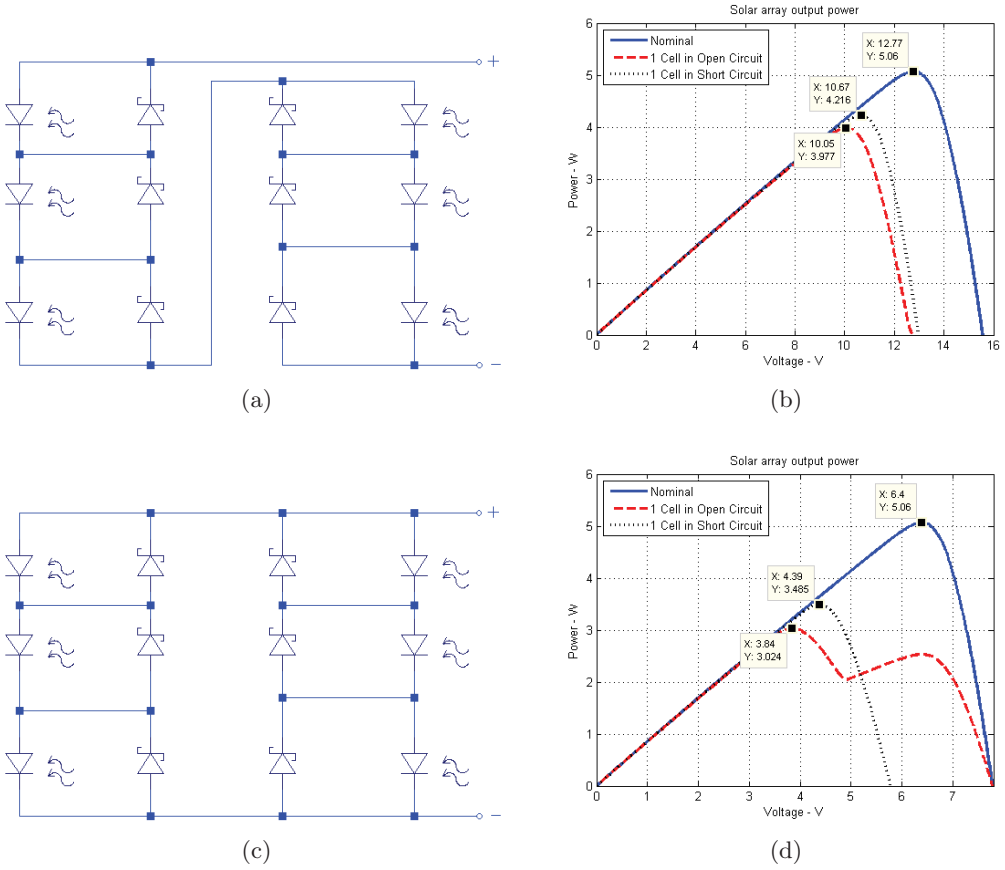


Figure 7.24: Solar cells configuration and fault tolerance (computed from SPICE simulations).

Since each tile is mounted on a different face, experiencing different light and thermal conditions, their output power is highly variable, thus requiring a separate power converter for each tile. A MPPT is responsible for tracking the maximum power point of the solar panel according to environmental conditions. The system is a switching-mode power supply which allows changing the load seen by the solar cells to get maximum power. A boost converter was selected because it causes a reduced current stress to solar panels and an hysteretic controller was employed because the control circuit is quite simple and it could be implemented without using CMOS integrated circuits, prone to latch-up in the space environment.

The basic schematic of this circuit is shown in Figure 7.26 where the boost converter can be seen on the top part, taking power from solar cells (only one is drawn for simplicity) and outputting it to the Power Distribution Bus. The MPPT is an hysteretic converter controlled from input voltage and current to get maximum power from the solar cells. To compute power a multiplication is needed and it is complex to implement it using an

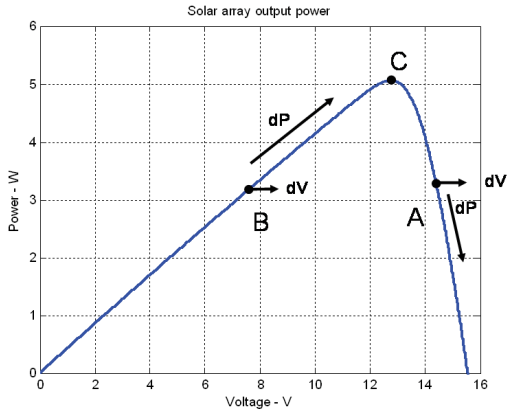


Figure 7.25: Perturbe&Observe algorithm: if solar cell operating point is A, and a positive voltage perturbation (dV) is applied to it, output power will experience a negative power variation; if the same perturbation is applied in B a positive power variation will happen. If dP is negative, the voltage perturbation should be complemented, while with a positive dP the voltage perturbation should not be inverted: in this way the point C (maximum power point) can be reached.

analog circuit. It was thus decided to implement it with a digital circuit, and in particular with a micro-controller: analog signals, sampled as housekeeping values, are acquired and the average power is computed by multiplying solar cells output voltage and current. The problem with this solution is that the micro-controller is not latch-up resistant, so power control cannot be available 100% of the time. To overcome this limitation, an hybrid solution has been employed by using an analog controller, which cannot compute the maximum power but makes solar cells operate around a fixed point which can be corrected by the digital controller, thus achieving full MPPT operations. This solution is also useful in case the micro-controller gets damaged because the analog system will stil be operating, achieving sub-optimal performances; anyway a limited power loss is better than a complete power loss.

The hysteretic comparator (see Figure 7.26) is designed such that it requires an external control voltage (V_{REF}) to be compared with instantaneous solar cells voltage (properly divided to fit into comparator input dynamic). In this way, by generating the proper reference signal, solar cells can operate at their maximum power point, where conversion efficiency is maximized. The comparator has an hysteresis used to limit input voltage ripple to few percent not to waste power in useless ripple. The reference signal is computed by the micro-controller using a Perturbe&Observe strategy [90]: a reference signal is generated and solar cell average power is computed, then this reference signal is perturbed (increased or reduced) and the output power is computed again. If power increased after the perturbation, a further perturbation is applied (in the same direction as the previous one) again, otherwise, the new perturbation will go in the opposite direction. In this way the system will converge to the top of the power curve (see Figure 7.25).

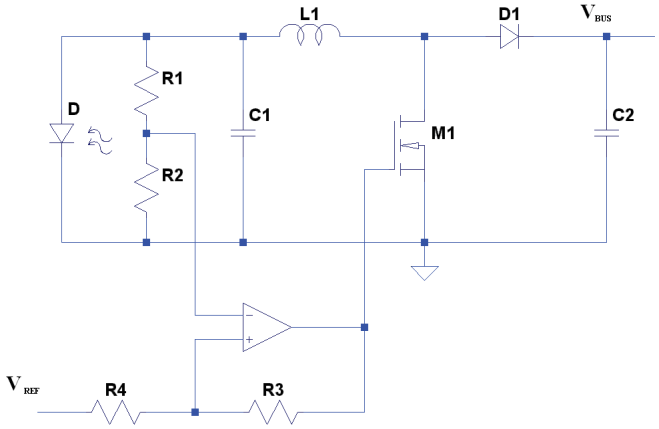


Figure 7.26: Maximum Power Point Tracker schematic circuit.

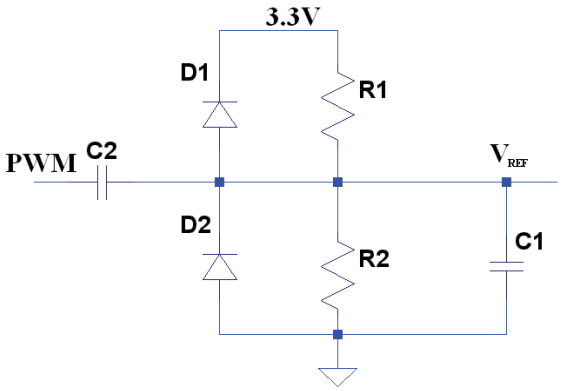


Figure 7.27: Fault tolerant reference voltage generator.

The micro-controller generates the reference signal using a PWM generator: by analyzing the possible faults that can happen the most probable is that this generator gets stuck (either to 1 or to 0) and if the reference signal is generated only by low-pass filtering the output, this signal will be stuck to 0 or 1 too. To prevent this, a high-pass filter has been employed (see Figure 7.27) that will remove the DC component, then a voltage divider is used to force a value near to theoretical maximum point. In case of faults in the micro-controller then, the analog reference value is set, while the digital controller acts to refine it and improve power generation.

7.3.2 Power Storage

Power storage is achieved using Li-Ion commercial batteries: we are using laptop cells (actually 18650 sized cells [91]) because of their high specific capacity and easiness of procurement. Lithium-ion cells show in general a longer life when compared to Lithium Polymer ones and this was a key parameter for the selection. Furthermore, these cells have a liquid electrolyte (a lithium salt in an organic solvent) and a metallic cathode: since lithium is highly reactive with water a sealed package is needed, so they are enclosed in a metallic can. Even if this makes the weight increase, there is no need for further protection against vacuum. Lithium polymer cells are instead packaged in a plastic bag with a residual internal pressure of approximately 15 mbar which makes them inflate at low pressure and increases the risk of cell damages in vacuum.

Each Power Management Tile houses two cells connected in series, thus giving an output voltage of 7.2 V, 2.5 A peak current and a stored energy of 15 Wh, weighting approximately 100 g. These cells are in general employed in notebooks because of their high endurance: their expected life is around 500 complete cycles with a maximum degradation limit of 70% of their initial capacity but this number can rise up to 1000 cycles if the discharge cycle is only partial. These numbers have been determined with many long run tests performed on different cells from many brands (see [92]).

Lithium batteries are quite critical since they can be easily damaged if not charged or discharged in the proper way: voltage across the battery should stay between 3 V and 4.2 V, while current should not be greater than the rated maximum value both in charge and discharge. Furthermore the charging cycle is quite complex:

- a constant charge current should be used until cell voltage is below 4.2 V,
- a constant voltage (equal to 4.2 V ± 1% over temperature) should be applied until current drops below maximum current divided by 10,
- after charge current dropped below the threshold, charge should be stopped not to overcharge (and damage) the battery;
- the charge process should be stopped in case battery temperature is above maximum operating temperature (60 °C).

Taking into account all these constraints, it was selected to use a commercial battery charger IC not to spend time in implementing and testing this complex charge mechanism. Many devices were analyzed but the most promising was the Linear Technology LTC4008: many reasons lead to this selection but the most important one was that this device has already been used in space in the SSETI Express mission, developed by the European Space Agency [93]. Furthermore, this IC features a synchronous buck converter with external transistors that make it easy to protect against latch-up. Since power supply to the IC comes from a separated wire than the charging current, the latch-up protection circuit current threshold could be set to the optimal value (about twice the power supply current) without considering charging current (which is a hundred times higher). Furthermore, the

synchronous rectifier employed has a high efficiency (96% measured at ambient temperature) and the application circuit was already designed, thus requiring a short development time.

7.3.3 Power dissipation

During normal operations, solar generated power is used to charge on-board batteries and power the system, but if a surplus power is available, it should be dissipated somewhere. For this purpose a shunt device has been added to every power tile: this device should be able to dissipate as much power as it could be generated on a single power tile (so 5 W) plus a reasonable safe margin. The resistor can be enabled from the OBC to select which tile should be heated but the amount of energy dissipated in the shunt depends on available power on the bus (the control mechanism will be better described in Section 7.4). This system acts also like an over voltage protection system on the bus, thus limiting spurious peaks that can damage the system.

7.3.4 House Keeping

House keeping is ensured by a micro-controller (an MSP430 [72]) which is responsible for system monitoring (voltages and currents in all the critical points in the Power Management Tile and temperatures of power devices) and system control (turning on and off all important switches to change the configuration of the bus).

This micro-controller is not involved in any decision regarding power control strategy, which is demanded to the satellite OBC. Bus voltage and current drawn from each tile is controlled automatically by the bus itself, by using its differential impedance. The tile controller can connect each tile power output to one of the two available power buses (or both in case of particular needs) or automatically disconnect the tile in case of anomalies.

Housekeeping values can be of different types:

- instantaneous values, used by the OBC to take decisions on current bus usage strategy,
- peak, average values and standard deviation, to get an overview of system operating conditions, mainly used from ground.

7.3.5 Fault Tolerance

The power distribution and management system is modular and scalable, since it is composed by this tile, replicated as many times as needed to get the desired power level. All these tiles work together and this massively redundant solution helps also from the fault tolerance point of view, making the system able to tolerate multiple faults and allowing a graceful performance degradation. Loosing one tile causes a reduction in generated power, but the system will still be operative (see section 7.4). A redundant power bus is used to increase reliability and to allow different loads to be connected and switched: power can be directed towards one or both buses according to specific needs. The second bus

allows to eliminate single points of failure because if one bus gets damaged, the second one can still be used. Each Power Management Tile is equipped with an output current protection system to prevent damages in case bus short circuit and with an under-voltage protection that disconnects the tiles in case bus is outside specifications. Output current limitation is achieved with the switching controllers, while output short-circuit protection is achieved with a series diode. In case bus voltage exceeds maximum allowed value, the shunt resistor is activated to avoid damages to the whole system.

7.4 Power Distribution Bus

The most important feature this power management system should satisfy is the flexibility, such that it can fit from small applications, like a small nano-satellite weighting only 5 kg up to bigger systems weighting ten or twenty times more. The innovative aspect of this power bus is an inherent modularity to allow simple Plug&Play interconnection of as many tiles as required to fit specific needs in terms of power and operating voltage.

Power distribution and management is a tough real-time tasks which requires response time well below 1 ms and this very fast reaction time may involve many Power Management Tiles and communication via the On Board Data Bus and therefore cannot be handled reasonably by the On-Board Computer.

It has therefore been decided to use a different distribution and management technique which does not involve the OBC, except for occasional reconfiguration tasks, which are not expected to take place more frequently than few times every orbit.

The bus is connected to four main devices:

- primary power sources, which are the solar panels of all the tiles,
- batteries, which can be seen both as sources and loads (while charging / discharging),
- protection devices, used to keep bus voltage and current between specifications,
- power loads, which are all the other satellite sub-systems.

7.4.1 Bus characteristic

The power distribution bus was developed as an unregulated bus, with a voltage varying from 12 up to 18 V. The unregulated solution has been selected to allow different devices with different output characteristic such as the MPPT (which can be modeled as a constant current source) or a battery (modeled as a constant voltage source) to be connected together in an easy way. The characteristic of the bus is depicted in Figure 7.28: current is considered positive when entering in devices connected to the bus and negative when exiting from them. It was selected not to employ a constant voltage bus because, as in the philosophy of the whole satellite, the bus should be modular and scalable. In this way, all the devices should have the same output voltage which is difficult to achieve since components can slightly degrade due to radiation thus generating small variations from device to device. These small variations would make load sharing between them quite

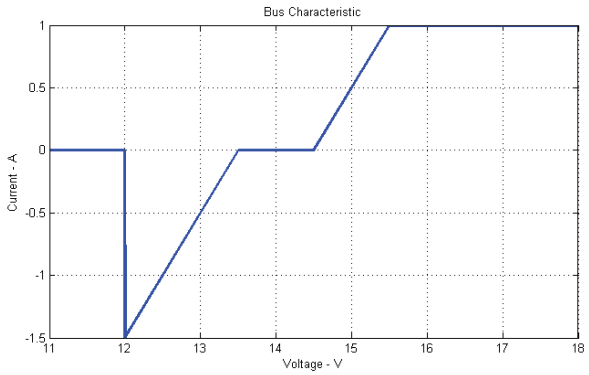


Figure 7.28: Bus characteristic. Current is considered positive when entering in devices connected to the bus and negative when exiting from them.

difficult, since the device with the highest output voltage will supply the highest current. To simplify power sharing it was decided to provide a high input and output resistance on every device connected to the bus. This resistor is equal to 1 Ohm (actually it is a virtual resistance, not to waste power on it). In this way, the device providing the highest output current will have the lowest output voltage and this will allow devices not supplying current to balance the generated power between multiple devices. Around 14 V the input / output current is zero to avoid that one device could oscillate between the source / sink function due to small variations in the voltage references. The bus structure and stability will be further discussed in the next section.

7.4.2 Bus stability

To model all these contributions as a function of time, bus voltage has been selected as a state variable and current balance has been computed taking into account the contributions from all the connected devices.

Solar panels (and to be more precise, their MPPT) can be modeled as a constant power source since there is no regulation on the output voltage or current. To follow the maximum power point the controller should only track input voltage and current, thus leaving the output not controlled. In a boost converter, when the load is too weak, output voltage will increase in order to keep output power constant while when the load is too strong, output voltage will go down until it reaches the input voltage, thus stopping boost operations, and making the output diode always conduct current. This can be clearly seen in Figure 7.29 where the output characteristic of the MPPT controller is plotted for different power input: in the right side, the constant power curve can be clearly seen, while it saturates to solar cells maximum current in the left part.

All the loads connected to the power bus will be switching converters and this means that all the loads will have an hyperbolic input characteristic, and coupling it with the MPPT output characteristic can generate problems. The equivalent circuit schematic

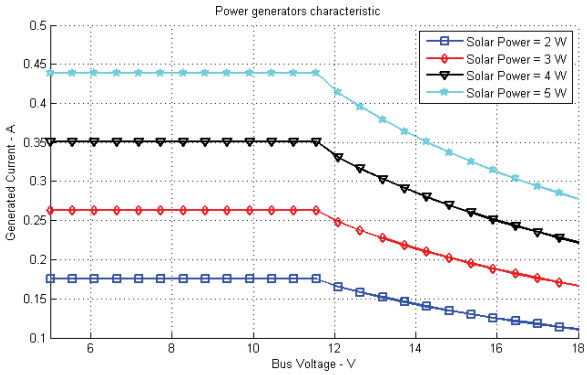


Figure 7.29: MPPT output characteristic for different values of solar power.

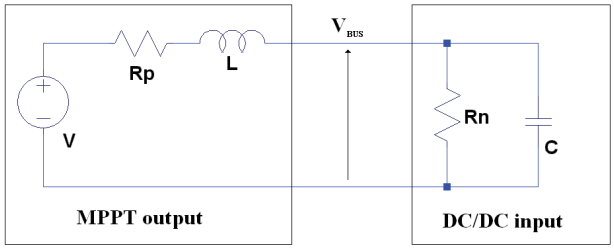


Figure 7.30: Cascaded switching converter equivalent schematic.

is shown in Figure 7.30 where the MPPT output stage is represented on the left, with its equivalent output resistance and the parasitic inductance (used only when one tile is driving a load on a different tile, thus modeling cable inductance) and on the right there is the input stage of another DC/DC converter. Since this is a second order circuit it can be unstable for particular values of the components. The transfer function, from the MPPT voltage source to the V_{BUS} voltage can be computed as follows:

$$G = \frac{R_p}{s^2 CLR_p + s(CR_nR_p + L) + (R_n + R_p)} \tag{7.21}$$

Stability can be ensured by requiring that poles lie on the left part of the complex plane, which means that they should be real and negative. The denominator is a second order equation with two roots, s_1 and s_2 and from second order equation properties it can be shown that:

$$s_1 s_2 = \frac{R_n + R_p}{CPR_p} > 0 \tag{7.22}$$

$$s_1 + s_2 = \frac{CR_nR_p + L}{CLR_p} > 0 \tag{7.23}$$

and in order to have both roots real and negative, the previous inequalities should give a positive result, thus implying that:

$$R_p \ < \ -R_n \tag{7.24}$$

$$C \ > \ -\frac{L}{R_n R_p} \tag{7.25}$$

In order to evaluate system stability, all the previous parameters should be computed for the different AraMiS configurations. In the $1 \times 1 \times 1$ configuration we can suppose to have only one tile generating solar power and only one using it and, in order to evaluate the worst case, we are going to use the two farthest tiles. The cable used to connect the tiles has a section of $2.5\,\mathrm{mm}^2$ and is approximately 50 cm long: from these values the approximate cable inductance can be computed in:

$$L \ = \ l\left(log\left(\frac{4l}{d}\right) \ - \ 1\right) \ 200 \ \frac{nH}{m} \tag{7.26}$$

where l is cable length and d is cable diameter, giving an inductance of 730 nH. R_p is computed as the derivative of MPPT characteristic and the worst case value is 24 Ohm; R_n is instead the derivative of loads input hyperbolic characteristic in the worst case, so with minimum input voltage and maximum output power, it is equal to -5 Ohm. From the previous equations, it can be clearly seen that the system is unstable, so proper solutions should be employed. The first condition that should be satisfied is given by equation 7.24 which requires lowering R_p but this is impossible because it would require to have an higher input power, so a bigger tile, but this would prove to be unstable in low Sun conditions. The alternative is making R_n higher, better if it could be made positive, thus removing any requirements for stability. If the switching converters employed as loads had a feed forward input (so they control the output voltage also using input voltage) this can be easier: the input voltage signal can be modified in order to take also information about the input current thus creating a virtual positive input resistance. This solution is depicted in Figure 7.31 and the circuit output voltage can be computed as follows:

$$V_{OUT} \ = \ V_{BUS} - \left(V_{BUS} \ - \ V_{DC/DC}\right) \ = \ (1 \ + \ k)\, V_{BUS} \ - \ kV_{DC/DC} \tag{7.27}$$

$$k \ = \frac{R_1}{R_3} \tag{7.28}$$

$$R_{eq} \ = \ R_2 k \tag{7.29}$$

$$R_4 = \frac{R_1 R_3}{R_1 \ + \ R_3} \tag{7.30}$$

R_2 is used to sense I_{BUS} current and to simulate input resistance but for power dissipation reasons this resistor could not be too big: an operational amplifier is thus used as a resistance multiplier featuring a gain of k while R_4 is used for compensating the amplifier bias current. With this circuit, it can be clearly seen from Figure 7.33 that battery charger

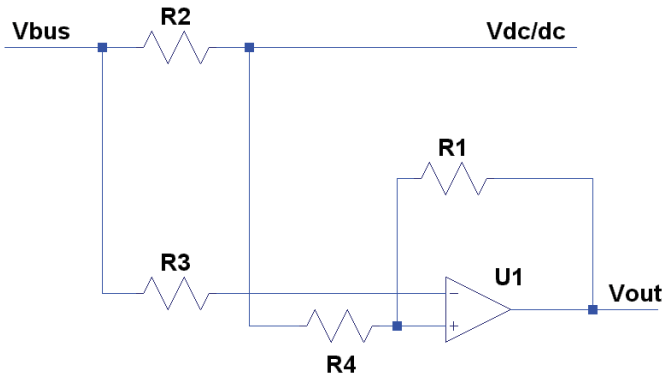


Figure 7.31: DC/DC converter input impedance control circuit.

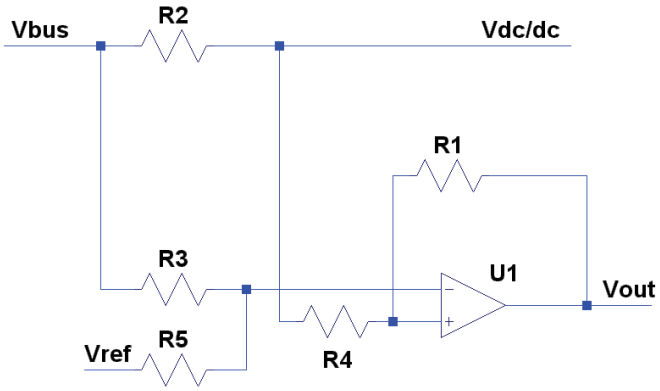


Figure 7.32: DC/DC converter input impedance control circuit with reference voltage.

impedance is now positive up to 15.5 V, while for higher bus voltage the resistance is still negative. The above circuit still has one problem: the input characteristic is resistive from 0 V up to intersection between the resistive characteristic and the constant power characteristic. But this circuit still has one problem: the battery charger is still drawing current when input voltage is below 14.5 V. This problem can be solved by adding a reference voltage and subtracting it from the output, such that the operational amplifier output voltage is below zero (so saturating at the negative rail of the operational amplifier when it is supplied without the negative supply). The transition point depends on battery charge current that depends on charge status and temperature: the input impedance can only be modified by reducing input current with respect to a constant power source, because increasing it would introduce losses and this cannot be tolerated. The higher is the value of input resistance and the wider will be the voltage range where a resistive characteristic is generated, so selecting an input resistance of 1 Ohm, that range will be from 14 V to 15.5 V, with a maximum input current of 1 A, that gives a maximum battery

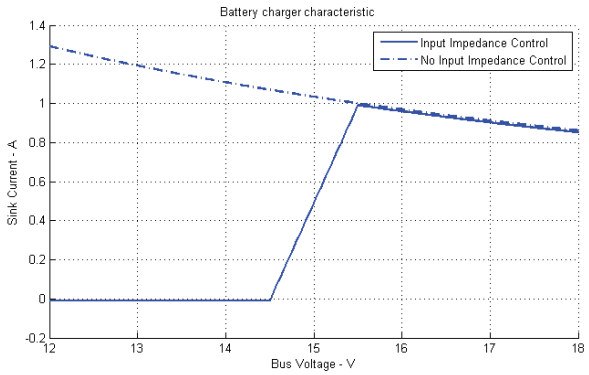


Figure 7.33: Battery charger input characteristic with and without input resistance control circuit.

charge current of 2 A.

With a positive input resistance on the loads side, the bus is stable but stability can only be ensured for the previous mentioned region, before the knee point, that depends on battery charge current as solar input power. When bus voltage passes the knee point, the characteristic is again composition of two hyperbolas, thus unstable. In this point the bus can be modeled as two constant power sources, one actually generating power (the MPPT) and the other sinking it (the load) with a capacitor modeling the bus itself (see Figure 7.34). This circuit can be easily solved with the following equation:

$$\frac{dV_{BUS}}{dt} = \frac{I_1 - I_2}{C} \tag{7.31}$$

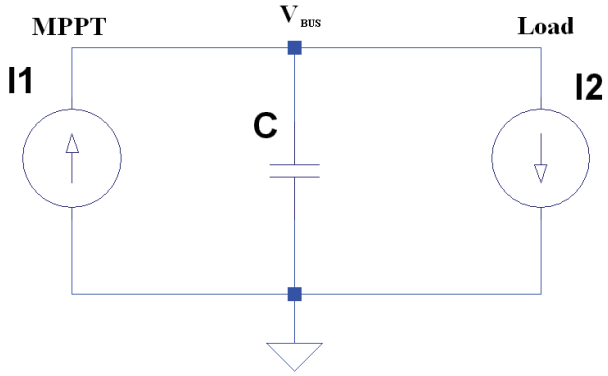


Figure 7.34: Bus equivalent circuit modeling instability region above loads knee voltage.

So bus voltage will keep on rising if input power is higher than load power and it will fall if input power is lower than loads power: the latter case gives no issues to stability,

since voltage will drop until the system will be back in the stable region. The former case will see instead the voltage rise without limit: this can be prevented by adding an over-voltage protection system that acts like a load, with an input positive resistance, but with a higher intervention threshold, as can be seen in Figure 7.35.

The bus characteristic has two stable regions and a non-stable one between the previous two which can make voltage oscillate from one stable point to the other: this requires two complex-conjugate poles in the transfer function but the system is only a first order circuit, so instability is going to happen.

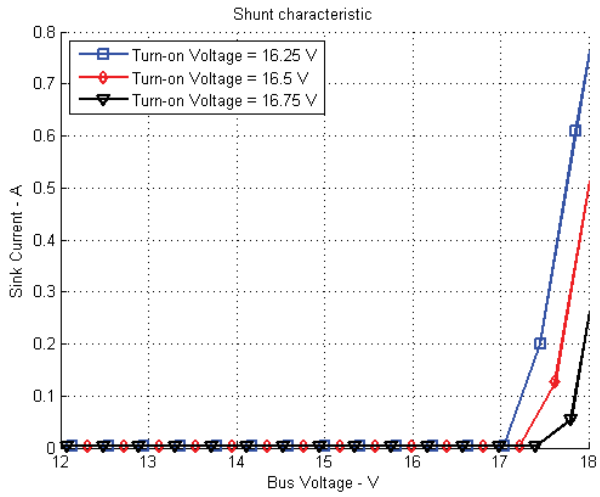


Figure 7.35: Voltage protection circuit characteristic with different intervention thresholds due to component tolerances.

Batteries can be also seen as power sources for the bus when in discharge mode, but their energy should be used only for power loads, and not for charging other batteries because this will simply waste energy. By having a look at the MPPT and charger characteristic it can be seen that batteries should not output current when bus voltage is higher than charging threshold, so above 14.5 V. If the battery discharger characteristic would be similar to the MPPT one, this behavior could not be achieved and it was thus decided to model them with a voltage generator with a series resistor. The value of this resistor is again 1 Ohm as it was for the loads: in this way an effective load sharing between multiple sources is achieved because the power source supplying the highest current will also be the one with the lowest voltage, thus allowing other sources to start providing current to the loads.

Again a real resistor could not be placed at the output of the battery discharger because it will waste too much power: since again the battery discharger will be a switching converter, the same approach used for the battery charger can be used and the battery charger characteristic is depicted in Figure 7.36.

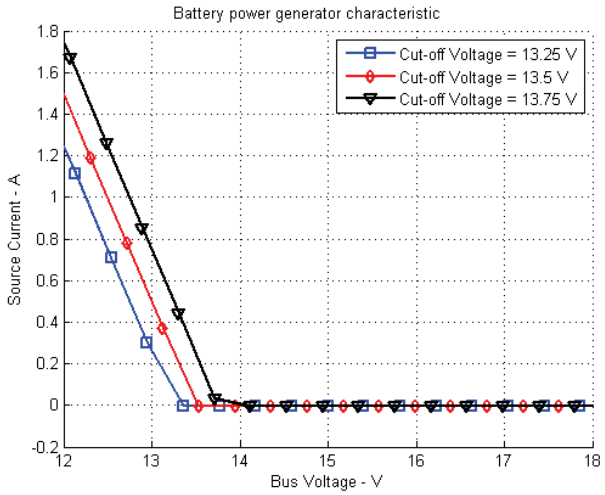


Figure 7.36: Battery discharge circuit characteristic.

Configuration	**Number of tiles**	**Size -** cm	**Voltage** V	**Current** A	**Power -** W	
					Avg	**Peak**
$1 \times 1 \times 1$	5	$16.5 \times 16.5 \times 16.5$	14	10	30	140
$2 \times 2 \times 2$	20	$33 \times 33 \times 33$	28	20	120	560
Cylinder	18	$\oslash\ 33 \times 50$	28	18	108	500
Flat 6×10	60	100×165	84	20	360	1680

Table 7.3: Different AraMiS configurations performances.

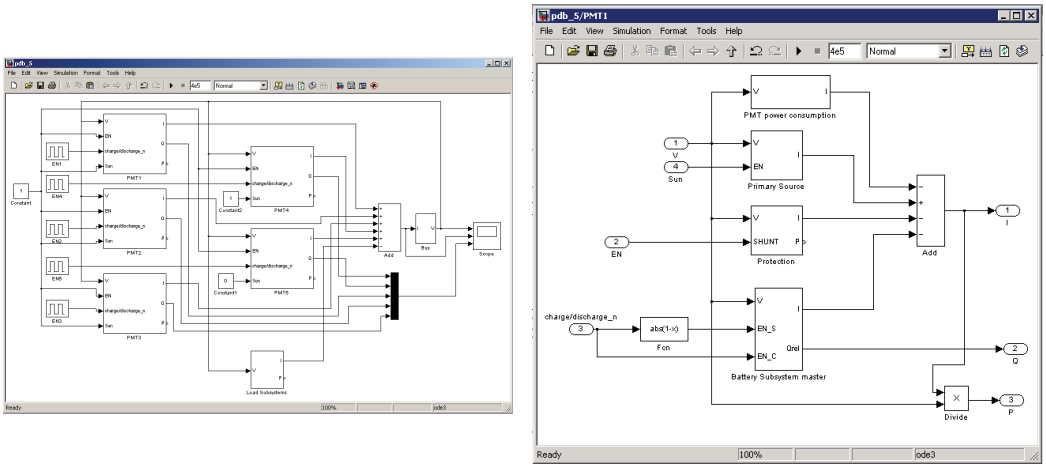
Since the bus is automatically controlled, the OBC is only involved in high-level decisions about battery charge / discharge and loads management. This is a great achievement since the on-board computer can save time for satellite management. Loads are controlled from the Power Management Tiles, where power switches are housed, but the control is performed on the OBC via the data bus.

Different configurations can be then built by composing the Power Management Tiles to fit specific needs: a survey of some of the possible configurations is described in Table 7.3.

7.4.3 Performances and Results

From the previous static models, a dynamic one was developed to verify system stability and evaluate the performances: software simulations have been performed using a Simulink model that is represented in Figure 7.37.

The configuration that was analyzed consisted of five Power Management Tiles, in the basic AraMiS configuration ($1 \times 1 \times 1$ cubic configuration) orbiting at 800 km with a spin rate of 6 ° degrees per second. The five on-board batteries have a 50% stored charge (6.25 Ah over a total of 12.5 Ah) at the beginning of the mission. Different loads



(a) (b)

Figure 7.37: Power Distribution Bus (a) and Power Management Tile (b) Simulink models.

were also modeled to better reproduce flight conditions: a constant power load (modeling Power Management Tiles internal power consumption) and three other loads, used too for modeling the other components of the satellite (on-board computer, transceiver board and the payload): peak and average power consumption are summarized in Table 7.4 and 7.5.

Load	**Power -** W **Power -** W	**Duty Cycle -** %	**Period** s	**Average Power -** W
1	0.5	300	7	0.035
2	6	1200	6	0.36
3	10	2800	5	0.5
Total	16.5			0.9

Table 7.4: Satellite loads average and peak power.

Single Tile W	**Load 1** W	**Load 2** W	**Load 3** W	**Total** W
0.42	0.035	0.36	0.5	3
0.56	0.035	0.36	0.5	3.7
0.7	0.035	0.36	0.5	4.4

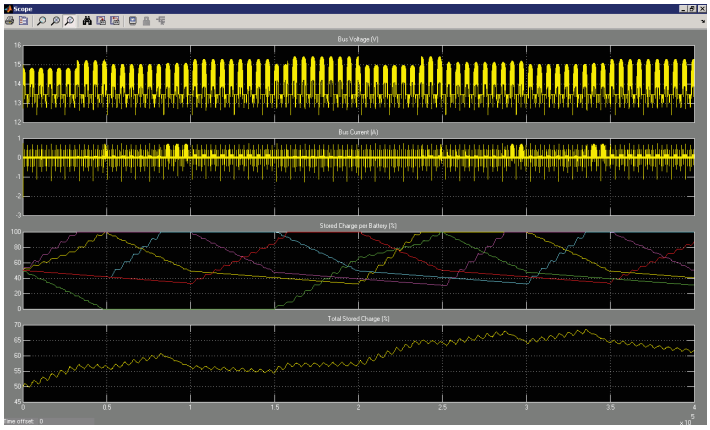
Table 7.5: Total satellite power consumption in different cases.

As can be seen from Figure 7.38, bus voltage is stable and follows Sun power trend during the orbit with peaks near to 16 V where the protection circuit starts operating, while current trend shows negative peaks due to loads power consumption. The most interesting trends are the last two, which represent battery charge status and total stored

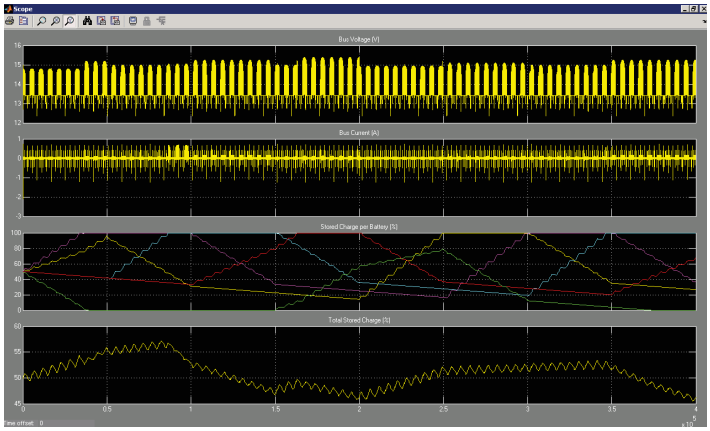
charge. Two different charging trends can be identified from these two plots: a fast charging trend, (when the curve has the highest $\frac{dC}{dT}$) which happens when only one battery is charged at once and a moderate one, which takes place when two batteries are charged at the same time because the a single battery is not able to draw all the available power. Also two discharging trends can be identified: a fast discharging trend takes place when the selected battery is powering the whole satellite electronics while a moderate one takes place due to battery self-discharge.

In can be clearly seen hat with a Power Management Tile internal power consumption of 420 mW total stored charge is continuously increasing, thus showing that the satellite can safely operate in space with this power budget. It could also be seen from Table 7.5 because total power consumption is 3 W while input solar power is around 5 W. In the second case (with a Tile power consumption of 560 mW) the power budget gets more critical and the average of total stored charge is not monotonically growing, even if it is still above 50%. The power net budget in this case is a bit narrow (3.7 W gainst 5 W) and this makes the battery charging and discharging policy a bit more complex since a higher efficiency is required not to waste precious power. The latter case is the most critical because the power budget is almost 1:1 (4.4 W against 5 W): this tells that stored charge will monothonically fall if sub-optima power usage strategies are not employed because the safety margin is too narrow. These simulations thus helped in defining the maximum Tile power consumption for safe operations in space but the anyway showed that the system is stable and, if properly sized, allows to achieve good performances.

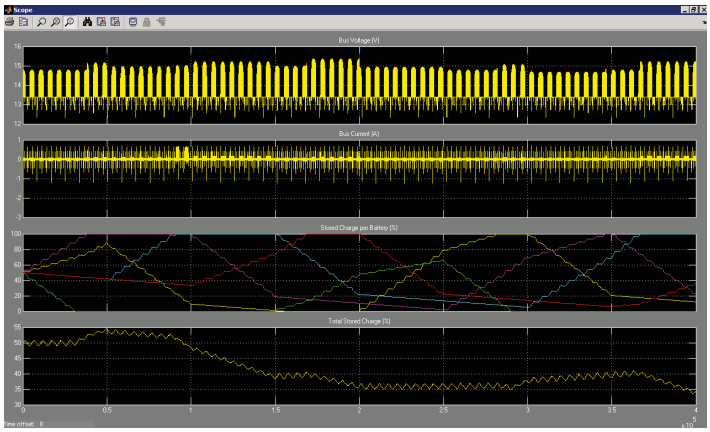
From simulation, maximum battery charge time could also be determined: this value is around 1 day and it will help a lot in determining the best power usage strategy.



(a) Per Tile power consumption 420 mW



(b) Per Tile power consumption 560 mW



(c) Per Tile power consumption 700 mW

Figure 7.38: Power Distribution Bus simulations of time evolution: X-axis is in seconds, and the whole time period in 4.6 days.

Chapter 8

Testing and quality assurance

Testing is extremely important to verify that the developed system is be able to meet the requirements defined. It can be performed at different levels in the design process, starting at component level, for example, to verify that components can stand mission environmental requirements, it can be employed at various stages of the development or at the end to validate and qualify the design.

Testing can be divided in three main categories, according to the goal it is pursued:

- functional test,
- qualification test,
- acceptance test.

Functional tests are employed during design phases to ensure that the system meets the requirement it was designed for. Testing a battery charger circuit to verify if it is able to actually charge a battery is an example of functional test.

Qualification tests are instead used to verify if a device or the whole system is able to operate without problems in the environment it was designed for and this is in general performed after functional tests. Verifying the effective operating temperature range for the battery charger is an example of qualification test, since it is used to qualify the device for operations up to (or down to) the specified temperature. In general qualification tests are performed on a slightly wider range than the one specified during the design to keep a safety margin. Tests are then performed on an engineering model because these kind of tests may damage the device.

Acceptance tests are instead performed on the produced devices, that are going to be shipped to the customer: the design had already been qualified with proper tests on an engineering model so the test does not need to be performed up to every specification limit. Furthermore the test should not cause stresses to the device, so test will not have to excite the device up to its maximum specifications.

The previous example of the battery charger can be used to clarify the previous discussion about test limits: from design specification, our example battery charger should have an operating temperature range of -20° C - 80° C and test conditions in the different phases will be the following ones:

- **Functional test**: performed on the engineering model at 25° C to verify that the system is working properly,
- **Qualification test**: performed on the engineering model from -30° C to 90° C to ensure operations in the full temperature range and slightly beyond it,
- **Acceptance test**: performed on the flight model from 0° C - 50° C to verify the device is ready to fly.

After this brief introduction, the following sections will analyze deeper some tests that were performed on various parts of the AraMiS satellite: the tests were all used for qualification so, the flight limits had to be exceeded to verify the quality of the design and the components.

8.1 Thermal and Vibration tests

Thermal tests can be divided into two main categories according to their goal: static and dynamic. Static tests are performed to verify system functionalities at a certain temperature, which is usually the maximum or minimum end of the operating range. Dynamic tests are instead used to verify if thermal gradients due to heating or cooling can generate cracks or, more in general, malfunctioning. Vibration tests are instead performed to ensure that the satellite is going to stand the launcher vibrations without damages; the device under test is mounted on a shaker (it is basically a high power loud speaker without the paper cone used to generate sound but with a shaft to attach to the device tested). The test is used to simulate launcher vibrations and it is divided into two halves: in the first part a sinusoidal stimulus is applied to the shaker and frequency and amplitude are swept to match launcher characteristics (see section 2.4 and Figure 8.1) while in the second half a random stimulus is applied to the shaker. Sinusoidal test is used to find a resonance

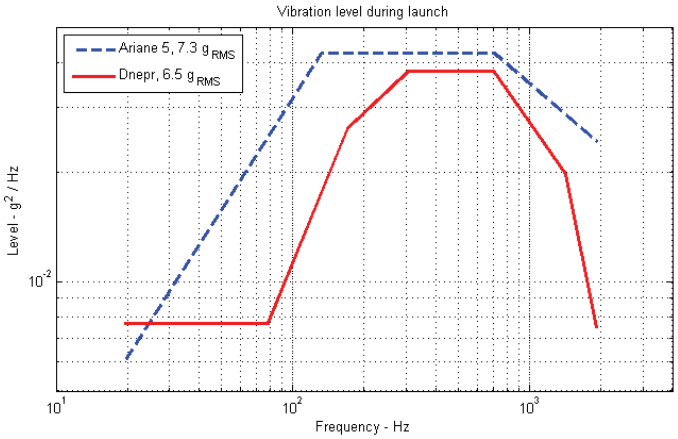


Figure 8.1: Vibration level during launch for Ariane 5 and Dnepr rocket.

frequency in the system that could, if excited lead to rupture. Random test is instead used to better reproduce launch vibrations by mixing different frequency components at once.

The tests that were performed were used to verify if the solar panel assembly was able to stand space thermal gradients and launch vibrations: for this reason a solar panel has been assembled and all the six cells had been characterized (voltage, current and power generated under a solar simulator) to verify if any change could be detected leading to anomalous cell stress generated during one of the thermal cycles or under vibrations. The cell that were mounted are classified "mechanical grade" because their efficiency is quite poor (on average around 15%) and they should be used only for mechanical tests.

The thermal cycles were performed over a temperature range wider than the operative one, in particular from -30° C to 90° C: the two temperatures were generated in an environmental chamber set to -40°C and in small oven set to 100° C. To speed-up the test, the solar panel assembly was manually moved from one of the two chambers to the other and left inside until the temperature reached the before mentioned values. To measure the temperature, a NTC resistor was installed in the gap between the internal PCB and the Aluminum panel so that internal temperature could be measured. The panel required about 20 minutes each time to pass from the minimum to the maximum temperature and vice versa, achieving an average thermal gradient of 6 degrees pr minute with a peak of 30 degrees in the first minute.

After a full characterization of all the cells, the system had been assembled on the shaker for the vibration test. The test has been performed according to the worst case amplitude in Figure 8.1 (the maximum of the two curves) to ensure that AraMiS could be flown on both rockets. The frequency sweep has been performed with a rate of an octave

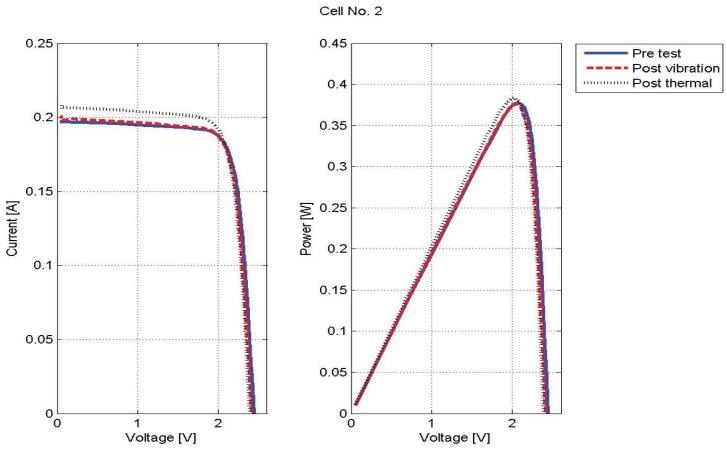


Figure 8.2: Cell characteristic before and after thermal and vibration test showing almost no change (after 6 thermal cycles the only difference that can be noticed is a small improvement in the characteristic due to annealing).

every minute to ensure that all the proper oscillatory modes have been excited and none of them generated a resonance. The random vibration test has been performed for 5 minutes for being sure no problem could arise during launch. By having a look at Figure 8.2, the pre-test curve and the curve after vibration test could not almost be distinguished, thus showing that during vibrations no problem arise for the cells.

After that, the solar panel was exposed to six complete thermal cycles and afterwards a complete cell characterization cycle was performed again, showing almost no change again. A small improvement can anyway be noticed in Figure 8.2 between the vibration and thermal test because the cells experienced a thermal annealing.

8.2 Radiation tests

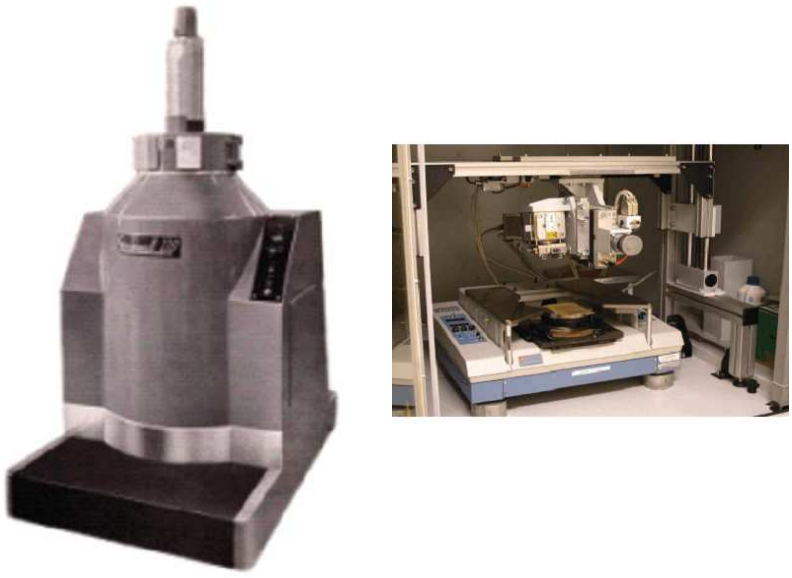
Radiation tests are in general performed on single components to verify that they could stand the space radiation level and afterwards the components are employed in the design. As it was specified in section 3, different effects can arise due to different types of radiation, but in general the effects are these three:

- total ionizing dose,
- displacement damage,
- single event effects,

each of them requiring a particular procedures and equipments.

Total dose testing involves testing the device under a particle source and verifying device functionalities as a function of total absorbed dose. There are two approaches that can be used to characterize the response: step stress and in-flux testing. Step-stress testing is performed by first characterizing the electrical performances of the device, exposing it to a fixed dose of ionizing radiation, and then measuring again the electrical parameters to determine their change. To determine the device response versus total dose, the test is performed with different samples of the same type at a number of accumulated dose levels. In-flux testing is performed instead by continually measuring the device response as it is being irradiated. The step stress approach is usually more convenient and much more widely used.

The main particles used in this kind of test are photons, which have a low deposed charge but they are quite easy to generate and handle. Photons actually do not depose energy in the device under test, but they interact with lattice atoms, actually stripping electrons from the outer atoms: these electrons, also called δ-rays, are responsible of induced dose. Cobalt-60 is the most used source: emitted photons generated by β decay have an energy of 1.25 MeV which allows them to penetrate through a thick shielding (about 8 mm through a lead layer). This high penetration generates some issues from the radiation protection point of view because photons can be harmful. The ^{60}Co source is depicted in Figure 8.3a: all the Cobalt rods are enclosed in a shielding barrel where devices under test are also held. These kind of testing machine is called γ-cell, and it is able to generate an almost isotropic particle flux inducing approximately 50 rad per



(a) (b)

Figure 8.3: Photons sources used for TID test: a γ-cell (a) and an X-ray machine (b).

second. Lower doses can be produced by partially shielding the device under test. The other photon source usually employed for testing is a X-ray machine (see Figure 8.3b): the energy in this case is much lower, from 7 keV to 45 keV, which gives an extremely low penetration, usually in the range of few microns. Photons are generated by a hot Tungsten cathode and dose rate is around 15 - 120 rad per second. This source is quite simpler to use, since it can be turned on and off by command and requires (while under operation) a thin shield to eliminate particle spill risk. Since penetration with X-rays is so low, devices have to be delidded because the package would completely stop radiations, while with γ-rays there is no problem in testing devices with the package, even the metallic ones.

Test procedure is specified in the MIL-STD-1019.4 standard, which requires the following steps:

1. ^{60}Co irradiate devices under worst case bias conditions to the specified dose at a dose rate of 50 - 300 rad(Si)/s;
2. Remove bias and maintain zero bias between irradiation and test;
3. Complete functional and parametric tests within 2 hours after irradiation;
4. If the devices pass all tests in (3), irradiate the devices again under the conditions of (1) to an additional dose equal to 0.5 times the specification;
5. Bake the devices at worst-case static bias for 168 hours at 100° C, or under conditions that have been demonstrated in characterization tests to cause equal or greater degradation in the parameter(s) of interest (e.g. speed, timing, and/or output drive);

6. Repeat the tests of (3).

In the test procedure, the dose rate is suggested to be in the range 50 - 300 rad per second which is much higher than the dose rate in orbit (in LEO it is around 10^{-3} rad per second) and this can cause differences between the behavior under test and during the operative life. In MOS devices, the higher the dose rate is under testing and the higher is the effect on the device. But the test can give false negative results: a MOS-based device can be in fact tested with an high dose rate on ground and fail at a much lower total dose than it would happen in orbit. To overcome this effect an alternative test method has been proposed [94][95]:

1. ^{60}Co irradiate devices under worst case bias conditions at a dose rate of 50 - 300 rad per second to the specified dose;
2. Remove bias and maintain zero bias between irradiation and test;
3. Complete functional and parametric tests within 2 hours after irradiation;
4. If the devices pass the tests in (3), go to step (7).
5. Perform a 25° C (± 6° C) anneal under the bias conditions of (1). The total, cumulative anneal time shall not exceed a time equal to the ratio of the total system dose to the maximum system dose rate, as defined:

$$T_{annealing} = \frac{Total\ Ionizing\ Dose}{Dose - Rate} \tag{8.1}$$

6. Perform functional and parametric testing. If devices pass, go to step (7). If devices fail, return to (5) or fail test.
7. Steps (8)-(10) may be waived if the system total dose is less than 5 krad(Si) and the MOS gate oxide thickness is less than 100 nm, or if it has been demonstrated that long-term interface-trap buildup does not occur in the device or technology under test and that process control has been maintained, or if the parties to the test agree that characterization testing and process control data are sufficient to demonstrate that devices will not experience interface-trap related failures in space. Otherwise, proceed to step (8).
8. Irradiate the devices again under the conditions of (1) to an additional dose equal to 0.5 times the specification.
9. Bake the devices with worst-case static bias for 168 hours at 100° C, or under conditions that have been demonstrated in characterization tests to cause equal or greater degradation in the parameter(s) of interest (e.g. speed, timing, and/or output drive).
10. Final functional and parametric test.

Tests described up to now are also defined as high dose-rate tests, and they are suited for MOS-based devices. Bipolar devices instead suffer the ELDRS (see section 3.3 for further details) thus they require an extremely low dose rate to have reliable test results. Dose rates in this case are around 10 millirad per second (or 36 rad per hour) which makes them quite time consuming: this dose rate showed the best matching between simulated radiation degradation and actual space degradation in LEO [30].

Displacement damage testing is performed by simply characterizing the electrical performances of the device under test, exposing it to an irradiation source, without bias, to a fixed particle fluence and characterizing it after irradiation to determine the parameter degradation. However, for photonic devices the degradation may be application dependent, so these devices may require to be active during irradiation. The radiation source used is generally a mono-energetic proton beam, and the part is irradiated to a fluence greater than the mission equivalent fluence. It is very important to choose adequately the test energy. For low thin shielding a low energy (for example 10 MeV) is adequate because it best represents the environment. For higher shielding, a higher energy is needed (for example 60 MeV), because they have an higher penetration. For solar cells the radiation source is generally a mono-energetic electron (1 MeV is a standard value) or proton (10 MeV is a standard value) beam, and the part is irradiated to a fluence greater than the mission equivalent fluence computed from environmental models. When high energy proton are employed (energy higher than 50 MeV), nuclear reactions can take place inside the target lattice and neutrons can be generated due to spallation (the incoming proton triggered a nuclear reaction that generated the neutron): this can make the device radioactive, so proper handling precautions should be taken.

Single event effects are usually simulated with low energy ions available in particle accelerators. The measurement unit for particle energy is not anymore the elettronVolt (eV) which expresses the kinetic energy the particle acquired but the LET which measures the energy deposed per unit length. Usually particles from accelerators have the same LET as galactic ones, but they have a much lower energy, thus a far lower penetration. These tests are then performed under vacuum and the device package in front of the die is removed. For each value of LET the upset rate is measured by counting a statistically significant number of errors and from this number the cross section, or the upset probability is computed. Upsets can be of multiple types, from memory bit flips, to transient errors in combinatorial or analog circuits, to latch-ups. It is also important to test the part at the application frequency. As feature sizes are reduced with advanced microcircuit technology, the cells become faster, and transients are more likely to be passed as logic signals. A test at lower frequencies may underestimate the device output errors.

Other alternative radiation sources are used for Single Event studies like Californium or high power pulsed lasers. The former generates fission products from ^{252}Cf that has a LET distributed primarily in the range of 41 - 45 MeV cm^2/mg and a range of about 10 um. Due to this low penetration range, Californium could only be used to get a qualitative estimate of SEE sensitivity. This test is useful to check the test set-up before accelerator testing and it is a cheap way to compare SEE sensitivities for part type pre-selection.

Laser light is also an easy way to create SEEs in devices in the laboratory: this technique has been proved to be very useful for SEE mechanism studies and part hardening but

it can also be used for part assurance screening to compare the sensitivity of different manufacturing lots. Laser light and heavy ion charge deposition processes are fundamentally different but their results can be compared [77] thanks to empirical formulas.

Different radiation tests were performed to ensure components survivability in space: in particular a total ionizing dose test campaign was carried out to evaluate radiation tolerance of an operational amplifier (Maxim MAX4092) and a current monitor IC (Texas Instruments INA138). Furthermore a test campaign was carried out to evaluate commercial photo-diodes and LEDs tolerance to displacement damage. The results of these tests are presented in the following sections.

8.2.1 Displacement damage on commercial photo-diodes and LEDs

Displacement damage is the most prominent radiation effect on optical devices and the only way to test device tolerance to it is to evaluate device behavior under a radiation source. Displacement damages in space are generated by protons, in particular by solar protons. Their spectrum in space is a continuous function of energy modeled by the JPL model (see section 2.3.2 for further details) but for ground testing such a wide energy and fluence range is impossible to use. Particle accelerators are able to generate only mono-energetic beams so an equivalence between a continuous spectrum and a mono energetic beam should be computed. The best way to correlate these spectra is to compute the equivalent number of lattice damages induced: this can be achieved by using the NIEL that quantifies the energy lost by a particle generating lattice displacements (see section 3.5 for details). Considering S(E) the protons spectrum as a function of energy (as can be seen for example in Figure 2.15) and NIEL(E) proton NIEL (as can be seen in Figure 3.23), the equivalent mono-energetic proton fluence from a beam of energy E_{beam} can be computed as follows:

$$Fluence \;=\; \frac{\int S(E)\; NIEL(E)\; dE}{NIEL(E_{beam})} \tag{8.2}$$

From the previous formula, for example, the equivalent 10 MeV proton fluence can be computed from proton spectrum inside the satellite (see Figure 6.7) and it is equal to 10^9 particles. The only source available for performing these tests was a 2 MeV proton accelerator at INFN in Padova (IT): these protons have a penetration of only 47 um in plastic so they could only be used to test plastic packages transparency variation under radiation. From the previous equation the equivalent 2 MeV protons fluence was computed in 3×10^8 particles.

From several papers in literature, it was decide to employ double-heterojunction GaAs LEDs and photo-diodes because they showed good tolerance to radiation [34] [55]. The following LEDs were tested: Osram SFH-4502, SFH-4258, SFH-4650 and Vishay TSHG8400. Few photo-diodes were also tested: Advanced Photonics PDI-C172SMF, PDB-C142 and PDB-C142F. Almost no information was anyway found about plastic packages degradation in space, so this research was performed to evaluate it. 2 MeV protons penetration in plastic is only 47 um, while the package is 5 mm tick: since optical degradation of the plastic package is directly proportional to particle fluence and penetration, total optical

Figure 8.4: LED mounted on the particle accelerator component tray.

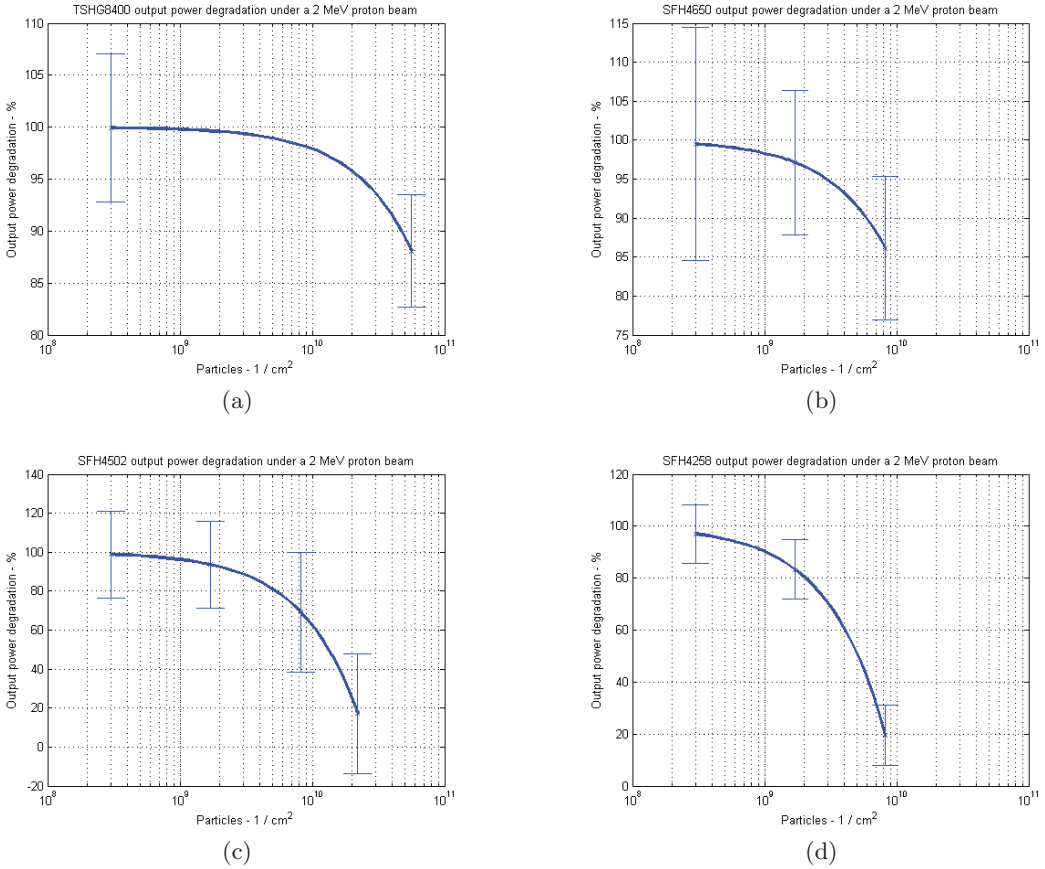


Figure 8.5: LED proton tests results.

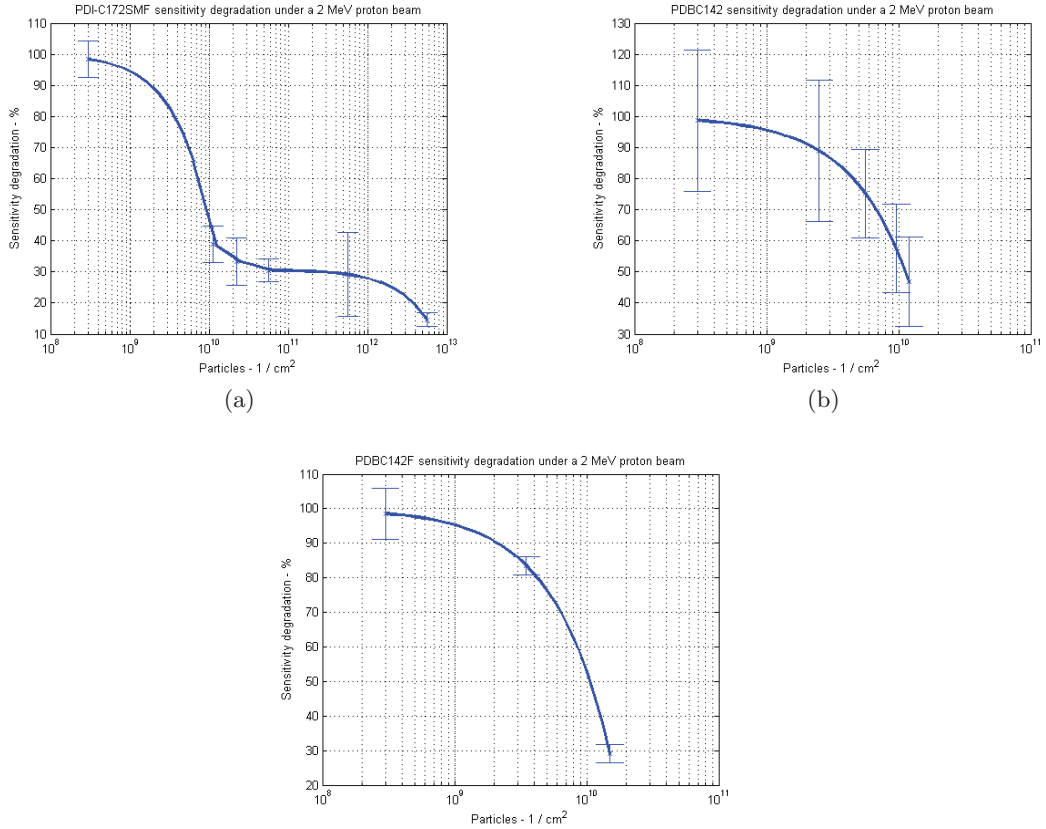


Figure 8.6: Photo-diodes proton tests results.

power reduction could be estimated by multiplying the transmission coefficient reduction times the ratio between package thickness and proton penetration to compute the effective reduction. Anyway, further analysis should be performed with higher energy protons in order to verify the hardness of device die itself.

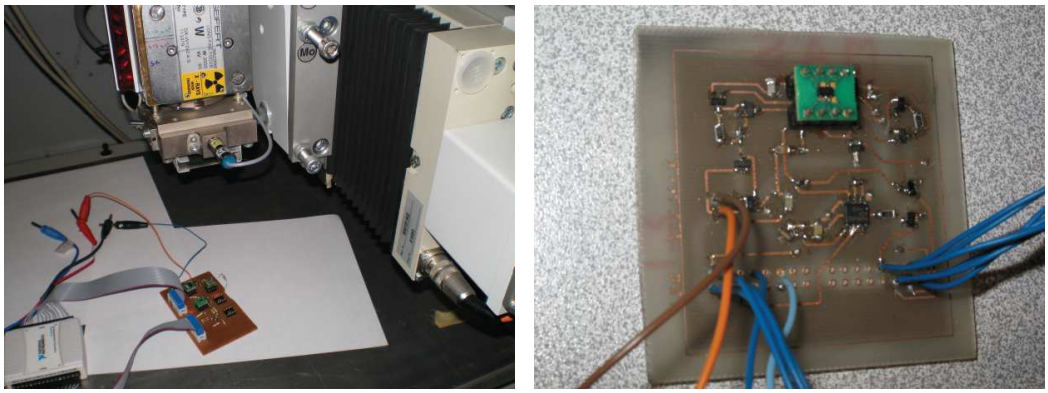
The device had to be mounted on the component tray in the particle accelerator (see Figure 8.4) that is mounted in a vacuum bell at the end of the accelerator beam line. After closing the bell, vacuum is formed inside the bell and then the particle beam could be turned on. Given the accelerator particle fluence and the desired particle number hitting the device, exposure time could be calculated. After an exposition session, the device was removed from the component tray and measured using a test equipment and then it was inserted again in the accelerator for further irradiation. The test equipment is measuring LED output power and photo-diode sensitivity: the results are plotted in Figure 8.5 and 8.6. The error bars represented in these figures represent variations from device to device: these plots were in fact computed from by irradiating three devices per each type to have an average behavior that could represent the whole device lot behavior.

8.2.2 Total Ionizing dose tests

Total ionizing dose tests were also performed to evaluate the tolerance of some components; unfortunately the only available source was an X-ray machine at INFN in Padova (IT) and this is not the best kind of source to be used for bipolar components tests because of the high dose rate that could lead to self-annealing effects (see section 3.3). The test was carried out anyway since it was a good opportunity to start practicing with this kind of tests. Anyway further tests are programmed using a ^{60}Co source.

Before performing this kind of tests, devices needed to be delidded since X-rays penetration is only about 30 um. This procedure is quite complex and dangerous since it requires 90% concentration Nitric Acid to etch the plastic package and it should be handled with particular care and under a chemical fume hood (see section 8.3 for further details).

Once delidded, the two devices have been put under the X-ray source and parameters degradation had been measured using a data acquisition board. The parameters that were measured are: supply current, offset current, offset voltage, bias current and open-loop gain. Unfortunately the devices showed no degradation up to 120 krad which is clearly motivated by the high dose-rate insensitivity of bipolar transistors. Anyway this test helped in refining all the procedure for performing radiation tests. IC delidding procedure will for sure be useful also when single event effects tests will be performed on these ICs since the required know-how for delidding will be already acquired.

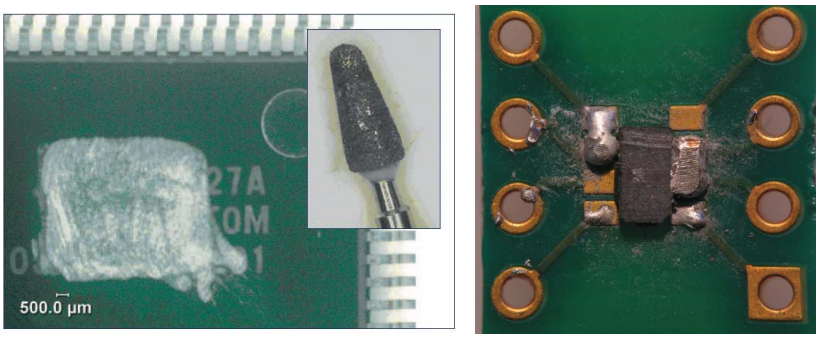


(a) (b)

Figure 8.7: (a) Maxim MAX4092 and (b) Texas Instruments INA138 testing equipment.

8.3 IC delidding

An operative procedure for successful chip delidding has been developed over a period of few weeks that allowed, when tuned, to delid every IC without wastes. This procedure was particularly developed for small ICs (SOT23 or SO8 packages) but can be easily adapted to bigger packages because fewer steps would be required.



(a) (b)

Figure 8.8: Package milling for removing the top layer ((a) is from Philip Yu, EMPF, USA).

The first step is mounting the IC on a small holder, like a small PCB (as in Figure 8.8b) that makes handling a bit easier; this procedure is not particularly needed for big packages (QFP40 or similar) since they can be already handled easily. The first part of the delid should be performed with a milling machine. The mechanical milling is important because the plastic package top layer experienced a particular chemical treatment that made it particularly strong against acid corrosion but it is also a way to faster reduce package thickness. The milled trench should not be deeper than few tenth on millimeter not to damage bonding wires, but this can widely vary from package to package. A simple way to determine how deep to go with the mill is to perform a device electronic test after every milling: if the device is still operating it is safe to continue. In this way the first device milled will be probably damaged, but afterwards trench depth is determined and it will be easier to perform the chemical etch. On SO8 or SOT23 packages the trench should not be deeper than 0.3 mm while for other packages it can be also 1 mm deep. The trench should be shaped to be a bit bigger than die size and should be placed above it: this is to be done in big packages but is almost impossible for SOT23 ones. The second purpose of this trench is to hold the hot acid above the die and prevent it to go etching other parts of the package, like the pins that will be eroded in almost no time (see Figure 8.8a). For small packages a different techniques should be used: since the trench is so small that cannot completely hold a drop of acid, an external container should be built as can be seen in Figure 8.9. The package has been covered with siliconic resin (QSil550SB from ACL Silicones) and it is only accessible from the top, exactly above the die: in this way the resin will act as a container and will prevent acid drops to erode the pins: this solution is particularly important since acid drops cannot be poured with precision during the etching process and errors are common, but with the device completely covered they will not create problems. A siliconic resin has been selected since it is extremely hard to acid corrosion, while epoxy resin will be eroded in few minutes, so it should not be used. Siliconic resin anyway has some bonding problems and may generate small gaps between the package and the resin: proper curing is fundamental to eliminate this problem since

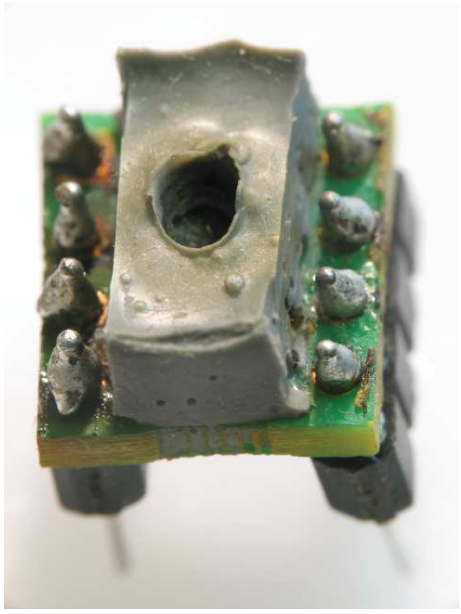


Figure 8.9: Siliconic resin cover for a SO8 package.

it could cause small quantities of acid to flow under the resin for capillarity and etch the pins.

At this stage the device is ready for the chemical etch. The following items are needed for properly completing this phase:

- a programmable heating plate able to reach a temperature of 100° C - 110° C and covered with an anti-acid layer (Alumina or ceramic is perfect);
- anti-acid gloves and clothes;
- anti-acid mask or at least glasses;
- a chemical fume hood;
- anti-acid tweezers;
- 3 anti-acid drop dispenser (glass made ones are perfect);
- 90% pure nitric acid (HNO_3) or sulfuric acid (H_2SO_4); even a mix of the two is good but acids with lower concentration should be avoided because etching time is considerably longer;
- acetone ($(CH_3)_2CO$);
- de-ionized water;
- 3 Becker for containing water, acid and acetone;

- a magnifier or a microscope;
- an ultrasonic cleaner.

Before beginning the etching phase, it is necessary to get dressed with acid protections. Then the heating plate should be put under the hood and heated up to 100° C (temperature can be slightly increased if the acid protecting layer is thick). Acid, acetone and water should be poured in the Becker with extreme caution and the three drop dispensers assigned to each of the three Becker and they should not be exchanged, to avoid liquid contamination.

Now the device can be put on the heating plate and, after few minutes needed to heat it up, etching can start. A single drop of acid should be poured onto the device (acid should immediately start boiling, so maximum caution is needed, but if this is not happening, heating plate temperature should be increased). After about one minute the corrosion should be stopped using one or two drops of acetone. This procedure should be repeated as many times as needed until the whole die becomes visible and almost clean in surface. The device should now be cleaned with water and observed under the microscope: if the surface is almost clean from plastic, final ultrasonic cleaning can be performed, otherwise other etching steps are required.

Ultrasonic cleaning should be performed by pouring acetone in the cleaner and by leaving the device in the cleaning machine for about five to ten minutes. If the surface is completely clean from plastic wastes the etching is complete, otherwise some further etching steps are required.

After the device has been cleaned, the siliconic resin can be removed: the simplest solution is to simply pull it apart from the device with a tweezer: it is important not to touch the bonding wires since they are extremely fragile. If this process is completed without problems, the etched device should look like in Figure 8.10.

The etching process can also be performed by automated machines, called jet etchers, but the are generally quite expensive (30k - 100k euros).

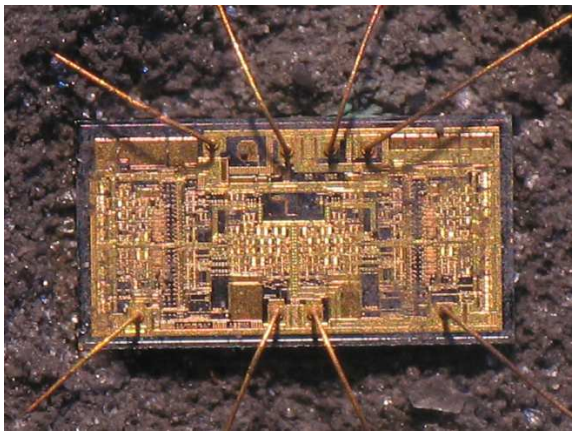


Figure 8.10: Maxim MAX4092 die exposed with the etching process.

Chapter 9

Conclusions

The work performed to complete this thesis allowed to achieve a deeper understanding of space systems complexity and how to tackle many practical problems related to complex systems development.

A low cost approach to space development is currently pursued by many institutions around the world to allow a broader access to space, even to organizations with a low budget. Cost reduction is a complex problem and it was analyzed from two points of view, product development and production costs. Two approaches were proposed with this thesis to address these problems and they were applied to a practical case, the development of an innovative satellite architecture. In particular, modularity and scalability were proposed to greatly reduce development costs by sharing them between multiple missions and production costs were addressed by proposing the use of commercial components instead of military / space-born ones. Single device reliability can be significantly lower in this case, but this disadvantage can be overcome with proper design techniques, as it was shown in this work.

The main contributions to the development of the AraMiS architecture were multiple: in a first phase, the space environment has been analyzed and conditions inside the mechanical structure were computed, with particular attention to cosmic particle fluxes. The development phase was then centered around several satellite sub-systems: a latch-up protection system, a wireless data communication bus and a power management and distribution system.

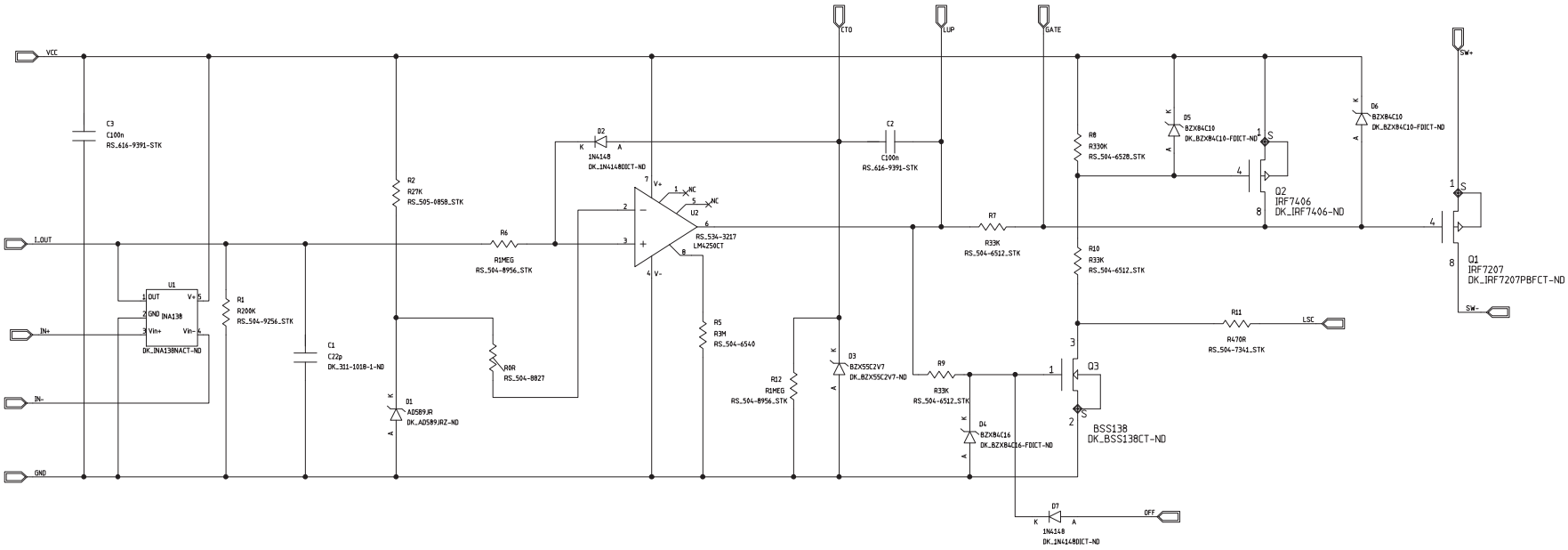
The latch-up protection system was developed for limiting galactic particles bad effects on commercial ICs and the final device was also produced commercially by Neohm Componenti. The other sub-systems were developed to be part of the AraMiS architecture and to solve two critical problems, on-board data transfer and power management and distribution.

Beside the development of these sub-systems, qualification test were also performed to ensure that the developed systems could stand the harsh space environment.

Appendix A

Latch-up protection circuit

This section includes the datasheet of the 1B127 module, used to protect commercial devices from latch-up. This device was developed as part of the work presented in this thesis (see section 7.1 for further details), and it was afterwards produced by Neohm Componenti [96][73].



DWG. NO.	SH.	REV

REVISIONS

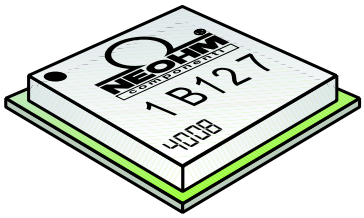
REV	DESCRIPTION	DATE	APPROVED

CONTRACT NO.		Schematic: anti_latchup Root: anti_latchup
APPROVALS	DATE	
DRAWN	11/04/2010 15.34.13	Project name: anti_latchup Lib: R:\AraMIS_Mentor_Lib\AraMIS_Mentor_Lib.lmc
CHECKED		SIZE: D; FSCM NO.; DWG. NO.; REV
ISSUED		SCALE; SHEET: 01/01

H=BND, V=4LA

MCM

Latch-Up and Overcurrent Protection System



1B127

FEATURES

- Wide Supply voltage range: 2.7 – 36V
- Low power consumption
- Maximum current: 2 A
- Programmable current threshold
- Only one external resistor required
- Wider voltage and current range with external transistor
- Load disable input
- Current monitor output
- Programmable off time *(with external capacitor)*
- Filters out inrush current *(with external capacitor)*
- Controlled slew-rate *(with external capacitor)*
- Guaranteed radiation tolerance up to 30krad TID
- Wide temperature range: –40°C to +125°C

APPLICATIONS

- Latch-up monitoring and protection systems
- Overcurrent protection
- Load switch with over-current monitoring
- Load switch with slew-rate limitation
- Protects low-cost devices from latch-ups

DESCRIPTION

Latch-up is a catastrophic phenomenon which affects CMOS devices. It is caused by high energy particles or ionizing radiations or other causes which trigger the parasitic transistors, mostly in CMOS devices that are not radiation-hardened.
The 1B127 is a latch-up and overcurrent protection system suited for high radiation level environments, such as Low Earth Orbit satellites.

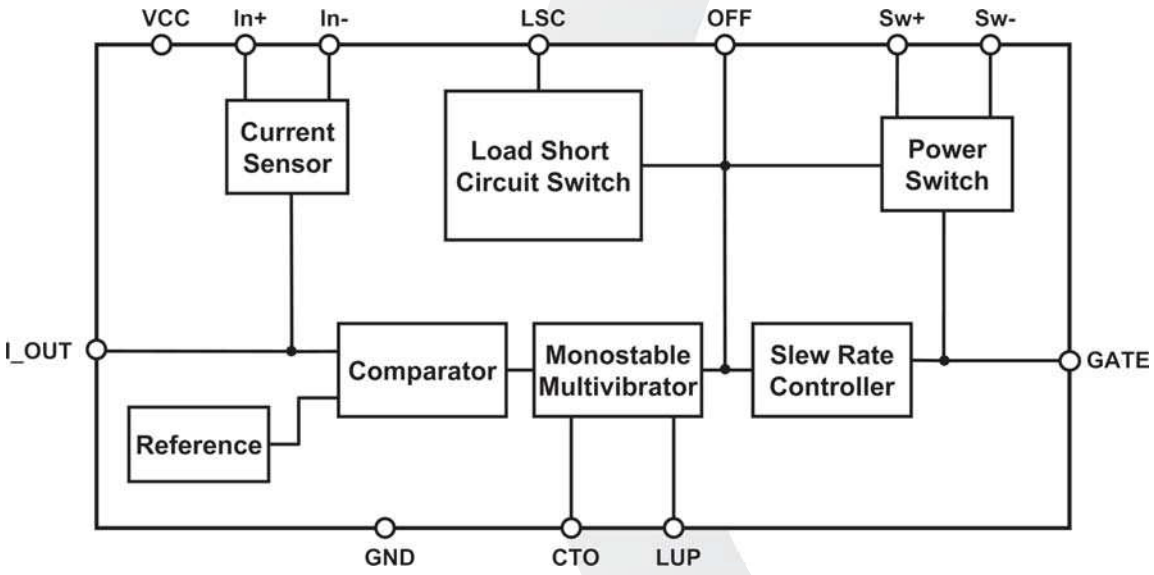


Figure 1. Block diagram.

NEOHM COMPONENTI S.r.l.

Via Torino 217
10040 Leinì (TO) Italia

Tel: +39 011 9974022
Fax: +39 011 9974982

www.neohmcomponenti.com
info@neohmcomponenti.com

UNI EN ISO 9001: 2000

It allows using low-cost devices (COTS) in radiation environments like space applications, high-energy physics experiments and biomedical equipment, by protecting them against the effects of latch-up. The system monitors the current flowing through an external sense resistor (Rs) and whenever a threshold value is reached the internal pass transistor switches off the load. After a user-defined recovery time, the pass transistor automatically turns on again restoring power supply to the load, once the latch-up has faded away.
The device can also be used as a current-limited load switch by sending a TTL signal to the OFF pin. For higher voltages or currents, an external PMOS pass transistor can be used.
Inrush current effects or short current spikes can be filtered away. Allowed inrush charge and recovery time can be set using external capacitors.
In addition, the slew-rate while turning on load supply voltage can be limited using an external capacitor to program the inrush current.
The device is available in a small 13x13 mm plastic 100BGA package and it is available for the extended temperature range (-40 °C to +125 °C). It has a low quiescent current, which makes it suitable for low power systems.

ABSOLUTE MAXIMUM RATINGS

Supply voltage (VAL)	-0.3V ÷ 36V
Analog Input Vin+ , Vin-	Vcc-0.3V÷Vcc+0.3V
Differential voltage	-40V ÷ 2V
OFF input	-0.3V ÷ 36V
Internal Load Switch Max V	18V
Load Curr. with Internal Switch	2A
Max current sink	300mA
Operating Temperature	-40°C ÷ +125°C
Storage Temperature	-50°C ÷ +150°C
Junction Temperature	+150°C

OPERATION

The 1B127 device (see block diagram in fig. 1) monitors the current on the high side by measuring the differential voltage across a shunt resistor (pins IN+ and IN-).
The floating differential voltage is transformed into a ground-referenced voltage on the I_OUT pin. An external capacitor connected to the I_OUT pin allows filtering short current spikes, while an external resistor connected to the same pin allows changing the gain of the current monitor.
A comparator detects when the current overpasses a threshold value, triggering a monostable for a fixed time period. An external capacitor across pins CTO and LUP allows to increase that time period.
The output of the monostable (which is available as a logic signal on the LUP pin) turns off the internal power switch (across SW+ and SW- pins), disconnecting the load from the power supply. An additional output LSC optionally sinks current from the load, when load supply is removed, to discharge capacitors which might be present on the load or to sink leakage currents which might come from other input/output connections of the load.
After the monostable resets to its normal condition, the power switch is turned on again to supply power to the load. The slew-rate controller optionally limits

The OFF pin may optionally turn load power off under user control.
The control of the power switch is also available via the GATE pin to drive an external power transistor, augmenting load voltage or load current capabilities.

Figure 2 shows the basic configuration of the circuit,

$$R_S = \frac{V_{th}}{I}$$

Where I is the current through the sense resistor (that is, load current), and Vth is the device threshold voltage. The power rating of this resistor shall be at least:

$$P_{R_S} = \frac{V_{th}^2}{R_S}$$

The I_OUT voltage can be calculated as:

$$V_{I\_OUT} = I \cdot R_S \cdot Gain$$

where nominal Gain without ext components is 30.

The Gain can be modified (together with threshold voltage) as shown in fig. 3, by adding an external resistor Rth:

$$R_G = \frac{R_{th} R_{OUT}}{R_{th} + R_{OUT}}$$

$$Gain = \frac{R_G}{5k\Omega}$$

$$V_{th} = \frac{600mV}{Gain}$$

where ROUT is the output impedance of the I_OUT output (nominal value 150k).

By adding the CIR capacitor, as shown in Figure 3, the I_OUT is filtered by a low-pass filter with cut-off frequency of:

$$f_T = \frac{1}{2\pi (C_{IR} + 100pF) \cdot R_G}$$

allowing an inrush charge without triggering of:

$$Q = \frac{(C_{IR} + 100pF) \cdot 3kV\Omega}{R_S}$$

Adding the CTC capacitor, as shown in Figure 3, the recovery time can be increased according to the following equation:

$$C_{TC} = \frac{T_{OFF}}{100k\Omega} - 100nF$$

NEOHM COMPONENTI S.r.l.

Via Torino 217
10040 Leinì (TO) Italia

Tel: +39 011 9974022
Fax: +39 011 9974982

www.neohmcomponenti.com
info@neohmcomponenti.com

To control the slew-rate of load voltage to prevent high inrush currents into the load, an external capacitor (C_{SR} in Figure 3) can be used and its value is given by:

$$C_{SR} = \frac{T_{rise} - 165us}{33k\Omega}$$

In more than 2A or more than 18V supply voltage are needed, an external PMOS pass transistors can be used (as shown in Figure 4). The external pass transistor is driven by the SW+, SW- and GATE pins shorted together.
In this case the CSR capacitor should be calculated according to the following equation (CG is the external PMOS gate capacitance):

$$C_{SR} = \frac{T_{rise}}{3.3 \cdot 10^4 \Omega} - C_G$$

Adding an external diode and transistor as shown in fig. 5, locks the the monostable in this way, once the device has been triggered by an overcurrent, only turning the external MOS on resets the 1B127 device to its normal operation.

Figure 6 shows a fault-tolerant configuration where a pair of 1B127 devices are paralleled to increase fault tolerance. A short circuit on one of the pass transistors (internal or external) and a stuck-open fault on the 1B127 mos driver can be tolerated. In case of stuck-closed on the 1B127 driver the latch-up protection is no more effective but the load can still be turned on.

NEOHM COMPONENTI S.r.l.

Via Torino 217
10040 Leinì (TO) Italia

Tel: +39 011 9974022
Fax: +39 011 9974982

www.neohmcomponenti.com
info@neohmcomponenti.com

ELECTRICAL CHARATERISTICS

All characteristics at T_A= 25 °C, V_{CC} = 5 V, unless otherwise noted. $^*V_{sense}$= V_{in+} - V_{in-}*

Parameter	Condition	1B127			Unit
		Min	Typ	Max	
INPUT					
Differential Input Voltage			20	500	mV
Common-Mode Input Range		Vcc -100mV	Vcc	Vcc +100mV	V
Common-Mode Rejection	V_{in+} = 2.7 V to 30 V, V_{sense} = 20 mV (*)	100	120		dB
Threshold voltage V_{th}	without external components	29	30	31	mV
over temperature range		28		32	mV
temperature drift			1		µV / °C
drift vs. common mode (V_{in+})	V_{in+} = 2.7 V to 36 V		0.04	4	µV / V
Input Bias Current			2		µA
over temperature range				10	µA
POWER SUPPLY					
Supply Voltage	Vcc	2.7		36	V
Supply Current	Vcc = 3 V			200	µA
	Vcc = 5 V			360	µA
	Vcc = 15 V			700	µA
	Vcc = 36 V			2	mA
POWER SUPPLY					
Continuous Load Current				2	A
Peak Load Current	1 ms pulses, DC < 1%			3	A
Supply Voltage		2.7		20	V
Leakage Current	-40 °C < T_A < 125 °C			100	µA
TIMING					
Turn-off time:	Inrush current = 2 I_{th}			100	µs
	C_{IR} = 0				
	C_{IR} = 2.7 nF			1	ms
Recovery Time	without external components		10		ms
	C_{TC} = 1 nF		110		ms
SINK CURRENT					
Sink Current	Vcc = 2.7 V	5			mA
	Vcc = 15 V ÷ 36 V			36	mA
Peak Current	1 ms pulses, DC < 1%			500	mA
I_OUT					
Gain, without external components		29	30	31	
	over Temperature	28	30	32	
Output Impedance R_{out}			150		KΩ
Output Swing			0.6		V
THERMAL					
Thermal Resistance between power switch junction to balls	Thermal Resistance θ_{JC}		12		°C / W
Thermal Resistance between power switch junction to ambient	Thermal Resistance θ_{JA} Mounted on std 1.6mm PCB		34		°C / W
ESD					
All Pins	HBM 1.5KΩ, 100pF	2			KV

NEOHM COMPONENTI S.r.l.

Via Torino 217
10040 Leinì (TO) Italia

Tel: +39 011 9974022
Fax: +39 011 9974982

www.neohmcomponenti.com
info@neohmcomponenti.com

APPLICATION DIAGRAMS

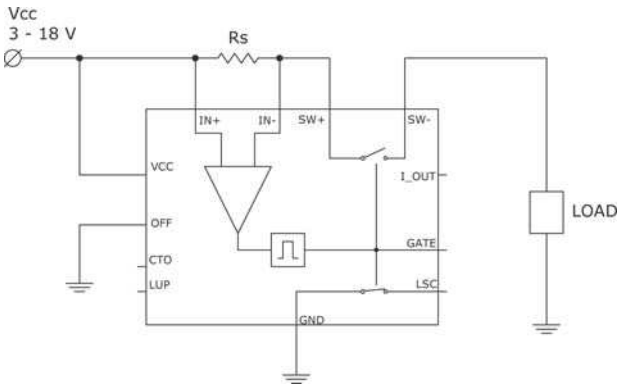


Figure 2. Basic device configuration.

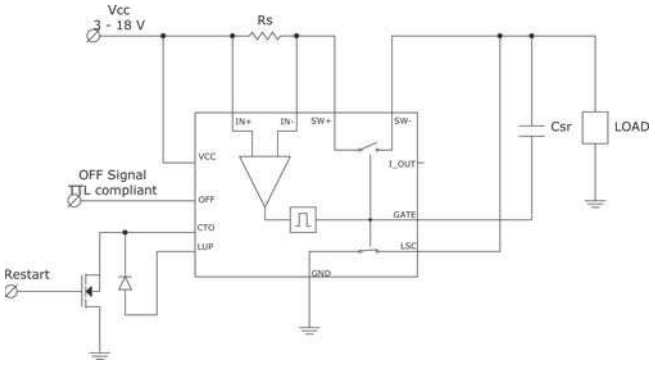


Figure 5. Current protection with restart signal.

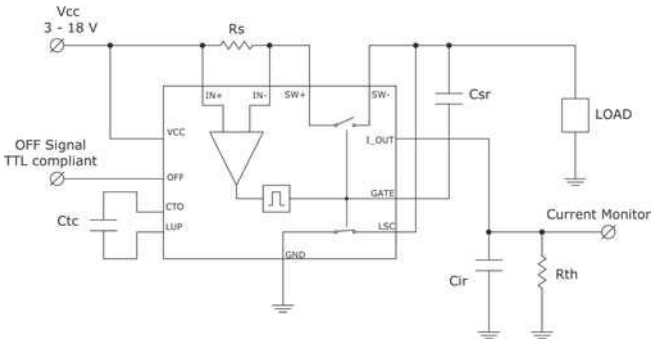


Figure 3. Circuit with T_{OFF} and Slew-Rate control.

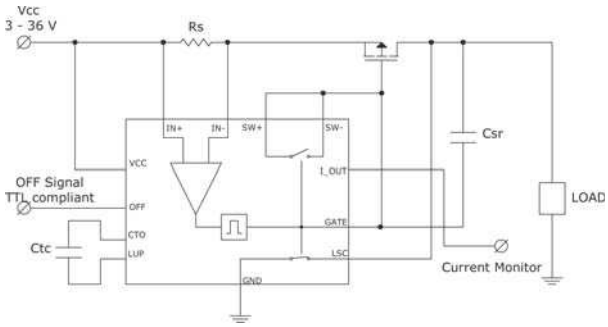


Figure 4. Circuit with an external Pass Transistor to increase load current.

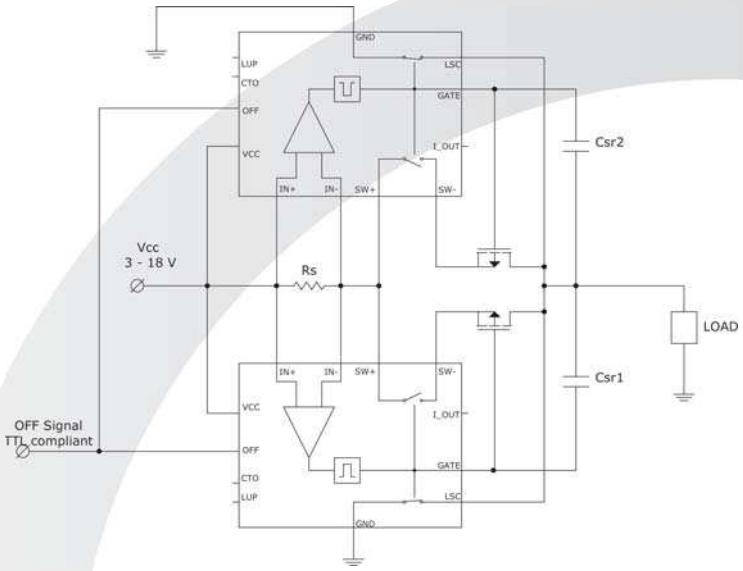


Figure 6. Fault-tolerant configuration with two parallel branches and two series transistors.

NEOHM COMPONENTI S.r.l.

Via Torino 217
10040 Leinì (TO) Italia

Tel: +39 011 9974022
Fax: +39 011 9974982

www.neohmcomponenti.com
info@neohmcomponenti.com

TYPICAL CHARACTERISTICS

All characteristics at T_A = 25 °C, V_{CC} = 5 V, unless otherwise noted.

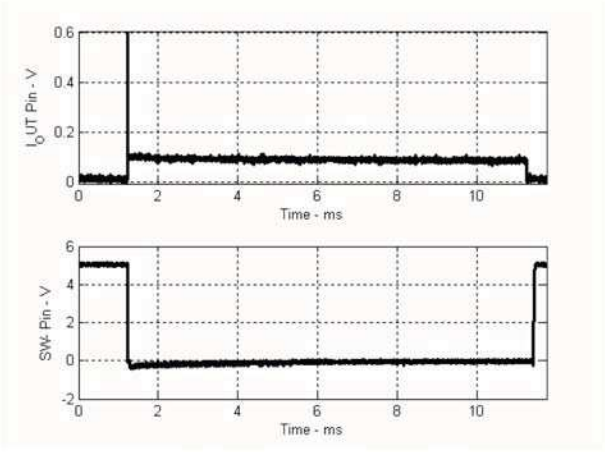


Figure 7. Latch-up event with V_{CC} = 5 V.

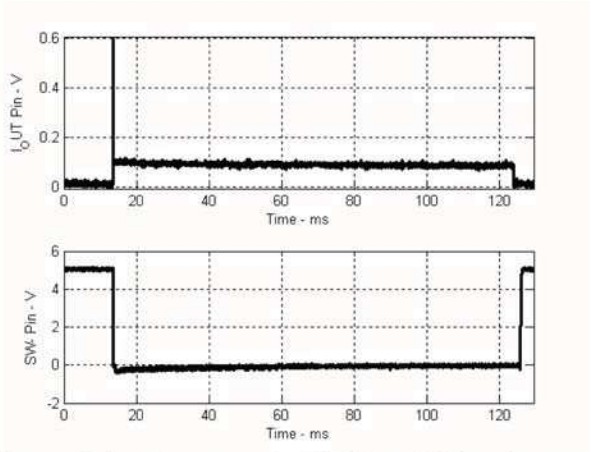


Figure 8. Latch-up event with V_{CC} = 5 V and C_{TC} = 1 uF.

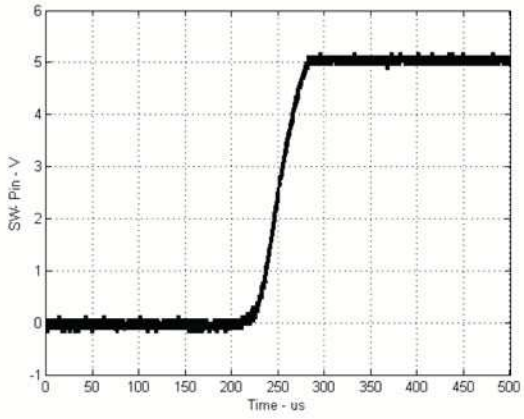


Figure 9. Turn-on time with V_{CC} = 5 V and no external component.

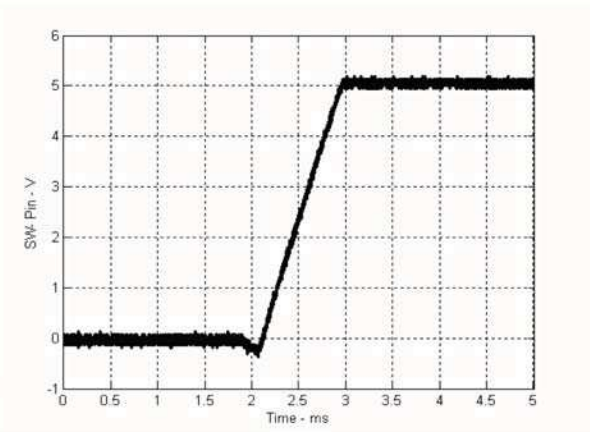


Figure 10. Limited slew-rate with V_{CC} = 5 V and C_{SR} = 2.7 nF.

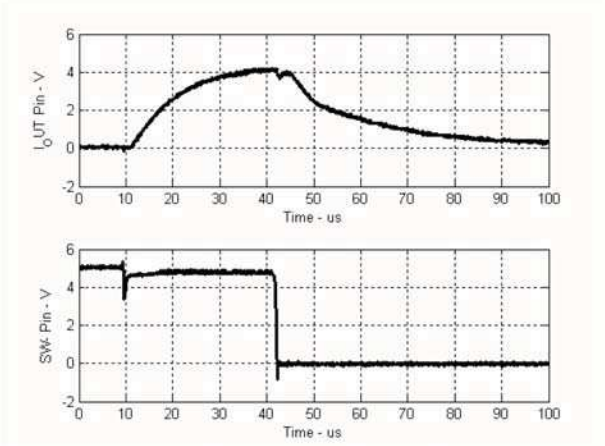


Figure 11. Turn off time with V_{CC} = 5 V.

NEOHM COMPONENTI S.r.l.

Via Torino 217
10040 Leinì (TO) Italia

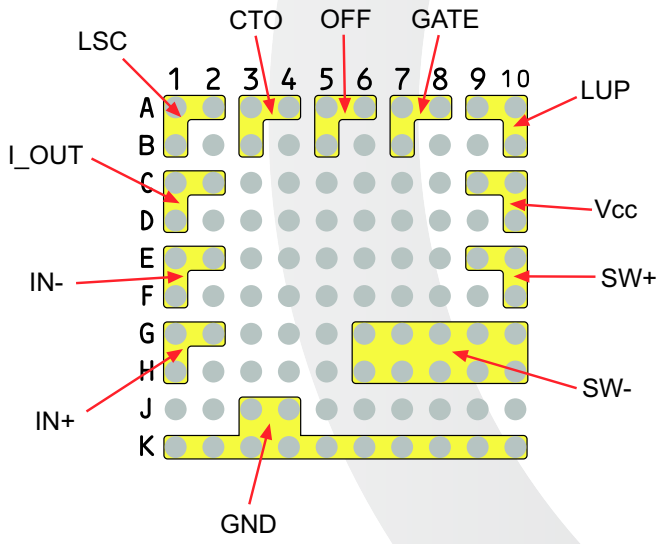
Tel: +39 011 9974022
Fax: +39 011 9974982

www.neohmcomponenti.com
info@neohmcomponenti.com

PINS DESCRIPTION

PIN Name	Documentation
GND	GROUND pin for the device
GATE	OUTPUT: to be connected to the gate of the external PMOS transistor, when used. It is also used to connect an optional external capacitor to limit slew-rate of the **SW-** pin. It can also be used to connect an optional external resistor to reduce slew-rate.
LSC	POWER: load sink. Sinks a current from load during a latchup condition.
VAL	SUPPLY pin for the device
CTO	INPUT: positive terminal of capacitor to set the turn-on time. The other terminal should be connected to **LUP** pin.
SW-	POWER: negative side of load power switch. It should be connected to the load, except when an external PMOS transistor is used. It can also be used to connect an optional external capacitor to limit the slew-rate of the **SW-** pin.
SW+	POWER: positive side of load power switch. It should be connected at the negative side of external shunt resistor, except when an external PMOS transistor is used
LUP	OUTPUT: set at logic-1 when a latchup or overload condition has been detected. Remains high for the whole turn-on time. Logic-1 level is coincident with supply voltage. It can also be used to connect an optional external capacitor to set the turn-on time.
I_OUT	OUTPUT: Ground-referenced voltage proportional to the load current. When load power is turned off due to a latchup or overload condition, output voltage is unpredictable. It can also be used to connect an optional external capacitor to set the **allowed inrush current.**
OFF	INPUT: load enable. When at logic-1 enables load power supply. When al logic-0 turns load power off
IN-	INPUT: negative side of external shunt resistor. It should be connected to the Load via the Power Load Switch.
IN+	INPUT: positive side of external shunt resistor. It should be connected to power supply generator.

PINS LAYOUT (Top View)



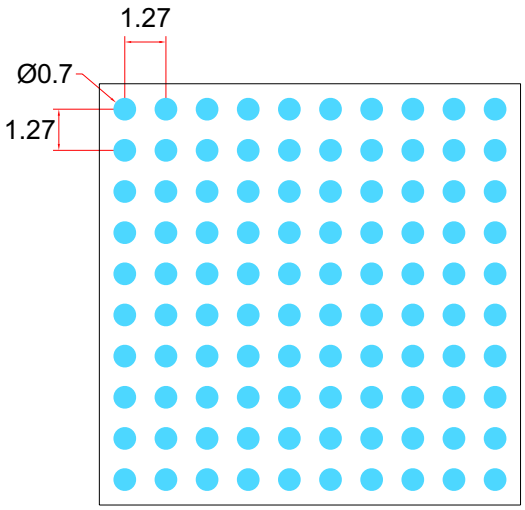
NEOHM COMPONENTI S.r.l.

Via Torino 217
10040 Leinì (TO) Italia

Tel: +39 011 9974022
Fax: +39 011 9974982

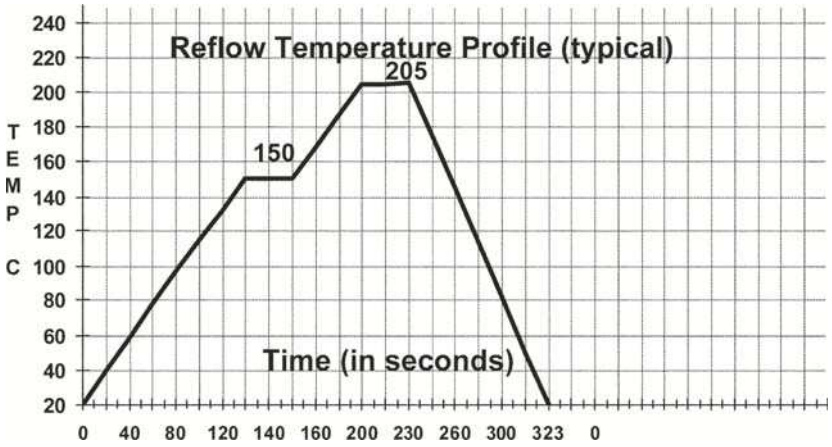
www.neohmcomponenti.com
info@neohmcomponenti.com

SOLDERING PATTERN



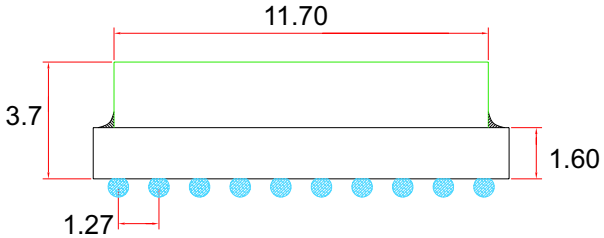
SOLDERING PRECAUTIONS

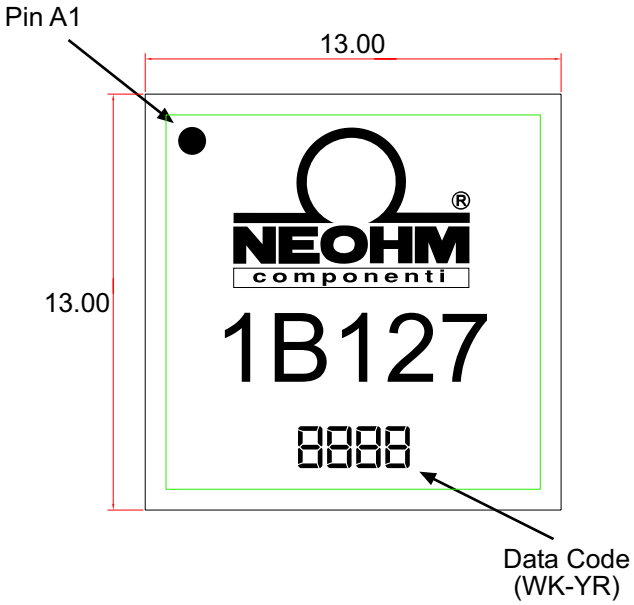
Before soldering put the 1B127 in oven @ 50° for 30 minutes



PACKAGE OUTLINE

(All dimensions are in mm)





ESD SENSITIVITY

The device can be damaged by ESD: we suggest to handle all integrated circuits with appropriate precautions.

Total or partial failure of the system can be originated by wrong installation and handling procedures.

ACKNOWLEDGEMENTS

This device has been designed in cooperation with the Space Avionics Technology Center of Dept. of Electronics of Politecnico di Torino, and is supported by project MAESS funded by Regione Piemonte (I).

NEOHM COMPONENTI S.r.l.

Via Torino 217
10040 Leinì (TO) Italia

Tel: +39 011 9974022
Fax: +39 011 9974982

www.neohmcomponenti.com
info@neohmcomponenti.com

Bibliography

[1] D. Biesecker, "Solar cycle 24 consensus prediction," *Space Weather Workshop*, April 2007.

[2] Wehrli 1985 AM0 Spectrum. [Online]. Available: http://rredc.nrel.gov/solar/spectra/am0/wehrli1985.new.html

[3] MSIS-86 at uk solar system data centre.

[4] P. R. G. et al., "Earthshine Observations of the Earth's Reflectance," *Geophysical Research Letters*, pp. 1671–1674, 2001.

[5] Space Environment Information System. [Online]. Available: http://www.spenvis.oma.be

[6] D. Del Corso, C. Passerone, L. M. Reyneri, C. Sansoè, M. Borri, S. Speretta, and M. Tranchero, "Architecture of a Small Low-Cost Satellite," *10th Euromicro Conference on Digital System Design*, pp. 428–431, August 2007.

[7] M. Caldera, S. Corpino, M. Masoero, and N. Viola, "Analisi termica di un nanosatellite universitario," in *Proceedings of XVII Congresso Nazionale AIDAA*, Volterra, Italy, September 2005.

[8] T. F. LAT and F. G. Collaborations, "Fermi observations of high-energy gamma-ray emission from GRB 080916C," *Science*, pp. 1688–1693, March 2009.

[9] H. Hoerlin, "United States High-Altitude Test Experiences: A Review Emphasizing the Impact on the Environment," Los Alamos Scientific Laboratory, Tech. Rep. LA-6405, October 1976.

[10] M. Walt, *Introduction to Geomagnetically Trapped Radiation (1st ed.)*. New York: Cambridge University Press, 1961.

[11] J. I. Vette, "The AE-8 trapped electron model environment," National Space Science Data Center, Tech. Rep. WDC-A-R&S 91-24, 1991.

[12] S. D. M. and J. I. Vette, "AP-8 trapped proton environment for solar maximum and solar minimum," National Space Science Data Center, Tech. Rep. WDC-A-R&S 76-06, 1976.

[13] M. Kruglanski, "Engineering tool for trapped proton flux anisotropy evaluation," *Radiation Measurements*, vol. 26, pp. 953–958, Novenmer 1996.

[14] J. Wertz and W. Larson, *Space Mission Analysis and Design*. New York: Spriner, 2007.

[15] J. Feynman, G. Spitale, J. Wang, and S. Gabriel, "Interplanetary proton fluence model: Jpl 1991," *Journal of Geophysical Research*, pp. 281–294, 1993.

[16] L. Rosenqvist, A. Hilgersa, H. Evans, E. Daly, M. H. R. S. ad R. Zwickl, S. Bourdarie, and D. Bosher, "A toolkit for updating interplanetary proton cumulated fluence models," *Journal of Spacecraft and Rockets*, no. 6, November-December 2005.

[17] J. Feynman, A. Ruzmaikin, and V. Berdichevsky, "The jpl proton fluence model: an update," *Journal of Atmospheric and Solar-Terrestrial Physics* -, pp. 1679–1686, 2002.

[18] J. H. A. Jr., "Cosmic Ray Effects on MicroElectronics," Naval Research Laboratory, Tech. Rep. Memorandum Report 5901, 1986.

[19] Cubesat Community Website. [Online]. Available: http://cubesat.calpoly.edu

[20] E. Perez, *Vega User's manual*, March 2006.

[21] SRIM, The Stopping and Range of Ions in Matter. [Online]. Available: http://www.srim.org

[22] C. Leroy and P. Rancoita, *Principles of Radiation Interaction in Matter and Detection*. World Scientific Publishing, 2009.

[23] Monte Carlo Simulations of Electron Trajectories in Solids. [Online]. Available: http://www.gel.usherbrooke.ca/casino/index.htmlg

[24] GEANT4. [Online]. Available: http://geant4.cern.ch/

[25] Large Hadron Collider, CERN. [Online]. Available: http://lhc.web.cern.ch/lhc/

[26] MULASSIS. [Online]. Available: reat.space.qinetiq.com/mulassis/

[27] ESTAR, Electrons Stopping power and Range. [Online]. Available: http://physics.nist.gov/PhysRefData/Star/Text/ESTAR.html

[28] T. Bortfeld and W. Schlegel, "An analytical approximation of depth-dose distributions for therapeutic proton beams," *Physics in Medicine and Biology*, vol. 41, no. 8, pp. 1331 – 1339, August 1996.

[29] S. T. Pantelides, L. Tsetseris, S. Rashkeev, X. Zhou, D. Fleetwood, and R. Schrimpf, "Hydrogen in MOSFETs - A primary agent of reliability issues," *Microelectronics Reliability*, vol. 47, no. 6, pp. 903 – 911, June 2006.

[30] T. Turflinger, A. Campbell, W. Schmeichel, R. Walters, J. Krieg, J. Titus, M. Reeves, P. Marshall, and R. Pease, "Eldrs in space: an updated and expanded analysis of the bipolar eldrs experiment on mptb," *Nuclear Science, IEEE Transactions on*, vol. 50, no. 6, pp. 2328 – 2334, dec. 2003.

[31] Jet Propulsion Laboratory, *The NASA ASIC Guide: Assuring ASICS for Space*. Pasadena, CA: Jet Propulsion Laboratory, California Institute of Technology and National Aeronautics and Space Administration, 1993.

[32] T. May and M. Woods, "Alpha-particle-induced soft errors in dynamic memories," *Electron Devices, IEEE Transactions on*, vol. 26, no. 1, pp. 2 – 9, jan 1979.

[33] J. Srour, C. Marshall, and P. Marshall, "Review of displacement damage effects in silicon devices," *Nuclear Science, IEEE Transactions on*, vol. 50, no. 3, pp. 653 – 670, june 2003.

[34] S. Messenger, R. Walters, E. Burke, G. Summers, and M. Xapsos, "Niel and damage correlations for high-energy protons in gallium arsenide devices," *Nuclear Science, IEEE Transactions on*, vol. 48, no. 6, pp. 2121 – 2126, dec 2001.

[35] NASA Advanced Missions Cost Model. [Online]. Available: http://cost.jsc.nasa.gov/AMCM.html

[36] M. N. Sweeting, "Space at surrey: micro-mini-satellites for affordable access to space," *Air & Space Europe*, vol. 2, no. 1, pp. 38 – 52, Jan - Feb 2000.

[37] Pumpkin CubeSat Kit. [Online]. Available: http://www.cubesatkit.com/

[38] ISIS CubeSat Shop. [Online]. Available: http://www.cubesatshop.com/

[39] Surrey Satellite Technology. [Online]. Available: http://www.sstl.co.uk/

[40] A. Toorian, K. Diaz, and S. Lee, "The cubesat approach to space access," *IEEE Aerospace Conference*, pp. 1–14, March 2008.

[41] Space Transportation Costs: Trends in Price Per Pound to Orbit. [Online]. Available: http://www.futron.com/pdf/resource_center/white_papers/FutronLaunchCostWP.pdf

[42] NASA, "Space radiation effects on electronic componets in low-earth orbit," *Preferred reliability Practices no. PD-ED1258*, 1996.

[43] J. Bongim, A. Sutton, R. Diestelhorst, G. Duperon, J. Cressler, J. Black, T. Haeffner, R. Reed, M. Alles, R. Schrimpf, D. Fleetwood, and P. Marshall, "The application of rhbd to n-mosfets intended for use in cryogenic-temperature radiation environments," *Nuclear Science, IEEE Transactions on*, vol. 54, no. 6, pp. 2100–2105, Dec. 2007.

[44] F. Faccio and G. Cervelli, "Radiation-induced edge effects in deep submicron cmos transistors," *Nuclear Science, IEEE Transactions on*, vol. 52, no. 6, pp. 2413–2420, Dec. 2005.

[45] W. R. D. J. et al., *Hardening semiconductor components against radiation and temperature*. Park Ridge, NJ: Noyes Data Corp., 1989.

[46] C. S. Dyer and G. R. Hopkinson, "Space Radiation Effects For Future Technologies and Missions," QinetiQ Space Departement, Tech. Rep. QINETIQ / KI / SPACE / TR010690 / 1.1, 2001.

[47] N. S. Saks, M. G. Ancona, and J. A. Modolo, "Generation of interface states by ionizing radiation in very thin mos oxides," *Nuclear Science, IEEE Transactions on*, vol. 33, no. 6, pp. 1185–1190, Dec. 1986.

[48] M. DeLaus and W. Combs, "Total dose and SEU results for the AD8001, a high performance commercial op-amp fabricated in a dielectrically isolated, complementary bipolar process," in *Radiation Effects Data Workshop, 1994 IEEE*, July 1994, pp. 104–109.

[49] H. Hjalmarson, R. Pease, S. Witczak, M. Shaneyfelt, J. Schwank, A. Edwards, C. Hembree, and T. Mattsson, "Mechanisms for radiation dose-rate sensitivity of bipolar transistors," *Nuclear Science, IEEE Transactions on*, vol. 50, no. 6, pp. 1901–1909, Dec. 2003.

[50] M. Shaneyfelt, R. Pease, J. Schwank, M. Maher, G. Hash, D. Fleetwood, P. Dodd, C. Reber, S. Witczak, L. Riewe, H. Hjalmarson, J. Banks, B. Doyle, and J. Knapp, "Impact of passivation layers on enhanced low dose rate sensitivity and pre-irradiation elevated temperature stress effects in bipolar linear ics," *Nuclear Science, IEEE Transactions on*, vol. 49, no. 6, pp. 3171–3179, Dec 2002.

[51] ESA, "European Standard for Space Electrical and Electronic Equipments," European Space Agency, Tech. Rep. ECSS-Q-ST-30-11C, July 2008.

[52] *Centro Elettrotecnico Sperimentale Italiano*. [Online]. Available: http://www.cesi.it

[53] J. Srour, J. Vendura, G.J., D. Lo, C. Toporow, M. Dooley, R. Nakano, and E. King, "Damage mechanisms in radiation-tolerant amorphous silicon solar cells," *Nuclear Science, IEEE Transactions on*, vol. 45, no. 6, pp. 2624–2631, Dec 1998.

[54] Qioptiq. [Online]. Available: http://www.qioptiq.com/

[55] R. Reed, P. Marshall, C. Marshall, R. Ladbury, H. Kim, L. X. Nguyen, J. Barth, and K. LaBel, "Energy dependence of proton damage in AlGaAs light-emitting diodes," *IEEE Transactions on Nuclear Science*, vol. 47, no. 6, pp. 2492–2499, December 2000.

[56] P. Tenga, T. Weidbergb, M. Chua, T. Duhc, I. Gregord, L. Houa, S.-C. Leea, P. Songc, and D. Su, "Radiation hardness and lifetime studies of the VCSELs for the ATLAS SemiConductor Tracker," *Nuclear Instruments and Methods in Physics Research*, vol. 497, no. 2-3, pp. 294–304, Feb 2003.

[57] S. S., R. D, R. L.M, S. C, T. M, P. C, and D. C. D, "Optical-based cots data communication bus for satellites," *ACTA ASTRONAUTICA*, vol. 66, pp. 674–681, 2010.

[58] D. McMorrow, S. Buchner, M. Baze, B. Bartholet, R. Katz, M. O'Bryan, C. Poivey, K. Label, R. Ladbury, M. Maher, and F. Sexton, "Laser-induced latchup screening and mitigation in cmos devices," in *Radiation and Its Effects on Components and Systems, 2005. RADECS 2005. 8th European Conference on*, Sept. 2005, pp. C22–1 –C22–7.

[59] M. DeLaus, D. Emily, B. Mappes, and R. Pease, "Converting a bulk radiation hardened bicmos technology into a dielectrically isolated process," *Nuclear Science, IEEE Transactions on*, vol. 40, no. 6, pp. 1774–1779, Dec 1993.

[60] Radiation Effects Data Workshop (REDW) Record. [Online]. Available: http://nsrec.com/redw/

[61] GSFC Radiation Data Base. [Online]. Available: http://radhome.gsfc.nasa.gov/radhome/RadDataBase/RadDataBase.html

[62] ESA Radiation Effects Database. [Online]. Available: https://escies.org/ReadArticle?docId=228

[63] Pumpkin CubeSa kit. [Online]. Available: http://www.cubesatkit.com/

[64] T. Vladimirova, C. Bridges, G. Prassinos, X. Wu, K. Sidibeh, D. Barnhart, A.-H. Jallad, J. Paul, V. Lappas, A. Baker, K. Maynard, and R. Magness, "Characterising Wireless Sensor Motes for Space Applications," *Second NASA/ESA Conference on Adaptive Hardware and Systems*, August 2007.

[65] Canadian Advanced Nanospace Experiment. [Online]. Available: http://www.utias-sfl.net/nanosatellites/CanXProgram.html

[66] T. Miyahira and G. Swift, "The use of advanced flash memories in satellite applications: Radiation responses and mitigation choices," in *Proceedings of the Military and Aerospace Applications of Programmable Devices and Technologies Conference*, September 1998.

[67] M. Elghefari and S. McClure, "Radiation Effects Assessment of MRAM Devices," EJet Propulsion Laboratory, Tech. Rep. 08-19 5/08, May 2008.

[68] D. Nguyen and L. Scheick, *TID testing of ferroelectric nonvolatile RAM*, August 2003.

[69] M. Borri and L. M. Reyneri, "Mechanical Subsystem of the AraMiS Architecture," Politecnico di Torino, Tech. Rep. 1B6, September 2008.

[70] U. S. D. of Defense, "Mechanical Subsystem of the AraMiS Architecture," United States Departement of Defense, Tech. Rep. MIL-DTL-5541F, 2009.

[71] GENSO, Global Education Network for Satellite Operations. [Online]. Available: http://www.genso.org

[72] Texas Instruments. [Online]. Available: http://www.ti.com

[73] 1B127 datasheet, Neohm Componenti. [Online]. Available: http://polimage.polito.it/aramis/products/1B127/22.06.20091B127DraftA.pdf

[74] A. Johnston, G. Swift, T. Miyahira, and L. Edmonds, "A model for single-event transients in comparators," *Nuclear Science, IEEE Transactions on*, vol. 47, no. 6, pp. 2624 –2633, dec 2000.

[75] Y. Boulghassoul, L. Massengill, A. Sternberg, R. Pease, S. Buchner, J. Howard, D. McMorrow, M. Savage, and C. Poivey, "Circuit modeling of the lm124 operational amplifier for analog single-event transient analysis," *Nuclear Science, IEEE Transactions on*, vol. 49, no. 6, pp. 3090 – 3096, dec 2002.

[76] R. Pease, A. Sternberg, Y. Boulghassoul, L. Massengill, S. Buchner, D. McMorrow, D. Walsh, G. Hash, S. LaLumondiere, and S. Moss, "Comparison of sets in bipolar linear circuits generated with an ion microbeam, laser light, and circuit simulation," *Nuclear Science, IEEE Transactions on*, vol. 49, no. 6, pp. 3163 – 3170, dec 2002.

[77] D. McMorrow, J. Melinger, S. Buchner, T. Scott, R. Brown, and N. Haddad, "Application of a pulsed laser for evaluation and optimization of seu-hard designs [cmos]," *Nuclear Science, IEEE Transactions on*, vol. 47, no. 3, pp. 559 –565, jun 2000.

[78] T. Hirao, H. Morib, J. S. Lairda, S. Onodab, and H. Itoha, "Comparison of the ion induced charge collection in Si epilayer and SOI devices," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 210, pp. 221–226, September 2003.

[79] D. Roascio, "Sottosostema di comunicazione tollerante ai guasti per satellite modulare AraMiS," Master's thesis, Politecnico di Torino, Nov 2008.

[80] A. Tai, S. Chau, and L. Alkalai, "Cots-based fault tolerance in deep space: Qualitative and quantitative analyses of a bus network architecture," *High-Assurance Systems Engineering, 1999. Proceedings. 4th IEEE International Symposium on*, pp. 97–104, 1999.

[81] C. Temple, "Avoiding the Babbling-Idiot Failure in a Time-Triggered Communication System," in *FTCS '98: Proceedings of the The Twenty-Eighth Annual International Symposium on Fault-Tolerant Computing*. Washington, DC, USA: IEEE Computer Society, 1998, p. 218.

[82] Infrared Data Association. [Online]. Available: http://www.irda.org/

[83] L.-S. Tsai and D. shan Shiu, "Equalization and Decoding for IRDA RLL-Coded Systems," *IEEE Singapore International Conference on Communication systems*, pp. 1–6, October 2006.

[84] Vishay. [Online]. Available: http://www.vishay.com

[85] Avago Technologies. [Online]. Available: http://www.avagotech.com

[86] TFDU8108 datasheet. [Online]. Available: www.vishay.com/ir-transceivers/list/product-82558/

[87] Selective Photodiodes from EPIGAP Optoelektronik GmbH. [Online]. Available: http://www.epigap-berlin.com/index.php?id=199&L=1

[88] R. Meitzler, W. Schneider, and R. Conde, "IRCOMM: Spacecraft Free-space Optical Bus Development," *Aerospace Conference, 2005 IEEE*, pp. 1583–1588, March 2005.

[89] M. W. Savage, J. E. Seiler, G. Dunham, and D. Platteter, "Analog Transients in the National VIP-10 and VIP-50 Process," *Radiation Effects Data Workshop, 2006 IEEE*, July 2006.

[90] N. Femia, G. Petrone, G. Spagnuolo, and M. Vitelli, "Optimizing duty-cycle perturbation of p&o mppt technique," *Power Electronics Specialists Conference, 2004. PESC 04. 2004 IEEE 35th Annual*, vol. 3, pp. 1939 – 1944 Vol.3, june 2004.

[91] J. P. Fellnera, G. J. Loebera, and S. S. Sandhub, "Testing of lithium-ion 18650 cells and characterizing/predicting cell performance," *Journal of Power Sources*, vol. 81-82, pp. 867–871, 1999.

[92] M. Borri and L. M. Reyneri, "Battery Aging Tests Report," Politecnico di Torino, Tech. Rep. 1B114, June 2008.

[93] C. Del Vecchio Blanco, M. D'Errico, and A. Conticchio, "A Distributed Design Architecture for Li-ion Battery: Integration of COTS ICs, uCs and CAN Bus," *Aerospace Conference, 2006 IEEE*, pp. 1 – 12, 0-0 2006.

[94] C. Poivey, "Radiation Hardness Assurance for Space Systems," SGT-Inc., Tech. Rep. NSREC02, 2002.

[95] D. Fleetwood, P. Winokur, and T. Meisenheimer, "Hardness assurance for low-dose space applications mos devices," *Nuclear Science, IEEE Transactions on*, vol. 38, no. 6, pp. 1552 – 1559, dec 1991.

[96] Neohm componenti. [Online]. Available: http://www.neohmcomponenti.com/