Università degli studi di Firenze
Facoltà di Scienze Matematiche Fisiche e Naturali
Dipartimento di Fisica

Development of tracking system acquisition electronics and analysis of first data for the PAMELA experiment

Ph.D. Thesis

Sergio Bruno Ricciarini

Advisor: Prof. Oscar Adriani

Dottorato di Ricerca in Fisica, XVIII Ciclo

Settore disciplinare: Fisica Nucleare e Subnucleare (FIS04)

Coordinator: Prof. Alessandro Cuccoli

31 December, 2005

ai miei genitori

Contents

Introduction

The present Ph.D. thesis work has been developed within the framework of the PAMELA experiment, whose main purpose is the study of the antiparticle component in cosmic rays outside the earth atmosphere, by means of a particle telescope which will operate for 3 years, starting from the first half of 2006, on board of the Russian satellite Resurs-DK1.

PAMELA is an international collaboration and is a part of the Italian-Russian cooperation program RIM (Russian Italian Mission), generally aimed at the study of the space environment in the vicinity of the earth, with specific interest for the particle physics items.

People participating in the PAMELA experiment belong to various Sections of the Italian National Institute of Nuclear Physics (INFN) and Physics Departments of Italian Universities (Trieste, Florence, Rome "Tor Vergata", Naples, Bari), to the INFN National Laboratories (Frascati) and the Institute of Applied Physics "Nello Carrara" (Florence). Besides, several foreign institutions are part of the PAMELA collaboration: the Physics Department of the Siegen University (Germany); the Royal Technology Institute of Stockholm (Sweden); the Cosmic Rays Laboratory of the Moscow Engineering and Physics Institute; the Laboratory of Solar and Cosmic Ray Physics of the "Lebedev" Physics Institute of Moscow; the "Ioffe" Physico-Technical Institute of St Petersburg; the NASA Goddard Space Flight Center (USA); the Particle Astrophysics Laboratory of the New Mexico State University at Las Cruces (USA); the Tata Institute of Fundamental Research at Bombay (India).

The PAMELA apparatus is composed of a set of detectors whose primary purpose is the identification of the incoming particle species and the precise measurement of the absolute value of the momentum $|\vec{p}|$. The components of PAMELA are a magnetic spectrometer, an electromagnetic imaging calorimeter, a time-of-flight (TOF) plastic scintillator system, plus other auxiliary detectors. The instrument is optimized to study the antiprotons and positrons fluxes in the kinetic energy range 80 MeV $-$ 190 GeV and 50 MeV $-$ 270 GeV respectively.

The magnetic spectrometer has been developed by the Florence group of the PAMELA collaboration, and is composed of a permanent magnet and a tracking system with 6 planes of double-sided silicon microstrip detectors, characterized by a high spatial resolution (better than 3 μm).

The final qualification tests of the PAMELA apparatus have been completed in March 2005; the apparatus has then been delivered to the TsSKB laboratories (Samara, Russia) for the integration within the satellite Resurs-DK1. The launch is foreseen in the first months of 2006 from the Baikonur Cosmodrome (Kazakhstan).

The Author's Contribution

I participated in the development of the PAMELA experiment, during the three years of Ph.D. course (January 2003 - December 2005), with several contributions.

An important part of my work was aimed at the completion of the data acquisition electronics for the tracking system and is described in chapter 3 of this thesis. Besides, I have developed methods for the precise measurement of important experimental characteristics of the PAMELA apparatus, in particular for the determination of the geometric acceptance of the instrument and for the analysis of the collected data to measure the light attenuation lengths and trigger efficiencies of the TOF scintillator system; chapters 4 and 5 are dedicated to these items.

I had already took part in the development of the electronics for the PAMELA tracking system during my Physics Degree work (2001-2002), focused on the implementation, test and characterization of the first stages of the data acquisition chain, for the amplification and digital conversion of the signals produced by the microstrip Si detectors.

During the Ph.D. course I have continued this work with my participation in the realization of the main control electronics of the tracking system, whose basic functionalities are to manage the acquisition of the detector signals, operated by the first electronics stages, to process the digitally converted data and to transmit them to the central control system of PAMELA. I have contributed to the design, optimization and test of this electronics and to its integration within the PAMELA apparatus; in particular I took care of the realization of the required logical control functionalities on programmable logics devices (FPGA). The main task was to obtain a system guaranteeing a high reliability for the whole 3 years duration of the experiment and a sufficiently fast operation to process the large number ($\approx$ 37000) of read-out channels, characteristic of the microstrip detectors used, at the same time satisfying the strict limits on available power and volume imposed

by the satellite mission. In particular the reliability of the electronic devices constitutes a major problem for space applications, because the highly ionizing radiations, which are present in the extra-atmospheric environment where PAMELA will operate, imply serious risks of damage for these devices.

During the development work I participated in the beam test of the PAMELA apparatus which took place in September 2003 at the SPS accelerator (CERN, Geneva) with the design of a preliminary version of the control logics for the tracking system, used for data acquisition during the test.

From July 2004 to February 2005, I directly contributed to the final preparation of the PAMELA apparatus at the laboratories of "Roma 2" INFN Section and Physics Department of Rome "Tor Vergata" University, where the various detectors, the associated electronics and several auxiliary systems were assembled together; various mechanical, electrical and data exchange interfaces between the PAMELA instrument and the Resurs-DK1 satellite were also defined, implemented and tested. During this integration phase the PAMELA apparatus has also undergone specific mechanical qualifications at the IABG facilities (Munich, Germany) to verify its ability to survive the strong accelerations foreseen during the launch.

This phase of work has involved a big effort of research and development to face and solve different problematics concerning the compatibility and reliability of the various parts of the system, ending with the final qualification tests performed with the apparatus in its "flight" configuration and the subsequent delivery of PAMELA to the Russian firm TsSKB (Samara) for the integration within the Resurs satellite.

After the end of this "hardware" work, the major effort of the collaboration has been directed to the precise determination of the experimental characteristics of PAMELA in the "flight" configuration, with the analysis of the cosmic-ray data collected during the final qualifications.

As already mentioned, I have in particular developed a method for the calculation of the geometric factor of PAMELA as a function of the magnetic rigidity of the incident charged particle; this method, based on a numerical approximation specifically optimized for the present application, is characterized by a high precision, with calculation uncertainties of less than 1% over the whole range of interest, reached within relatively short computation times (of no more than few days).

I have then performed a characterization of the TOF scintillator detectors of PAMELA, after their integration within the apparatus. The TOF system has the fundamental task

of generating the fast *trigger* signal to start the data acquisition for the other detectors; it is composed of 6 layers of plastic scintillator strips, read-out by photo-multiplier tubes at both ends, for a total of 24 strips. I have determined the attenuation lengths of the scintillation light for the strips, which are related not only to the characteristics of the scintillator material but also to the optical properties of the lateral surfaces of the strips. After the optical characterization I have used the same cosmic-ray data to measure the trigger efficiencies of the single strips and of the 6 scintillator layers and to calculate the overall trigger efficiency of the TOF system in the "flight" configuration.

Outline of the Thesis

The first two chapters of the present thesis constitute a general introduction and framework for the more specific discussion contained in the following chapters.

Chapter 1 illustrates the main characteristics of the PAMELA experiment and of the satellite mission, focusing on the description of the experimental capabilities of the apparatus, of the overall structure of the acquisition system, and of the three main detectors (TOF scintillators, magnetic spectrometer and calorimeter); particular attention is devoted to the design and operation of the TOF and tracking systems and of their electronics, since these parts of the apparatus are directly connected to the work described in the last three chapters.

Chapter 2 is dedicated to a general discussion of the problems that have to be faced in the development of electronics for a space experiment, and of the specific solutions adopted for the PAMELA apparatus. In particular this chapter describes the outstanding features of the programmable logic devices (FPGA) used for the implementation of the control functionalities.

The following chapters illustrate the main parts of my personal contribution to the PAMELA experiment.

Chapter 3 contains the detailed description of the structure, operation and test of the control electronics for the tracking system, with specific attention to the design of the control logics on FPGA devices.

Chapter 4 presents the method for the calculation of the geometric factor of PAMELA.

Chapter 5 describes the analysis of the cosmic-ray data acquired at earth for the determination of the attenuation lengths and trigger efficiencies of the TOF scintillator system.

Chapter 1

The PAMELA experiment

The PAMELA experiment (*a Payload for Antimatter/Matter Exploration and Light Nuclei Astrophysics* [1] [2]) was conceived to study the antiparticle component of primary cosmic rays outside the atmosphere by means of a satellite-borne apparatus.

After many years of intensive work PAMELA is now completed and its launch is foreseen in the first months of 2006 from Baikonur Cosmodrome (Kazakhstan) on board of the Russian satellite Resurs-DK1.

This first chapter presents the general characteristics of the PAMELA mission and instrumentation, followed by a detailed description of the main detectors of the apparatus.

1.1 The PAMELA mission

The term *cosmic rays* usually indicates all the charged particles that reach the atmosphere from the external space.[1]

The cosmic rays are composed of stable particles coming from outside the solar system (*galactic cosmic rays*) or emitted by the sun. Their composition varies with the energy and depends on the solar activity; as a raw estimate, it can be said that the cosmic rays are mainly composed of protons (p, $\approx 87\%$), helium nuclei (He, $\approx 10\%$), *heavy nuclei* with atomic number $Z > 2$ (1%) and electrons (e^-, $\approx 2\%$); also the presence of antiprotons $\overline{p}$ and positrons e^+ has been observed ($\approx 0.01\%$).

Precise measurements of the properties of the cosmic rays, such as relative distribution of the various particle species, energy spectrum, direction of incidence and time dependence of the fluxes, can be compared with expectations from different theoretical models explaining their production mechanisms, the propagation in the interstellar medium and their interaction with the solar-system environment.

[1] A more general definition considers also neutrinos and gamma rays as part of the cosmic-ray spectrum. For a review of the subject see [3].

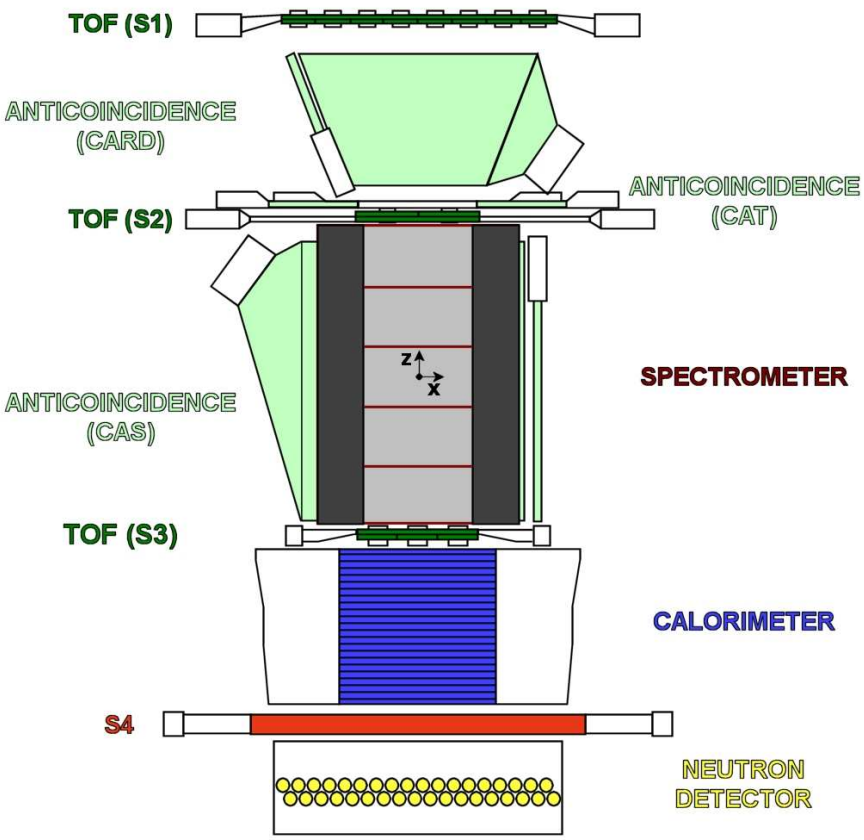

Figure 1.1: scaled drawing of the PAMELA apparatus.

Particularly interesting is the study of the antimatter component of cosmic rays (antiprotons $\overline{\mathrm{p}}$, positrons e^+, antinuclei); in fact there is no confirmed evidence up to now of possible galactic or extra-galactic primary sources of antiparticles, and the observed fluxes of $\overline{\mathrm{p}}$ and e^+ do not significantly differ from the expectations based on the secondary production by interaction of primary cosmic rays with the interstellar medium. In particular the observation of antinuclei in cosmic rays would necessarily imply the presence of primary sources of antimatter, since the contribution of secondary mechanisms is largely below the sensitivity of the present experiments.

The PAMELA experiment is primarily aimed at extending the known energy spectrum of both antiprotons and positrons and obtaining significantly smaller statistical uncertainties in the measured intensities with respect to those affecting the data collected by the experiments performed up to now. This goal has been achieved with the design of an apparatus characterized by a high sensitivity and efficiency in the identification of $\overline{\mathrm{p}}$ and e^+ and in the measurement of their energies, coupled to the choice of operating it on board of a satellite, thus making it possible to collect a high number of primary antiparticle

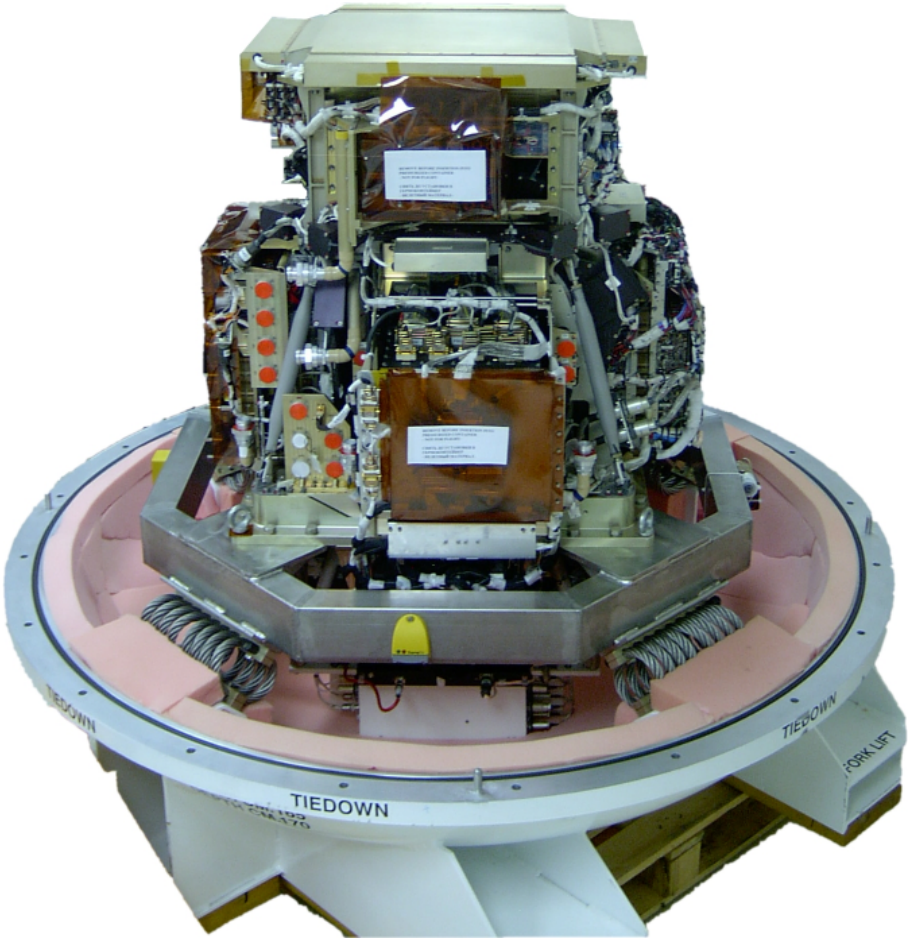

Figure 1.2: the finished PAMELA apparatus on its transport container before the integration on satellite.

events in a very clean environment.

The operation outside the atmosphere (on an elliptic orbit ranging from 350 to 600 km of altitude) will practically eliminate the background of secondary particles produced by the interactions of primary cosmic rays in the atmosphere; the orientation of the apparatus, with the acceptance[2] window toward the outer space, will simplify the identification and rejection of the residual background formed by secondary particles escaping from the atmosphere (known as *albedo*). The satellite Resurs-DK1 will guarantee a duration of the flight of at least 3 years, much longer than for the experiments of comparable sensitivity operated on board of balloons, whose flight cannot extend for periods longer than some days or weeks at maximum.[3]

Another important factor which influenced the choice of the satellite is the strong

[2]The acceptance of the PAMELA apparatus is discussed in chapter 4.
[3]For a synthetic review of balloon techniques see for example [4].

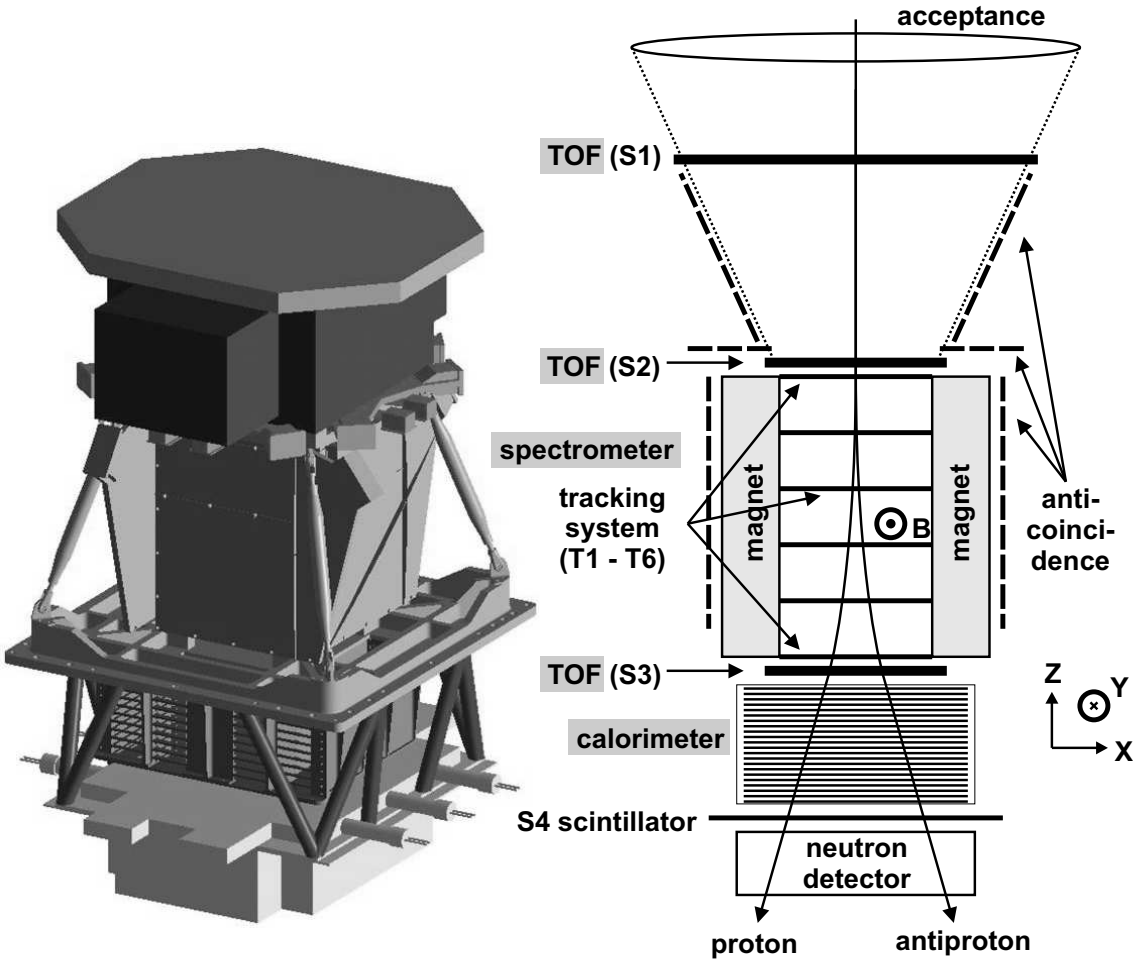

Figure 1.3: three-dimensional view and section of the PAMELA apparatus, with illustrated the principle of discrimination between particle and antiparticle in the spectrometer.

reduction of the intensity of incident particles observed moving from the polar regions to the equator; this is an effect of the terrestrial magnetic field (known as *geo-magnetic cut-off*), which has a quasi-dipolar shape along a main magnetic axis inclined of 11° with respect to the rotation axis of the earth; moving away from the magnetic poles the field tends more and more to reflect away charged particles coming from outside; this effect is stronger for particles of lower energy. To maximize the number of cosmic rays incident on the apparatus, a satellite characterized by a quasi-polar orbit (70.4° inclination on equator) has been chosen, allowing the precise measurement also of the low-energy part of the spectrum.

The main drawbacks of the operation on satellite are the strict limits in mass, volume and power consumption, the strong mechanical stresses foreseen during the launch phase and the presence of highly ionizing radiations outside of the atmosphere which could cause serious damages to the electronic systems of PAMELA. These aspects and their effects on

the design and development of the apparatus will be treated with more details in various parts of this work.

Given the characteristics of the mission mentioned above and taking into account the experimental capabilities of the apparatus, discussed in the following sections, it is expected that PAMELA will be able to:

- measure the energy spectrum of $\overline{\mathrm{p}}$ in the kinetic-energy range 80 MeV $-$ 190 GeV (expected $\approx 3 \cdot 10^4$ antiprotons in 3 years);

- measure the energy spectrum of e^+ in the kinetic-energy range 50 MeV $-$ 270 GeV (expected $\approx 3 \cdot 10^5$ positrons in 3 years).

Note that no antiparticles have been directly detected in cosmic rays until now for kinetic energies greater than 50 GeV.

1.2 Experimental apparatus

The main purpose of the PAMELA apparatus (shown in fig. 1.1, 1.2 and 1.3) is the efficient identification of the incident particles under study and the precise measurement of the absolute value of their momentum $|\vec{p}\,|$, and in particular the selection of $\overline{\mathrm{p}}$ and e^+ in presence of the background from p and e^-. This is achieved by means of a set of specialized detectors, whose principle of operation is schematically explained in what follows.

- A system of 3 double planes (S1, S2, S3) of plastic scintillators, divided into strips read-out by pairs of photo-multiplier tubes, gives a fast and efficient *trigger* signal to synchronize the data acquisition for the whole apparatus.
  With the timing information provided by the scintillator planes it is possible to identify and reject the upward going particles (*albedo* background).
  Besides, the time-of-flight (TOF) measurement can be used for the determination of the velocity v of the incident cosmic ray, useful for particle identification in the case of non-ultra-relativistic particles; typically this measurement contributes significantly to the rejection of the proton background from the e^+ sample (and analogously of electrons from the $\overline{\mathrm{p}}$ sample) for kinetic energies below ≈ 1 GeV.
  This scintillator system will be indicated as *TOF system* to distinguish it from the other scintillators employed in the apparatus, introduced below.

- A magnetic spectrometer, formed by a permanent magnet with a rectangular cavity and a tracking system composed of 6 planes of silicon microstrip detectors, determines the charge sign and momentum $\vec{p}$ of the particle through the reconstruction of the curvature of the trajectory in the magnetic field; the absolute value of the charge is determined from the measurement of the ionization produced in the Si layers.

- At kinetic energies greater than $\approx$ 1 GeV a dedicated electromagnetic imaging calorimeter performs the discrimination between particles of equal momentum and charge (e^+ and the background of p, or $\overline{p}$ and e^-), based on the different characteristics of the shower of secondary particles produced traversing thick absorber layers; at lower kinetic energies this discrimination becomes less efficient and it is better done with the direct measurement of velocity obtained from the TOF system, exploiting the fact that for fixed momentum the velocities of $e^\pm$ and $p/\overline{p}$ are significantly different given the difference in their mass[4].
  The calorimeter also performs an independent measurement of the energy of the incident electron or positron, that can be cross-checked with the momentum measurement operated by the spectrometer.

- Other detectors are used to improve the overall performance:[5]

  - an anti-coincidence (AC) system, composed of various scintillator planes, read-out by photo-multiplier tubes, detects particles entering the apparatus from outside the acceptance and possibly interacting with the materials surrounding the detectors (magnet, mechanical support parts, electronic systems);

  - a further scintillator plane (S4), positioned below the calorimeter, detects particles escaping from it and provides additional information on the extension of the electromagnetic or hadronic shower;

  - a neutron detector (ND), with ^{3}He proportional counters and polyethylene moderator, complements the calorimeter in the lepton/hadron discrimination,

[4] The following relativistic relation holds: $v/c = |\vec{p}|c/E$, with $E = \sqrt{(mc^2)^2 + (|\vec{p}|c)^2}$.

[5] The original design of the PAMELA apparatus also included a transition radiation detector (TRD) with straw-tube gas counters; the aim of this detector was to improve the performance of PAMELA in the discrimination between $e^\pm$ and $p/\overline{p}$. During the development phase of the TRD it became clear that the implementation of a reliable system for the control of the gas flow in the counters, complying with the strict limits and requirements imposed by the satellite mission, would have caused a delay incompatible with the schedule foreseen for the completion of the other parts of the apparatus.

exploiting the fact that a much greater number of neutrons are produced by hadronic showers with respect to the electromagnetic ones.

The PAMELA apparatus has a total weight of 470 kg, a total height of about 120 cm and a maximum cross-sectional diameter of 102 cm; the power consumption is less than 360 W.

Fig. 1.3 also shows the *PAMELA reference system*, defined as the orthogonal right-handed frame with the origin in the center of the magnetic cavity and axes oriented as shown: the Z axis is directed along the longitudinal dimension of the apparatus, toward the incoming particles; the Y axis is directed opposite to the main direction of the magnetic field inside the spectrometer. All the positional information given in the present work is referred to this frame for clarity; $\widehat{\imath}$, $\widehat{\jmath}$, $\widehat{k}$ will indicate the unit vectors of X,Y,Z axes respectively.

The following sections describe the general characteristics of the acquisition system of PAMELA and of the main detectors, with specific attention for the tracking and TOF systems which are directly involved in the present work.

1.3 General operation and acquisition system

The PAMELA apparatus will be housed inside an aluminium container, filled with N_2 pressurized at 1 atm, and integrated within the satellite structure.

The container is supported by a movable arm; for the launch phase and the orbital manoeuvres the container is placed in the "initial" position along the side of the satellite, as shown in fig. 1.4. During the operating phases, the arm will put the instrument in the optimum position with respect to the rest of the satellite, in such a way that no material, aside from the thin wall (2 mm) of the container, will be present inside the acceptance window of PAMELA, and with the spectrometer pointing toward the zenith part of the sky.

The necessary power for the operation of PAMELA is provided by the Resurs-DK1 satellite, and is produced by two solar panels; the supply system of PAMELA distributes the incoming power to all the parts of the apparatus.[6] The operation of PAMELA is entirely and independently managed by the *PAMELA storage and control unit* (PSCU), specifically designed and qualified for this space mission and based on a dedicated microprocessor; the operating system and configuration parameters of the PSCU are perma-

[6] A more detailed description of the power supply system of PAMELA will be given in sect. 3.9.

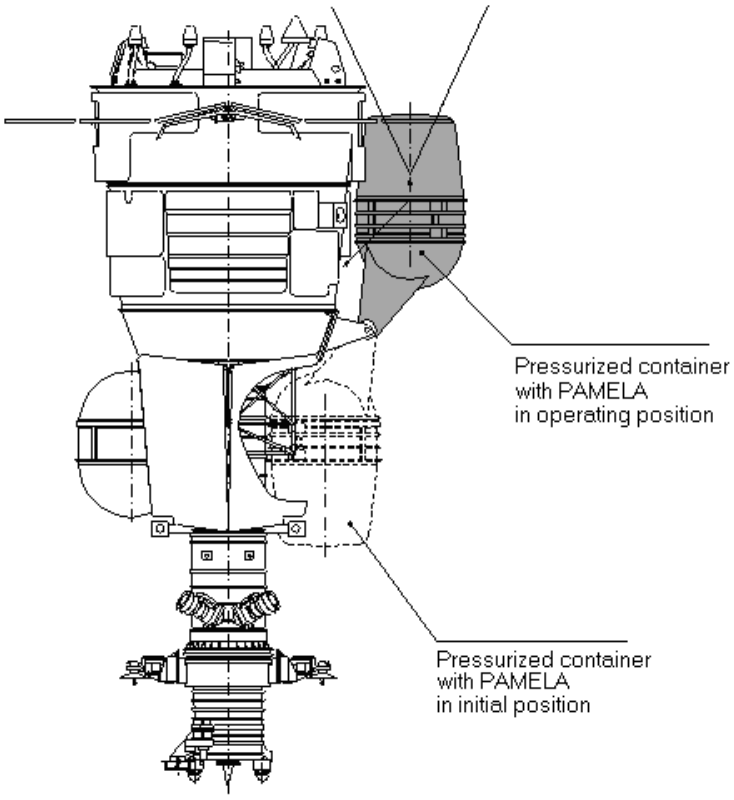

Figure 1.4: illustration of the initial and operating positions of PAMELA in the Resurs-DK1 satellite.

nently stored in non-volatile memory elements (PROM and EEPROM[7]).

The PSCU automatically performs all the procedures to assure the required functionalities of PAMELA and acts as interface with the telecommunication systems of the Resurs satellite for data exchange with the ground-based control station. The data continuously collected by the apparatus are stored in the mass memory of the PSCU and periodically downloaded to ground (a few times per day); the configuration parameters of the system can be changed with telecommands transmitted from the control station. The mass memory is composed of a 2 Gbyte SDRAM[8]; to increase the reliability of the system the memory is duplicated and error correction codes are stored with the data.

A schematic drawing of the control system of PAMELA is shown in fig. 1.5.

The data acquisition system for each detector (TOF, tracking system, calorimeter, AC, S4, ND) can be generally divided into a *front-end* (FE) part, including all the circuitry

[7]The *programmable read-only memory* (PROM) and *electrically erasable PROM* (EEPROM) are solid-state memories maintaining the stored information also in absence of power supply.

[8]The *synchronous dynamic random access memory* (SDRAM) is a volatile solid-state memory, where the information is stored on capacitors and refreshed with a proper rate, to avoid that the data are lost due to the discharge of the capacitors.

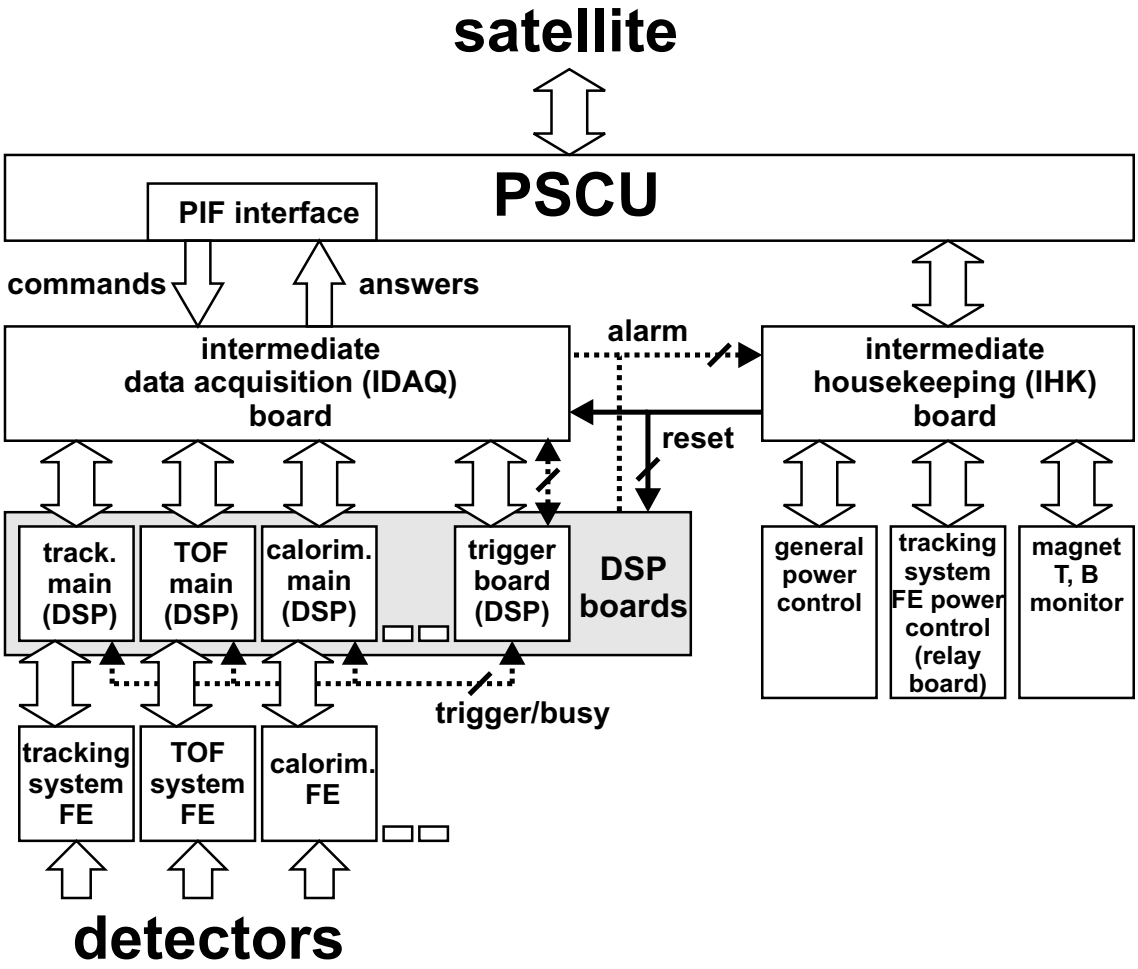

Figure 1.5: simplified scheme of the control system of PAMELA; on the left the data acquisition part, on the right the housekeeping part.

necessary for the read-out and digital conversion of the detector signals, and a *main* part that manages the operation of the FE electronics and in particular the transfer of the acquired data to the PSCU; both the front-end and main electronics are composed of one or more printed-circuit boards. The digital circuits, commonly indicated as *control logics*, that perform the general management tasks within a board, are typically implemented on programmable logic devices (FPGA, described in sect. 2.6).

The main boards usually contain one or more *digital signal processors* (DSP)[9] for the elaboration of the acquired data; for this reason the main boards are also indicated as *DSP boards*. The various DSP boards can receive commands from the PSCU and send back data (answers) when required; the data exchange is managed by a fast *intermediate data acquisition* (IDAQ) board with direct independent access to two dedicated temporary data

[9] A DSP is a microcomputer chip optimized for calculation tasks on digital data; the specific DSP chip employed in the PAMELA apparatus will be described in sect. 3.1.1.

buffers, for commands and answers respectively, that reside in the part of PSCU acting as the interface toward the IDAQ board and known as *program interface* (PIF); this approach is required for the minimization of the dead time associated with the read-out of data produced for each physics event.

The commands to be executed by the IDAQ board, including the ones to be passed to the single DSP boards, are organized in sequences (*queues*) that are prepared by the PSCU and stored in the command buffer of the PIF interface; when the IDAQ board is enabled by the PSCU to start execution, it automatically reads the first command of the queue on the PIF, operates accordingly and directly writes the answer data on the corresponding answer buffer. Then the same cycle is repeated for all the other commands in the queue.

During typical operation the DSP boards and IDAQ board, with the execution of suitable command queues, are initialized, configured and finally set in a state waiting for a trigger signal; the generation of this signal, as a consequence of a valid physics event, activates on all the DSP boards the automatic read-out of data, their on-line processing and their storage in temporary memories. The trigger pulse is also received by the IDAQ board, which waits a convenient time and then starts executing a specific command queue to transfer these data into the answer buffer of the PIF interface; then the apparatus is again ready to start a new acquisition cycle. The PSCU will later copy the content of the answer buffer of the PIF into the mass memory.

One of the fundamental tasks of the PSCU is to continuously check for the correct operation of the apparatus and, if an anomaly is detected, to start the proper recovery procedure. Only in the case that the automatic recovery were not sufficient to restore the standard conditions, the PSCU would stop operation, turn off all the subsystems and wait for specific instructions from the ground control station.

The PSCU checks the conditions of the apparatus partly directly, through a set of monitoring lines (voltage and current sensors in the power supply system, distributed temperature sensors etc.) and partly through a dedicated *intermediate housekeeping* (IHK) board which collects a number of *alarm* signals generated by the various DSP boards.

The IHK board is operated by the PSCU with specific command queues, in a similar way to the IDAQ board. The IHK board, after a request from the PSCU, can send a common *reset* pulse to all the DSP boards for the initialization of their digital components, typically during a recovery procedure or after the turning on of power to the apparatus. The IHK also manages those subsystems that do not take part in the fast event read-out

and that are generally indicated as *slow control* units:

- the general control of the power supply system for the whole apparatus;
- the control of power supply for the front-end electronics of the tracking system, that is performed by two dedicated boards (*relays boards*, described in sect. 3.9) due to the great number of independent sections;
- the monitoring of temperature and magnetic field through a set of sensors placed on the external walls of the support structure of the spectrometer.

1.4 TOF scintillator system

The TOF system [5] (see fig. 1.6 and 1.7 and table 1.1) is composed of 3 double planes of scintillator detectors, read-out by photo-multiplier tubes (PMT): S1, positioned on top of the PAMELA apparatus, S2 and S3, immediately above and below the spectrometer respectively.

The timing information of the signals produced by the TOF system at the passage of an ionizing particle is extracted and used for the generation of the trigger pulse, for the discrimination of upward-going particles (*albedo*) and, in combination with the information on the trajectory provided by the tracking system, for a measurement of the time of flight across the PAMELA apparatus.

The main requirement of the TOF detectors is the efficient production of output pulses characterized by short delays and small uncertainties with respect to the time of passage of the incident particle, thus achieving on one hand a quick generation of the trigger signal, and on the other a high resolution in the determination of the crossing times at each plane. For this reason a plastic scintillator material characterized by prompt response and relatively high light output has been selected and coupled to fast PMT tubes and front-end electronics.

Each of the 3 planes S1, S2, S3 is formed by two superposed rectangular scintillator layers, thus having two independent timing measurements. The 6 layers, whose areas match the geometric acceptance window of the spectrometer, are divided into strips whose number and dimensions vary from layer to layer, for a total of 24 strips. At each of the two ends of a strip a light guide conveys the scintillation light toward the input window of a PMT, which transforms the light pulse into an electric current signal, which is then processed by the front-end electronics; a total of 48 PMT are present for the read-out of the 24 scintillator strips. The strips of the upper layer of each plane (S11, S21, S31) are

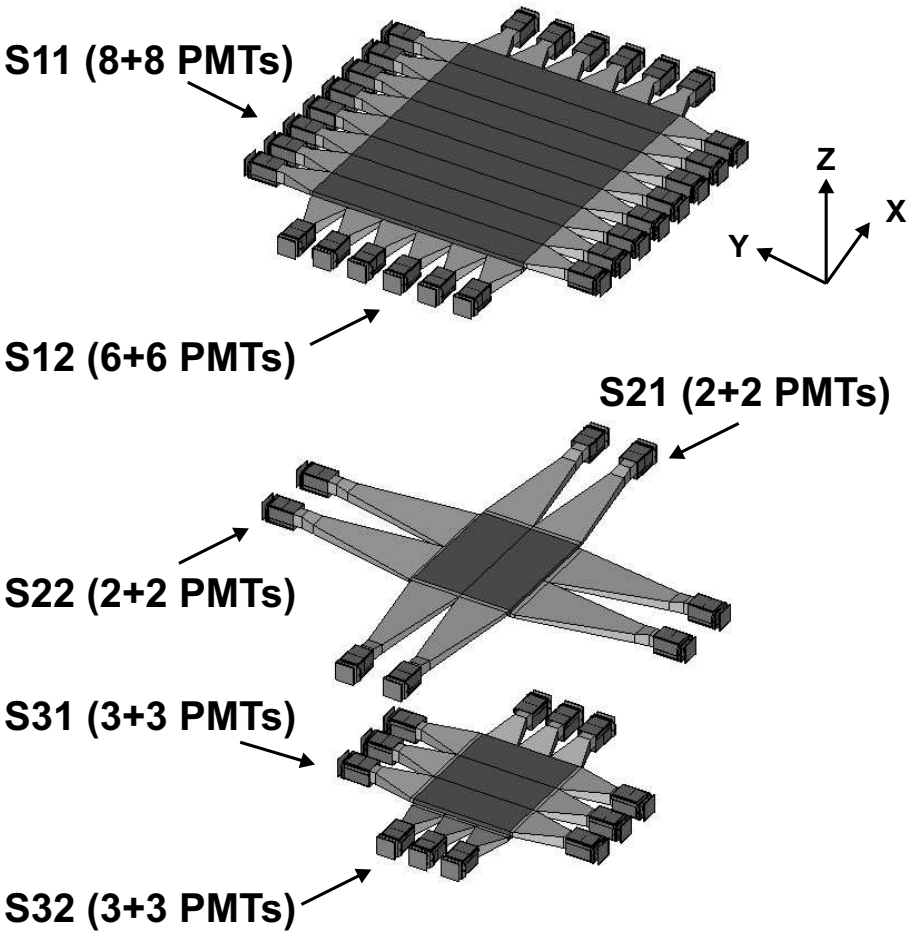

Figure 1.6: layout of the TOF scintillator system.

plane	layer	number of strips	Z (cm)	thickness (mm)	strip dim. X (cm)	strip dim. Y (cm)	strip cross-section (cm^2)
S1	S11	8	+53.74	7	5.1	33.0	3.57
	S12	6	+53.04	7	40.8	5.5	3.85
S2	S21	2	+23.94	5	18.0	7.5	3.75
	S22	2	+23.44	5	9.0	15.0	4.50
S3	S31	3	-23.49	7	6.0	15.0	4.20
	S32	3	-24.34	7	18.0	5.0	3.50

Table 1.1: table resuming the main geometrical characteristics of the TOF scintillator system; X,Y,Z axes are those of the PAMELA reference system.

orthogonal to the ones of the corresponding lower layer (S12, S22, S32 respectively), thus coupling to the timing information also a spatial information on the crossing point in the plane.

The principle of measurement of the time of particle crossing on a scintillator strip, of

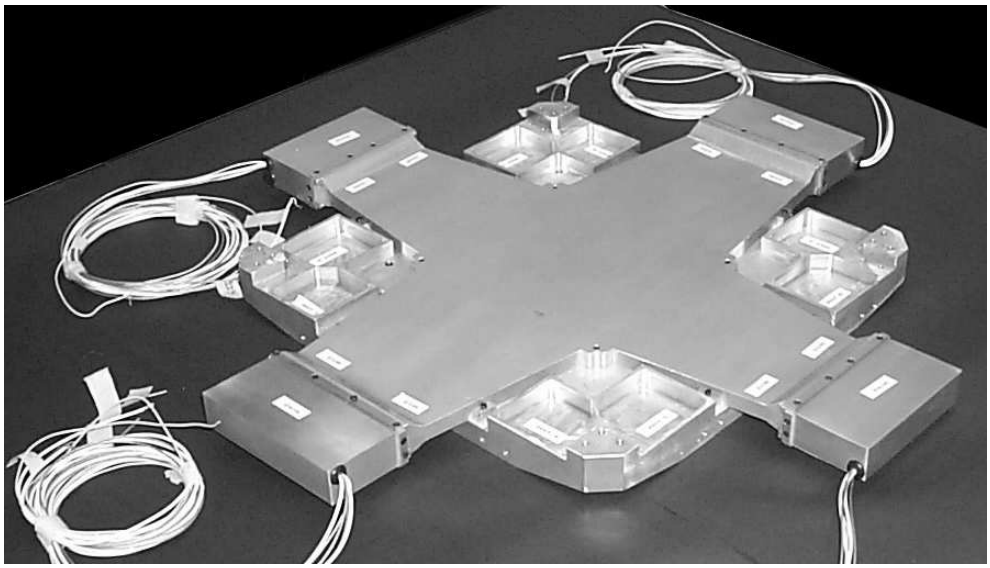

Figure 1.7: the finished S2 scintillator detector, inside its metallic support container, before the integration within the PAMELA apparatus.

length L and oriented along the X axis, is schematically illustrated in fig. 1.8. When an ionizing particle traverses the strip at time t_0 and coordinate x_0, the scintillation light, originally isotropically directed, produced along its path, propagates toward both ends mainly through a sequence of total reflections[10] on the walls of the scintillator strip and of the light guide until it enters the PMT input window. Considering one of the two PMT (PMT1), the output anode signal is produced at a time t_1 given by:

$$t_1 = t_0 + \frac{x_0}{v} + D_1 \tag{1.1}$$

where v is the velocity of the light pulse in the scintillator and D_1 is a delay introduced by the light guide and the PMT. A similar relation holds for PMT2:

$$t_2 = t_0 + \frac{L - x_0}{v} + D_2 \tag{1.2}$$

If the quantities L, v, D_1 and D_2, characteristic of the system, are known, the measurement of t_1 and t_2 allows to determine t_0 and x_0.[11]

The performances of the TOF system of PAMELA are characterized by a typical delay in the generation of the event signals toward the trigger board of the order of 10 ns and an uncertainty in the determination of the instant of particle crossing on a plane of the order of 150 ps (standard deviation at 68% confidence level). The specific characteristics of the various components of the system will be described in what follows.

[10] The mechanism of collection of the scintillator light will be discussed in sect. 5.1.

[11] A supplementary (and more precise) spatial information on the impact point x_0 is given by the tracking system.

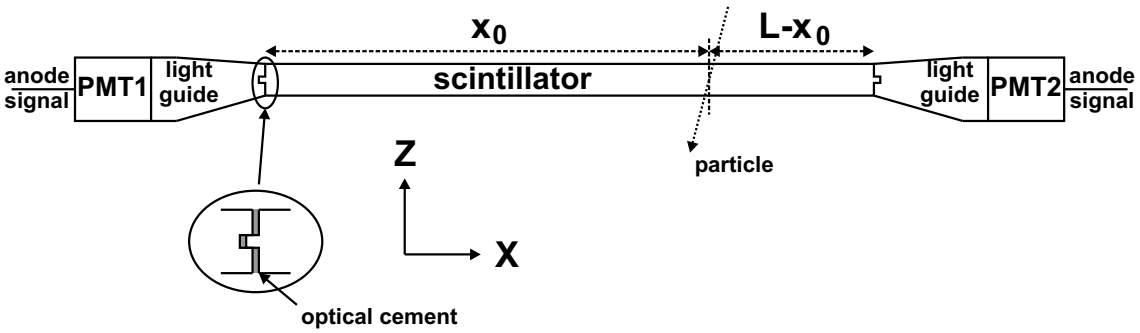

Figure 1.8: principle of operation of a strip of the TOF system.

1.4.1 Scintillators and light guides

The scintillating material selected for the TOF system is the BC-404 [12], a plastic organic scintillator with a relatively high light output (68% of Anthracene) and optimum timing performances: the scintillation light produced by an ionizing particle in the BC-404 is characterized by a rise time of 0.7 ns, a decay time of 1.8 ns and a *full width at half maximum* (FWHM) of the pulse of 2.2 ns. Besides, this kind of material is light and can be relatively easily machined into the desired shapes.

The light guide is made of plexiglas, characterized by a light transmission window matching the emission spectrum of the BC-404 (wavelength of maximum intensity: 408 nm) and thus minimizing the fraction of absorbed light. The contact surface between strip and light guide has been shaped in the form of a "tooth" (see fig. 1.8), with the contact between the two materials mediated through an optical cement[13]. The presence of the "tooth" results in a harder structure along the Z axis, useful to avoid problems due to the mechanical stresses foreseen during the launch phase.

The external walls of the strip and its two light guides are optically insulated by means of a 23 μm thick aluminized mylar foil. The fraction of light escaping reflection at the interface of the wall can be reflected back by the Al layer, thus allowing a partial recovery of the escaped light and preventing it from being collected by the adjacent strip, which would possibly generate a spurious pulse in the corresponding PMT pair.

1.4.2 Photo-multiplier tubes

A photo-multiplier tube is basically composed of a photo-sensitive semiconductor layer (*photo-cathode*), which emits electrons when hit by light in a characteristic range of wave-

[12] Technical information on this material, manufactured by Bicron, can be found at [6].
[13] BC-600 [6], with a transmission factor above 95% for the wavelengths of interest.

lengths, coupled to an electron multiplier chain, obtained with a set of intermediate electrodes (*dynodes*) and a final anode output electrode, kept at increasingly high voltages with respect to the photo-cathode; the whole structure is sealed in a vacuum metallic box with a glass window for the incoming light in correspondence of the photo-cathode. The e^-, emitted by the photo-cathode, are accelerated by the strong electric field toward the first dynode stage; when striking the dynode surface, these accelerated primary electrons cause the emission of a greater number of secondary electrons, which are then accelerated toward the next diode stage; finally, a relatively large number of e^- is collected at the anode, forming a current pulse which is processed by the following stages of electronics.

The PMT model chosen is the R5900, manufactured by Hamamatsu [7], whose characteristics best match the various requirements for the application in the TOF system of PAMELA; this device was already successfully employed in space environment [8]. The sensitivity spectrum of its photo-cathode matches the light emission band characteristic of the scintillator BC-404, with a *quantum efficiency*[14] reaching its maximum value ($> 21\%$ for the selected specimens) in the region from 350 nm to 450 nm.

The timing response of the R5900 PMT is comparable to that of the scintillating material BC-404, with a fast and short current pulse: the typical *transit time*[15] of the signal from the photo-cathode to the anode output is 8.8 ns, with a *transit-time spread* (defined as the FWHM of the transit time distribution of the photo-electrons) of 260 ps and a rise time of the anode pulse of 1.5 ns. These quantities are not significantly affected by the number of photons collected on the photo-cathode.

Besides, the R5900 is a device of limited size (housed in a metal parallelepiped of dimensions 3 cm $\cdot$ 3 cm $\cdot$ 2.6 cm) and weight (25 g), rugged enough to survive the mechanical stresses of the launch.

The interface between PMT and light guide is obtained through a soft silicone optical pad (BC-634A [6], characterized by a transmission factor $> 90\%$ for the wavelengths of interest). The PMT is fixed to the metallic frame supporting the scintillator plane by means of two helicoidal steel springs (see fig. 1.9) hooked on one side to the frame and on the other side to an insulating fiberglass plate positioned at the back of the PMT. This coupling system avoids a direct mechanical contact between the plexiglas of the light guide and the thin glass window of the PMT, at the same time guaranteeing a good

[14] The quantum efficiency Q, for a given wavelength of the incident light, is defined as the ratio between the number N_e of electrons emitted by the photo-cathode and the number N_γ of incident photons, $Q = N_e/N_\gamma$.

[15] The transit time of a PMT is defined as the average time between the arrival of a photon on the photo-cathode surface and the collection of the subsequent electron burst at the anode output.

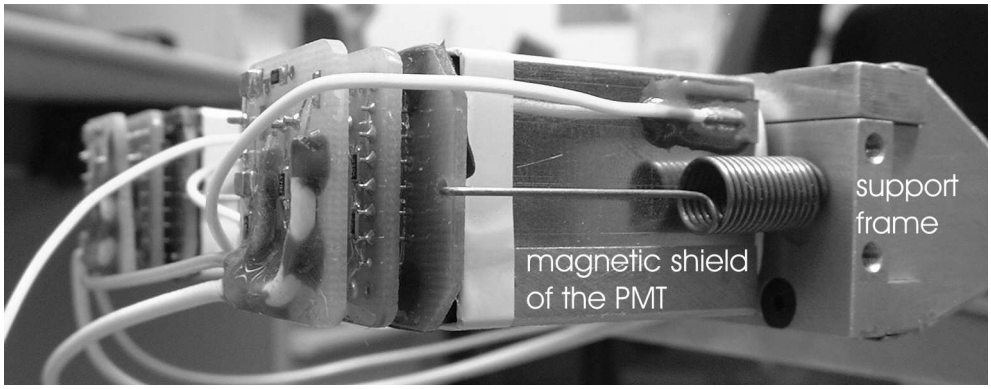

Figure 1.9: detail of the assembly of a PMT with two helicoidal springs. The tube is surrounded by a μ-metal magnetic shield.

optical contact and hardness of the structure.

The residual magnetic field outside of the spectrometer of PAMELA, if not correctly shielded, can affect the trajectories of the electrons inside the PMT, worsening the electric focusing properties of the dynode chain and consequently the performances of the tube both in transit time spread and overall gain[16].

This problem is especially important for the S2 and S3 planes, which are positioned in the vicinity of the permanent magnet. For this reason all the PMT are enclosed in a 1 mm thick μ-metal magnetic shield; the residual field intensity inside the screen is at most a few mT, with effects on the PMT performances of no more than $\approx 1\%$.

1.5 Front-end electronics of the TOF system

The front-end (FE) electronics of the TOF system is composed of 6 identical printed-circuit boards which simultaneously perform the read-out of the 48 PMT channels (8 anode signals per board); the association between PMT and boards is arranged in such a way to minimize the loss of read-out channels for each of the 6 layers of the TOF system in case of failure of a board. The FE board (see fig. 1.10) is segmented into two halves, each independently powered and managing 4 anode signals, plus a common interface, implemented on the FE controller FPGA, toward the TOF DSP board[17].

The 4 anode signals entering half FE board are processed in parallel by a *time section* and a *charge section* of the electronics, and finally by a time-to-digital converter (TDC).

[16]The gain G of a PMT is defined as the ratio between the number of e^- collected at the anode N_{an} and the number of e^- emitted by the photo-cathode N_{ph}, $G = N_{an}/N_{ph}$.

[17]The TOF DSP board houses the main electronics of the TOF system and interfaces with the PSCU through the IDAQ board, as mentioned in sect. 1.3.

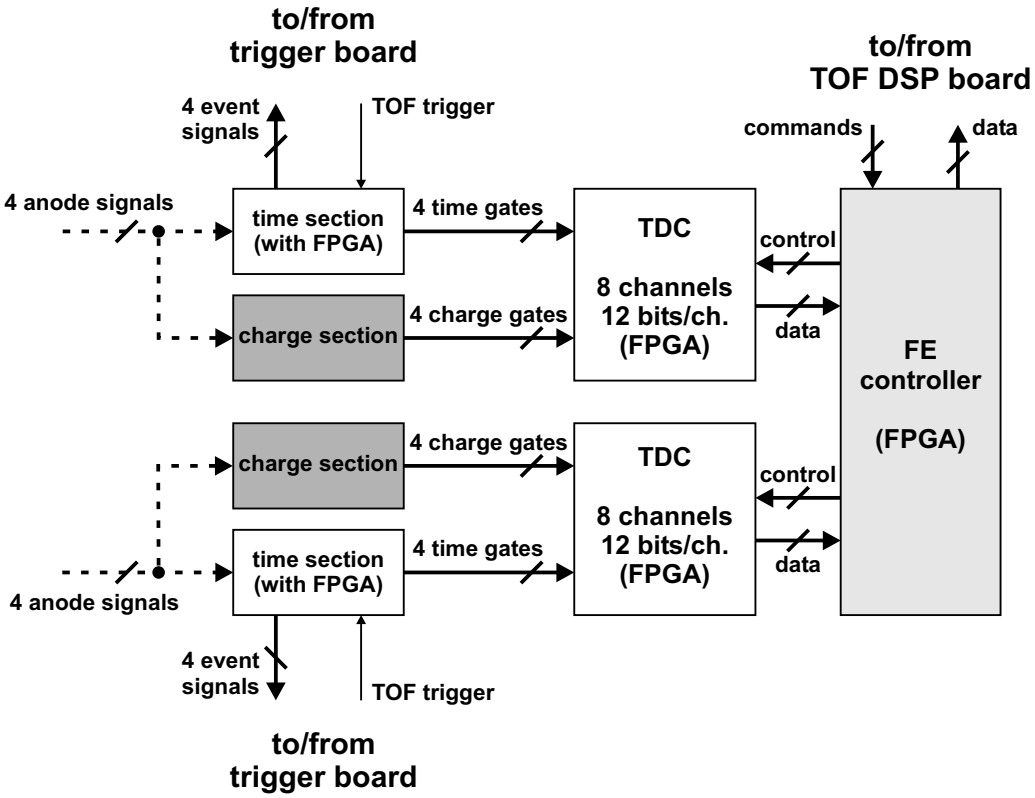

Figure 1.10: scheme of one of the 6 FE boards for the TOF scintillator system; the board manages 8 PMT anode signals.

In the time section, each of the 4 input channels is discriminated by a fast discriminator with a settable threshold and associated with a corresponding *event signal* output line toward the trigger board, used for the generation of the trigger signals for the whole apparatus (see sect. 1.6); in particular the trigger board sends back a fast *TOF trigger* signal to all the 6 FE boards, which is used as a common reference for the timing measurement of the anode pulses.

A *time gate* line is also associated with each input channel; the duration of the *time gate* pulse is proportional to the elapsing interval between the anode pulse and the *TOF trigger* pulse. Each of the 4 *time gates* is processed by the time-to-digital converter (TDC) producing a 12-bit digital value proportional to the length of the *time gate* pulse. The time section is described in greater detail in sect. 1.5.1.

On the other hand, the charge section produces 4 *charge gate* pulses, whose lengths are proportional to the integrated current forming the anode signals; the *charge gates* are similarly processed by the TDC. The integrated charge, having a linear dependence on the intensity of light collected at the photo-cathode of the PMT, can be correlated to the

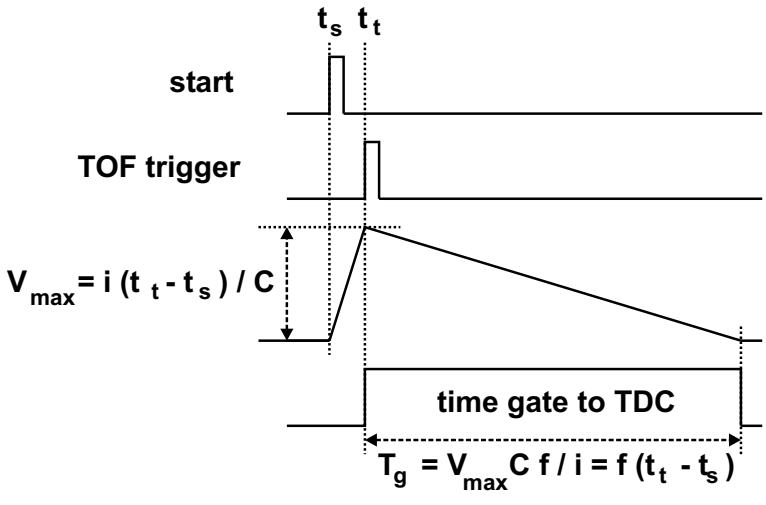

Figure 1.11: principle of operation of the TAT converter.

initial intensity of the scintillation light, as shown in chapter 5. The charge section is discussed in sect. 1.5.2.

The values of all the digital conversions are then transferred to the TOF DSP board; the transfer is managed by the FE controller FPGA according to a command sequence sent by the DSP board, which synchronizes its operation on the trigger signal, waiting a sufficient time for the completion of the operation of the TDC.

1.5.1 Time section

Each of the 4 anode inputs is processed in the time section by a dedicated chain, with a double-threshold discriminator and a time-to-amplitude-to-time (TAT) converter (or *time stretcher*).

The anode pulse enters the fast double-threshold discriminator[18]; the transition of the output of the discriminator causes the generation of the *event signal* toward the trigger board and of the internal *start* pulse activating the time stretcher (see fig. 1.11). When started, the time stretcher begins to charge a low-loss capacitor C with a constant current i generated by a high-stability current source; with the transition on the *TOF trigger* input, the *time gate* line is asserted while the same capacitor starts being discharged with a constant current i' of opposite sign and $f \approx 200$ times smaller, $i' = i/f$; the *time gate* is finally reset by the internal logic when the voltage across the capacitor reaches a minimum value. Therefore the length of the *time gate* pulse T_g depends linearly on the time between

[18]The reason for the employment of a double-threshold discriminator instead of a single-threshold one is explained at the end of the present subsection.

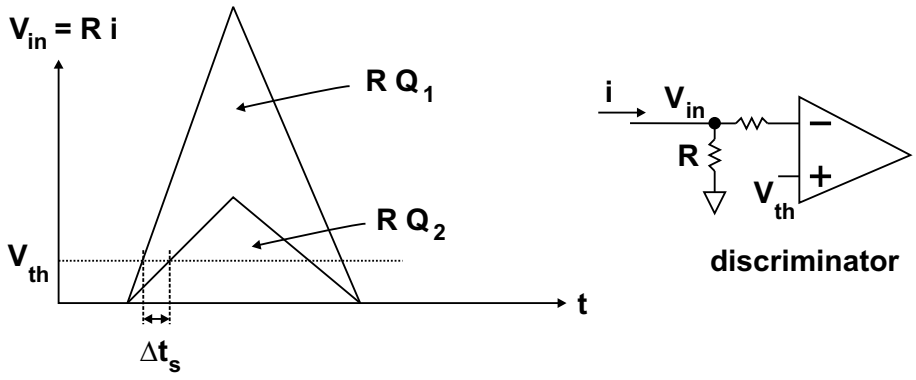

Figure 1.12: illustration of the time-walk effect.

the arrival of the anode pulse t_s and that of the trigger signal t_t: $T_g = f \cdot (t_t - t_s)$.

The time expansion is an efficient method to achieve a sufficiently small contribution to the uncertainty in the time measurement from the digital conversion operated by the TDC; with this method the maximum quantization error introduced by the TDC (± 5 ns) in the measurement of T_g, fixed by its operating frequency (100 MHz), results in an effective error of 5 ns/f $\approx$ 25 ps for the measurement of the physical time interval $t_t - t_s$.

The choice of the time expansion method, instead of a direct measurement of the physical interval $t_t - t_s$ by means of a fast TDC, operated at 20 GHz, satisfies the request of reducing the power consumption of the system (which increases linearly with the operating frequency). The main drawback of this method is in general the dead time introduced with the elongated *time gate* pulse; this, on the other hand, does not constitute a problem in the case of the PAMELA experiment because the expected mean interval between successive events ($\approx$ 83 ms, corresponding to a rate of 12 Hz) is several orders of magnitude greater than the duration of the *time gate* (~ 20 μs).

The use of a double-threshold discriminator allows the minimization of the *time-walk* effect, which constitutes another important source of uncertainty in the time measurement. Consider the case of a single-threshold discriminator, illustrated in fig. 1.12. The rise time and full width at half maximum (FWHM) of the anode current pulse is approximately constant for a given PMT, while the total charge Q forming the pulse, and therefore the amplitude of the anode signal, varies from event to event; as a consequence V_{in}, given by the voltage drop caused by the anode current i across the input resistance R of the discriminator, traverses the fixed threshold V_{th} with different delays respect to when the light is collected on the photo-cathode of the PMT. The time-walk effect is the systematic shift thus introduced in the determination of t_s (generation of the *start* pulse) that can be

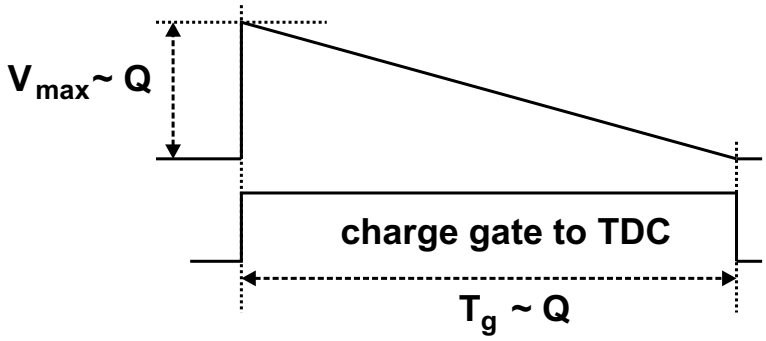

Figure 1.13: principle of operation of the charge-time conversion.

corrected in the data analysis phase with the information on Q given by the charge section, at the expense of a further contribution to the uncertainty of the time measurement.

The double-threshold discriminator employs two independent discriminators characterized by different (low and high) thresholds. The low threshold is kept at a very small value to minimize the time-walk effect on the generation of the *start* pulse; the second discriminator, with the high threshold safely greater than the typical noise amplitude on the anode line, is used to validate the *start* signal thus reducing the amount of spurious pulses due to noise.

The high- and low-threshold voltages for all the discriminators of a time section are generated by means of an on-board 8-bit digital-to-analog converter (DAC); the output of the DAC can be adjusted through specific commands sent to the FE board by the TOF DSP board.

1.5.2 Charge section

Similarly to the time section, the charge section is divided into 4 independent signal processing chains corresponding to the 4 anode inputs; each chain contains a fast charge preamplifier (current integrator) and an amplitude-to-time converter.

The anode current is integrated by the preamplifier, whose output voltage rises rapidly (in a time of the order of 10 ns) to the maximum value, proportional to the total integrated charge Q (see fig. 1.13); this output voltage enters the amplitude-to-time converter, where it is stored on a capacitor C which is then slowly discharged through a constant current source, in a similar way to the time stretcher described above. In the meantime also the output of the preamplifier independently returns to its quiescent value.

The internal logic of the amplitude-to-time converter asserts the *charge gate* signal when the voltage across the capacitor C exceeds a minimum value; with this method,

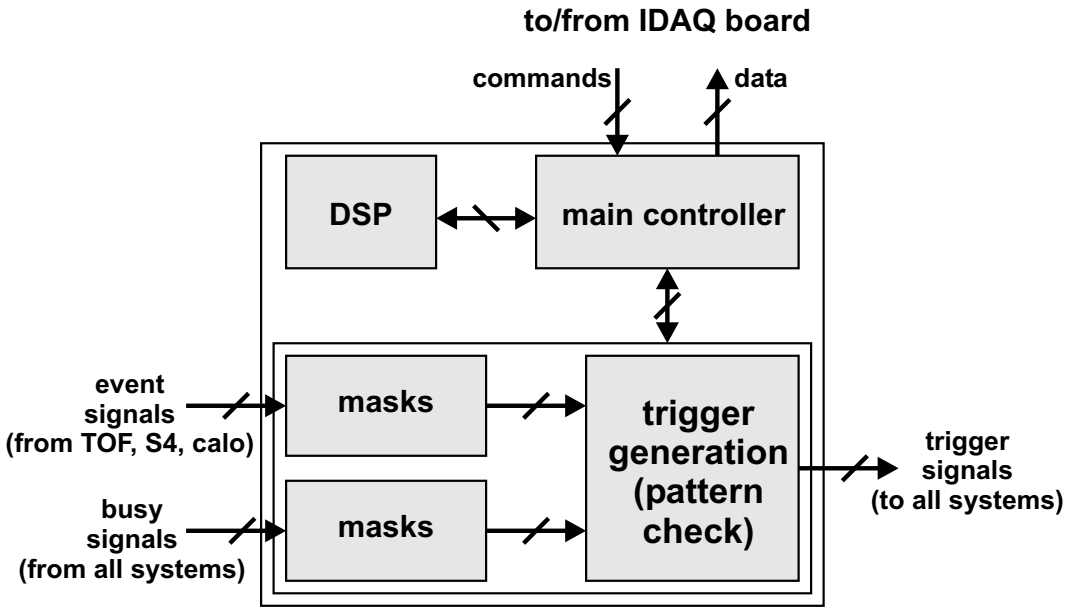

Figure 1.14: functional scheme of the trigger board.

the duration of the *charge gate* pulse depends linearly on the charge Q integrated by the preamplifier.[19]

1.6 Trigger board

The trigger board manages the synchronization and activation of the data read-out for the whole PAMELA apparatus, generating trigger pulses with suitable timings for the electronics of the various detectors, the PSCU and the IDAQ board, when a certain number of detectors, used in the trigger logic, generate signals above the respectively fixed thresholds.

The functional scheme of the trigger board is shown in fig. 1.14; the required control functionalities are implemented on 9 programmable logic chips (FPGA, see sect. 2.6.).

The board receives *event signals* from the 48 read-out channels of the TOF scintillator system[20], from the *autotrigger* part of the calorimeter (see sect. 1.9) and from the S4 detector; these signals enter a *mask* stage that makes it possible, for example, to ignore a single channel in case of its permanent failure (presence of spurious pulses because of short-circuit or too noisy line). After the mask stage a *trigger generation* stage logically

[19]The contribution of the short rise time to the total length of the *charge gate* can be neglected if the discharge of C is sufficiently slow.

[20]The thresholds for the generation of the event signals by the time sections of the TOF FE boards can be adjusted by means of suitable commands; see sect. 1.5.1.

combines the selected channels according to the currently active *trigger pattern*, to decide whether the condition for the generation of the trigger signals is met or not. The trigger pulses are sent to the various systems with different timings (from tens of ns to about 1 μs), according to the characteristic shaping times of the signals produced by the different detectors.

The applicable trigger patterns (see tables 1.2 and 1.3) are given by the possible OR combinations of the S4 signal, the *calo* (calorimeter) signal and one of 7 different TOF patterns, thus having 31 available trigger patterns.

During standard operation the selected active pattern is:

$$\text{TOF} = (\text{S11 OR S12}) \text{ AND } (\text{S21 OR S22}) \text{ AND } (\text{S31 OR S32})$$

meaning that the incoming particle must have produced an event signal on at least one of the two layers forming each of the 3 double planes S1, S2 and S3 of the TOF scintillator system. If this condition is met, then there is the timing information concerning the crossing of at least 3 of the 6 scintillator layers along the particle trajectory; besides, since the geometrical dimensions and positions of the planes S1, S2 and S3 are designed to match the geometric acceptance of the spectrometer, the probability that the particle has crossed all the 6 planes of the tracking system is maximized.

The other trigger patterns are intended for specific tasks. For example, including only the planes S1 and S2 in the pattern would allow to detect and study also low-energy particles that loose all their energy before reaching the S3 scintillator;[21] clearly in this case both the information on timing and trajectory is less precise than for particles satisfying the standard trigger pattern.

On the other hand, by considering in the pattern the *calo* signal, it will be possible to detect also primary particles producing showers in the calorimeter without traversing the TOF and tracking system detectors (for example entering the calorimeter from a lateral wall); this is particularly useful to increase the effective acceptance of the apparatus for the study of high-energy electrons (up to more than 1 TeV) which constitute relatively rare events.

The trigger board collects a set of *busy* signals from all the subsystems of PAMELA, to detect when the apparatus is ready to process a new event. The standard operation implies that every subsystem sets its *busy* line to 1 after reception of the trigger, and resets the line to 0 only after the end of the internal processing of the read-out data. The trigger board does not generate other trigger pulses, even in presence of a combination

[21] The minimum kinetic energy for protons to reach S3 is about 80 MeV.

general trigger patterns
TOF
TOF OR S4
TOF OR *calo*
TOF OR S4 OR *calo*
S4
calo
S4 OR *calo*

Table 1.2: the seven general trigger patterns allowed. The TOF pattern is selected among seven possible patterns, shown in table 1.3.

specific TOF patterns
(S11 + S12) * (S21 + S22) * (S31 + S32)
(S11 * S12) * (S21 * S22) * (S31 * S32)
(S21 + S22) * (S31 + S32)
(S21 * S22) * (S31 * S32)
S12 * (S21 * S22)
(S11 + S12) * (S31 + S32)
(S11 * S12) * (S31 * S32)

Table 1.3: the seven possible TOF patterns (+ means OR, * means AND). Sxy (x=1,2,3; y=1,2) represents one of the two layers that form each of the 3 planes of the TOF scintillator system.

of event signals matching the active pattern, until all the *busy* lines have returned to 0. Clearly also the *busy* inputs can be internally masked to assure the correct functionality in case of failure of one or more of them.

In practice the last *busy* line to be reset is the one generated by the IDAQ board, which has to complete the transfer of all the acquired data from the DSP boards of the various detectors to the PIF interface of the PSCU. The time elapsing from the assertion of the trigger signal to the end of this *busy* pulse represents the *dead time* of the apparatus, during which PAMELA is not able to process further physics events. The dead time varies for each single acquisition and is directly measured by the internal logic of the trigger board on an event-by-event basis; a discussion of the main contributions to the dead time of the apparatus is presented in sect. 3.7.1.

The main controller FPGA of the trigger board, with the help of an on-board DSP, stores significant parameters of the currently processed event, such as the dead time (defined above), the *live time* between the reset of the *busy* lines and the following trigger, or the pattern of input pulses generating the trigger. Besides, at each new event it

updates global counters storing information such as the total number of triggers from the last initialization of the board, or the pulse rates for single TOF planes or specific patterns of photo-multiplier tubes.

A bidirectional data link toward the IDAQ board allows to change the internal configuration, in particular to set the mask and active trigger patterns, and to read the data collected during the operation; in response to specific commands, trigger pulses can be generated even in absence of event signals and used during the calibration of the apparatus.

1.7 Magnetic spectrometer

The magnetic spectrometer, designed and developed by the Florence group of the PAMELA collaboration [9] [10] [11] [12], is conceived to give a very precise measurement of momentum modulus $|\vec{p}|$ and charge of the incident particle.

It is composed of a permanent magnet[22], with an internal rectangular cavity, and a tracking system with 6 planes of double-sided Si microstrip detectors, uniformly positioned along the cavity; each plane measures both the X and Y coordinates of the crossing point of an incident ionizing particle, for the reconstruction of the trajectory within the cavity.

The principle of operation of the magnetic spectrometer can be described as follows (see fig. 1.3 for reference): a particle of charge $\mathrm{q} = \mathrm{z} \cdot e$ (z positive or negative, e electronic charge) with given initial downward going momentum $\vec{p} = -|\vec{p}|\,\widehat{k}$ entering the cavity, characterized by the presence of a strong magnetic induction $\vec{B} = -|\vec{B}\,|\,\widehat{\jmath}$ (supposed uniform for clarity), moves along an arc of circumference in the XZ plane with curvature k[23] given by:

$$\mathrm{k} \approx 0.3 \cdot |\vec{B}\,| \cdot \frac{\mathrm{z}}{|\vec{p}|} \tag{1.3}$$

where k is expressed in m^{-1}, $|\vec{B}\,|$ in Tesla, $|\vec{p}|$ in GeV/c [3].

From the knowledge of $|\vec{B}\,|$ and the measurement of k, operated through the reconstruction of the trajectory, it is possible to determine the ratio $\rho = |\vec{p}|/\mathrm{z}\,e$, known as the *magnetic rigidity* of the particle:

$$\rho \approx 0.3 \cdot \frac{|\vec{B}\,|}{\mathrm{k}} \tag{1.4}$$

[22] The choice of a standard electromagnet would have resulted in a greater magnetic field, but at the expense of a too large power consumption; on the other hand, the construction and operation of a superconducting magnet would have been quite complex for a satellite mission.

[23] The absolute value of the curvature k is given by 1/R, where R is the radius of the circumference; the sign of k is defined as positive when the particle moves clockwise along the circumference in the XZ plane, negative for counter-clockwise motion.

where ρ is expressed in GV/c, $|\vec{B}|$ in Tesla and k in m^{-1}. The sign of ρ is equal to the sign of z, which is fundamental in the discrimination between a particle and the corresponding antiparticle (proton from antiproton, electron from positron). From the measurement of the ionization energy loss ΔE in the silicon detectors, it is possible to determine the absolute value of z ($\Delta E \propto z^2$) and hence $|\vec{p}|$.[24]

In practice, in the actual reconstruction of the trajectory, we must take into account that the magnetic field is not exactly uniform and that the momentum of the incoming particle is not necessarily directed along the Z axis.

The performance of the spectrometer in the measurement of the rigidity depends critically on the precision with which the tracking system can determine the curvature of the trajectory, that is strongly affected by the intrinsic spatial resolution of the detectors; the silicon microstrip detectors adopted satisfy quite well this requirement, as will be discussed in sect. 1.7.2.

On the other hand, a stronger magnetic field in the cavity volume enhances the bending effect on the incoming particles and improves the overall performance of the spectrometer. Once fixed the mass of the magnetic material surrounding the cavity[25], the field intensity can be increased by reducing the cross-section of the cavity, with the drawback that in this way the geometric acceptance of the system is reduced and consequently also the number of particles that can be detected. Hence the dimensions of the magnetic cavity for the PAMELA spectrometer has been chosen for an optimum compromise between a sufficiently high number of particles entering the geometric acceptance of the system and the possibility of achieving a precise measurement of their rigidity.

The precision in the measurement of rigidity obtainable by a magnetic spectrometer becomes worse for higher values of $|\rho|$, since the trajectory of the particle is less affected by the magnetic field and tends toward a straight line. This in particular implies that it is more and more difficult to determine the sign of the rigidity, with a higher probability that a particle is wrongly identified as the corresponding antiparticle and vice versa.

This effect is known as *spillover* and is particularly important in the case of the PAMELA experiment, which aims to explore the high-energy spectrum of $\overline{p}$ and e^+, since the ratio between the flux of a particle species and the corresponding antiparticle is high ($\sim 10^4$ protons per $\overline{p}$, $\sim 10^2$ electrons per e^+); the resulting contamination of particles in the antiparticle sample must be properly estimated and subtracted for the measurement

[24] Independent determinations of z through measurements of the ionization energy loss are also performed by the TOF system and by the calorimeter.

[25] In the case of PAMELA the mass of the permanent magnet has been fixed to 120 kg on the basis of the limit imposed by the transport capabilities of the satellite.

of the intensity of incident antiparticles.

The quality of the measurement operated by a spectrometer, in the limit of high absolute values of the rigidity, is usually expressed in terms of the *maximum detectable rigidity* (MDR), defined as the absolute value of ρ for which the relative error in the measurement is equal to 100% or, in other terms, for which $\Delta\rho = |\rho|$.

On the other hand, the discrimination of antiparticles from the corresponding particles can be performed up to a value $|\rho_{\rm sp}|$ for which the *spillover* background and antiparticle signal are comparable; $|\rho_{\rm sp}|$ depends not only on the intrinsic resolution of the spectrometer but also on the expected intensities of the various particle species and usually is much smaller than the MDR.

The PAMELA spectrometer has been designed to achieve an MDR of at least 740 GV/c for particles with both positive and negative charges and, basing on the predicted particle spectra, to identify the antiproton component up to $|\rho| \approx 190$ GV/c and the positron one up to $|\rho| \approx 270$ GV/c.

1.7.1 Magnet

The permanent magnet is composed of 5 superposed identical modules shaped as right parallelepipeds with a central rectangular cavity (see fig. 1.15 and 1.16).

Each module is obtained by gluing together several magnetized elements, shaped as right prisms with triangular section, and arranged in such a way to produce inside the cavity a quasi-dipolar magnetic field $\vec{B}$ with practically all the strength concentrated along the Y axis (the X and Z components are less than 10% of the Y component) and rather uniform; the presence of a preferential direction of the magnetic field greatly simplifies the algorithms for the reconstruction of the trajectory of the charged particle, whose curvature develops almost entirely in the XZ plane (*bending plane*).

To maximize the intensity of the field, a Nd-Fe-B alloy (Vacodym 335 HR [13]) was selected as magnetic material, because of its high residual induction (≈ 1.32 T).

Each module is housed inside a light and diamagnetic support frame in Al; the free surfaces are covered with a thin (0.3 mm) Al layer to protect them against external agents. The 5 aluminium frames are fixed together to form a tower characterized by an internal cavity of area 16.14 cm $\cdot$ 13.14 cm and height 43.66 cm. Between two adjacent modules a vertical gap of 8 mm allows for the insertion of the silicon detector planes.

The magnetic field $\vec{B}$ inside the cavity shows a good uniformity: in the 80% of the volume it does not change by more than 10%. The maximum value of the field is 0.48 T

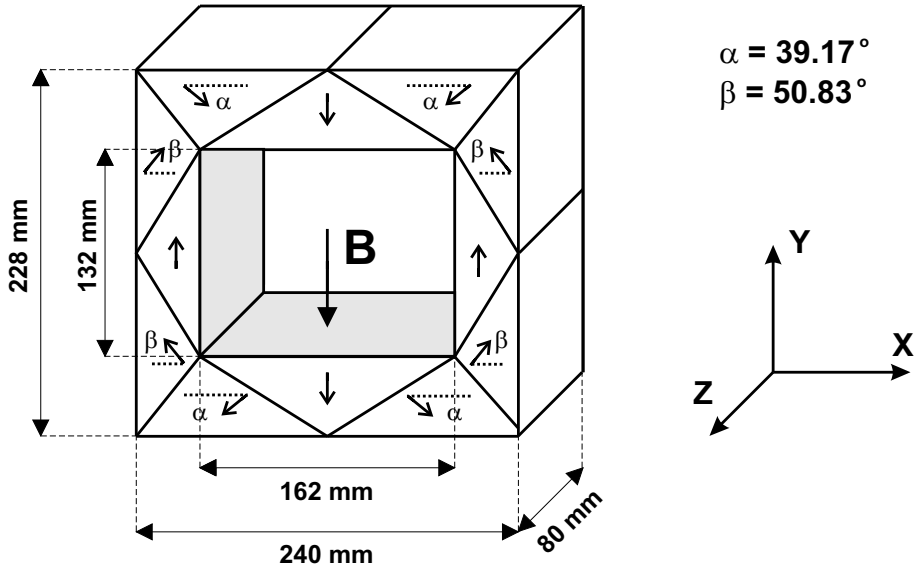

Figure 1.15: one of the 5 modules of the PAMELA magnet; the direction of the residual magnetic induction of the elements forming the module is also shown.

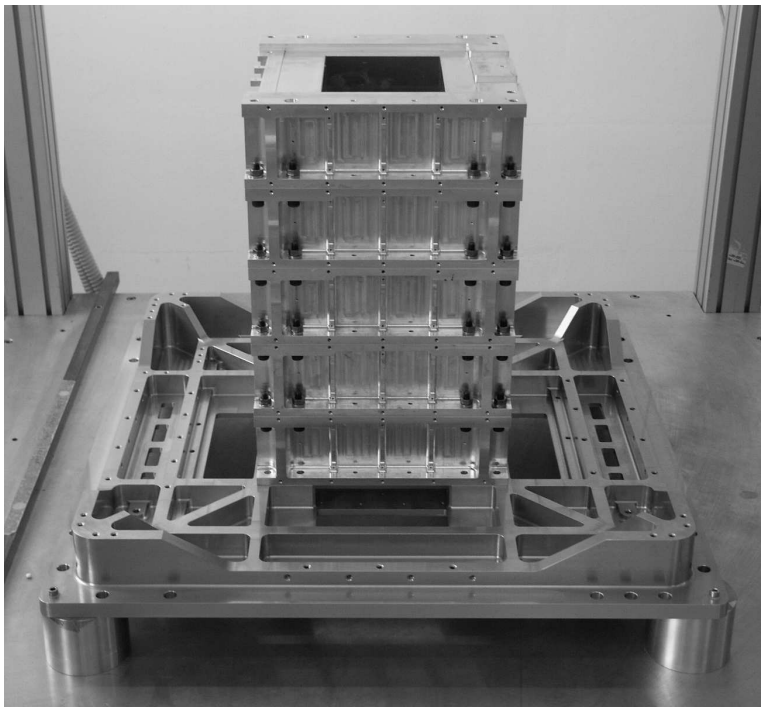

Figure 1.16: the magnetic tower after assembling; each of the 5 modules is fixed inside the corresponding Al support frame.

at the centre of the cavity, the average value is 0.43 T.

The spatial dependence of the magnetic field inside the cavity was measured in Florence with a three-axial Hall probe (FW-Bell 9950 [14]) driven by a motorized machine in 70000 points with steps of 5 mm; analogous measurements were done in the region

outside the cavity, in particular between the S1 and S2 planes of the TOF system, with a spatial resolution of 3 cm.

1.7.2 Tracking system: detectors

The tracking system of the PAMELA magnetic spectrometer is composed of 6 planes of high-precision double-sided Si microstrip detectors, positioned between the 5 magnetic modules and on the top and bottom of the magnetic tower, with equal spacing of 8.9 cm. The planes are indicated as T1 to T6 starting from the one positioned at the highest Z coordinate in the reference frame of PAMELA.

The main characteristic of these detectors is the presence of a sensitive layer both on the upper and lower side (or *view*) of the Si plane, with the implanted strips of one side orthogonal to those of the other, so as to achieve the measurement of both the X and Y coordinates of the crossing point of an incident ionizing particle. Other distinctive features, illustrated in what follows, are the presence of an integrated capacitive coupling for read-out and of a double metallic layer on one of the views, to have the metallic read-out strips parallel on both sides. The use of double-sided detectors, instead of pairs of single-sided ones, offers several advantages, in particular the smaller thickness of material traversed by the incident particle[26] and a simplified support structure for the planes.

Each detector plane (see fig. 1.17, 1.18 and 1.19) is divided into three identical independent sections (known as *ladders*) along the X axis and housed in an aluminium support frame; each ladder is formed by two rectangular (5.3 cm $\cdot$ 7.0 cm) silicon sensors and a *hybrid* circuit, containing the first stage of the front-end electronics, realized partly with implantation and partly with standard printed-circuit board techniques (hence the name) on a Al_2O_3 insulating layer. The three components of a ladder are directly glued together; the 3 ladders constituting the plane are alternated and glued to 4 carbon-fibre rails, that stiffen the whole structure and guarantee the mechanical contact of the plane with the aluminium frame, through many glue points.

Each silicon sensor (manufactured by Hamamatsu [7]) is formed by a substrate of thickness 300 μm, with residual n-type doping and high resistivity; on one side 2035 p^+ parallel strips are implanted to form p^+-n junctions with the substrate, separated each other by a 25.5 μm pitch. On the opposite side there are 1024 n^+ strips, with 67 μm pitch, alternated to p^+ *blocking strips* (see below); the n^+-n contact is of ohmic type

[26] This reduces the probability of *multiple scattering* of the particle inside the traversed material. The possible changes of direction caused by multiple scattering in the detector planes must be taken into account for the correct reconstruction of the trajectory in the magnetic cavity, contributing to the overall uncertainty in the measurement of the curvature.

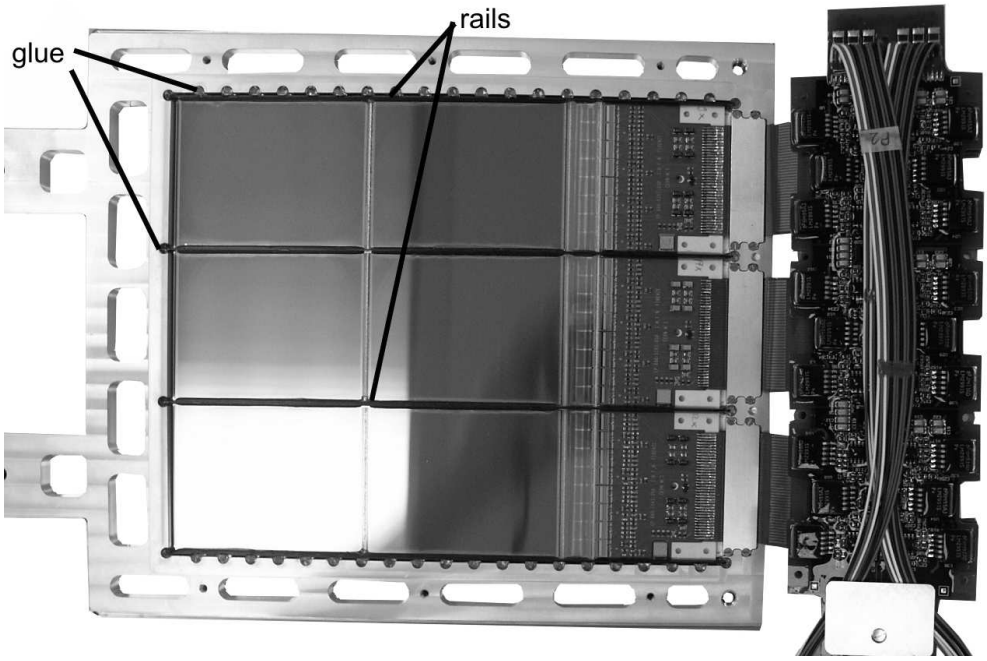

Figure 1.17: one of the 6 planes of the tracking system, completed with the front-end electronics on separate boards. Note the subdivision of the plane into 3 identical ladders, alternated to 4 carbon-fibre rails, and the glue points fixing the rails to the aluminium frame: 21 for the external carbon-fibre rails, 2 for the internal ones.

i. e. non-rectifying.

Since, as explained below, the junction view is characterized by a much better spatial resolution, it is used to measure the X coordinate of the impact point of the charged particle, which is by far the most important for the determination of the curvature of the trajectory, because the particle tends to bend in the XZ plane, normal to the main component of the magnetic field. For this reason the junction view is also indicated as *X view* or *bending view*; on the other hand, the ohmic side, measuring the Y coordinate, is indicated as *Y side*.

During standard operation a voltage difference of +80 V (*bias*) is applied between the n^+ strips of the Y side and the p^+ ones of the X side so as to reverse-bias the p^+-n junctions and completely deplete the substrate. When an ionizing particle traverses the sensor, the e^-/hole pairs, created in the depletion region, drift along the electric field lines toward opposite directions; holes are collected by the p^+ strips on the X side, e^- by the n^+ ones. From the distribution of the ionization charge collected by the strips it is possible to reconstruct the coordinate of the crossing point in the direction normal to the strips; besides, from the total collected charge on each side it is possible to determine the ionization energy loss of the particle across the Si layer.

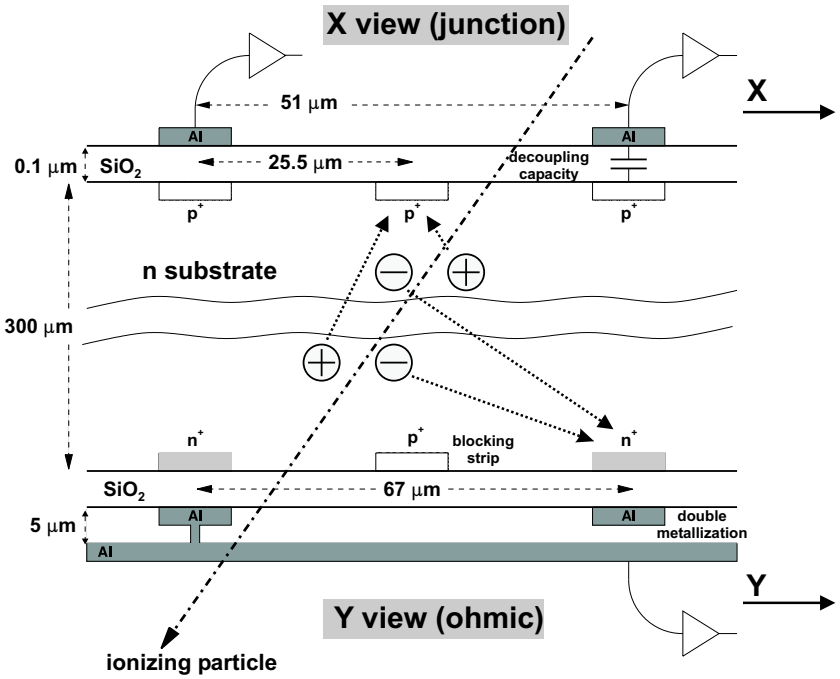

Figure 1.18: section of a silicon sensor, with a draft of the charge collection mechanism. To better illustrate the structure, the Y side is represented rotated by 90° with respect to the X side.

The charge read-out is performed for one out of two strips (i. e. 1018) on the junction side and for all the 1024 n^+ strips on the ohmic one, by means of a dedicated Al electrode placed above the strip for its whole length and separated from it by a thin (0.1 μm) insulating layer of SiO_2, thus obtaining an integrated capacitive coupling of ≈ 20 pF/cm.

The integration of the coupling capacity directly inside the sensor has the advantage of greatly simplifying the read-out electronics, avoiding the presence of about 3000 capacitances on each side of the hybrid. On the other hand, to avoid applying the +80 V bias across the thin SiO_2 integrated capacities on the detector planes, the reference ("ground") voltages of the front-end electronics have been short-circuited with the voltages applied to the strips of the corresponding sides; therefore the reference voltage for the Y-side electronics is +80 V above the one for the X-side read-out, which coincides with the ground of the PAMELA apparatus. This has made necessary the introduction of a dedicated inductive high-pass filter on the electric connections between the front-end electronics and the following stages, as explained in sect. 3.3.

At the interface between Si and SiO_2 a thin layer of fixed positive charges develops inside the oxide near the separation surface with the Si substrate; these charges attract free e^- of the substrate toward the oxide/Si interface with the effect of locally increasing the conductivity; on the ohmic side, where the n^+ strips do not form junctions with the

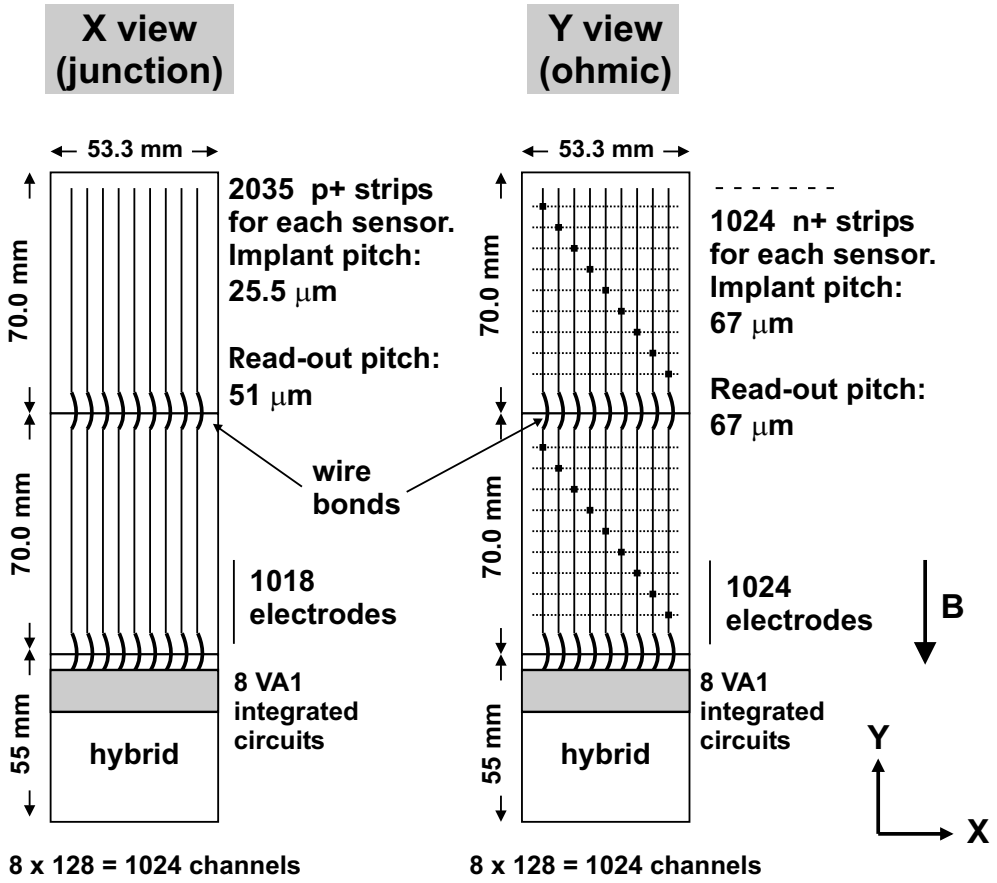

Figure 1.19: scheme of the X and Y view of a ladder; note on the Y side the association of two n^+ strips, 7 cm apart from each other, to the same metallic electrode.

substrate, the inter-strip resistance would be of a few kΩ, too low to avoid the mutual redistribution of the collected charge between adjacent strips and practically preventing the measurement of the Y coordinate. The purpose of the blocking strips p^+, implanted between each pair of n^+ strips and forming p^+-n junctions with the substrate, is precisely to break the inter-strip electric continuity; in fact with this solution the effective inter-strip resistance becomes of the order of 10 MΩ.

On the X side of the ladder the corresponding electrodes of the two sensors are soldered together through wire bonds of 17 μm diameter, thus having 1018 electrodes of double length (see fig. 1.19). On the Y side the n^+ strips and the electrodes above them are directed along the X axis; a second level of 1024 metallic tracks (*double metal*), oriented along the Y axis, brings the signals toward the same edge of the ladder as for the junction side. The tracks of the second level are insulated from the underlying level of electrodes by a thick (5 μm) SiO_2 layer, except for the interconnections between the two levels, associating each of the 1024 tracks with two electrodes, one per sensor and 7 cm apart

from each other.

The ambiguity thus introduced in the measurement of the Y coordinate can be easily solved by using the position information from other detectors (TOF system and calorimeter). On the other hand, bringing all the signals from the X and Y views toward the same edge of the ladder makes it possible to have a much more compact and stiff overall mechanical structure, because the read-out electronics for both views can be concentrated in a unique double-sided hybrid.

With this method there are 1024 signal channels per ladder and per view, and a total of 36864 channels for the 6 planes of the tracking system.

The performances of the detectors, in the determination of the crossing point of the ionizing particle, depend mainly on the inter-distance between read-out strips (*read-out pitch*: 51 μm for the X side, 67 μm for the Y side) and on the noise affecting the acquisition chain; in fact the algorithm for reconstructing the distribution of ionization charges along the array of strips is based on the identification of a group (*cluster*) of adjacent strips whose signals significantly differ from the value expected considering the combined contribution of the quiescent level (*pedestal*) and of the average noise.

The noise performance of the microstrip detectors can be characterized in terms of the mean *signal/noise ratio* $< \mathrm{S/N} >$ for minimum ionizing particles (MIP), defined as the mean value of the ratio between the total ionization charge collected by the strips of a *cluster* and the average noise amplitude for those strips. A complete characterization of the noise performances of the tracking system has been done[27], obtaining $< \mathrm{S/N} >$ of 56 and 26 for the X and Y views of the first plane under test and similar values (within 10%) for the other planes.

The smaller values of $< \mathrm{S/N} >$ for the Y views are mainly due to the presence of the double metallic layer, resulting in a higher effective capacitance of the detector, $\mathrm{C_{det}}$, seen at the input of the first stage of the acquisition chain (charge preamplifier); in fact the noise on the input signal increases with the detector capacitance (in first approximation, it varies linearly with $\mathrm{C_{det}}$), while the signal itself is substantially independent from $\mathrm{C_{det}}$.

The spatial resolution [17] achievable for the X (bending) views, which is critical for the determination of the curvature of the trajectory, is better than 3 μm; for the Y views the resolution is better than 13 μm.

[27] This item is described in [15]; see also [16].

1.8 Front-end electronics of the tracking system

The fundamental requirement of the front-end electronics for the tracking system is the minimization of the contribution of noise on the signal lines, while keeping the power consumption at an acceptable level for a satellite application.

Moreover, the architecture of the whole electronics for the tracking system (FE and DSP parts) has been designed to be able to process the 36864 channels within a sufficiently short time, thus limiting the contribution to the overall dead time of the PAMELA apparatus to an acceptable amount (see sect. 3.7.1). This has been achieved by setting the operating frequency and the number of sections working in parallel in such a way to meet the limits in power consumption and available volume imposed by the satellite; the total power consumption for the tracking system is less than 65 W, corresponding to ≈ 1.8 mW/channel.[28]

The front-end electronics of the tracking system is housed partly on the hybrid circuits integrated within the detector planes and partly on dedicated printed-circuit boards positioned in the vicinity of the planes.

The first stage of the FE electronics for each side of a ladder is formed by 8 dedicated integrated circuits (VA1[29]), placed on the corresponding side of the hybrid and aligned along the X axis. A VA1 chip (see fig. 1.20) contains 128 low-noise and low-power independent acquisition chains and an internal multiplexer and shift-register for the sequential read-out of all the input channels through a unique output differential line; the digital inputs lines (*hold*, *reset*, *clock*, *shift-in*) allow the control of the internal operation of the VA1. The VA1 occupies an area of 6.2 mm $\cdot$ 4.5 mm and has 128 input pads spaced by 47 μm, looking toward the adjacent sensor; the pads are soldered to 128 metal electrodes of the sensor through wire bonds.

The second FE stage is formed by two printed-circuit boards per detector plane; the boards are functionally identical: one is dedicated to the X side and the other to the Y side. Each board is divided into 3 independent sections, corresponding to the 3 ladders of the plane, and containing the electronics for the amplification and digital conversion of the output signals of the VA1 chips and the control logics (*FE controller* FPGA) for the management of the acquisition process and the transmission of the conversion results to the DSP board. Since each FE section operates in parallel with the others, the total time required for the read-out of the whole tracking system is reduced by a factor 36, equal to

[28] A detailed description of the operation of the FE electronics of the tracking system and of the various aspects of its development can be found in my Physics Degree Thesis [15].

[29] The VA1 was developed at CERN (Geneva) and manufactured by Ideas (Norway) [18].

the number of FE sections (6 for each of the 6 planes).

The two FE boards for a plane, plus a third board containing the corresponding power circuits, are superposed on each other in a compact structure ("sandwich") and fixed to the external wall of the Al support structure of the spectrometer (see fig. 1.17 and 1.22). The signal exchange between one side of a hybrid and the corresponding section of FE board is done through a flat flexible connection obtained by depositing parallel copper tracks between two thin kapton[30] layers 5 cm long.

1.8.1 First stage: VA1

Each of the 128 input pads of the VA1 chip is internally connected to an independent read-out chain, composed of a charge preamplifier (current integrator), a CR-RC shaper (passband filter) and a track-and-hold (T&H) circuit (see fig. 1.20).

When an ionization current pulse[31], collected by an electrode across the SiO_2 coupling capacitance, enters the VA1 through the corresponding pad, the output of the shaper V_{sh} rises to a maximum value, depending linearly on the charge Q integrated by the preamplifier, with a characteristic rise time of ≈ 1 μs; V_{sh} then falls again to the quiescent voltage within a time of the order of 10 μs.

Normally (*hold* control input kept to 0) the T&H is the *track* state with its output $V_{T/H}$ reproducing V_{sh} through a buffering amplifier stage; when $V_{T/H}$ reaches the maximum value, the FE controller FPGA, synchronized with the trigger pulse generated by the TOF system and hence with the instant of particle crossing, asserts the *hold* line forcing the T&H in the *hold* state. The *hold* line is kept asserted until all the 128 $V_{T/H}$ have been digitally converted in the FE board; during this time (2 ms) the value of $V_{T/H}$ remains sufficiently stable thanks to the high decay constant characterizing the discharge of the input capacitance of the T&H circuit through the input resistance of the amplifier stage.

1.8.2 Second stage: FE boards

As mentioned above, the front-end board is divided into 3 identical sections, each dedicated to a corresponding group of 8 VA1, for a total of $8 \cdot 128 = 1024$ input channels; the schematic structure of a section of FE board and of the 8 VA1 chips on the hybrid is

[30] Kapton is an organic polymeric plastic film with optimized physical properties, produced by DuPont (http://www. dupont. com) and widely employed as a high-resistivity insulating substrate for flexible printed circuits.

[31] The duration of the current pulse is practically given by the maximum drift time across the Si substrate for holes on the junction (X) side, kept at 0 V and for e^- on the ohmic (Y) side, at +80 V; these are 30 ns and 8 ns respectively.

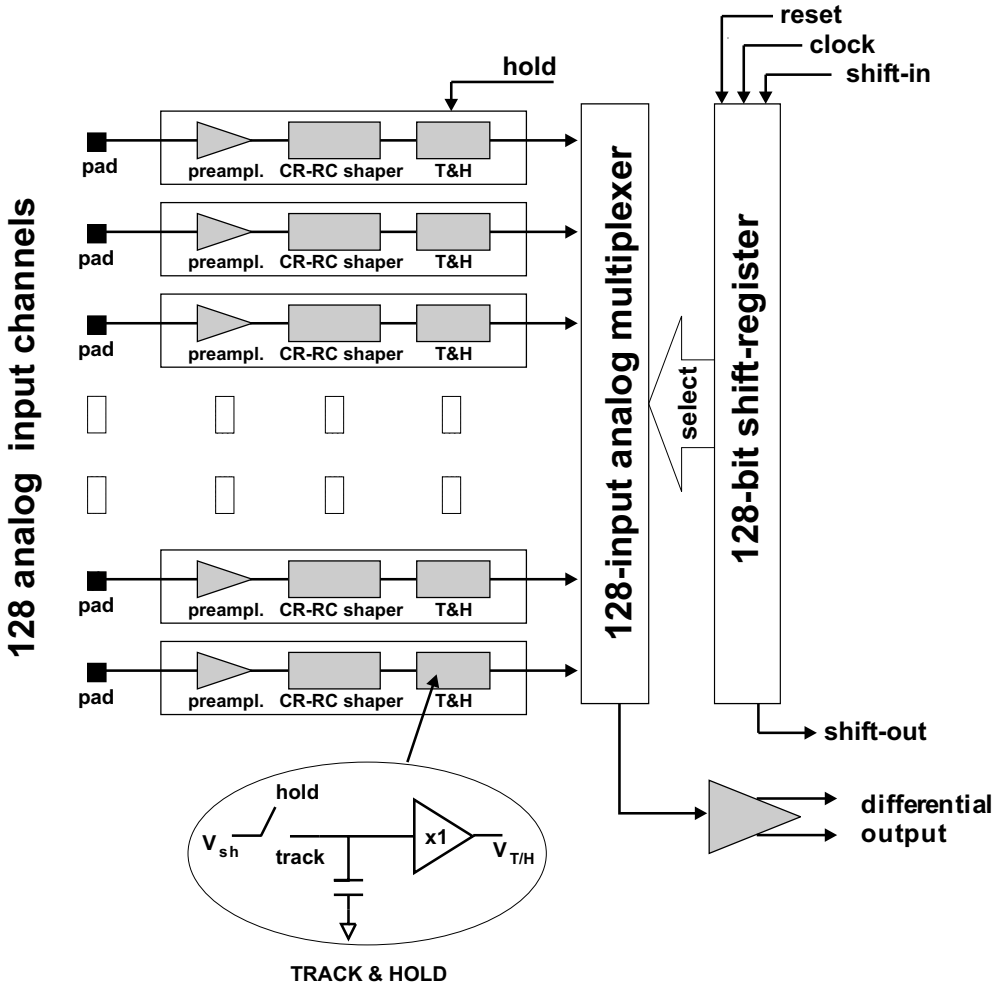

Figure 1.20: basic scheme of the VA1 chip.

represented in fig. 1.21.

An analog-to-digital converter (ADC) sequentially processes the 1024 channels, while the FE controller FPGA manages all the phases of operation both of the ADC and of the VA1 chips and constitutes the interface with one of the two DSP boards of the tracking system.[32]

The differential outputs of the 8 VA1 reach the unique ADC input through a differential amplifier stage[33], followed by an analogical switch (ADG701 [19]) and a final summing

[32]The two DSP boards of the tracking system, dedicated to the X and Y sides of the detectors respectively, will be described in detail in chapter 3.

[33]Note that the set of 8 VA1 is segmented into 4 independent pairs; on the hybrid circuit, the positive and negative differential outputs of a VA1 are respectively short-circuited with those of the other VA1 of the same pair; this is possible because, when quiescent, the VA1 holds its output in a high-impedance state. This solution is necessary to reduce the number of required amplification chains from 8 to 4, because of the limits in the available area both on the FE board and on the flat connections toward the hybrid.

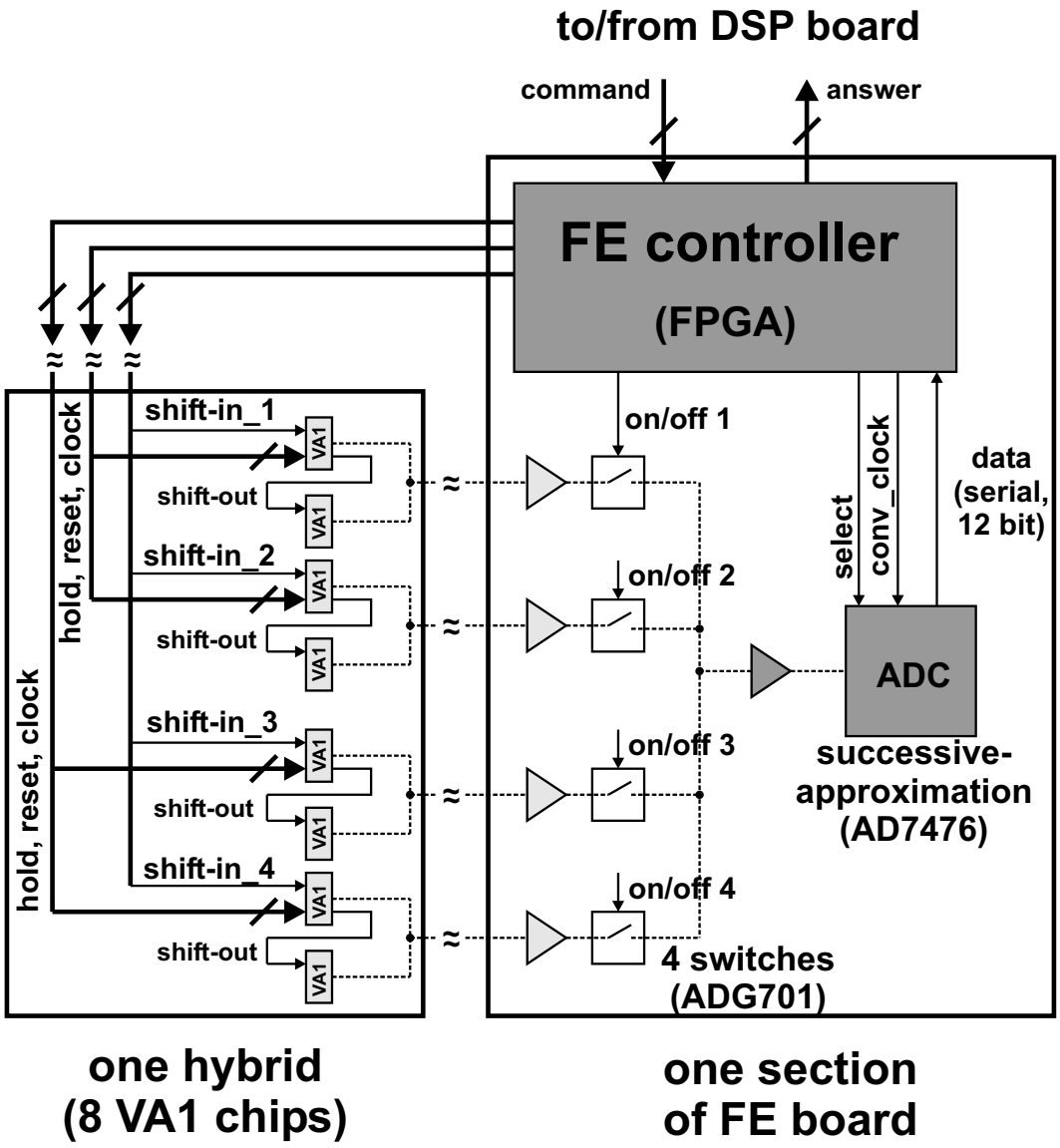

Figure 1.21: scheme of one section of a front-end (FE) board of the tracking system, together with the corresponding 8 VA1 chips on the hybrid circuit.

amplifier stage. The function of the analogical switch is to break the electrical continuity of the corresponding line even in the case of a failure (e.g. permanent short-circuit on the output of a VA1).

After the generation of the trigger signal, the control logics of the DSP board starts sending a command sequence to the FE controller (see sect. 3.7), which immediately asserts the *hold* signal for the 8 VA1, and then enables the shift register of the first VA1 through a *shift-in_1* pulse. The subsequent 128 *clock* pulses select one after another all the input channels of the first VA1 to be reproduced on its differential output; when the 128-th *clock* pulse is received by the first VA1, it automatically generates a *shift-out* pulse

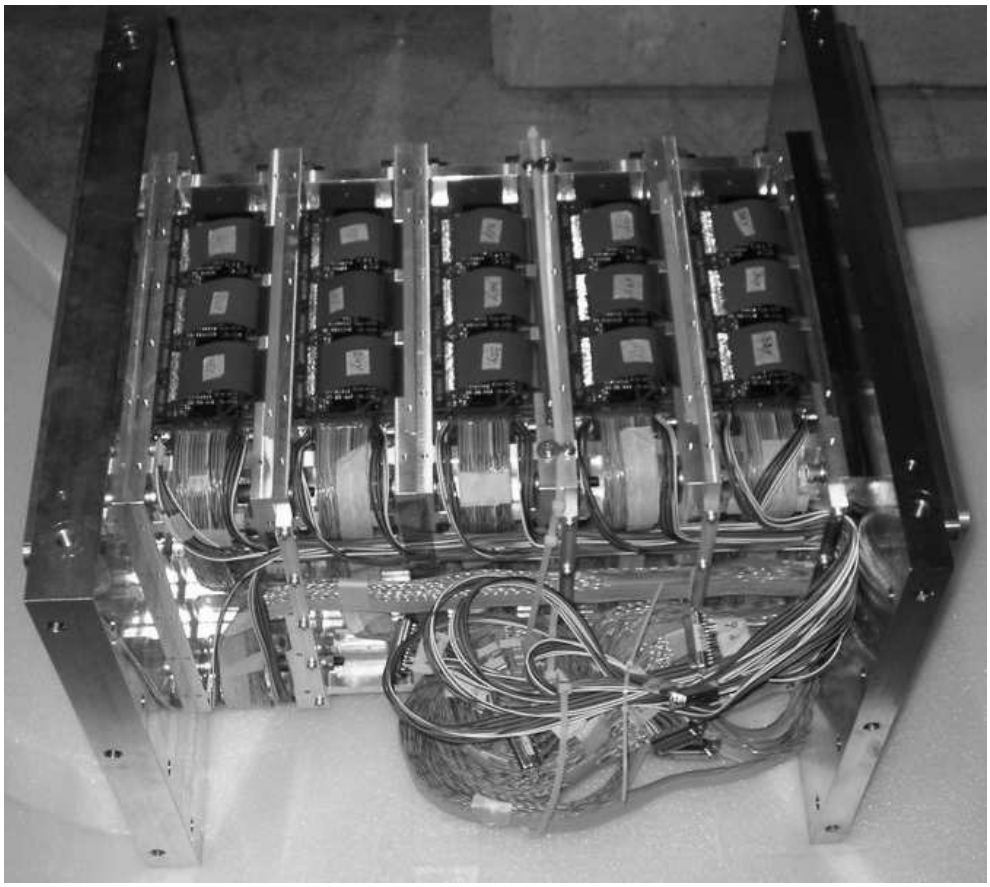

Figure 1.22: the completed spectrometer before its integration within the PAMELA apparatus.

that is used as *shift-in* to enable the second VA1; the following *clock* pulse brings the first VA1 into its quiescent state (differential output in high-impedance) while the second VA1 begins its read-out cycle. The same sequence of operations is repeated for the 4 VA1 pairs; finally the *hold* line is negated and a *reset* pulse is sent to all the VA1 chips.

The ADC model employed in the FE board, AD7476A [19], is a 12-bit ADC, composed of a T&H input stage followed by a successive-approximation digital converter, characterized by low noise, limited power consumption and small area occupation (3 mm $\cdot$ 3mm). The choice of the number of conversion bits and of the specific ADC model was based on the requirement that the maximum quantization error, $\pm$ 0.5 *least significant bits* (LSB), and the maximum non-linearity error, $\pm$ 1 LSB[34], introduced by the ADC should be significantly smaller than the total noise contribution of the read-out chain, and at the same time that the ADC should not saturate for signals within the dynamic range of the

[34]The linearity characteristics of the AD7476A have been verified during my Physics Degree work [15] with dedicated tests at the sampling frequency (0.5 MHz) actually employed in the read-out of the tracking system, confirming the manufacturer specifications: absence of missing codes and differential non-linearity within $\pm$1 LSB over all the dynamic range.

VA1 chips (equivalent to 10 MIP). With the adopted solution (12-bit ADC for a full scale of 4096 LSB) the noise of the read-out chain typically is 4.5 LSB for the X views and 10 LSB for the Y views; given the average measured signal/noise ratios for MIP (56 for the X sides, 26 for the Y sides), the 10-MIP signal corresponds to about half ADC scale ($\sim 2500 - 2600$ LSB) for both views, as desired.

The ADC conversion cycle begins when the FE controller FPGA asserts the *select* control line that puts the input stage of the ADC in the *hold* state (the input is sampled); the subsequent 16 *conv_clock* pulses drive the successive steps of the internal conversion machine. The ADC serially produces the 12-bit result of the conversion on a single *data* output line; this is read by the FE controller that immediately transmits the 12-bit value to the DSP board. The length of the conversion cycle has been set to 2 μs (corresponding to a sampling frequency of 0.5 MHz), for a total time of ≈ 2 ms for the sequential read-out of the 1024 input channels; this is also the time necessary for the complete read-out of the tracking system, since the 36 FE sections work in parallel.[35]

1.9 Imaging calorimeter

The imaging calorimeter of PAMELA [20] (fig. 1.23) is composed of 22 equal superposed modules, with squared section of 24 cm $\cdot$ 24 cm and total height of 18 cm, each containing a central W (tungsten) absorber plate of thickness 0.26 cm between two planes of single-sided Si strip detectors (thickness 380 μm), whose operation is similar to that of the microstrips employed in the tracking system.

Each Si detector plane is formed by an array of $3 \cdot 3$ square sensors of area 8 cm $\cdot$ 8 cm; on each sensor 32 strips are implanted with 2.4 mm pitch[36]. The 9 sensors of a plane are arranged in 3 ladders in a similar way to that discussed for the tracking system, thus giving a total $32 \cdot 3 = 96$ read-out channels per plane. The strips of the upper and lower plane of a module are orthogonal to each other, providing two-dimensional spatial information.

A primary particle crossing the W absorber layers can produce a cascade (or *shower*) of secondary particles, which release energy by ionization in the Si detectors; the spatial information given by the Si strips allows to detect the presence of a shower and to finely reconstruct its longitudinal and transversal profiles (the *image*). The *electromagnetic*

[35] In sect. 3.7.1 this and the other contributions to the overall dead time of the PAMELA apparatus are discussed.

[36] This horizontal pitch, together with the vertical inter-distance between adjacent modules, is sufficient to assure the required granularity of the detector for the optimum separation between $e^{\pm}$ and $p/\overline{p}$.

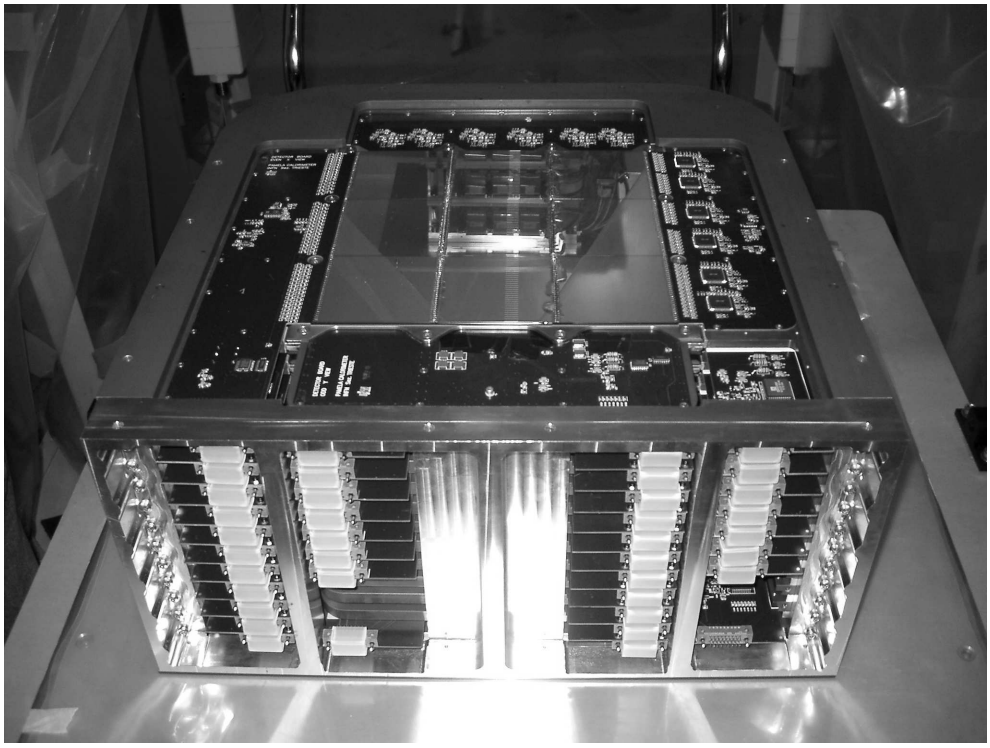

Figure 1.23: the imaging calorimeter of PAMELA; the first silicon plane and its support board housing the corresponding FE electronics are visible on the top.

shower produced by an e^- or e^+ and the *hadronic shower* generated by a proton or antiproton are characterized by different interaction mechanisms that are responsible for the formation of the cascade of secondary particles.[37] This difference is exploited to distinguish an incoming primary e^+ from a p of equal momentum (or an $\overline{p}$ from an e^-).

The electromagnetic shower generated by a high-energy (at least several GeV) e^- or e^+ traversing a thick layer of material is caused by an initial *bremsstrahlung* interaction with emission of a high-energy photon; this photon can then annihilate into a new high-energy pair $e^\pm$ which can interact again by bremsstrahlung and so on, with the creation of a large number of secondary $e^\pm$ and photons. The number of particles in the shower reaches a maximum value when the average fraction of primary energy carried by the secondary particles approaches the *critical energy* (≈ 8 MeV for W), below which the ionization energy loss becomes the dominating mechanism of interaction for $e^\pm$ and the Compton scattering for photons.

The mean profile of the longitudinal energy deposition in the absorber for an electromagnetic shower is well described in terms of the *radiation length* X_0[38]: for typical

[37] See for example [3] for a general review of calorimeter techniques and [21] for a detailed discussion of the properties of the imaging calorimeter of PAMELA.

[38] The radiation length X_0 is defined as the mean thickness of traversed material after which a high-

primary energies (from some GeV to hundreds of GeV) the maximum energy deposit is located at a thickness t^{el}_{max} between 4 and 10 X_0 (increasing logarithmically with the primary energy) and 98% of the released energy is contained within

$$t^{el}(98\%) = t^{el}_{max} + 13.6\ X_0 \tag{1.5}$$

from the point of the first interaction. $t^{el}(98\%)$ thus typically ranges between 17.6 and 23.6 X_0 (6 to 8 cm in tungsten).

On the other hand the hadronic shower produced by an incident p or $\overline{p}$ is initiated by an inelastic nuclear interaction causing the excitation and fragmentation of a nucleus of the absorber material, with subsequent nuclear interactions of the recoiling fragments. In this case the average longitudinal profile is characterized by the *nuclear interaction length* λ_I[39]: the maximum energy deposit is between 1 and 2 λ_I, while most of the released energy is contained within

$$t^{hadr}(95\%) = t^{hadr}_{max} + 2.5\ \lambda_I \cdot (E/GeV)^{0.13} \tag{1.6}$$

from the primary interaction point. $t^{hadr}(95\%)$ ranges from 3.5 to 6 λ_I (34 to 58 cm in W) for the primary energies considered ($\approx 1 - 100$ GeV).

Analogous considerations can be done for the transversal profile of the shower, which, for a given primary energy, is more compact in the electromagnetic case, because of the smaller mean transverse momentum of the secondary particles.

The choice of the absorber material and of its total thickness for the calorimeter of PAMELA were operated with the purpose of maximizing the development of electromagnetic showers, at the same time minimizing the hadronic interactions of incoming $p/\overline{p}$.

The total thickness of the 22 tungsten layers (5.72 cm) is equal to 16.3 radiation lengths X_0 and only 0.6 interactions lengths λ_I, resulting in a high probability that an incident $e^{\pm}$ initiates an electromagnetic cascade, with a good longitudinal (and also transversal) containment of the shower. On the other hand, the probability that a hadronic shower is generated by an incoming $p/\overline{p}$ undergoing an inelastic nuclear interaction is given by $1 - \exp(-0.6) \approx 45\%$, thus greatly facilitating the hadron/lepton discrimination on the basis of the absence or presence of the shower.

---

energy $e^{\pm}$ looses all but 1/e of its initial energy by bremsstrahlung interactions; the residual energy of the particle is described by the law $E(\Delta z) = E(0)\exp[-(\Delta z/X_0)]$, where Δz is the traversed thickness. X_0 depends on the material (atomic number Z and atomic mass A) and, for a residual energy E of the primary particle greater than ≈ 1 GeV, is independent of E. X_0 is commonly expressed in g/cm^2; for tungsten, $X_0 = 6.8$ g/cm^2, corresponding to a thickness of 3.5 mm.

[39] λ_I represents the mean free path for nuclear inelastic interactions; λ_I depends on the atomic mass A of the material and is commonly expressed in g/cm^2. For tungsten, $\lambda_I = 185$ g/cm^2, corresponding to a thickness of 9.6 cm.

With these characteristics it is expected that the calorimeter will be able, for primary energies greater than ≈ 1 GeV, to discriminate $\overline{\mathrm{p}}$ from the e^- background with a *rejection factor* of more than 10^4 (one e^- out of 10^4 is identified as an $\overline{\mathrm{p}}$) at a selection efficiency for $\overline{\mathrm{p}}$ of more than 90%; a similar performance is expected in the rejection of the p background for e^+.

The good containment of the electromagnetic showers will also make it possible the direct measurement of the energy of primary $\mathrm{e}^\pm$ or photons up to energies of about 2 TeV.

Additionally, the calorimeter is equipped with an *autotrigger* capability, i. e. it will generate a *calo* event signal for the trigger board when a specific energy distribution is detected in predetermined planes of the lower half of the instrument. This feature will allow the direct energy measurement of high-energy electrons (from ≈ 300 GeV to more than 1 TeV) with a geometric acceptance about 30 times larger than the standard one, limited to the input window of the spectrometer; the enhanced acceptance is fundamental to increase the accumulated statistics, since these high-energy events are quite rare with respect to those entering the standard acceptance of the PAMELA apparatus.

Chapter 2

Electronics of PAMELA: general characteristics

In this chapter the most significant characteristics of the electronics operating in PAMELA are illustrated, with specific attention to the digital control circuits which will constitute the argument of the following chapter. In the first part (sections 2.1 to 2.5) the problems and solutions concerning the reliability of the electronic systems will be discussed, and in particular the effects of the ionizing radiation present in the space environment where PAMELA will operate. The second part of the chapter (sections 2.6 to 2.8) is focused on the technological characteristics and development methods involved in the design and implementation of the control logics in PAMELA.

2.1 Reliability of electronics

The main requirement that has to be satisfied by the electronics of a space experiment is to assure a high reliability during the whole duration of the mission (3 years in the case of PAMELA) and a sufficiently low power consumption.

In a practical design it is necessary to reach an optimum compromise between these two partially conflicting requirements and also other important constraints, such as the request of limiting the volume occupation. The constraints on the available power and volume are particularly important in the case of PAMELA, because of the strict limits imposed by the satellite capabilities[1]; also costs play an important role.

The main aspects characterizing the reliability of electronics are the stability of the operating conditions and the probability of temporary or permanent loss of a subsystem because of failure of a component. Failures can be caused by mechanical, thermal or

[1]The available volume for the PAMELA apparatus is about 0.5 m^3, most of it being occupied by the detectors, the magnet and their mechanical support system. The available power is ≈ 360 W.

electric stresses and by the ionizing radiations which are present in the extra-atmospheric environment where PAMELA will operate.[2]

2.1.1 Stability of operation

Generally speaking, the functional characteristics of any electronic device are sensitive to the *working conditions*, represented by temperature and supplied voltage; typically the manufacturer of the device guarantees that these functionalities are correct within a given range of voltages and temperatures.

Stable operating voltages are commonly obtained with regulated power supplies; the main drawback of this method is that for the generation of the required potential across the load, V_{load}, the regulator stage needs a higher input voltage V_{supply}; therefore for a given current i absorbed by the load, the power consumption (which would be $i \cdot V_{load}$ in absence of the regulator stage) is $i \cdot V_{supply}$, with a relative increase equal to $V_{supply}/V_{load} - 1$.

For example, the DSP boards of PAMELA, housing the digital control electronics of most detectors, operate according to the LVTTL[3] electrical standard, requiring the devices to be powered with a voltage $V_{load} = (3.3 \pm 0.3)$ V; in this case the regulator stage for each section of electronics is represented by an on-board integrated circuit, whose output is stable within a few percent and is characterized by a small dependence on temperature and output current. The input voltage to this LVTTL regulator stage, supplied from the upper level of the power system of PAMELA, and required for correct operation, is typically $V_{supply} = 3.6$ V, with a 9% increase of the power consumption respect to the unregulated case.

The stabilization of temperature conditions is another important aspect in the operation of PAMELA. The apparatus will be housed inside an aluminium shell filled with N_2 at atmospheric pressure, where the local increase of temperature in the vicinity of powered devices is favoured by the lack of thermal convection due to the practical absence of gravity in space.

To achieve an efficient heat exchange with the external environment, a cooling-loop pipe has been designed to run across the apparatus and be in good thermal contact with its various parts; the heat absorbed by the pipe through thermal conduction is subtracted with the continuous flow of a refrigerating fluid (liquid iso-octane) inside the loop; the heat is then released to the cooling system of the satellite, where it is radiated in the space.

[2] For its complexity the subject of ionizing radiation will be treated separately (sect. 2.2 to 2.4).

[3] *Low voltage transistor-transistor logic*: this standard is widely employed in digital circuits characterized by relatively low power consumption and fast operation.

With this system, and taking into account the different conditions of sun lighting along the orbit, the temperature inside the shell will be limited in the range $5 - 40°$ C; on the other hand, most electronics components employed in PAMELA are commercial devices whose correct operation is guaranteed by the manufacturers within the temperature range from 0° to 70° C. Besides, the temperatures of various parts of the PAMELA apparatus are periodically monitored both by the PSCU and the satellite control system by means of two sets of temperature sensors.

2.1.2 Permanent failures

A general cause of permanent damage for an electronic device is the internal breaking of a component typically caused by repeated thermal and electric stresses of the microscopic structures; the probability of a failure increases with the overall operating time and also with the total number of power cycles performed, since when the device is powered on or off high transient fluctuations of both electric and thermal conditions can happen. This phenomenon is commonly known as *aging* of the component.

Besides, the random fluctuations during the fabrication processes of a given device, may lead to the presence of single specimens that are highly sensitive to microscopic stresses; the typical method to identify and remove these components is to expose them to repeated thermal cycles in the test phase, which will cause their breaking: the application of this method, known as *burn-in*, in the specific case of the control electronics of the tracking system, is described in sect. 3.8.

Another possible cause of failure, typical of a satellite mission, is the mechanical damage of a printed-circuit board (e. g. detachment of soldered components on the surface or breaking of a metal track in the inner layers of the board), which can be produced by the strong accelerations that the apparatus undergoes during the launch phase.

Similar problems concern also the detectors and all other parts of PAMELA. Mechanical qualification tests have been sustained both by single subsystems and by the whole apparatus; the last general qualification test took place on January 2005 at IABG facility in Munich [22].

2.2 Effects of cosmic rays on electronics

The electronics operating on board of satellites, outside the natural shield against primary cosmic rays provided by the terrestrial atmosphere, is exposed to a significant flux of ionizing particles and to the consequent risk of ionization-induced damages.

A detailed treatment of the complete phenomenology goes well beyond the scope of the present work; the effects of ionizing radiation on electronics, not only in space environment, constitute a vast field of study and the reader is referred to specialized publications for a comprehensive treatment (for an overview of the subject see for example [23]). Here a simple qualitative introduction of the main effects of interest for the PAMELA mission will be given, followed by an illustration of the various countermeasures adopted to limit the possible failures caused by ionizing radiation.

2.2.1 Ionizing radiation in the space environment

The ionizing components of cosmic rays of main interest for the PAMELA mission are:

- protons and nuclei of galactic origin;
- protons trapped in the *inner radiation belt*, which is crossed by the Resurs satellite during its orbits around the earth[4];
- protons and nuclei of lower energies with respect to the galactic component, generically defined as *solar energetic particles* (SEP), which are sporadically emitted by the sun during flares or coronal mass ejections.

The first two components are continuously present, while SEP emissions happen at a mean rate of 3 per year, with the actual rate increasing in proximity of the maximum solar activity, with a period of 11 years. A typical SEP event can last from several hours to some days, thus giving a mean overall duration of SEP emissions of the order of 10 days per year; nonetheless, given the much more higher intensity of SEP fluxes with respect to the galactic and trapped components, the effects of SEP are of the same order of those of the other two components.

Typically nuclei of higher atomic number Z are increasingly less abundant both in the SEP and galactic fluxes, but their effect cannot be neglected because, for a given energy of the incident particle and a specific irradiated material, the ionization[5] they produce is directly proportional to Z^2. For the same reason the direct ionization induced

[4]The inner radiation (or Van Allen) belt is mainly formed by protons that are produced by the interactions of high-energy cosmic rays with the atmosphere and then trapped by the terrestrial magnetic field; another more external belt (*outer radiation belt*) is similarly formed by electrons. The inner belt, in correspondence of the South Atlantic Anomaly, extends down to altitudes of about 500 km and will be crossed by the Resurs satellite whose orbit is between 350 and 600 km.

[5]The ionizing power of a particle species in a material is commonly represented by the *linear energy transfer* (LET), defined as the ratio between the energy released by ionization and the length of the path inside the traversed material. Usually this quantity is normalized to the density of the material and typically measured in $\mathrm{MeV} \cdot \mathrm{cm}^2/\mathrm{mg}$.

by the proton and electron components of cosmic rays represents a practically negligible contribution; on the other hand protons give a non-negligible contribution to the overall ionization, since they can interact with atoms of the irradiated material, with production of highly ionizing recoiling nuclei.

2.3 Radiation effects on MOS and CMOS devices

The radiation damage that can affect an electronic device is strictly related to the specific characteristics of the microscopic structures involved; for integrated circuits this means a dependence on the technology standard (MOS, CMOS, BJT etc.) and on the level of integration (geometric dimensions of the structure). Significant variations of the response to irradiation have been also observed between nominally identical specimens; these are related to the statistical fluctuations in the microscopic properties of each chip, even if obtained with the same fabrication process.

Since CMOS devices provide the best combined performance in terms of fast operation, small area occupation and low power consumption, they are generally the most widely employed components, both in digital and low-power analogical electronics; for example it can be estimated that the integrated circuits mounted on the DSP boards of the tracking system of PAMELA (described in chapter 3) are composed of some tens of millions CMOS transistors. As for the other standards, both BJT and MOS technologies are commonly employed in power circuits; MOS transistors are also used as switches, integrated in combination with CMOS circuits.

The most significant radiation effects in space involve the MOS and CMOS technologies. A first long-term effect of ionizing particles on MOS/CMOS electronics is the variation of the electrical characteristics related to the total absorbed dose. A second kind of phenomenology is correlated to single highly ionizing particles (*single-event effects*, SEE) that can instantaneously cause temporary or permanent failure of a device.

For a discussion of these effects and of the countermeasures adopted in PAMELA, it is necessary to briefly explain the characteristics of these technologies. In the following section the basic building block of the CMOS circuits, the CMOS inverter, will be described, with particular attention to its application in digital circuits to perform the NOT logical function, since the development of digital electronics is a major subject of the present work. Besides, in this example also the behaviour of a MOS transistor acting as switch will be implicitly described.

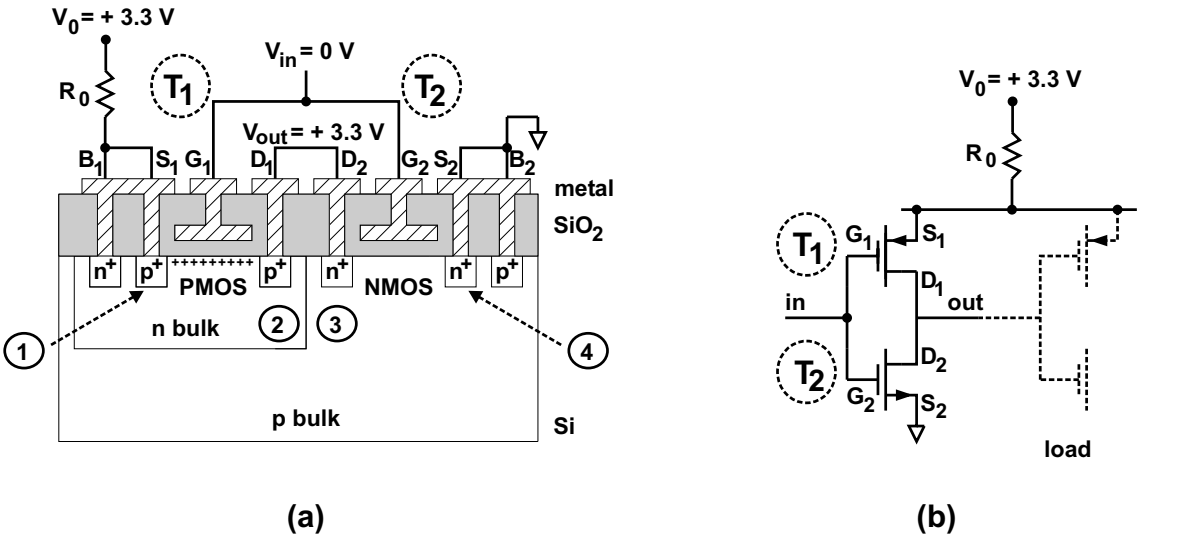

Figure 2.1: (a) scheme of a CMOS inverter. It is illustrated the case with $V_{in} = 0$ V and hence $V_{out} = +3.3$ V because of the formation of a conduction channel constituted of holes (indicated by the symbol +) in the PMOS T_1. The circled numbers 1-4 indicate the zones interested in the single-event latch-up (SEL, sect. 2.3.3). (b) circuital scheme of the CMOS inverter; the dashed lines represent the possible load which in particular can be formed by one or more CMOS structures.

2.3.1 CMOS inverter

The *complementary MOS* (CMOS) inverter is composed of two MOS transistors of opposite (complementary) doping. The *metal-oxide-semiconductor* (MOS) field-effect transistor is characterized by a thin layer of SiO_2 separating the metallic control electrode *gate* G from the doped Si bulk acting as conduction channel. Fig. 2.1 shows a CMOS inverter operated according to the LVTTL electrical standard, with power supply voltage of 3.3 V.

Consider the behaviour of the circuit in absence of load on the output terminal *out*. Assume that the voltage V_{in} of the input terminal *in* (corresponding to the *gate* terminals G_1 and G_2 of both transistors) is 0; then in the PMOS transistor T_1, whose *source* terminal S_1 is kept at 3.3 V, a thin layer of positive charge carriers (holes) is induced in the surface region of the Si bulk from the high-doping p^+ region of the source; T_1 is then said to be *on*. The channel formation happens if the voltage difference between *gate* and *source* exceeds a threshold value (typically ± 2 V, negative for the PMOS, positive for the NMOS). In the complementary NMOS transistor, where the *source* is kept at 0 V, no charge carriers are induced (T_2 is *off*). From this follows that the output terminal *out*, corresponding to the *drain* electrodes D_1 and D_2, is in ohmic contact with S_1 and electrically isolated from S_2, and necessarily is at 3.3 V; on the contrary, when $V_{in} = 3.3$ V, for an analogous mechanism $V_{out} = 0$ V.

These results remain valid also in presence of a load, represented by one or more other CMOS stages controlled through their input *gate* terminals: the static current absorbed by the loads is very small, because of the SiO_2 layer isolating the *gate* terminal from the rest of the structure, and this guarantees a sufficiently low voltage drop across the active conduction channel, thus keeping the value of V_{out} near 3.3 V or 0 V. It can be seen that for the same reason the static power consumption is very small, being proportional to the leakage current across the *gate* capacitors. In practical applications the main contribution to the power absorption of CMOS digital circuits is the dynamic consumption, during the switching, when the CMOS inverter crosses an intermediate regime of operation in which both transistor are conducting, and the load capacitors are charged or discharged depending on the transition. Clearly the power consumption in a given time interval is proportional to the number of switchings; in other terms, the power absorption of a CMOS digital circuit is directly proportional to the operating frequency.

2.3.2 Long-term effect: total ionizing dose (TID)

The *total ionizing dose* (TID) is defined as the accumulated energy released by the incident ionizing radiation per unit mass of the irradiated material[6]; an integrated circuit with MOS/CMOS transistors, illuminated by ionizing particles, shows a variation of its electrical characteristics that increases with the absorbed TID, finally causing the possible loss of functionality of the device.

The mechanism responsible for this alteration can be understood by considering that when an ionizing particle produces electron/hole pairs in the SiO_2 layer, a fraction of these charge carriers can reach the interface between the oxide and the Si bulk because of thermal diffusion or drift under the electric field applied across the dielectric; in particular holes can be trapped at the boundary between oxide and bulk for times even as long as years [24].

With subsequent ionizing events, the amount of holes accumulated at this boundary increases; this excess of positive charge induces a higher concentration of negative charge carriers (electrons) in the adjacent channel region. For a PMOS (T_1 in fig. 2.1) these electrons must be completely removed to achieve the formation of the conduction channel, thus increasing the absolute value of the negative voltage threshold; on the contrary for an NMOS (T_2) the positive threshold is reduced.

The variation of these voltage thresholds changes the dynamic properties of the tran-

[6]The TID is measured in gray (1 Gy = 1 J/Kg) or more commonly in krad for space applications (1 krad = 10 Gy).

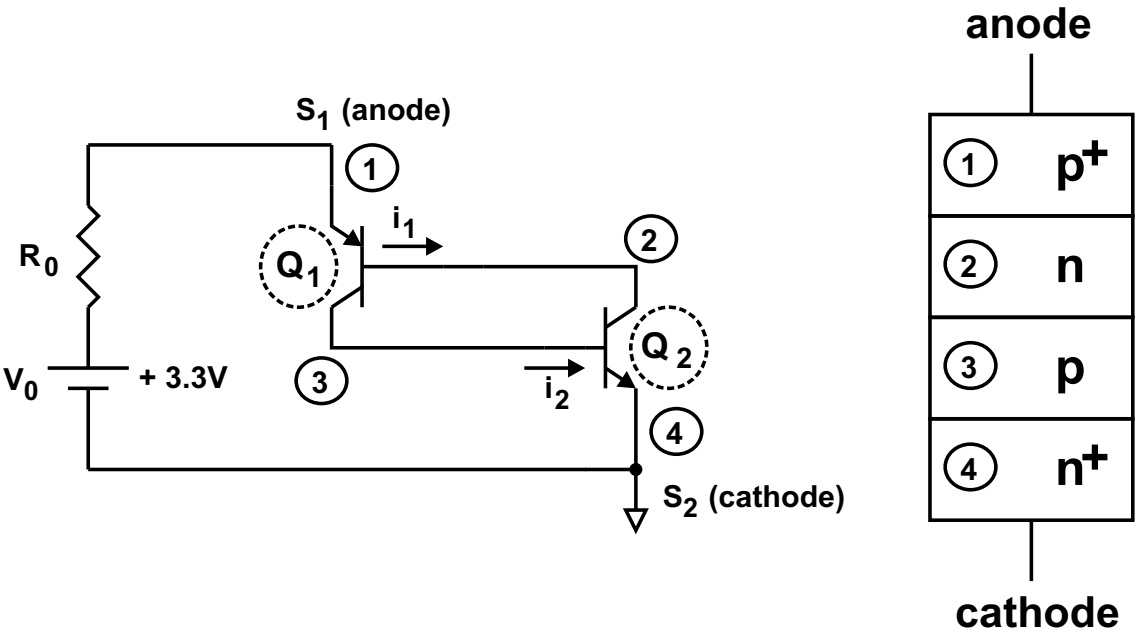

Figure 2.2: parasitic SCR circuit associated with the CMOS inverter architecture.

sistors (propagation time of a state transition, power consumption during the transition) and finally also the static ones: for example, in an NMOS, with increasing accumulated holes at the oxide-bulk boundary, the number of induced negative charges can become sufficiently large to form a permanent conduction channel, with the complete loss of functionality of the transistor which remains *on* for all the values of the *gate* voltage.

2.3.3 Single-event latch-up (SEL)

This potentially destructive effect is characteristic of the CMOS technology and is related to the presence of a parasitic circuit composed of the 4 Si layers with alternated doping between the electrodes S_1 and S_2 of a CMOS inverter (zones indicated by the circled numbers 1-4 in fig. 2.1), forming a direct conduction path between power supply (applied to S_1) and ground (S_2) that is normally inhibited but can be activated by anomalous conditions.

The parasitic circuit, schematically illustrated in fig. 2.2 and known as *silicon controlled rectifier* (SCR [25]), contains two coupled bipolar-junction transistors (BJT, indicated as Q_1 and Q_2 in figure) between its anode (represented by S_1) and cathode (S_2). Note that the p-n interface between the bulks of the two MOS transistors (zones 3 and 2 respectively) acts both as collector-base junction for Q_1 and as base-collector junction for Q_2.

During standard operation the 3 junctions 1-2, 2-3, and 3-4 are reverse-biased, because

both zones 1 and 2 are kept at 3.3 V while zones 3 and 4 are grounded, through the metallic electrodes S_1, B_1, B_2 and S_2 respectively.

Suppose now that, because of any perturbation, the base-emitter junction of one of the two parasitic BJT transistors (e.g. Q_2) is put into direct conduction with an initial current $i_2 = i_{in}$ entering its base; this, because of the characteristics of BJT transistors, would generate a much greater collector current $i_1 = \beta_2 i_{in}$ (with $\beta_2 \gg 1$); i_1, acting as base current for the other BJT Q_1, would similarly cause the activation of Q_1 and a further increase of i_2 which would pass from the initial value i_{in} to $i_{in} + \beta_1\beta_2 i_{in}$ (with $\beta_1\beta_2 \gg 1$), thus tending to raise indefinitely.

The voltage drop across the internal resistance R_0 of the power supply increases with the absorbed current thus reducing the voltage difference between anode and cathode of the SCR circuit (zones 1 and 4 respectively). On the other hand, the characteristic voltage between base and emitter of the active BJT Q_1 forces zone 2 at about 0.7 V below the anode potential; similarly for Q_2 zone 3 is forced at about 0.7 V above the cathode potential. In conclusion, the voltage difference between zones 3 and 2 increases until also this junction is put into direct conduction, thus saturating both transistors Q_1 and Q_2.

From this point on, the SCR circuit is *latched-up* with a stable voltage (≈ 1 V) between anode and cathode and a large current i_{latch}, flowing from the power supply, given by the relation $i_{latch} \approx (V_0 - 1\text{ V})/R_0$ and typically being of the order of 1 A.

The SCR self-increasing mechanism can be started by a highly ionizing particle crossing the CMOS inverter, as in the example shown in fig. 2.3, where the inverter is being normally operated with the input G_1/G_2 kept at 0 V and consequently with the output D_1/D_2 forced to 3.3 V in absence of external perturbations.

In this configuration the n^+-p junction between the D_2 terminal and the p bulk is reverse-biased and the 3.3 V drop is applied across the depletion zone developing almost completely into the low-doping bulk. When an ionizing particle traverses this depletion zone, the holes produced by ionization in this region, before being able to completely recombine, are transported by the electric field toward the p bulk (zone 3 of the SCR parasitic circuit) initially at 0 V, where they cause a temporary local increase of voltage. If the amount of ionization is sufficiently high, the p-n^+ junction between zones 3 and 4 of the SCR circuit can enter direct conduction with the injection of e^- from the n^+ zone toward the p bulk, thus activating the SCR mechanism.

The section of electronics powered by the same supply as the latched-up component will not work properly, because operated at about 1 V instead of the required 3.3 V; most

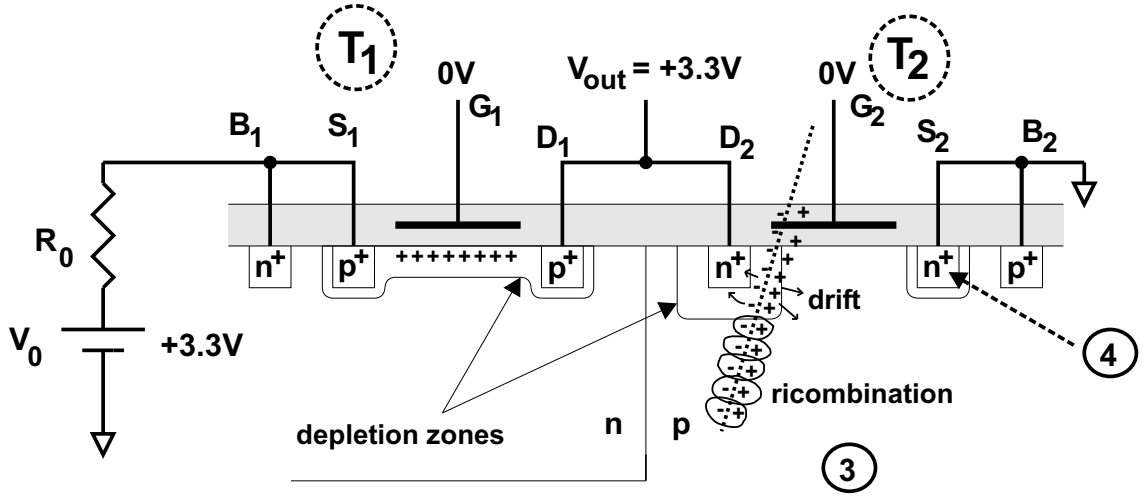

Figure 2.3: effects of an ionizing particle crossing a CMOS inverter, when the *out* electrode is initially kept at 3.3 V. The zones 3 and 4, interested by the activation of the SEL mechanism, are also indicated.

importantly, if the latch-up condition is not readily removed by turning off the power supply, the heating of the metallic terminals due to the large current can cause them to melt with a permanent damage to the circuit.

2.3.4 Single-event upset (SEU)

Another important effect of the passage of an ionizing particle through a CMOS device is the *single-event upset*, observable in digital memory elements, such as for example the static RAM (*random access memory*) cell illustrated in fig. 2.4, composed of two coupled CMOS inverters. The operation of this memory cell is based on the fact that the loop composed by the two CMOS inverters has only two stable configurations, corresponding to voltages of 0 V or 3.3 V on the *data_out* line and representing the digital values 0 and 1 respectively. The state of the cell can be forced (*written*) through the control line *data_in* with the input MOS switch kept closed; when the MOS switch is open, the value 0 or 1 remains stored in the cell until a new write operation is performed.

Consider now an ionizing event such as the one described in the previous section and illustrated in fig. 2.3, with the passage of a charged particle producing ionization pairs inside a reverse-biased depletion region of a CMOS inverter, specifically the region between the D_2 terminal and the p bulk of the MOS transistor T_2. While holes drift toward the p bulk where they can be at the origin of an SCR latch-up, electrons move toward the n^+ zone where their accumulation causes a temporary decrease of the voltage on the output terminal D_2; a transient current flows from the power supply through

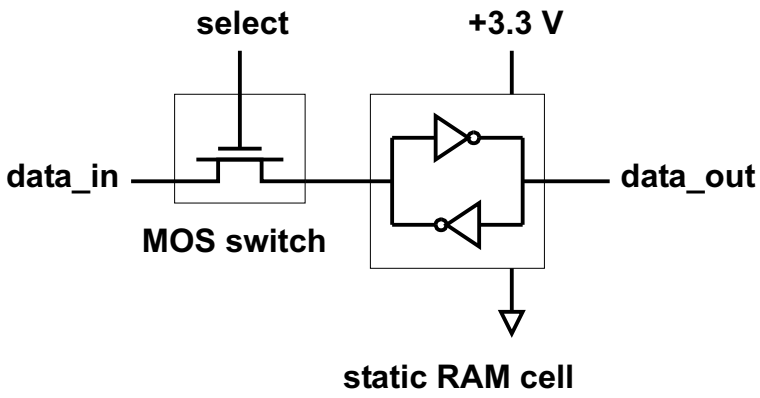

Figure 2.4: a static RAM cell, composed of a pair of coupled CMOS inverters.

the active conduction channel of T_1 (constituting the only possible current path in this configuration) to restore the local charge balance.

As a consequence of the ionizing event a temporary negative voltage pulse is observed at the output of the CMOS inverter; in particular this voltage pulse can be sufficiently high to cause the other inverter of the memory cell to switch and consequently the whole RAM cell to invert its logical state and remain latched in the new configuration (*upset*).

The introduction of the wrong logical value in a memory element generally causes a failure in the operation of a digital circuit. This kind of damage is not permanent, and correct operation can be easily restored through a power cycle or reset of the circuit.

2.4 Limiting the effects of radiation in PAMELA

Following widely adopted methods for qualification of electronic devices against radiation damage, all the components employed in PAMELA have been selected after specific tests of their susceptibility to radiation damage [26].

After having identified a set of candidates for all the various functional parts of the system (e.g. ADC, DSP, programmable logics, voltage regulators), the ones showing a sufficiently low sensitivity to ionizing radiation have been chosen for use. With this approach it has also been possible to qualify and employ commercial components not specifically designed for space applications, with a significant reduction of costs and improvement of performances.

To take into account the cumulative effect related to TID, the devices under test were irradiated with γ rays of known energy, from which it is possible to calculate the absorbed TID per unit time. Given that the expected TID after 3 years of operation of

PAMELA amounts to some krad, as estimated by means of the CREME96 simulation software [27] [15], the components were selected only if after the absorption of 30 krad their functional characteristics did not show significant variations.

On the other hand single-event effects (SEE) such as the latch-up (SEL) or upset (SEU) previously described, being random phenomena, cannot be completely avoided; nonetheless, their mean expected rate during the PAMELA mission has been reduced to an acceptable value, with the selection of devices characterized by sufficiently low cross sections for SEE; a specific application of this method will be showed in sect. 2.6.

Moreover the possible consequences of an SEE have been reduced with suitable solutions in the general design of the electronic apparatus and in particular of the data acquisition system of PAMELA (see sect. 1.3 and fig. 1.5 for reference).

Each voltage regulator stage is equipped with a current limiting circuit, setting the maximum current that can be supplied to the load, thus avoiding that also in presence of a latch-up condition a too high current could flow. The most critical systems, like the IDAQ and IHK boards, are equipped with dedicated *latch-up alarm* lines whose state changes in presence of anomalous voltage or current conditions on the board; these lines will be monitored by the PSCU during the operation of PAMELA.

For other electronic parts, like the front-end stages of the tracking system, we have foreseen a continuous consistency check of the acquired data, performed during the data processing phase by the corresponding DSP board on an event-per-event basis, to detect failures caused by SEL or SEU.

Single-event effects on one of the DSP boards, causing its passage to an anomalous logical state with full or partial loss of functionality, will be effectively detected by the IDAQ board during the data exchange procedures, since the implemented general communication protocol requires that when the IDAQ board sends a command, the addressed DSP board should reply with an answer data packet within a predefined time interval.

On the other hand the addressed DSP board will perform a consistency check of the received command, by means of control codes inserted in the data stream. The control codes, calculated according to a *cyclic redundancy check* (CRC) algorithm[7], are stored in the non-volatile memory of the PSCU together with the corresponding commands. When a command is sent by the IDAQ board to a DSP board, this can internally calculate the CRC code from the actual data and compare it with the code received. A more detailed

[7] A CRC algorithm (see for example [28]) adds to a given set of binary words a redundant check code obtained with the sequential cyclic application of a fixed logical operation to each word of the packet. CRC algorithms are international IEEE standards [29] and commonly adopted in data communication systems.

explanation of the data exchange protocols between the IDAQ board and the DSP boards of the tracking system will be given in sect. 3.5.

After the assertion of the alarm line by the system which has detected an anomalous condition, the PSCU will start a dedicated recovery procedure; in particular a possible latch-up state will be removed after at most few seconds from its occurrence by performing a complete power cycle of the apparatus, since the only way to avoid a permanent damage caused by SEL is the quick removal of the parasitic currents induced in the latched-up component, by means of the complete turning off of the power supply.

The integrity of the data on-line processed in the DSP boards, subsequently transferred to the mass memory of the PSCU and finally transmitted to ground, is guaranteed by multiple check methods: the DSP typically inserts a series of consistency indexes in the processed data packet, summarizing its main physical content; on the other hand, the control logics of the DSP board adds a CRC code during the transmission.

2.5 Segmentation and redundancy

To limit the consequences of permanent failures of electronic devices due to thermal, electrical, mechanical stresses, aging or radiation damage, the general design of the PAMELA apparatus is characterized by the adoption of two general methods to enhance reliability: *segmentation* and *redundancy*.

The segmentation is obtained by splitting a given electronic system into several sections, each working independently and with dedicated voltage regulators and current limiting stages, in such a way that the permanent loss of a section does not affect the correct operation of the others. Examples of segmentation can be found in the description of the front-end electronics of the tracking or TOF system (chapter 1).

The redundancy of a part of the electronics means its physical duplication so that an identical spare (or *cold*) copy is ready for use in the case of failure of the main or *hot* one.

The drawbacks of these two methods are the greater complexity of the mechanic support structures, electrical interfaces and control procedures; moreover since the resulting overall system has larger dimensions, because of the multiplication of the number of employed components and connections, the applicability of these methods is mainly limited by the available volume, which is particularly small for a satellite mission like PAMELA.

A practical application of the redundancy principle and its complications in PAMELA is illustrated in fig. 2.5, showing the configuration adopted for one of the two DSP boards of the tracking system (subsystem A), the trigger board (subsystem B), their interface and

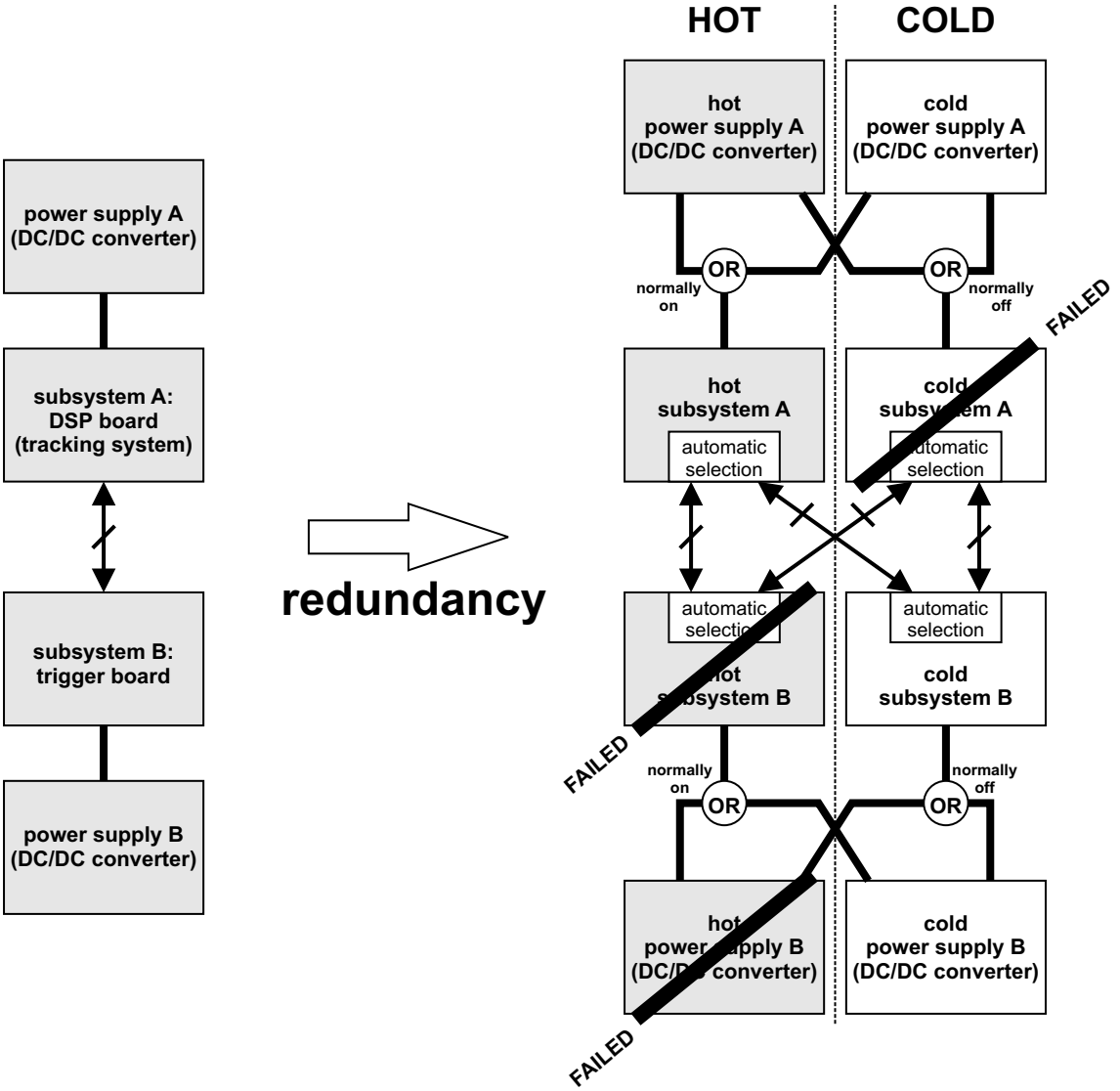

Figure 2.5: example of implementation of redundancy in PAMELA. In the redundant configuration it is illustrated a case of triple failure that does not affect the functionality of the system.

the corresponding power supplies (DC/DC converters). Without redundancy (left side of the figure) the global structure would be subject to complete loss of functionality with the failure of any of the four elements (boards or power supplies); on the other hand, with the redundant architecture shown on the right, in general it is possible to replace a damaged element with its spare copy without having to change any of the other active subsystems, and therefore only specific combinations of double or higher-multiplicity failures involving both the *hot* and *cold* sides would cause a complete loss.

In the redundant configuration, both the control electronics and their power supplies are doubled and all the possible combinations of active devices are made available by increasing the number of power lines and links for data exchange. In particular the on-

board control logics automatically recognizes the active data link, thus avoiding the need for a configuration procedure managed by the PSCU when the pattern of active devices is changed. Note also the insertion of a further stage that performs the analogical OR-ing of the *hot* and *cold* power lines for each subsystem.

2.6 Control logics in PAMELA: device selection

The digital structures, known as *control logics*, that perform the fundamental control tasks of the various electronics boards of PAMELA, are specifically required to assure the maximum reliability and stability of operation during the whole duration of the experiment.

Essentially two device standards have been available in the last years for the realization of high-quality digital components: ASIC and FPGA integrated circuits.

In an *application-specific integrated circuit* (ASIC) the functionalities are implemented and fixed during the fabrication process; the device can thus be optimized according to the requirements of the specific application (power consumption, area occupation, operating speed, radiation hardness), but costs and development times are rather high, mainly because changes in the logical design imply a different microscopic layout of the chip.

On the other hand, *field programmable gate arrays* (FPGA) contain a general-purpose logical structure that can be directly configured by the user to achieve the desired functionalities; thus a unique FPGA model can be employed in many different applications, with a drastic reduction of costs and development times. For these reasons and since the combined performances of these devices have shown to be quite satisfactory, FPGA chips have been chosen for the implementation of the control logics in PAMELA.

Among the available commercial FPGA technologies, the significantly different susceptibilities to SEU and SEL have played a major role in the selection of the device family.

The sensitivity of a digital component to single-event effects is usually described (see for example [30]) in terms of the measured cross section per memory element (or *flip-flop* or bit) as a function of the linear energy transfer (LET) of the ionizing radiation, since it has been observed that different logical structures containing the same number of flip-flops practically show the same response to single-event effects.

During a test session at GSI [31] (Darmstadt, Germany 2000) samples of the two best candidate FPGA families, Actel A54SX/SXA [32] and Quicklogic pASIC1 [33], with implemented the same logical structure, have been operated while irradiated with heavy

ions of known LET[8] [15] [26]. By monitoring the observed number of SEU and SEL per unit time and from the knowledge of the intensity of the incident beam, the cross sections per bit for SEU and SEL were measured as a function of the LET.

For the Quicklogic device a ratio SEL/SEU of the order of 1% was observed, while for the Actel FPGA no SEL were reported, with a measured cross section for SEU about one order of magnitude smaller than the one measured for the Quicklogic device. Given these results, the Actel A54SX/SXA family has been selected for use in PAMELA.

With the simulation software CREME96 [27] it is possible to predict the intensities of cosmic rays as a function of LET for the orbit of PAMELA and, from the measured cross sections per bit, calculate the expected number of SEU per bit for the Actel FPGA during the 3 years of operation. The steady contribution of the galactic and trapped cosmic rays is $1.7 \cdot 10^{-6}$ SEU/bit, while the order of magnitude of the contribution of the solar energetic particles can be estimated by assuming a total of 10 days of solar emissions characterized by the mean intensities of SEP registered in the 1989 during the week of maximum activity, thus obtaining $5.2 \cdot 10^{-6}$ SEU/bit.

Considering for example the total number of flip-flops (6054) implemented in all the FPGA constituting the control logics of the tracking system, the expected number of SEU during the 3-year mission is $1.0 \cdot 10^{-2}$ (caused by the steady components) plus $3.1 \cdot 10^{-2}$ (due to SEP and concentrated in few days of solar activity). Even considering a correction factor 10 as safety margin, it can be concluded that the control logics of the tracking system is practically free from SEU (and SEL); analogous considerations hold for the FPGA chips employed in the other subsystems of PAMELA.

2.7 Characteristics of Actel A54SX/SXA family

A typical FPGA integrated circuit of the Actel A54SX/SXA family contains, on an area of few cm^2, an array of thousands or more fundamental CMOS logical cells, whose input and output nodes are interconnected through a microscopic grid of metallic tracks, implementing a high number of potential links, normally interrupted by high-resistance *antifuses*.[9]

[8] During this test, beams of ^{131}Xe and ^{238}U with kinetic energies in the range from 100 to 800 MeV/nucleon were used. The corresponding LET has been calculated with algorithms based on the Bethe-Bloch formula.

[9] The Actel antifuse FPGA and other devices using similar technologies are a better choice for space applications than common *reprogrammable* FPGA chips, for which the configuration of the interconnections of the array of logical cells is achieved by means of MOS switches, whose states are stored in static RAM memory elements (see fig. 2.4). In fact the presence of these memory elements greatly increases the risk of radiation damage of the device, as discussed in sect. 2.3. On the other hand, a reprogrammable FPGA is ideal for the prototyping phase, since the logical structure is not permanent and can be changed

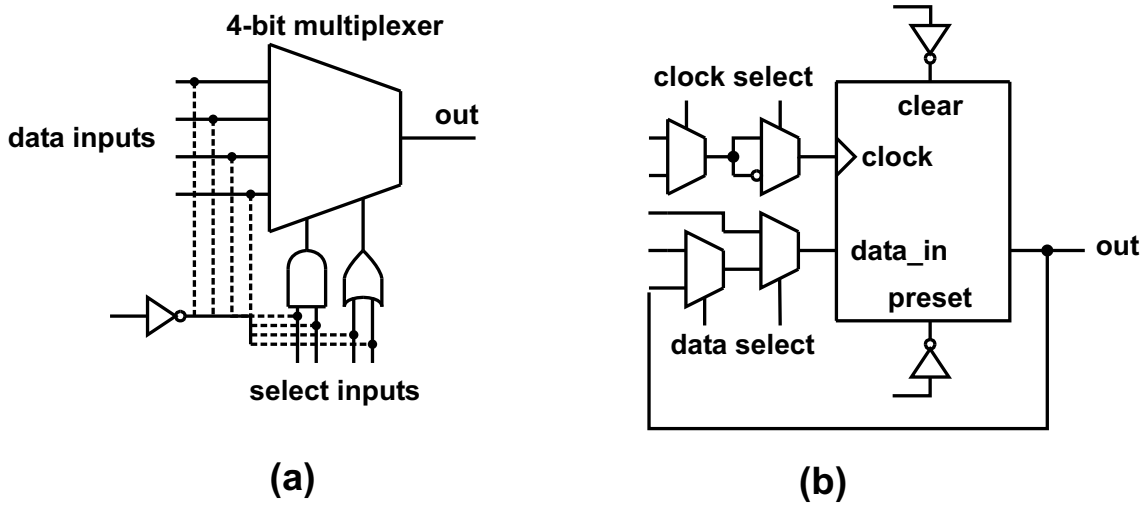

Figure 2.6: fundamental cells for A54SX/SXA family: (a) combinatorial cell; (b) sequential cell. Continuous lines indicate permanent connections, dashed ones represent programmable links. Note that all the external connections of a cell are programmable.

The metallic net is connected to the external pins of the device through special modules that can be configured as input, output or bidirectional elements.

The fundamental cells are of two kinds (see fig. 2.6), representing the two basic logical elements of digital circuits:

- *combinatorial* cells, whose output is a fixed logical function (combination of AND, OR, NOT etc.) of its inputs, without memory elements: the output generally changes after a transition of any of the inputs. The usual implementation of this cell is a multiplexer;

- *sequential* or memory cells: basically a 1-bit register (*flip-flop*); the output changes according to the *data_in* input value only in correspondence of a rising edge of the *clock* line (it is usually said that the input is *sampled* or *loaded*). Also independent *clear* and *preset* controls are available, to force the output to 0 or 1 respectively. In particular the source and polarity of *clock* can be independently selected for each cell when configuring the FPGA.

A typical example of Actel antifuse is shown in fig. 2.7: two metallic tracks are isolated by a thin layer of ONO dielectric material, constituted by three sub-layers of SiO_2, Si_3N_4 and SiO_2 for a total width of less than 10 nm. The antifuse can be "burnt" with the application to the metallic electrodes of a voltage difference (a few V) that generates an electrostatic discharge across the dielectric material; the flowing current (of the order of

without having to replace the FPGA chip, unlike in the antifuse case.

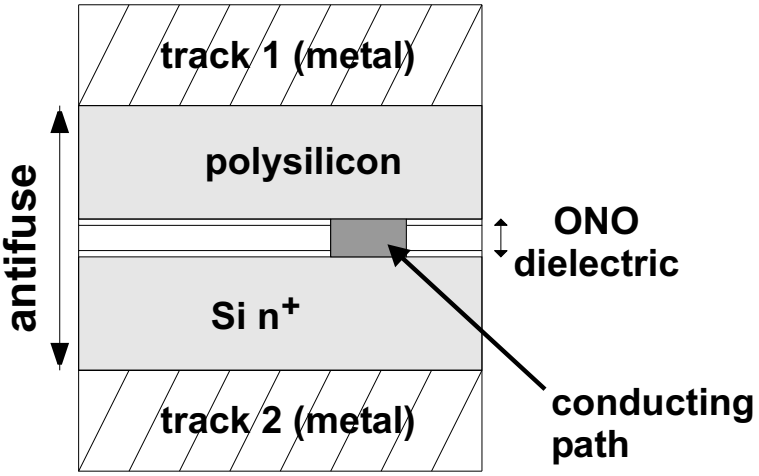

Figure 2.7: Structure of a typical Actel antifuse.

5 mA) leads to a local melting with the formation of a permanent Si conduction path through the ONO layer, with typical diameter 20 nm and resistance of the order of 10 Ω.

The configuration of the FPGA, for the implementation of the desired logical structure, is obtained by "burning" a suitable number of antifuses, to achieve permanent low-resistance paths between a set of input/output internal nodes and of external pins. This operation is done automatically by a specifically designed *programmer* device, with a socket to house the FPGA during the configuration; the current pulses to produce the desired pattern of burnt antifuses enter through the external pins of the FPGA.

2.8 Development of control logics on FPGA

This section contains general considerations on the design methods for the development of control logics on FPGA chips, with particular attention to the choices that allow to achieve a higher reliability.

2.8.1 Design and optimization

The design of a generic digital circuit (FPGA or ASIC) can be in principle done at an elementary level by defining a net of suitably interconnected fundamental cells, but for systems of medium or high complexity (thousands of cells[10]) it is better to use a high-level *hardware description language* (HDL), whose main advantage is the possibility of describing the desired functionalities with abstract logical constructions without necessarily having to specify the underlying array of cells.

[10]This is the case of the control logics of the tracking system, described in chapter 3.

This synthetic design method allows to reduce the development time and also makes the code much more easily readable and accessible for further developments. The translation of the high-level HDL architecture into a net of elementary cells is automatically done by a compiler (or *synthesizer*) software, which also checks that this low-level structure complies with the general consistency requirements of digital circuits.

An important task performed by the synthesizer is the minimization of the number of cells needed for the implementation of a given design. As an example consider two distinct nodes that have been defined in the HDL code to have two complicated and apparently different functional dependences f_1 and f_2 on the same set of inputs: if f_1 and f_2 are found to be logically equivalent, the two formally distinct nodes in the code can be replaced by a unique physical signal in the device.

After the synthesis step, a *place-and-route* software transforms the abstract logical net into a physical layout, i. e. an image of the structure to be implemented on a specific integrated circuit; in the case of the FPGA, a *configuration file* is produced which can then be used by the programmer device.

Writing HDL code is made more efficient by the availability of libraries containing descriptions of the most commonly used logical blocks (such as counters, shift-registers, multiplexers, decoders etc.); various manuals [34] [35] [36] [37] [38] provide examples of general-purpose code as starting point for original development in view of a particular application.

At present two hardware description languages are commonly employed: Verilog HDL and VHDL[11], basically equivalent for what concerns the main descriptive capabilities. Since the general trend in digital design seems to privilege VHDL, I have chosen this language for the development of the control logics described in chapter 3.

The correct functionality of a digital circuit obviously depends not only on the logical architecture itself, but also on whether the physical layout satisfies the timing requirements for each single flip-flop. The data input of a flip-flop (see fig. 2.8) must remain stable over a fixed voltage threshold to be correctly sampled at the clock rising edge; in the LVTTL standard the thresholds are $V_{high} = 2.0$ V for the high (1) state, $V_{low} = 0.8$ V for the low (0) state. This stability must last for a time interval which extends both before (*setup time* T_{setup}) and after (*hold time* T_{hold}) the clock edge. If the setup or hold requirements are not satisfied, either the wrong state is loaded (0 instead of 1 or

[11] VHDL stands for *very-high-speed-integrated-circuits hardware description language*, introduced since 1983 by the US Department of Defense. Both Verilog HDL and VHDL are international standards of IEEE [29].

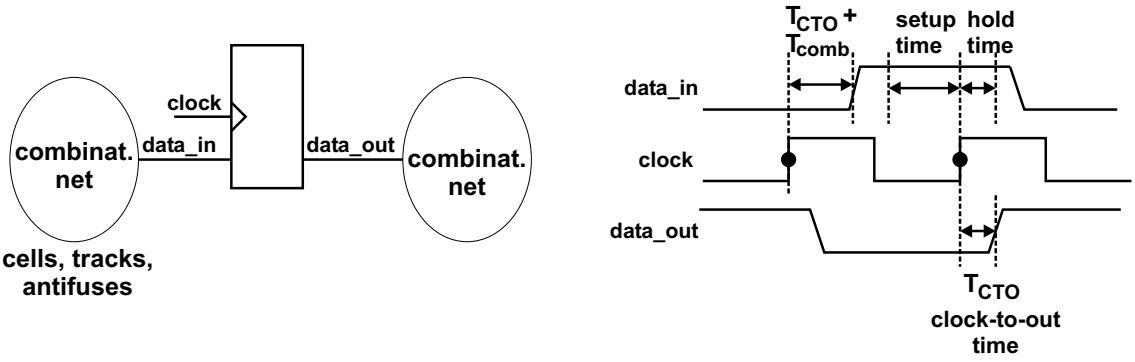

Figure 2.8: Timing characteristics for a flip-flop.

vice versa) or an undefined value is sampled (0.8 V $<$ V_{data_in} $<$ 2 V), this last condition causing the internal latch of the flip-flop to enter a *metastable state*, from which it decays into one of the two possible stable states after a characteristic time (of the order of 1 ns for the A54SX/SXA family)[12]. The induced wrong output level (*upset*) may propagate through the logical structure on successive clock rising edges causing further upsets of other registers and finally the loss of functionality of the whole design.

The setup and hold requirements for the flip-flops in an FPGA chip depend on many different parameters: device technology (with statistic fluctuations between different specimens), aging, absorbed TID, working conditions (temperature and supplied voltage). On the other hand the timing characteristics of a digital line, going through a combinatorial net from the output of a flip-flop to the input *data_in* of another one, vary with the electric properties (resistance and capacitance) of the specific implementation, determined by number and configuration of the combinatorial cells, length of the paths along the metallic tracks, number of traversed antifuses. The delay between the rising edge of the clock on the first flip-flop and the transition of the *data_in* input of the second flip-flop is given by the sum of the internal propagation time of the flip-flop (T_{CTO}, *clock-to-out time*) and of the delay introduced by the combinatorial net (T_{comb}).

The place-and-route software, for a given logical design and device type, automatically arranges the physical layout in order to maximize the time margins between *data_in* and *clock* transitions for all the flip-flops; besides the designer can further modify this layout with specific optimizations that the algorithms of the place-and-route software are not able to produce. Anyway the need to check the absence of timing violations for each single flip-flop in the circuit would nullify the advantage of having previously defined a synthetic high-level architecture.

[12]Specific application notes on the metastability problem for Actel devices are available at [32].

It is much more convenient to adopt a coding style at the HDL level that minimizes the dependence of the design functionalities on the timing characteristics of the signals. A common and relatively simple method that optimizes this aspect of the design is to use (when possible) a unique signal, generated by a high-stability oscillator circuit, as clock input for all the flip-flops in the device, thus obtaining a *synchronous* architecture. With this solution the setup-time requirement is easily met for any flip-flop in the device, since the clock period T_{clock} for usual applications is sufficiently great to satisfy the relation

$$T_{CTO}^{max} + T_{comb}^{max} + T_{skew}^{max} + T_{setup} < T_{clock} \tag{2.1}$$

meaning that *data_in* becomes stable sufficiently early with respect to its sampling, even taking into account the *skewness* between the actual clock rising edges on any pair of flip-flops, represented by the positive term $+T_{skew}^{max}$. For example the control logics of the tracking system works with clocks of periods greater than 50 ns, against typical values of the left-member sum of a few ns for an Actel FPGA.

A similar condition for the hold time is guaranteed to be satisfied by the device manufacturer of the FPGA for all the possible physical layouts:

$$T_{CTO}^{min} + T_{comb}^{min} - T_{skew}^{max} > T_{hold} \tag{2.2}$$

where the negative term $-T_{skew}^{max}$ represents the worst condition set by the skewness between the rising edges of the clock signals on two generic flip-flops. The clock skewness in the FPGA is minimized with the presence of dedicated fast distribution nets characterized by small propagation times of the clock signal across the chip.

2.8.2 Simulation and test

The functionalities of the layout, obtained after the place-and-route step are tested through a simulation software tool before the actual implementation on the FPGA. The user specifies the working conditions and device characteristics, builds timing sequences of input signals (usually known as *stimuli*), as far as possible reproducing all the different situations in which the chip will operate, and checks the behaviour of the circuit, reproduced by the simulator. During this phase it is possible to analyze the signal waveforms both at external pins and at internal nodes as well, which in the real device generally are not physically accessible.

The simulation aims to identify the two main sources of logical errors in the design: functional errors introduced by the designer and critical signal paths with violations of the setup or hold timing requirements and possibility of upset; in particular the simulation

software can provide a list of all the flip-flop inputs for which the automatic layout algorithm is inefficient and that present timing violations. For what said above, this second problem does not affect synchronous architectures, but can constitute the main source of failures when one or more asynchronous interfaces are required by the application; asynchronous architectures will be discussed in sect. 3.2 and 3.4.1.

After a number of cycles between simulation, changes and corrections to the design, synthesis and layout, when the designer is confident that the functionalities are correct, a real chip is configured, mounted on a board and the subsequent laboratory tests are started. Details on the tests performed for the qualification of the control logics for the tracking system will be given in sect. 3.8.

Chapter 3

Control electronics of the tracking system

The control electronics of the tracking system was developed by the Florence group of the PAMELA collaboration during the years 2003 and 2004.

I took part in the design, optimization and test of the various parts of this electronics; in particular I have been in charge of the development of the whole control logics on programmable FPGA devices. Most of this work has been done at the Florence INFN Section and Physics Department of Florence University; finally the various systems have been integrated with the rest of the PAMELA apparatus at the "Roma 2" INFN Section and Physics Department of Rome "Tor Vergata" University.

The control electronics of the tracking system (see fig. 3.1 and also refer to sect. 1.3) is composed of:

- a pair of *DSP boards* for the management of the data acquisition operated by the front-end electronics, the processing of the acquired data and their transfer to the PAMELA control system (PSCU) through the *intermediate data acquisition* (IDAQ) board. Each of the two DSP boards works independently from the other and is dedicated to half tracking system (board 1 to the 6 X views of the Si detector planes, board 2 to the Y views);

- a pair of *relays boards*[1] to control the power lines for the FE electronics and the bias for the 6 detector planes (board 1 controls the first 3 planes, T1-T3, board 2 the others); the two boards exchange data with the PSCU through the *intermediate housekeeping* (IHK) board.[2]

[1] The name comes from the presence of relay switches, which will be justified in the discussion of the board operation.

[2] The DSP and relays boards have been manufactured by CAEN Aerospace [39].

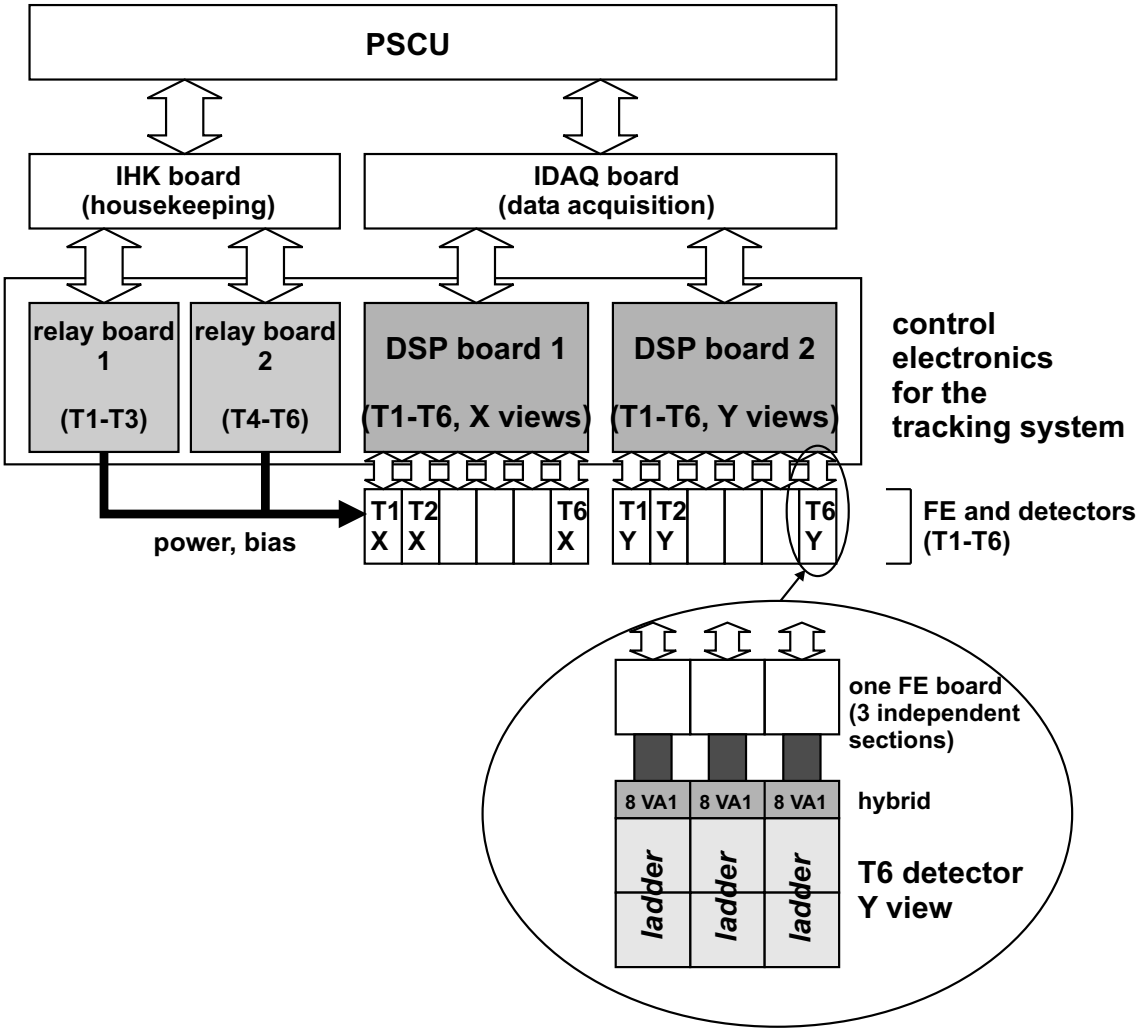

Figure 3.1: general structure of the control electronics of the tracking system.

Most of this chapter will be dedicated to the description of the structure and operation of the DSP boards, whose implementation has constituted a major effort, in view of the creation of a reliable, fast and flexible system. I will specifically concentrate on the aspects of this work that have involved my original contribution, in particular the development of the external digital links and data exchange protocols and the design of the internal architectures of the control logics on FPGA devices.

Sect. 3.1 contains a general description of the DSP board. The design techniques used for the development of the control logics are discussed in sect. 3.2.

Sect. 3.3 and 3.4 are dedicated to the problem of the implementation of a fast and reliable system for data exchange with the external devices (IDAQ board on one side, front-end boards on the other) and to the adopted solutions.

Sect. 3.5, 3.6 and 3.7 contain the description of the different functional parts of the

control logics of the DSP board and of their operation. Sect. 3.8 describes the various phases of development and the test methods adopted for the final qualification of the boards.

The last section (3.9) will be dedicated to the illustration of the main characteristics of the relays boards.

3.1 DSP board: general structure

The two DSP boards of the tracking system, dedicated respectively to the X and Y sides of the silicon detectors, are actually doubled for redundancy, thus achieving two *hot/cold* pairs of boards.

Each *hot/cold* pair forms a compact mechanical structure ("sandwich"), with the *cold* board superposed on the *hot* one; the "sandwich" is fixed to a common support frame in aluminium, inserted within a slot of the electronics crate of PAMELA; the gap between the *hot* and *cold* board is 5 mm.

Fig.3.2 and 3.3 respectively show a picture and the schematic structure of a board.

As mentioned in sect. 1.3, these boards are indicated as *DSP boards* for the presence of *digital signal processors* (DSP), dedicated to the elaboration of the acquired data. On the other hand, the control logics on FPGA chips[3] provides the general management and external interface of the board.

The main task of the DSP is a fast and efficient compression of the data produced by the silicon detectors at each event. The 36864 12-bit read-out channels of the tracking system (generating about 90% of all the physics data in PAMELA), will yield 54 kbyte/trigger[4] if no compression is applied; hence with the expected average trigger rate of 12 Hz, the whole tracking system will produce $\approx$ 50 Gbyte/day. On the other hand, the available bandwidth for data exchange with the ground control station, through the telecommunication system of the Resurs satellite, is limited to 10 Gbyte/day; therefore it is necessary to perform a compression of the read-out data before their transmission.

The program to be executed by the DSP contains a specific compression algorithm that basically eliminates from the originally acquired data the 12-bit read-outs associated with all the strips whose signals do not significantly differ from the value expected considering the combined contribution of the quiescent level (*pedestal*) and of the average noise for

[3]The FPGA models employed in the DSP boards, according to the selection criteria illustrated in sect. 2.6, are A54SX32A and A54SX16P manufactured by Actel [32].

[4]The present work follows the general convention according to which 1 kbyte = 1024 byte, 1 Mbyte = 1024 kbyte and so on.

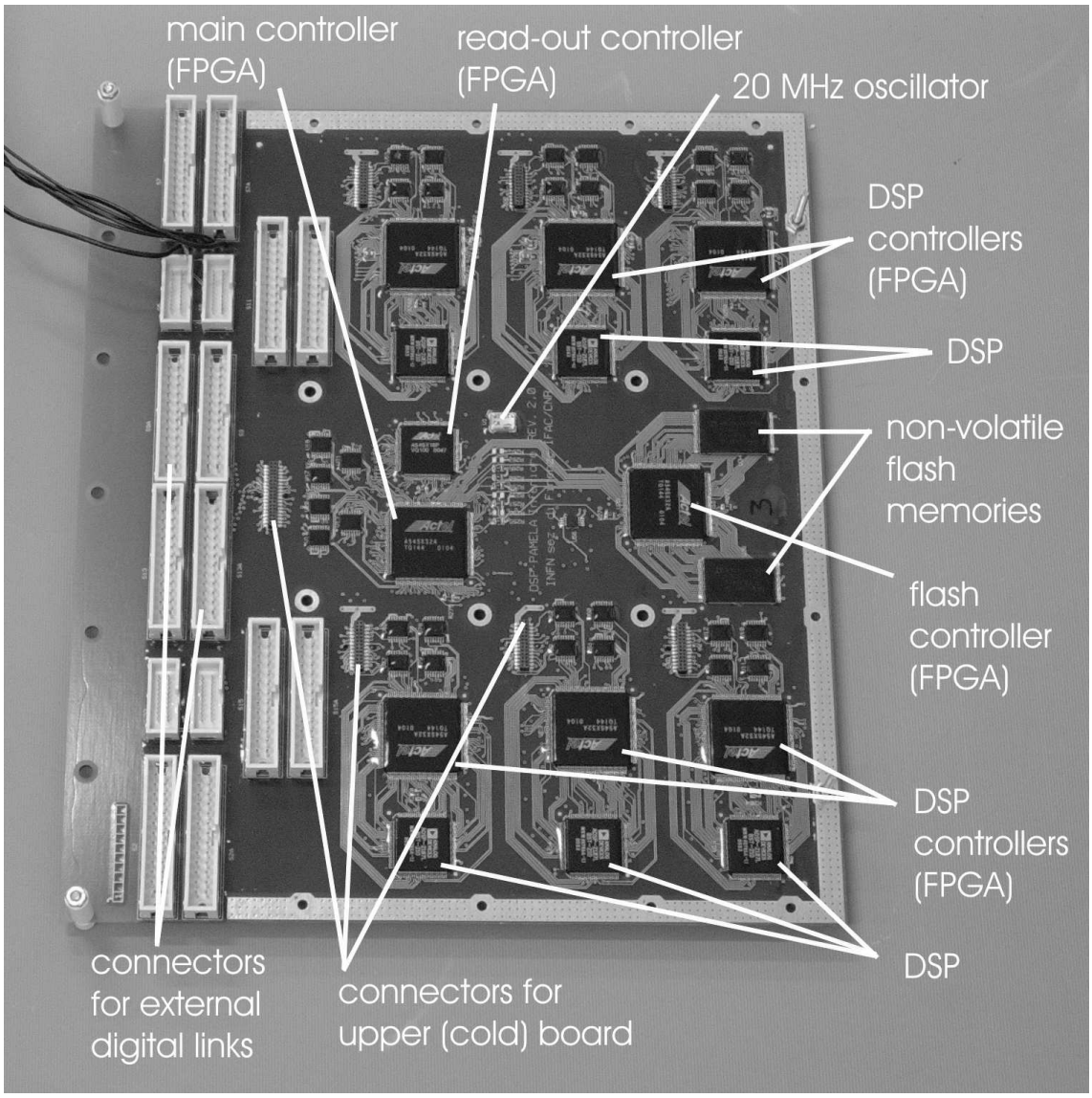

Figure 3.2: upper side of a *hot* DSP board featuring the main chips; the board dimensions are 20.0 cm · 23.2 cm. The corresponding *cold* board is superposed on top of this *hot* board.

that channel. Pedestal and noise parameters for all channels are determined during a dedicated calibration procedure of the compression algorithm, consisting in the acquisition of a fixed amount of events (1024) with triggers generated by the trigger board in response to specific commands. This calibration will be repeated periodically (typically every 45 minutes) during the operation of the PAMELA apparatus, to take into account variations of the parameters with working conditions, like temperature. With the implemented compression algorithm it is possible to achieve a typical *compression factor* (ratio of read-out data to compressed data) of 20 within a compression time of 1.1 ms[5].

[5]Note that both the compression factor and time depend on the actual pattern of signals on the silicon microstrips produced by the primary (and possibly secondary) particles crossing the apparatus; studies

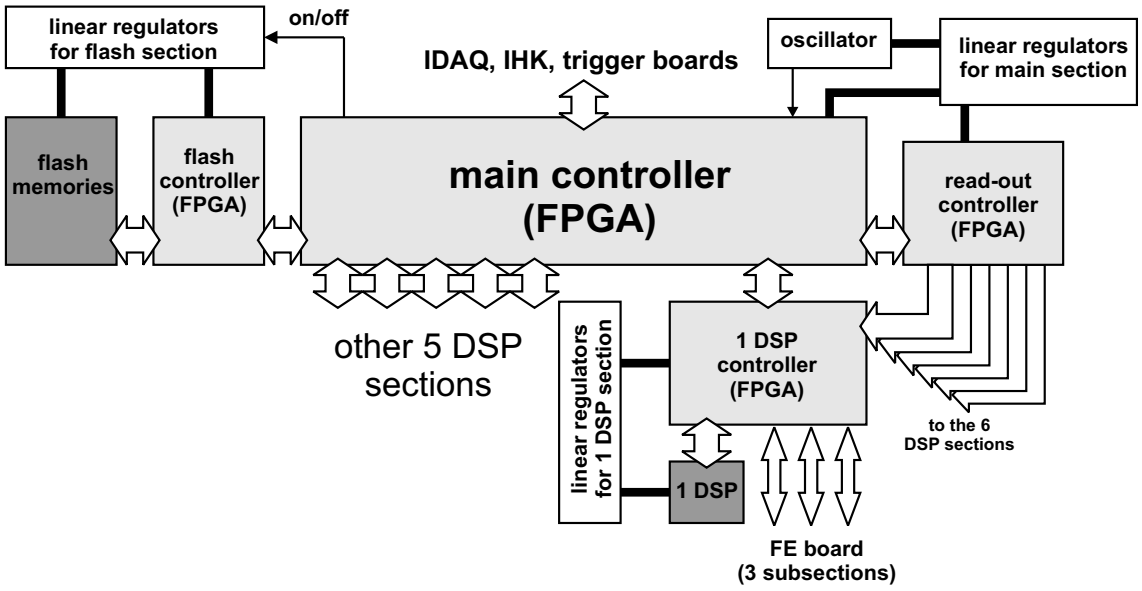

Figure 3.3: the structure of a DSP board, controlling half tracking system.

The DSP board is divided into 8 functional blocks, each with an independent power circuit:

- a main section, with:
  - the *main controller* FPGA, for the general management of the board operations and for interface toward the IDAQ, IHK and trigger boards[6];
  - the *read-out controller* FPGA, dedicated to the generation of the control sequence for the FE electronics after the arrival of a trigger pulse;
- 6 sections, each with a DSP and a *DSP controller* FPGA, for the simultaneous read-out and processing of data from 6 sides of the tracking system; one DSP section manages one FE board (3 independent subsections);
- a section with non-volatile *flash memories*[7], for the permanent storage of the code (about 10 kbyte) to be executed by the 6 DSP, and a dedicated *flash controller* FPGA[8].

by means of simulations and analysis of events acquired during tests of the apparatus have shown that the compression time follows a distribution with most of entries around 1.1 ms and a tail extending as far as 1.5 ms.

[6]These boards and the general structure of the data acquisition system of PAMELA have been introduced in sect. 1.3.

[7]A flash memory is a high-density and high-speed variety of EEPROM (*electrically erasable programmable read-only memory*) integrated circuit.

[8]For the implementation of the flash controller a design was used which had been previously developed and successfully tested for the flash section of the IDAQ board.

The power circuit for each section of the DSP board is based on two linear regulator chips (ADP3336 [19]) that transform the voltage (3.6 V nominal) supplied by the power system of PAMELA into two voltages characterized by high stability: an LVTTL-standard 3.3 V (nominal) used by all the on-board devices and the 2.5 V (nominal) for the inner cores of the FPGA chips, that with this method can achieve a lower power consumption (only the external interfaces of the FPGA chips are powered with 3.3 V for compatibility with the LVTTL standard). A dedicated current limiting stage sets the maximum current that can be absorbed by each regulator; the single limits have been chosen to be safely higher than the maximum dynamic current required by each section for proper operation.

The operating frequency of the board was set to 20 MHz, as the best compromise between the conflicting requirements of a sufficiently fast operation and a not too high power consumption. The clock signal is provided by a 20 MHz oscillator chip (CX03M [40]) based on a quartz crystal and guaranteeing high stability for all the working conditions: the relative deviation from the average frequency is less than $5 \cdot 10^{-5}$ in the temperature range $-10°$ C to $+70°$ C; the duty cycle (ratio between duration of the high state and clock period) is limited between 40% and 60%.

The operation of the DSP board (see fig. 3.4) is managed by the PSCU through the IDAQ board, with suitable sequences of commands within the general procedures involving the whole PAMELA apparatus. At the initialization of the board, the memory of the DSP chips is loaded with the program code to be executed and the parameters to configure the acquisition mode (basically to select partially different implementations of the compression algorithm); the program code is transferred from the on-board flash memories, while the configuration parameters are sent directly by the PSCU and can be varied either automatically, according to different scheduled operating modes of the PAMELA apparatus, or from the ground control station.

After the calibration of the compression algorithm, the board is ready for the acquisition phase; for an incoming trigger pulse, the FPGA chips control the FE electronics for the read-out of the detector and for the transfer of acquired data into the memory of the DSP chips. Each DSP waits for a fixed time interval to allow the complete storage of the data produced by one detector side on a reserved area of its memory, then starts the data compression procedure. Finally, after a suitable time to allow the end of data compression, the IDAQ board starts sending to the DSP board the command sequence to transfer the processed data into the mass memory of the PSCU for later transmission to ground.

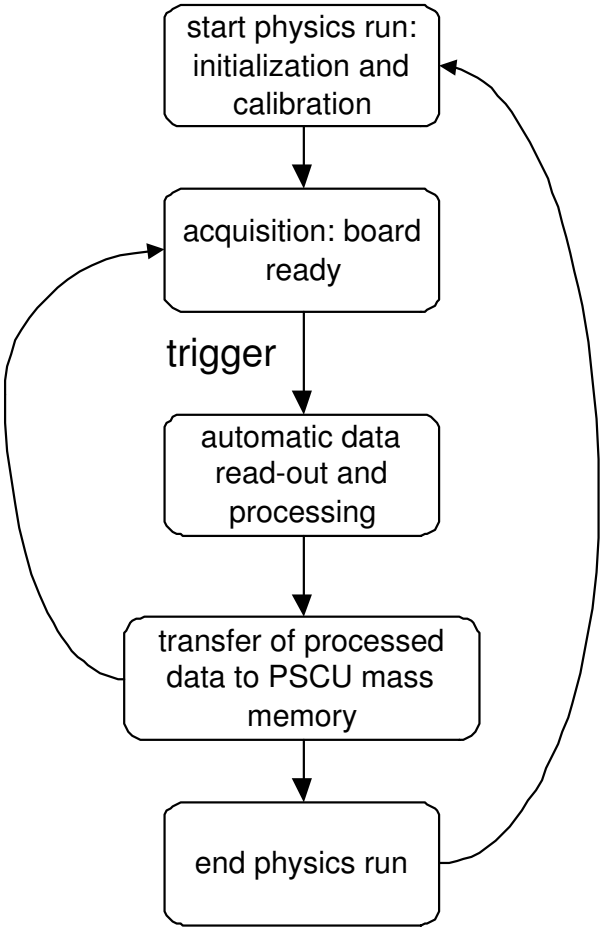

Figure 3.4: simplified scheme of the operation of the DSP board.

3.1.1 Characteristics of the DSP microcomputer

The DSP microcomputer selected for use in the control electronics of the tracking system (and also for data processing tasks in other subsystems of PAMELA) is the ADSP2187L [19]. This chip (1.4 cm · 1.4 cm area) integrates a 2100-architecture microprocessor and a basic set of peripherals, including a 160-kbyte RAM memory, two serial ports, a set of programmable input/output lines that can be used as interrupts or as control bits (*flags*) and an *internal direct memory access* (IDMA) port, granting the direct access to the internal memory of the DSP for an external device (in the present application represented by the DSP controller FPGA).

This DSP is capable of a maximum of 52 MIPS[9], corresponding to a time interval between successive operations of 19 ns. The operating frequency is fixed by the synchro-

[9]The MIPS (*mega-instruction per second*, 1 MIPS = 10^6 instructions/s) is a usual unit for measuring the timing performances of microprocessors. Here an instruction means a change of the internal logical configuration of the DSP according to the word of program code loaded in the instruction register.

nization signal (clock) entering a dedicated *clkin* input pin; each (positive or negative) transition on the *clkin* line causes an internal operation to be executed. For the DSP boards the *clkin* input frequency is set to 20 MHz thus achieving an operating frequency of 40 MIPS.

The internal RAM memory of the ADSP2187L is composed of a *data memory* and a *program memory*. Each of these two blocks contains 32 kword[10], with a 16-bit word length for the data memory and a 24-bit length for the program memory, the latter matching the instruction length of the 2100 microprocessor; hence the program memory is suitable for storing the binary code to be executed by the DSP, while the different word length of data memory provides a greater flexibility of use. The execution of program code, previously loaded in the program memory, starts automatically when an instruction is written on the location 0 of the program memory.

The program memory is divided into 4 pages or *overlays*; the first one is simply accessible by specifying the address, while the other 3 overlays have common addresses. The information on the currently active address is stored in a dedicated *address register* of the microprocessor, while the overlay is specified in an *overlay register*. The data memory is similarly structured.

The IDMA port is designed to allow an external controller to have a very fast access to the internal memory of the DSP, with a small number of control signals and a minimal data exchange protocol, such that a data transfer can be typically completed within few clock cycles, and completely managed by the external controller, even if the microprocessor itself is simultaneously operating on the memory during standard program execution.

The typical IDMA operation consists of one *address latch* to specify the address on which the subsequent transfer will operate (and possibly another similar latch to specify the overlay), followed by a *write* or *read* cycle on that location. The details of operation will be described in sect. 3.6.1.

A useful feature of the IDMA port is that the pointer to the memory location that will be affected by the next read/write operation, contained in the address register, is automatically increased by 1 after each data transfer; hence with a single address latch a continuous area of memory can be read or written sequentially.

[10] 1 kword = 1024 words.

3.2 DSP boards: control logics

The present section contains a description of the general design techniques that I have used for the implementation of the control logics on the FPGA chips of the DSP boards.

As mentioned in sect. 2.8.1, a synchronous logical machine, based on a unique clock signal propagating on a dedicated fast internal net inside the device, guarantees the maximum reliability against logical errors (upsets) caused by setup and hold time violations on data inputs of flip-flops. On the other hand, signals that are generated inside an FPGA and sent, for example, to another chip on a different board, are generally not guaranteed to satisfy the setup and hold requirements; in this case, when it is not possible to control the time interval between the transition of the input data line and the rising edge of the clock on the flip-flops of the external device, it is said that the sampling is *asynchronous*.

The consequence of an asynchronous sampling is that a metastable state and possibly an upset can be introduced in the signal chain. The impact of asynchronous sampling is worsened if the sampled signals are affected by the presence of short transient pulses (commonly known as *glitches*). Glitches are generally observed on the outputs of any combinatorial element; consider for example the simple 2-input AND gate in fig. 3.5: the delays between the *clock_a* rising edge and the transitions of the *input_1* and *input_2* lines are in general different, because the physical paths covered by the signals inside the chip are not the same. Therefore the output of the AND gate may pass through a transient spurious state when the values of both *input_1* and *input_2* change like shown in figure; the duration of the spurious pulse is given by the relative delay between the two transitions (order of ns). If the output of the AND gate is sent to a different device where it is asynchronously sampled with *clock_b*, characterized by a certain phase shift with respect to *clock_a*, there is the possibility that the sampling happens during the glitch and that the flip-flop output *data_out* decays in the wrong state, thus introducing a logical error in the signal chain.

On the other hand, glitch-free signals are naturally generated by a flip-flop sampling a synchronous signal, since its output performs no more than one single transition for each clock rising edge. With the introduction of a flip-flop stage (and a consequent fixed delay equal to one *clock_a* cycle) on the output of the combinatorial element (see fig. 3.6), the consequence of asynchronous sampling is limited to the introduction of a time uncertainty (commonly indicated as *jitter*) on the transition of *data_out*.

In fact it can be seen that in this configuration the presence of a glitch on the output of the AND gate would have no effect on the *data_out* line. On the other hand, the delay

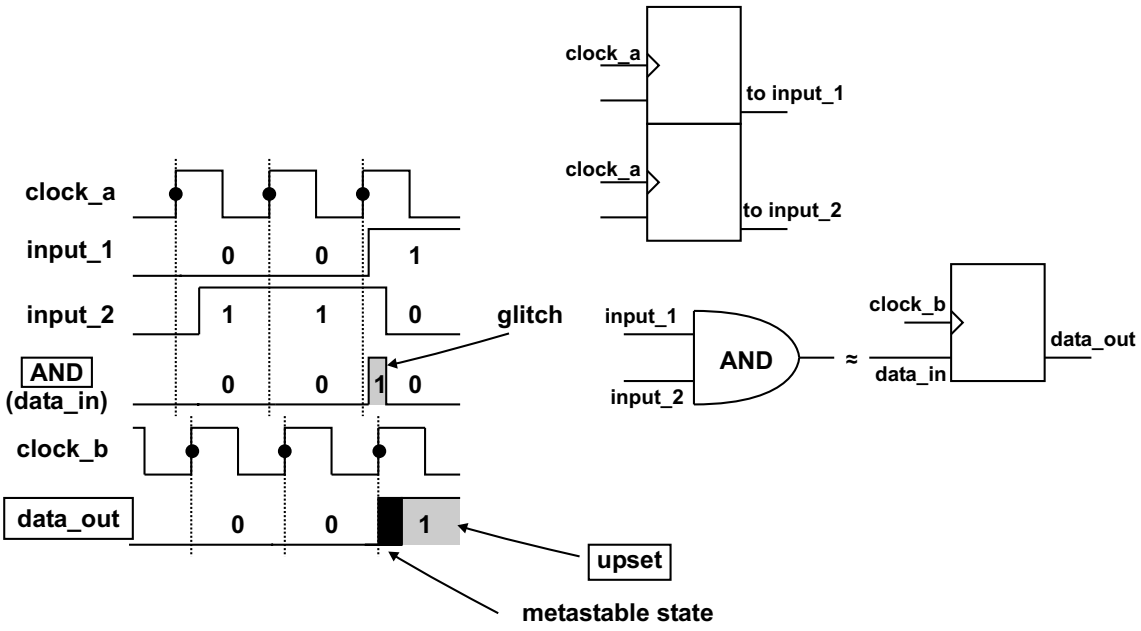

Figure 3.5: asynchronous sampling of a combinatorial signal (*data_in*), affected by possible spurious pulses (glitches) which would introduce wrong transitions (upsets) in the signal chain; in this example *data_in* is nominally kept at 0 while *data_out* is wrongly set to 1. The glitch on *data_in* is due to the different propagation times from the rising edge of *clock_a* to the transitions of the lines *input_1* and *input_2*.

between the transition of *data_in* and the correct change of value for *data_out* can vary up to one *clock_b* period, depending on the relative phase shift of the two clocks; note that even in the case that the asynchronous sampling of *data_in* causes a violation of setup or hold time with the subsequent decay of *data_out* in the wrong state, this does not introduce an upset but simply delays the correct update of *data_out* to the following rising edge of *clock_b*.

Basing on the previous general considerations, I have extensively used synchronous machines generating glitch-free signals. A typical, elementary example is shown in fig. 3.7: the rising edges of the clock cause a counter (array of flip-flops plus combinatorial net) to increase its output code by one; a decoder (combinatorial net), according to the counter output, sets the data inputs for an array of flip-flops, whose outputs are used as control signals for other devices or sections of the same FPGA. The layer of output flip-flops guarantees that each output signal is free from glitches and suitable to be asynchronously sampled.

When required by the application, a more general structure (commonly known as *state machine*, see fig. 3.8) has been implemented, differing from the previous one for the presence of external synchronous control signals and, instead of the counter, of an array of

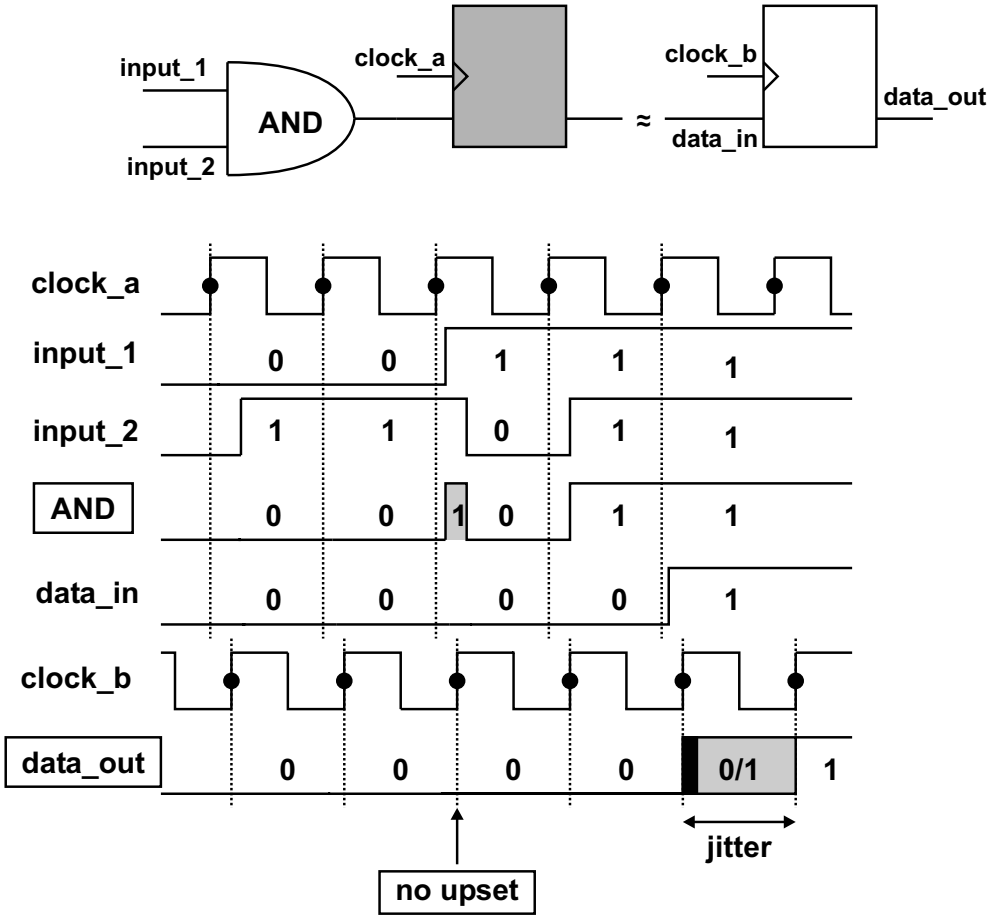

Figure 3.6: asynchronous sampling of a sequential signal (*data_in*), which is not affected by glitches thanks to the insertion of a further flip-flop stage (in grey) in the signal chain. With this system the output *data_out* replicates the nominal AND waveform without upsets and at the expense of introducing a delay of one *clock_a* cycle plus a jitter.

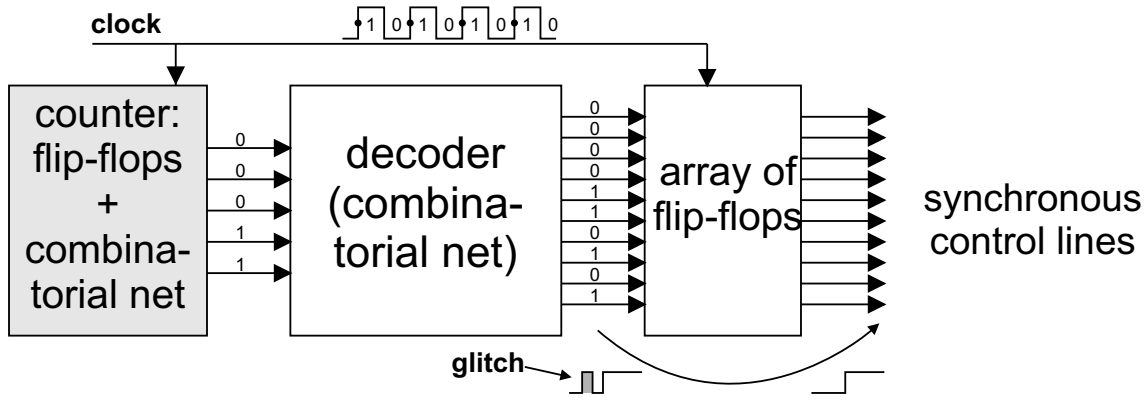

Figure 3.7: elementary synchronous machine based on a counter.

flip-flops whose configuration represents the logical state of the machine. Once started by an external *enable* input, the machine operates a full cycle finally returning to the initial

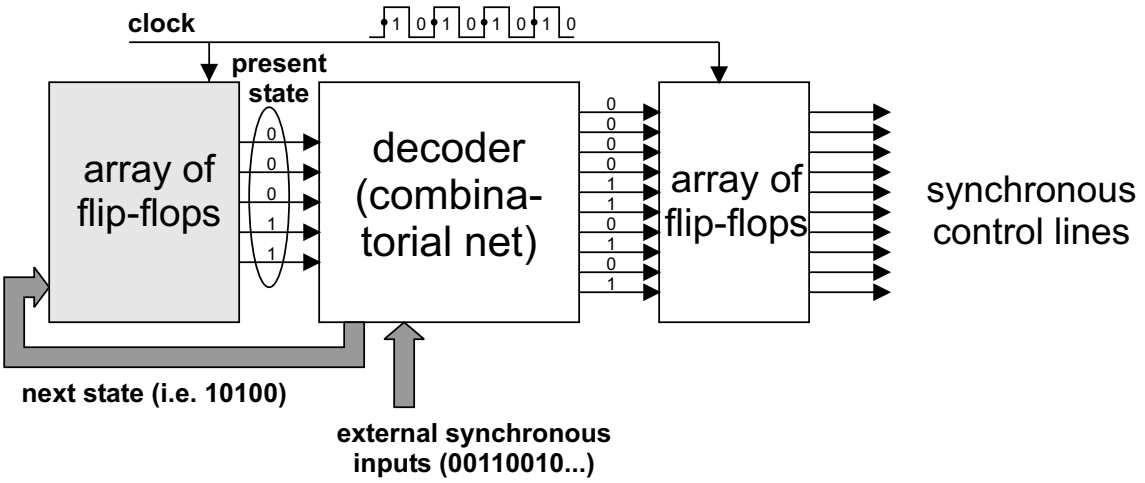

Figure 3.8: a more general synchronous machine (state machine).

no-operation (*idle*) state. While the elementary machine follows a fixed sequence of logical states, represented by the output codes of the counter, the generic state machine can jump from one state to another depending on both the current state and the external input pattern; also the output pattern for a given internal state can change depending on the external inputs. This flexible architecture is convenient when the sequence of operation to be performed or the number of times that a specific internal loop must be repeated is variable, for example according to the value of a parameter which is stored in a dedicated register.

3.3 DSP board: external digital links

Each digital transmission line connecting the DSP board with an external device (IDAQ, IHK, trigger, FE boards) is formed by a twisted pair of wires on which the signal is sent according to the *low voltage digital signalling* (LVDS, see fig. 3.9) electrical standard; this standard has been widely adopted in PAMELA since it allows the transmission of fast logical signals (up to 400 Mbps)[11] with relatively low power absorption.

An LVDS line is managed by dedicated driver and receiver integrated circuits; on the DSP board we have employed DS90LV047A drivers and DS90LV048A receiver chips [41], each capable of managing 4 independent lines. The operation of an LVDS line is as follows: the driver translates an input LVTTL voltage level into a steady output current; typically this current is 3.1 mA with sign depending on the logical state. The voltage

[11]The speed of a digital link is usually expressed by means of the number of bits (changes of logical level) performed by the line per unit time (*bit rate*, expressed in bit per second, bps); 1 Mbps = 10^6 bps.

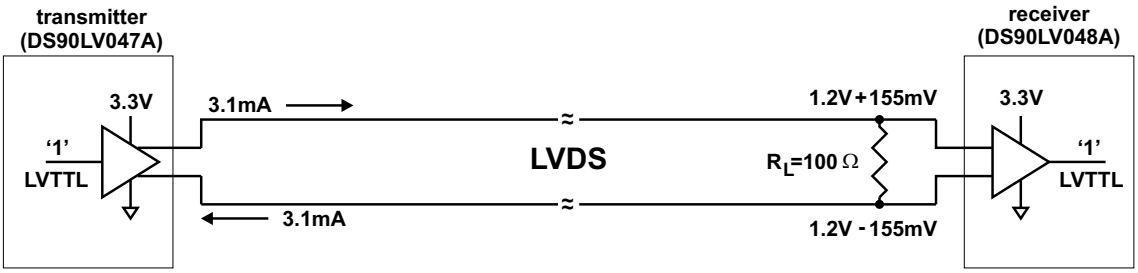

Figure 3.9: an LVDS digital link in the logical state 1.

drop across the load resistance R_L is detected by the receiver chip which produces an LVTTL-compatible output voltage. The value $R_L = 100\ \Omega$ is chosen to best match the characteristic impedance of the twisted-pair transmission line, hence minimizing the fraction of signal pulse which is reflected back along the line during the transient phase; the typical voltage difference across R_L is ± 310 mV, centered around 1.2 V (driver offset voltage).

The typical static current absorbed by a driver/receiver pair (4 LVDS lines) is 30 mA; the dynamic current, at the frequencies of transmission used in PAMELA (not more than 20 Mbps as will be explained in sect. 3.4.1) is practically the same.

By exploiting the fact that only one DSP board in a *hot*/*cold* pair is operating at a time, it is possible to reduce by one half the total number of external lines with the implementation of a shared configuration (shown in fig. 3.10). The external connectors for the flat cables with the twisted-pair transmission lines are placed along the front side of the *hot* board; a set of internal connectors assures the electric continuity between the external ones and the LVDS chips of the superposed *cold* board (see fig. 3.2).

The feasibility of the output configuration is based on the fact that the output stage of the driver is in a high-impedance state both when powered (ideal current source) and turned off. Suppose for clarity that the *hot* driver and receiver are powered (*hot_3.3V* on) and the *cold* ones are turned off: the static current generated by the *hot* driver goes entirely across the load R_L, generating the typical voltage drop. On the other hand, when the *hot* driver changes the polarity of its output current, half of the wave propagating along the line deviates at node A toward the *cold* driver where it is completely reflected back on the high impedance, reaching again node A; because of the consequent multiple reflections, the current across the load R_L reaches its final value after a time of the order of $\Delta t = 2 \cdot d/v$, where $v \approx 20$ cm/ns is the velocity of the electric signals and $d \approx 1$ cm is the typical distance between node A and one of the two drivers, minimized by keeping the

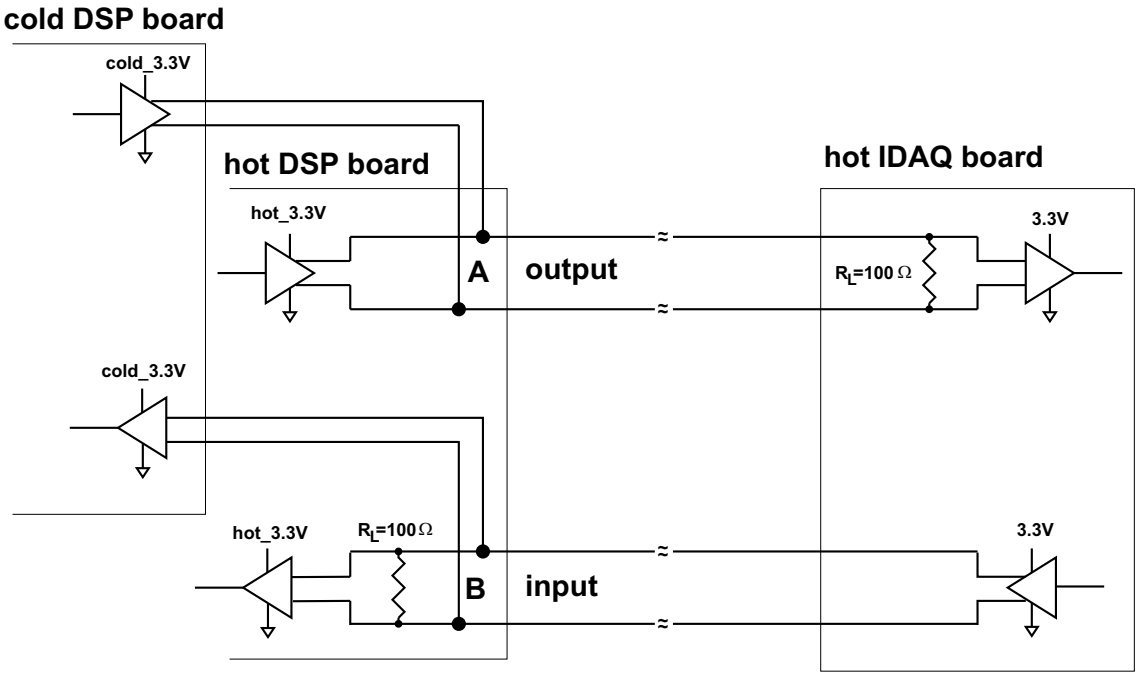

Figure 3.10: actual shared LVDS links between a pair of *hot*/*cold* DSP boards and the (*hot*) IDAQ board.

two boards as near as possible and by positioning the chips in the vicinity of node A. In this configuration $\Delta t \approx 0.1$ ns and no spurious pulses are observed on the LVTTL output of the receiver since for the LVDS standard the characteristic duration of a transition is of the order of 2 ns, much greater than Δt.

For the input configuration analogous considerations can be done, since also the input stage of the receiver can be represented by a high-impedance both when powered and turned off.

The interface toward the FE electronics is complicated by the fact that the reference of the Y-side FE electronics is +80 V above the PAMELA ground[12] which coincides with the ground of the X-side FE electronics and of the DSP boards.

The decoupling of the bias high voltage and low-frequency noise has been achieved with an active high-pass filter on the FE board, inserted between the LVDS stage and the rest of the circuit, as exemplified in fig. 3.11 for a signal going from the DSP board to the FE board: after the LVDS receiver stage a dedicated chip (ADuM1100 [19]) with an internal inductive decoupler (transformer) guarantees the low-frequency filtering. A transition in the LVTTL input of ADuM1100 is encoded by the internal logics as a current pulse in the transformer; the sign of the current depends on the logical transition. The

[12]This is necessary to avoid applying the 80 V bias across the thin SiO_2 integrated coupling capacities on the detector planes.

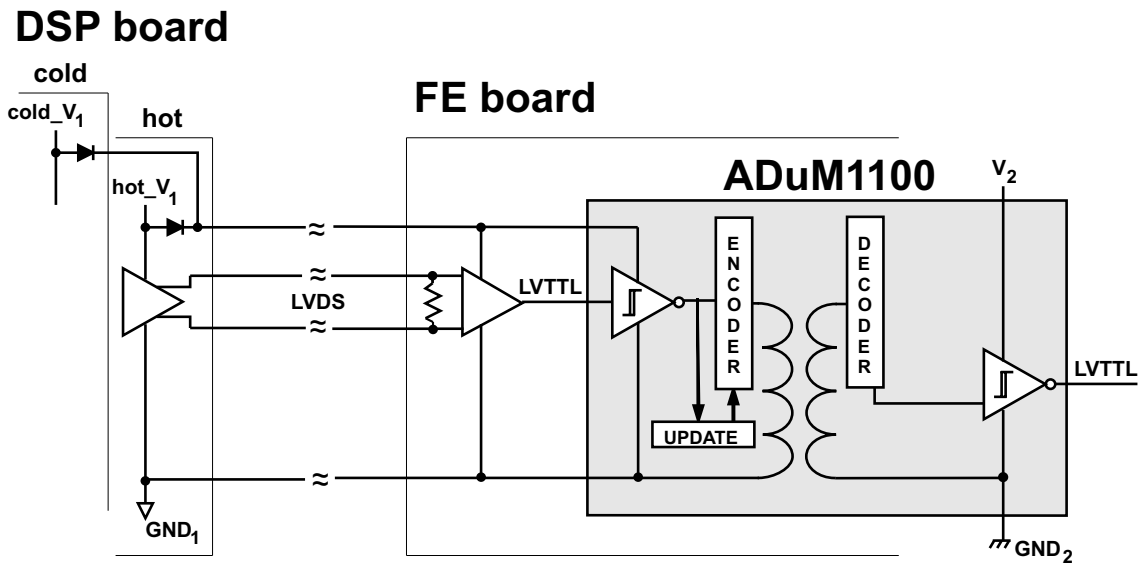

Figure 3.11: high-pass filter with ADuM1100 for the LVDS link from a DSP board to a FE board; an analogous configuration has been implemented for the other direction of the link.

decoder changes its logical state according to the sign of the current pulse.

Besides this, the update circuit of the ADuM1100 assures that, in absence of transitions on the input LVTTL line, the decoder output will remain in the correct state, by sending a current pulse of the appropriate polarity to the transformer every 2 μs from the last input transition.

This solution has some drawbacks. The use of an active component increases the power consumption (on average 0.6 mA/line). Additional wires must be added to bring the voltage from a DSP section to the LVDS stage and the corresponding side of the ADuM chip on the FE board. Besides, for safety reasons diode protection stages must be inserted to break the electric continuity between the power lines of each *hot* and *cold* pair of DSP sections. The voltage drop introduced by this diode stage must be sufficiently low to guarantee that the voltage supplied to the FE board is safely inside the specified LVTTL range (3.3 ± 0.3 V). With the employment of a Schottky diode characterized by a ultra-low internal drop and setting the voltage on the DSP sections at 3.35 V, the total drop has been limited to 0.25 V (of which 0.2 V due to the diode and 0.05 V due to the resistance of the wires connecting the FE boards to the DSP board) thus obtaining 3.1 V on the FE boards.

Anyway this solution is preferable to others employing passive decoupling methods (e.g. RC filters or simple transformers), since in those cases the digital line cannot remain too long in the same state, otherwise the output of the high-pass filter would decay

to its quiescent state, possibly causing a logical error. To avoid the decay of the filter output, the data transmission protocol would be required to assure a suitably short time interval between subsequent transitions, even in absence of changes of logical state of the transmitted data. A specific study [15] has led to the conclusion that similar protocols would imply to multiply at least 3 times the rate of transitions on the signal lines to achieve the same bit rate, with an increase in dynamic power consumption of about 3 mA/line (at the chosen bit rate of 10 Mbps/line).

3.4 DSP board: digital protocol for data exchange with external devices

The design and implementation of the external digital links connecting the various subsystems of PAMELA together has represented an important part of the integration work on the apparatus.

The presence of many electronics boards, typically segmented into several independent sections and duplicated for redundancy, and the consequent request of a high number of links, strongly conflicts with the strict limits imposed by the satellite on available volume and power. With the adoption of a serial protocol for data exchange, instead of a parallel one, a drastic reduction of the number of wires and connections has been achieved, with a significant improvement also of the reliability of the whole system and with the acceptable drawback that in this way, for a fixed operating frequency of the boards, the data transfers during the acquisition procedures give a larger contribution to the overall dead times, with respect to a solution characterized by parallel data exchange.

The *data/strobe* serial protocol has been specifically adopted for use in PAMELA, mainly because of its advantages in terms of dynamic power consumption with respect to other available serial standards. The data/strobe transmission needs only two lines for each direction; hence a bidirectional data/strobe link, with an LVDS physical layer, is obtained with 4 differential lines (8 wires).

The principle of operation of this protocol is illustrated in fig. 3.12, in the implementation adopted for PAMELA: the transmitter, on an FPGA chip operating at a certain *clock* frequency, sends the sequence of significant data bits on the *data* line, while the auxiliary *strobe* line performs a transition only in correspondence of those *clock* rising edges for which the *data* value remains unchanged. The *data* and *strobe* transmission lines are operated according to the LVDS electrical standard, as explained in the previous section. The receiver (an FPGA chip on another board) can reconstruct a periodic synchronization

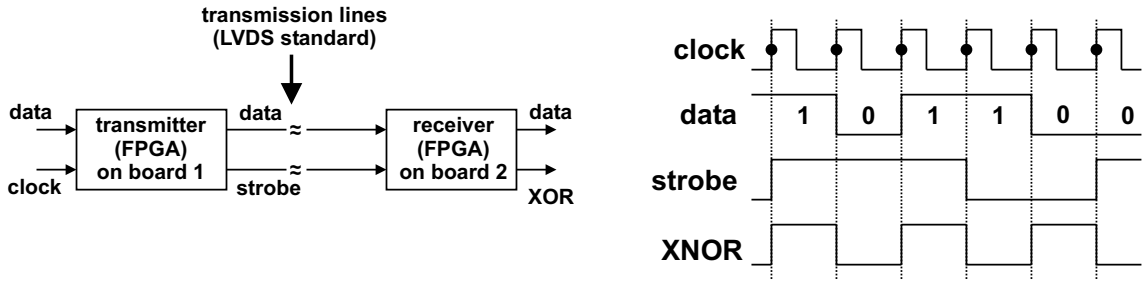

Figure 3.12: data/strobe protocol: principle of operation. A transition happens on the *strobe* line if and only if *data* does not change.

signal with a simple logical operation (XNOR) on the *data* and *strobe* values, and then use this signal for loading the *data* bit on an internal register.

The advantage of the data/strobe encoding, in terms of dynamic power consumption, is evident if compared to a traditional serial protocol, that simply transmits the *data* and *clock* signals, using the *clock* falling edge for data loading on the receiver. Considering the transmission, in a given time interval, of a generic sequence of N bits that causes N_d transitions to happen on the *data* line, the number of switchings on the *strobe* line is $N_s = N - N_d$, since for each new transmitted bit a transition happens on only one of the two lines. Therefore the dynamic power consumption on the *strobe* line (proportional to the number of transitions in the given time interval) is less than half that of the traditional case, where the switchings on the *clock* line are fixed to be $2 \cdot N$.

The application of this protocol to the data exchange between DSP boards and FE electronics of the tracking system (discussed in [15], p. 57), is simplified by the fact that the FE section works only during the read-out of the detector and the clock needed for all the internal operations can be directly derived from the XNOR output of the data/strobe receiver; for this reason no oscillator is present on the FE electronics, with the consequent minimization of the noise induced in the charge preamplifiers performing the read-out of the silicon sensors.

On the other hand, in the case of the communication between the DSP board and the IDAQ board, the transmitted data enter a board which is operating with a clock generated by a different oscillator chip and must therefore be synchronized with this internal clock to be made available for further use.

For this reason the design of a reliable receiver unit and the study of the timing characteristics of the data/strobe protocol have constituted an important part of my development work and are described in the following subsection. The detailed explanation of the internal structure of this unit is needed to discuss and justify its characteristics in

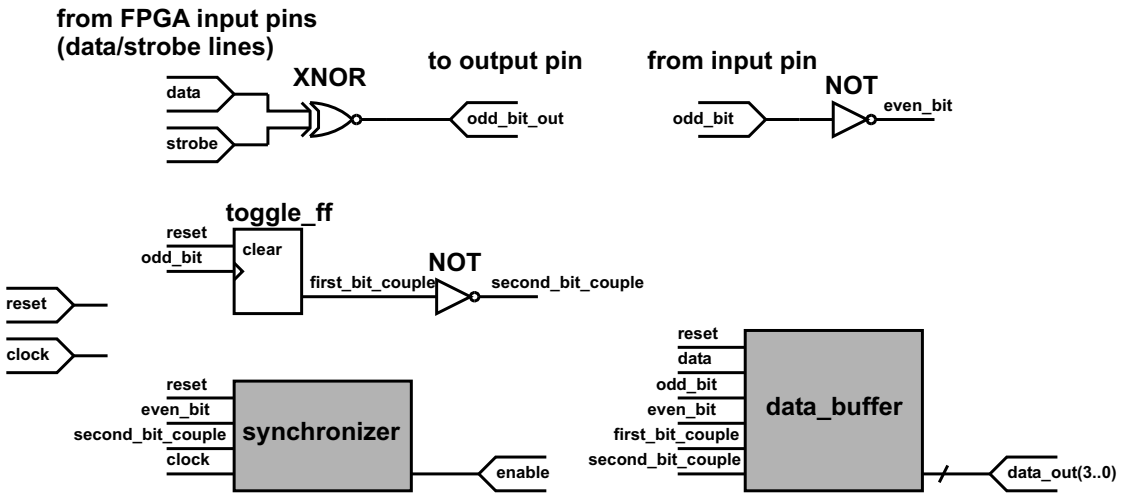

Figure 3.13: scheme of the 4-bit data/strobe receiver unit for the synchronization of the incoming data with the *clock* generated by the on-board oscillator.

view of the implementation of the external digital links of the DSP boards.

3.4.1 Design and test of an asynchronous data/strobe receiver

The structure of the receiver unit that I have designed is shown in fig. 3.13[13]. The main logical blocks are a *data_buffer* (fig. 3.14) with an internal shift-register for the transformation of the serial data stream into a sequence of 4-bit parallel words *data_out(3..0)*, and a *synchronizer* (fig. 3.15) that generates an *enable* pulse to indicate each time that the content of the *data_out* bus is valid and ready for use. The internal signals *odd_bit*, *even_bit*, *first_bit_couple*, *second_bit_couple* are used in combination to indicate the presence of the first, second, third or fourth bit of each word on the *data* input of the receiver.

The *data_buffer* works as an ordinary shift register with the difference that it exploits the four subsequent XNOR edges to load the 4-bit word on *data_out(3..0)*. Note that in this particular case it is assumed that *data* and *strobe* lines are initially in different states (their XNOR is normally 0), and that the bits forming the word are transmitted starting from bit 3, the most significant one (MSB) down to bit 0 (least significant bit, LSB).

The value of *data* input is loaded on the flip-flops of *data_buffer* in correspondence of the rising edges of the signals *odd_bit* and of its negated *even_bit*; *odd_bit* is derived from the XNOR line after inserting a sufficient delay to assure that no violation of the setup time can happen (note in fact that the transition of the *data* line at a flip-flop input is ideally simultaneous with the transition of *odd_bit* or *even_bit*). This delay is

[13]Here for clarity a 4-bit receiver is discussed, while, as explained below, for the DSP board a completely equivalent 8-bit version has been adopted.

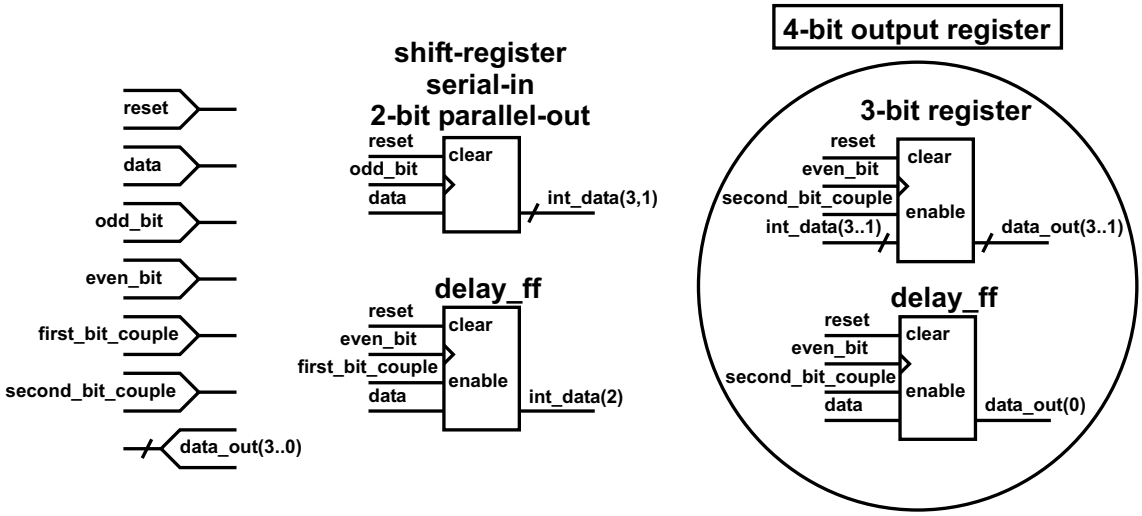

Figure 3.14: structure of the *data_buffer* block.

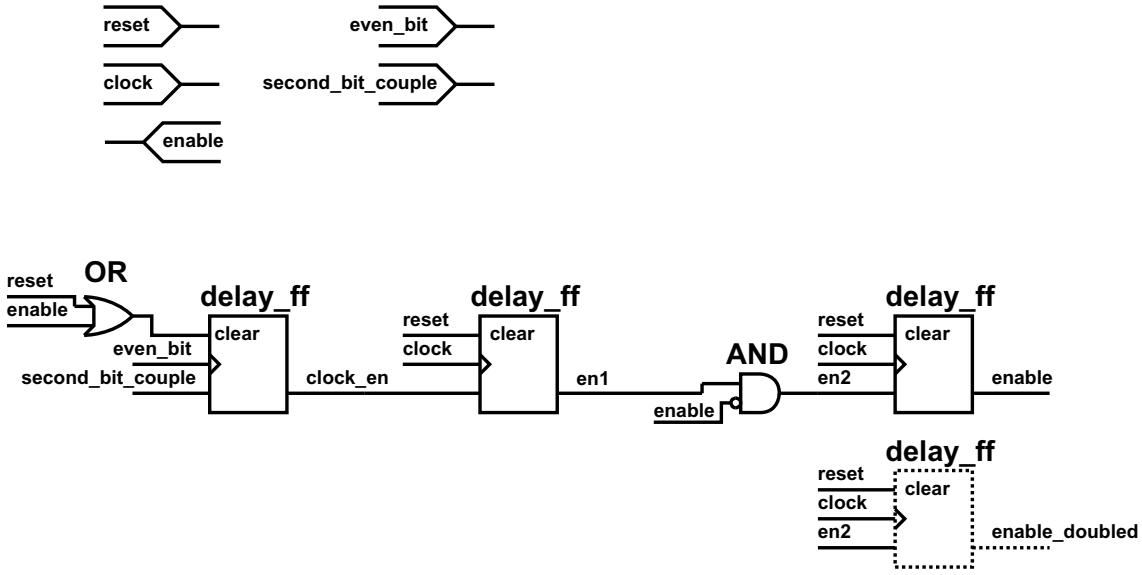

Figure 3.15: structure of the *synchronizer* block.

obtained by sending the XNOR signal to an output pin of the FPGA which is externally short-circuited with another input pin; this solution is feasible since the minimum delay introduced by the cascade of an input and an output stage of the FPGA is sufficient to guarantee that no setup time violation can happen in any working conditions.

With the fourth (last) data/strobe transition (represented by a rising edge of *even_bit* with *second_bit_couple* kept at 1) the content of the internal bus *int_data(3..1)*, constructed with the 3 previous transitions, together with the content of the *data* input, are loaded in the 4-bit output register, thus assuring the stability of *data_out* till the arrival of a new

fourth-bit transition.

The *synchronizer* unit (fig. 3.15) generates the *enable* pulse, synchronous with *clock*, in correspondence of each fourth transition on the data/strobe line (end of word). The fourth data/strobe transition causes the assertion of *clock_en* which is then asynchronously sampled by *clock*, the delay from the assertion of *clock_en* to the rising edge of *clock* being in general arbitrary between 0 and the *clock* period T_{clock}. In particular if the *clock_en* transition is sufficiently near to the *clock* rising edge, a setup or hold time violation can happen with *en1* entering a metastable state from which it decays to 1 or 0 in few ns (in case it goes to 0, the following *clock* rising edge will set *en1* in the correct 1 state). The effect of this mechanism is to introduce a jitter corresponding to one *clock* cycle in the assertion of *en1*; this jitter is unavoidable because of the asynchronous sampling.

The *enable* signal is generally used in different parts of the control logics containing the data/strobe receiver, to check when the content of *data_out* is valid and ready. This introduces a complication in the design of the *synchronizer*, since when the number of logical cells driven by a flip-flop is such that the total load capacitance causes a non-negligible delay in the rise time of its output, it is necessary to duplicate (or in general multiply) the flip-flop. If the flip-flop generating *en1* would be duplicated, the metastable state could decay differently in the two flip-flops, thus introducing a delay of one *clock* cycle between the two physically distinct *en1* lines which should instead represent the same logical signal. This problem has been avoided by fixing the load of *en1* line to one flip-flop and using the output of this further flip-flop stage as *enable* signal.

The assertion of *enable* causes the first flip-flop of the chain to immediately clear its output to 0; with the following *clock* rising edge, the whole chain returns to its quiescent configuration. The circuit takes at most 3 *clock* cycles to recover the initial configuration after the activation by the fourth data/strobe transition; this is the minimum possible delay, given the previous considerations. To assure safe operation, the interval between two subsequent n-bit words on the data/strobe link must be at least 4 *clock* cycles; therefore the interval T_{bit} between two consecutive bits must be greater than $(4/\mathrm{n}) \cdot T_{clock}$.

Besides, typically the parallel *data_out* bus must be serialized for transmission to other on-board FPGA chips; this implies that the time employed to transmit a word in the data/strobe link, $\mathrm{n} \cdot T_{\mathrm{bit}}$, must be greater than the time required for the serialization of the word, $(\mathrm{n}+1) \cdot T_{clock}$, given by n *clock* cycles plus 1 additional cycle to take into account the presence of the jitter in the *enable* signal, which could delay the serialization of the previous word by one *clock* cycle.

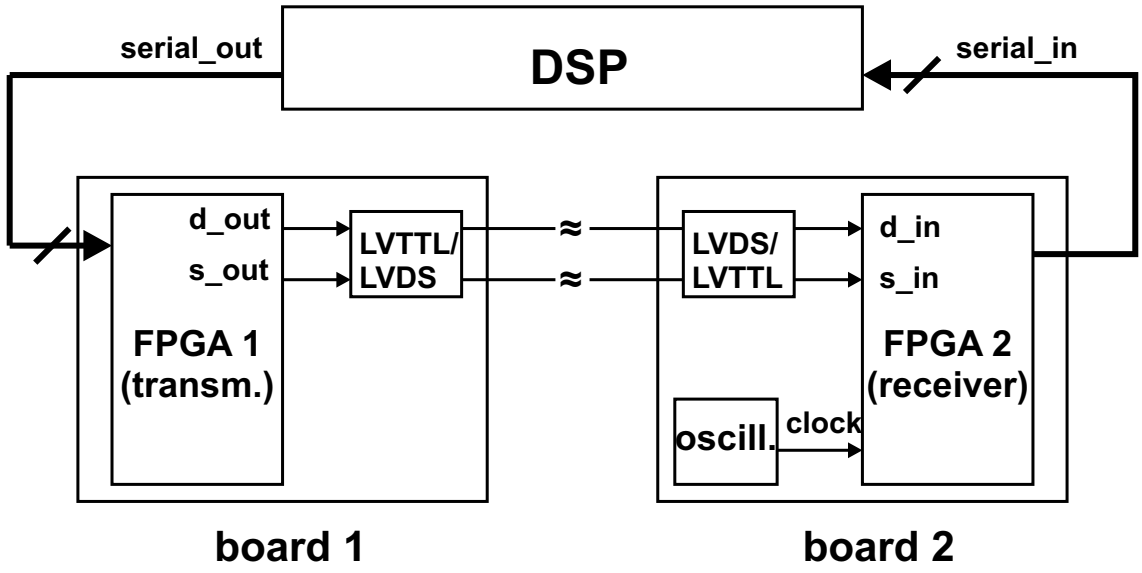

Figure 3.16: setup used for testing the functionality of the data/strobe receiver.

The previous considerations can be summarized in the following relation:

$$\mathrm{T_{bit}} > \frac{\max[4, (\mathrm{n}+1)]}{\mathrm{n}} \cdot \mathrm{T}_{clock} \tag{3.1}$$

To satisfy this requirement we set the word length to 8 bits, obtaining:

$$\mathrm{f}_{clock} > \frac{9}{8} \cdot \mathrm{f_{bit}} \tag{3.2}$$

and chose the frequencies f_{clock} of the on-board oscillators to 20 MHz for the DSP board and 16 MHz for the IDAQ board; the transmission from the IDAQ board to the DSP board is done at a bit rate $\mathrm{f_{bit}} = 16$ Mbps, while the transmission from the DSP board to the IDAQ board is done at 10 Mbps.

We performed dedicated functionality tests to verify the reliability of this design over all the voltage ($3.0 - 3.6$ V) and temperature ($0 - 70^\circ$ C) range recommended for operation by the FPGA manufacturer [32].

The test setup is shown in fig. 3.16. A DSP microcomputer produces a sequence of 8-bit words of value increasing cyclically from 0 to 255, which are encoded in data/strobe protocol by the transmitter on board 1 and sent to board 2 through an LVDS link; the receiver sends back the words to the DSP through another serial port. The DSP checks that every received word is equal to the word sent.

The test has been performed without registering any error for transfers of tens of Gbyte (equivalent to several days of operation of PAMELA).

3.5 DSP board: main controller FPGA and general operation

The main controller FPGA, whose scheme is shown in fig. 3.17, manages the general operation of the DSP board and acts as the unique interface toward the IDAQ, IHK and trigger boards.

The controller receives commands from the IDAQ board and sends back answers through bidirectional data/strobe links. The active IDAQ board (*hot* or *cold*) is automatically identified with the arrival of the first command byte on the corresponding data/strobe inputs; once selected the active IDAQ link, the other link is temporary disabled until the end of command execution.

The internal operation is based on an 8-bit parallel data bus, matching the data/strobe word length. The 20 MHz clock produced by the on-board oscillator is used internally and distributed to the other FPGA chips.

The main controller exchanges data with the 6 DSP controllers and the flash controller through bidirectional serial links. Serial protocols are also employed for the data exchange between the read-out controller and the 6 DSP controllers (see sect. 3.7); this aims to reduce the total amount of lines linking the various FPGA chips, to match the available number of external pins and the space for signal tracks within the board.

As an exemplification, the serial protocol adopted for command transmission from the main to the DSP controller FPGA is illustrated in fig. 3.18. The negative transition of *DSP_clock* is used by the receiver to load the content of *comm_data* on a flip-flop; the *comm_frame* line set to 1 signals that a valid bit is present and enables the data transfer. The use of a third line (*comm_frame*) is made necessary by the fact that the clock is continuously sent to the DSP controller for its internal operations.

The *status register* of the main controller (see table 3.1) stores all the relevant operating parameters and keeps track of the occurrence of various error conditions by means of 25 alarm bits; 6 mask bits indicate if the *busy* and *alarm* signals coming from a specific DSP section must be ignored (e.g. in case of permanent failure of the section); 2 other mask bits allow to ignore trigger transitions on either the *hot* or *cold* trigger input, to avoid that the presence of anomalously high noise on the line currently not active could be interpreted as trigger pulses. The OR of all the (non-masked) alarm bits forms the external *alarm* signal which is sent to the IHK board to indicate the presence of an anomalous condition.

Normally the main controller is in the *idle* state, waiting for a command sent by the IDAQ board. The first *enable* pulse produced by the internal data/strobe receiver,

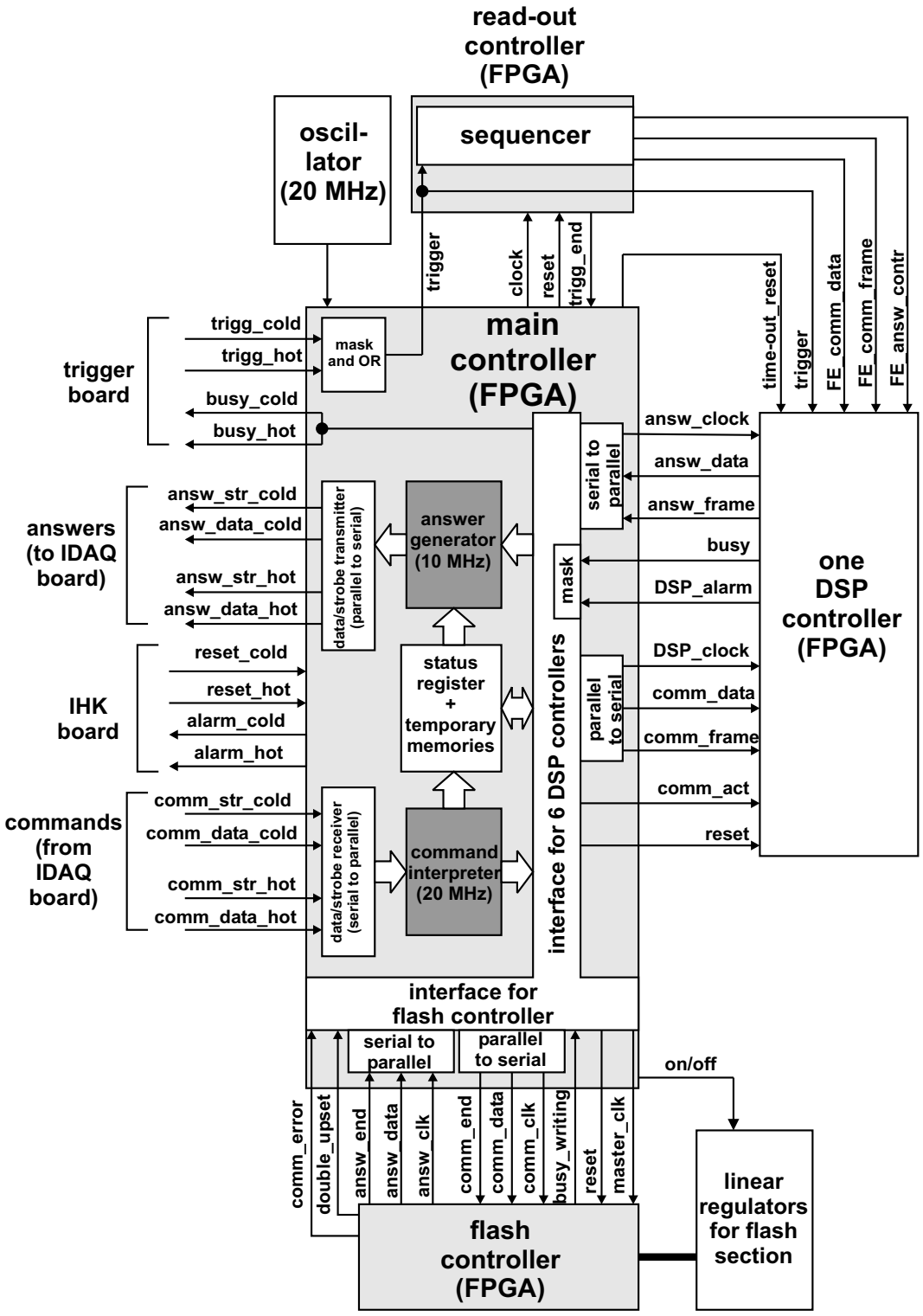

Figure 3.17: functional scheme of the main controller and of the read-out controller FPGA of the DSP board, with their external interfaces (for clarity only the interface with one of the 6 DSP controllers is shown).

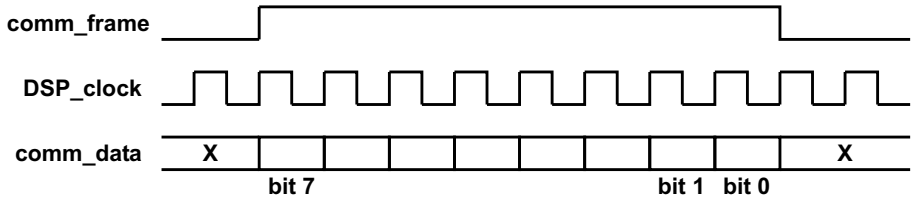

Figure 3.18: serial protocol used for data transfer from the main controller to the DSP controller FPGA.

Status bits	Meaning		
39..32	hot trigger mask	cold trigger mask	DSP mask (5..0)
31..24	flash off when addr.	flash on	DSP busy (5..0)
23..16	flash double upset	flash command error	DSP alarm (5..0)
15..8	interm. CRC error	final CRC error	IDMA err. dur. comm. (5..0)
7..0	unknown comm.	DSP comm. dur. r. -out	IDMA err. dur. read-out (5..0)

Table 3.1: map of the status register. Bits 0 to 23 and bit 31 represent the 25 possible alarm conditions; mask bits are 32 to 39.

corresponding to the reception of the first byte of a command, activates the *command interpreter* (CI, fig. 3.19) that sequentially decodes the byte stream and performs the required operations. The answer to the IDAQ board is produced by a dedicated *answer generator* unit (AG, fig. 3.20). Finally the main controller returns to the *idle* state, ready to execute a new command. While the CI works at the full 20 MHz clock frequency, the AG operating speed is set to 10 MHz to match the data/strobe link requirements, as discussed in sect. 3.4.1.

A list of all the possible commands is reported in table 3.2. The command execution may involve only the main controller (*internal command*) or also other peripheral FPGA chips of the board (one DSP controller or the flash controller). The general structure of a command and of an answer are shown in fig. 3.21 and 3.22 respectively.

The command is formed by a 4-byte header and a tail (final byte), with a possible intermediate data block. The header and tail contain information on the length (in bytes) of the command, the command identifier as in table 3.2 and two CRC[14] codes for consistency checks. Similarly the answer to a *read command* is composed of a data block and a 6-byte tail, with a copy of the command identifier, a board identifier, the value of the *alarm* line, the length and a CRC code; the answer to a *non-read command* is formed

[14]The CRC (cyclic redundancy check) has been introduced in sect. 2.4.

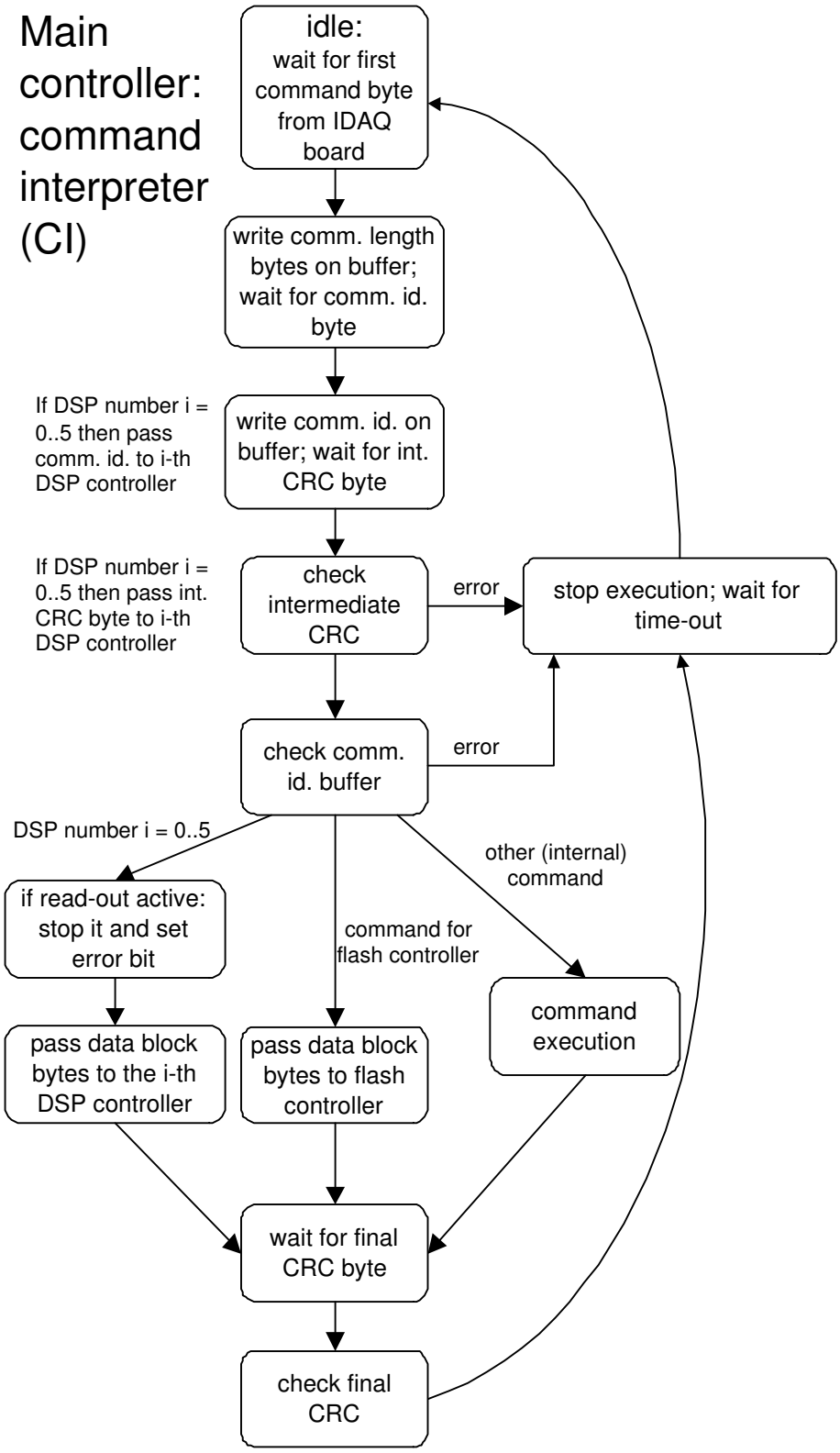

Figure 3.19: main controller: functional diagram of the command interpreter unit.

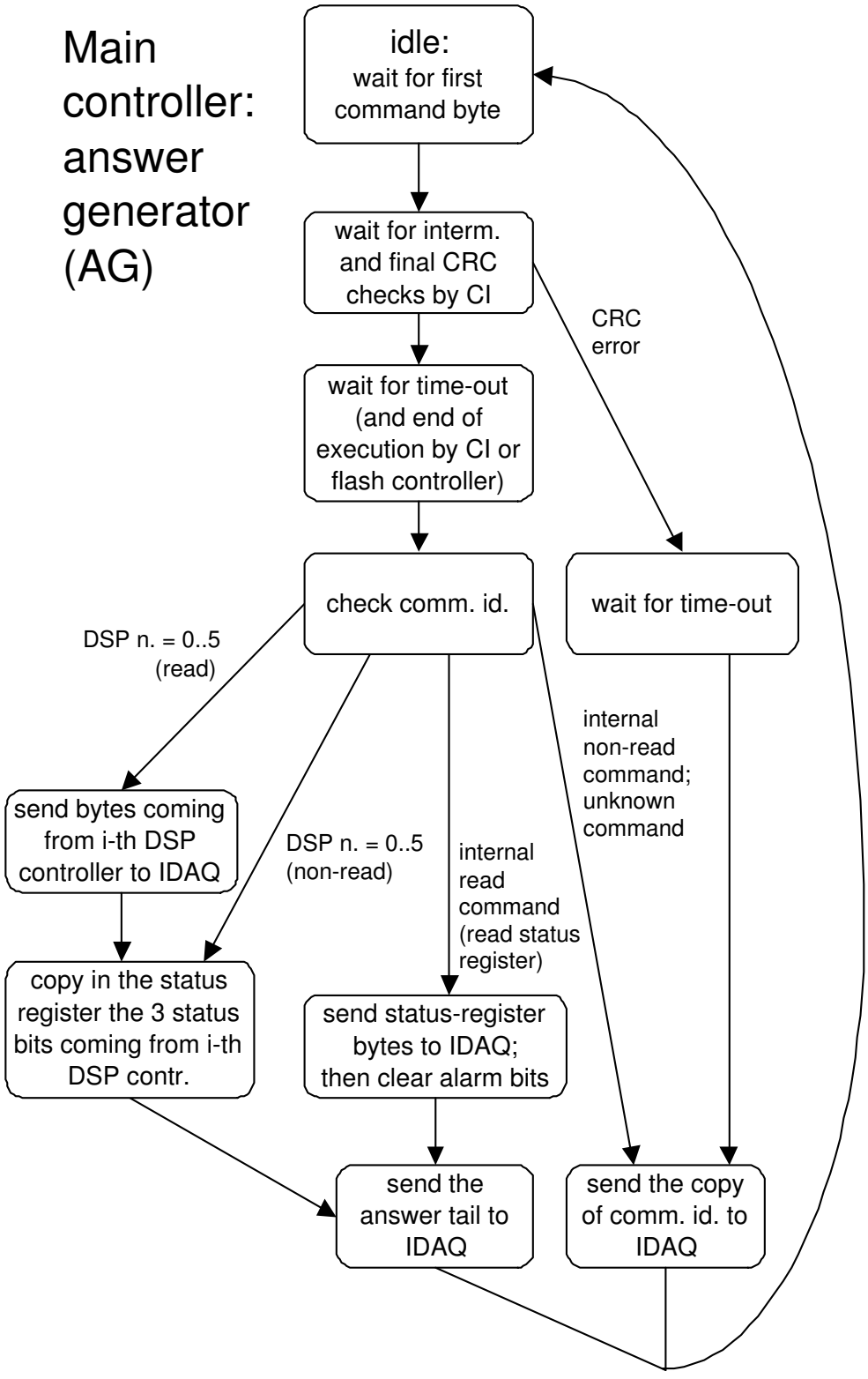

Figure 3.20: main controller: functional diagram of the answer generator unit.

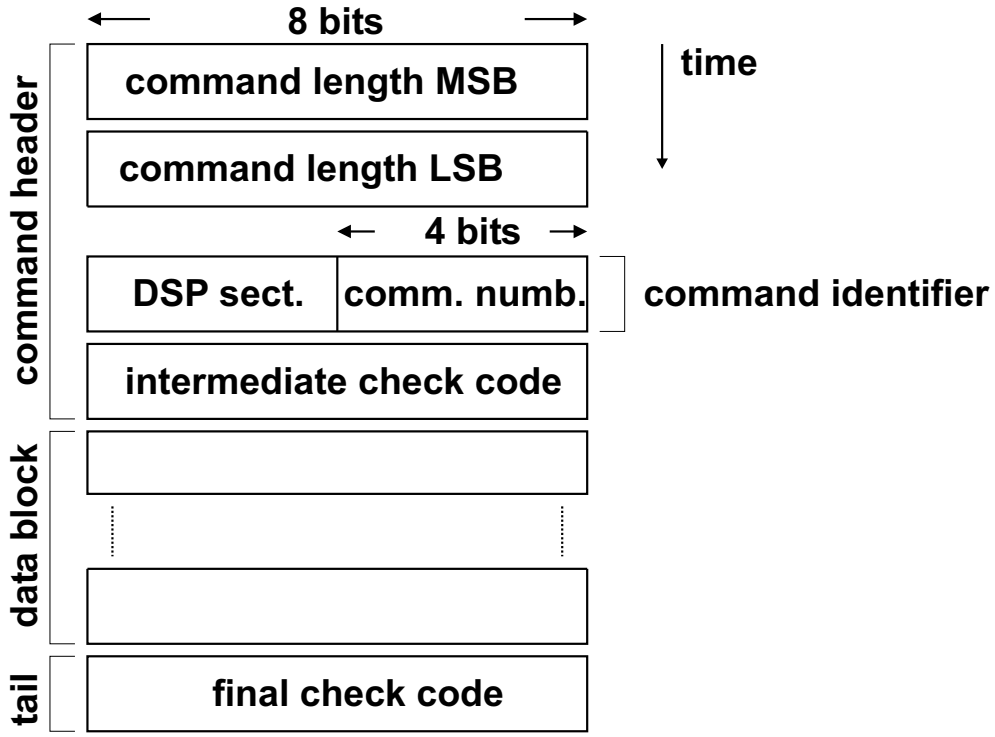

Figure 3.21: general format of a command for the DSP board.

Command identifier		Command name	Command type
DSP section	Command number		
0..5	0	Reset DSP(i)	non-read
0..5	1	Write memory block on DSP(i)	non-read
0..5	7	Transfer memory block from flash to DSP(i)	non-read
0..5	8	Read processed data on DSP(i)	read
0..5	9	Read memory block on DSP(i)	read
7	0	Write mask bits of status register	non-read
7	1	Turn on flash section	non-read
7	2	Turn off flash section	non-read
7	3	General reset	non-read
7	4	Command for flash controller	non-read
7	5	Reset flash controller	non-read
7	8	Read status register	read

Table 3.2: available commands for the DSP board. DSP section = 7 means that the execution involves only the main or flash controller (internal command).

by a simple acknowledge byte, containing the command identifier.

The command interpreter (CI) verifies the consistency of the block of bytes constituting each command by executing two CRC checks: an intermediate one, to verify the

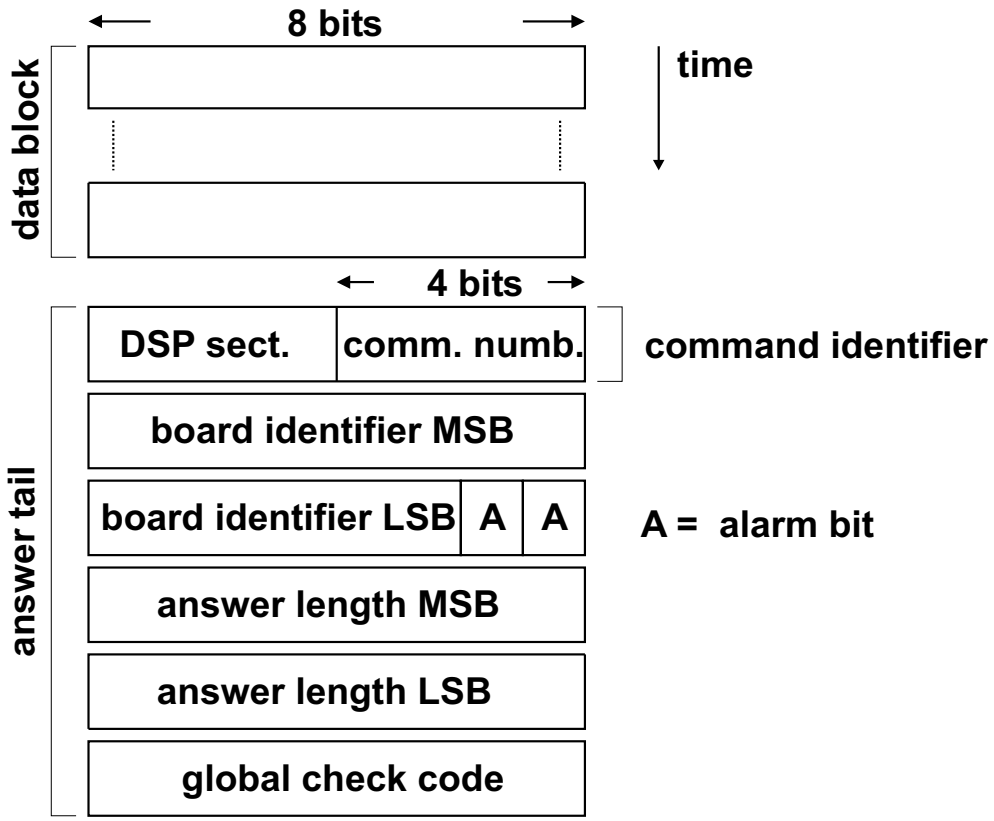

Figure 3.22: general format of the answer to a read command from the DSP board.

command identifier byte, and a final one, to check the integrity of the data block. If one of the two CRC checks fails, or if the command identifier does not represent one of the recognized commands, the corresponding alarm bit in the status register is set to 1 to signal the error condition.

The bytes forming each command are continuously sent by the IDAQ board; the actual end of the stream of bytes is acknowledged by the main controller through a time-out logical block, based on a counter which is started with the first transition on the command data/strobe lines; with the arrival of each new command byte, the counter is reset and started again. When the command stream stops, the counting continues until a predefined maximum value is reached; this causes a time-out pulse to be sent both to the CI and AG units, the former ending its operating cycle, the latter being enabled to prepare and transmit the answer bytes to the IDAQ board and then return to its initial *idle* state.

With this system the main controller can independently acknowledge the actual end of a command stream, and check that the length contained in the command header is consistent with the number of received bytes; if for example the time-out pulse is generated

before the arrival of the final CRC byte, meaning that the command stream has been wrongly interrupted, the corresponding alarm bit (*unknown command*) is set to 1, while the CI and AG units automatically end their operating cycles instead of remaining to wait indefinitely for new bytes.

3.5.1 Flash memories

The flash section is formed by two flash memory chips (Am29LV800B [42]) and a dedicated controller FPGA.

The two memories are managed in parallel: each location of the first memory contains a data byte, while the same location on the second chip stores the corresponding 5-bit check code plus a possible 3-bit end-block code, used to signal the end of the data block to be transferred during memory reading. When the flash controller writes a data byte on a memory location, it creates and writes also the corresponding check code, according to the Hamming standard which, with the subsequent reading of that location, allows the automatic correction of a single bit upset and the detection of a much less probable double bit upset in the data byte. The flash controller can erase a memory area, set the active address for a subsequent reading or writing operation, write a data byte (plus Hamming code) or an end-block code on the active location, read a continuous memory block starting from the active location until an end-block code is found.

During the final qualification of the board, the program code (about 10 kbyte of data) for the DSP chips has been loaded in the flash memories; with standard operation, the code will be transferred from the flash memories to the DSP memories after receiving a dedicated command in the initialization phase of the board. In case that a double bit upset is detected by the flash controller during the data transfer, an error line will signal this alarm condition to the main controller that will set to 1 the corresponding error bit in the status register. Another copy of the program code is stored for safety in a different area of the flash memory and also in the non-volatile memory of the PSCU. In case of detection of a double bit upset, the sequence of commands for the DSP board can be changed to switch to a spare copy of the program code.

The flash memories also contain a set of spare parameters (information on pedestals and noise) for the DSP compression algorithm, referring to a specific calibration cycle performed during the final qualification tests; this set will be possibly transferred to the DSP and used in case of failure of the calibration procedure during the on-flight operation of the apparatus. This emergency solution is feasible since it is expected that

the characteristics of the signal chain, from the silicon detectors to the digital conversion in the FE electronics, will vary only slightly with time; therefore the use of these spare calibration parameters will have the only drawback of slightly worsening the performances of the compression algorithm.

Since the flash memories will have to work only during the initialization of the DSP board, the linear regulators powering the flash section are normally off during board operation; the section is powered only when the main controller receives a dedicated command; after all the required data transfers have been performed, another dedicated command turns off the flash section.

3.6 DSP board: DSP controller FPGA

Within each of the 6 DSP sections of the board, the DSP controller FPGA (fig. 3.23) acts as the external interface of the DSP microcomputer, managing the data transfers to and from the internal memory of the DSP through the IDMA port; the DSP controller also exchanges data with the main controller (commands and answers), receives the control sequence for the FE electronics, generated by the read-out controller, transmits it to the 3 FE subsections, and finally receives the data coming back from them.

The execution of commands and generation of answers is achieved with a similar structure to that of the main controller FPGA, with a command interpreter CI (see fig. 3.24) and an answer generator AG (fig. 3.25). Commands sent by the IDAQ board and involving a specific DSP section are passed by the main controller only to the interested DSP controller.

The data block of a write command contains the IDMA overlay and starting address information followed by the data words to be written; the command interpreter (CI) performs the IDMA overlay and address latches and then starts sequentially writing the following words, as soon as they have been received, until the end of the command stream. From the point of view of the DSP controller, the same operations are required when the data block comes from the flash section instead of the IDAQ board, since both data blocks have the same fixed structure, varying only for the number of words to be written.

For a read command, the data block is composed of overlay, address and the number of words to be read. A dedicated command is foreseen to read the data processed by the DSP for each acquisition (trigger); the length of this data block, variable from event to event and depending on the number of strips which pass the cuts of the compression algorithm, is stored by the DSP in a specific memory location at the end of the compression; the

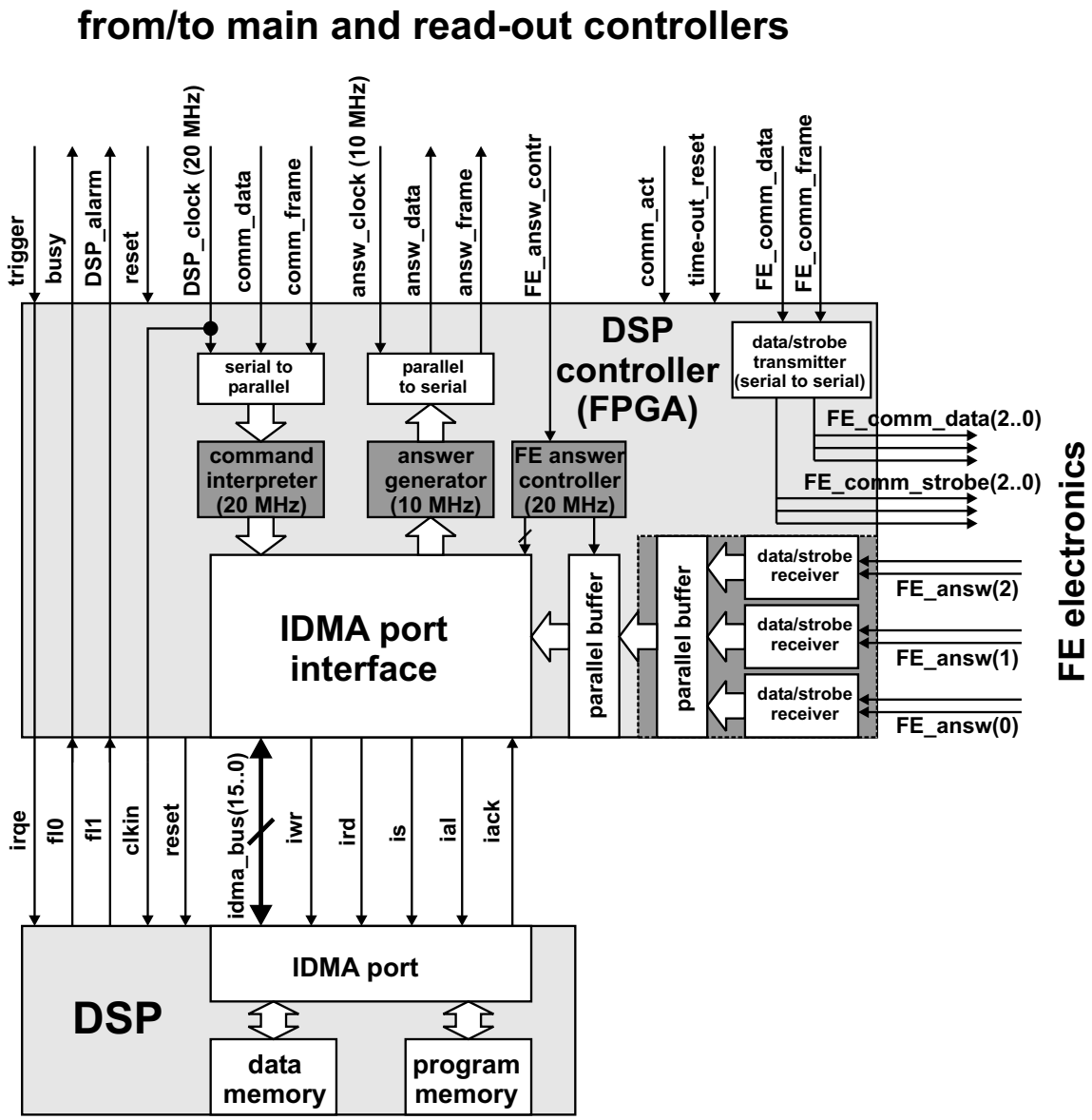

Figure 3.23: functional scheme of the DSP controller FPGA (for clarity only the significant signals toward the DSP are shown).

content of this location is read and used by the DSP controller FPGA to know the actual length of the block of processed data to be transferred toward the main controller.

The *comm_act* line from the main controller FPGA is kept to 1 for all the duration of the command stream; when the stream finishes, the line is reset to 0 and the CI returns to the *idle* state, while the AG begins operation. The *time-out_reset* line brings the time-out pulse (see sect. 3.5) to put the CI in its *idle* state also in the case that the command stream from the IDAQ board uncorrectly stops leaving the serial command receiver and CI machine in some intermediate state, waiting for further data.

The AG performs the required number of IDMA read cycles, simultaneously transmit-

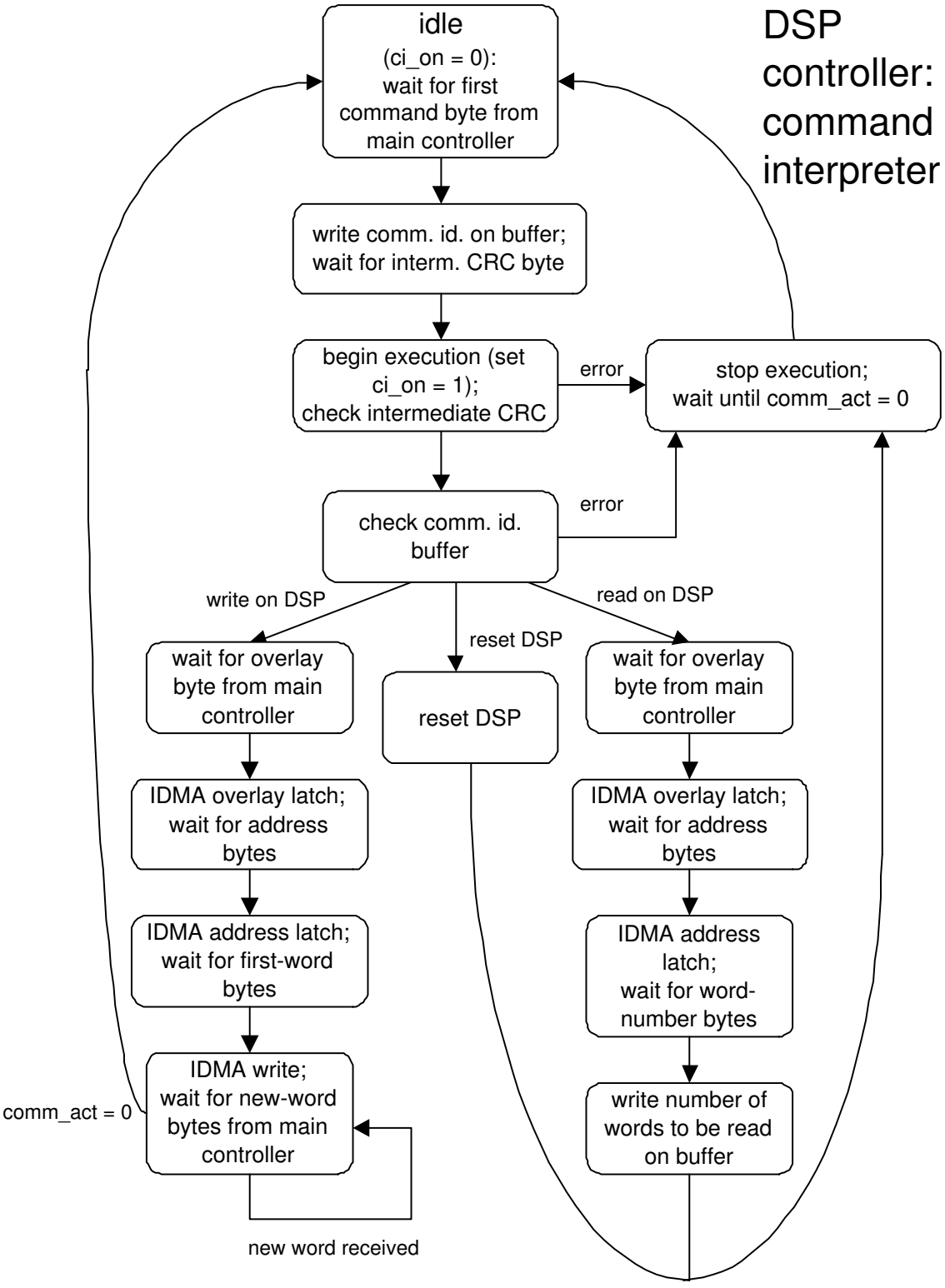

Figure 3.24: DSP controller: functional diagram of the command interpreter unit.

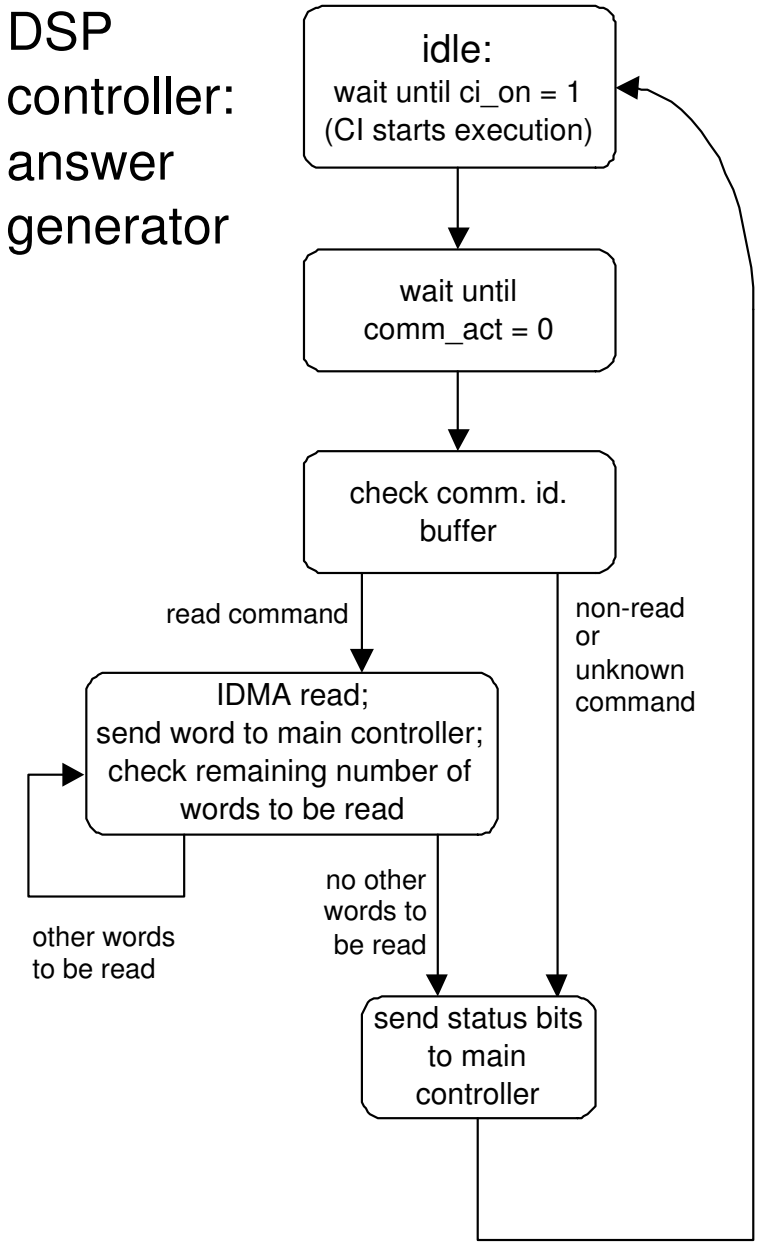

Figure 3.25: DSP controller: functional diagram of the answer generator unit.

ting the words to the main controller. In any case, at the end of the command execution, the AG transmits also three status bits, containing information on error conditions: unknown command, IDMA error during last command execution, IDMA error during FE read-out (trigger) operations since last received command. According to their values, the main controller updates the 3 corresponding alarm bits of the status register (with the possible assertion of the *alarm* line).

3.6.1 IDMA operation

The external interface of the IDMA port is composed of the following signals:

- a bidirectional 16-bit bus *iad(15..0)*, used by the DSP controller both for data input/output and for addressing[15];
- an output acknowledge line (*iack*), to indicate the availability of the DSP to operate a new IDMA address, write or read cycle;
- some input control lines, used by the DSP controller to select the type of transfer: *ial* for an address latch, *iwr*/*ird* for a write/read cycle. Also an auxiliary input (*is*) is present, for the selection of a single DSP in case that the other control lines would be shared by several DSP chips; in the present application each DSP has a dedicated controller FPGA and therefore the *is* signal is simply driven by the OR of *iwr*, *ird* and *ial* (DSP always selected).

An example of a typical sequence of IDMA operations is shown in fig. 3.26; the address latch, corresponding to the *ial* signal pulse, is followed, after an appropriate time interval, by a write cycle, started with the assertion of the *iwr* line and ended with the negation of *iwr*. For proper operation the *iack* line must be low before starting the address or the write/read cycle and must have returned low sufficiently early with respect to the end of the write/read cycle. If the whole set of timing requirements, corresponding to the internal setup and hold times of the IDMA port, are not satisfied, it is not guaranteed that the IDMA cycle has been successfully completed, with the possibility that a wrong address or data value has been written/read.

Therefore the IDMA interface of the DSP controller has been carefully designed and tested to avoid the possibility that a violation of the timing requirements can happen. In particular the value of the time interval T_{iack_low}, elapsing from the assertion of *iwr*/*ird* by the controller to the negation of *iack* by the DSP, depends also on the presence of simultaneous internal operations of the DSP, related to the program execution. The maximum value of T_{iack_low} has been determined from the knowledge of the program code and taken into account in the design of the IDMA interface. Besides, the DSP controller also checks the *iack* level before starting a new IDMA operation (assertion of *ial*/*iwr*/*ird*) and when closing a write/read cycle (negation of *iwr*/*ird*); the detection of an

[15]Note that to complete the transfer to/from the program memory of a 24-bit word, two subsequent write/read cycles are necessary.

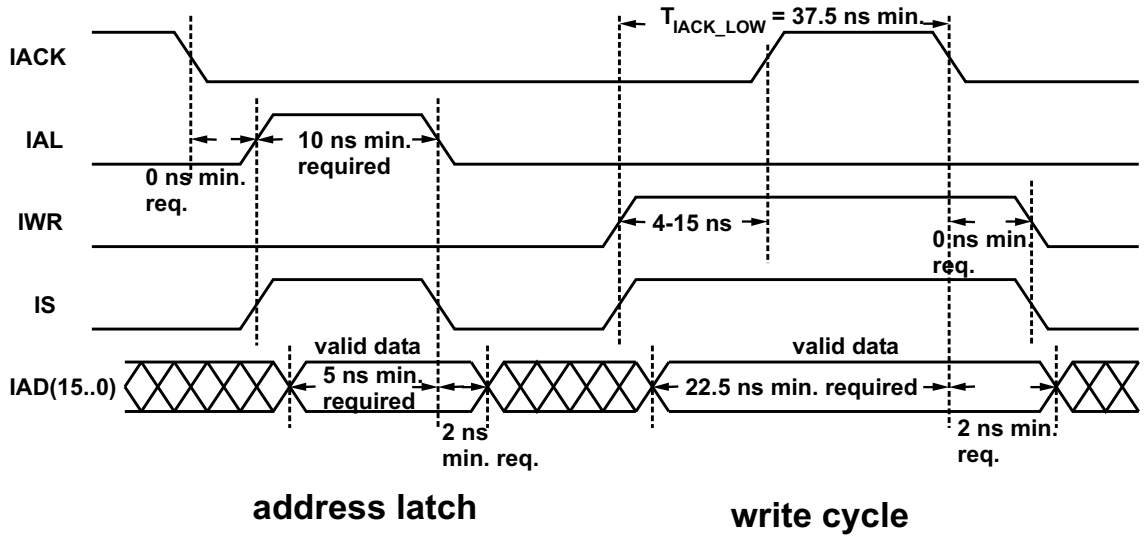

Figure 3.26: IDMA port: example of an address latch followed by a write cycle. Timing requirements to be satisfied by the input signals and timing characteristics of *iack* output refer to the actual configuration of the DSP board (*clkin* input of the DSP set to 20 MHz frequency).

error condition does not stop operation but is registered in a temporary buffer and then sent to the main controller as status bit.

3.7 DSP board: read-out of detector data

The read-out of the detector begins with the reception on the DSP board of a pulse on the trigger input line driven by the active trigger board. The signal reaches a dedicated *read-out controller* FPGA that starts sending two fixed command sequences to each DSP controller through the lines *FE_comm_data* and *FE_answ_contr* (see fig. 3.17 and 3.23), with the same serial protocol shown in fig. 3.18; the frame is given by *FE_comm_frame*, while the clock is the 20 MHz *DSP_clock* generated by the main controller.

Each DSP controller encodes the continuous bit stream, received on the input line *FE_comm_data*, in data/strobe format and sends it (with bit rate of 20 Mpbs) to 3 FE subsections, corresponding to one side of the tracking system. The input *FE_answ_contr* is used internally to control the transfer to the DSP memory of the read-out data transmitted by the FE.

On each FE subsection, the FE controller FPGA (see also sect. 1.8) decodes the data/strobe command stream, generating the *hold* signal for the 8 VA1 chips and then sequentially controlling the acquisition, digital conversion and transmission of the 1024 input channels.

The average delay between the edge of the trigger pulse entering the DSP board and

the assertion of the *hold* signal for the VA1 chips has been fixed (and measured) to be 225 ± 25 ns, taking into account also the contribution of the digital transmission lines between the DSP board and the FE electronics; the unavoidable uncertainty (jitter) is due to the asynchronous sampling of the trigger pulse to generate the synchronous control sequence. This choice, together with the fact that the delay between the time when the particle traverses the apparatus and the generation of the trigger pulse is 750 ± 50 ns, gives a total delay between the particle crossing and the sampling (*hold*) of the signals produced by the silicon detectors of 975 ± 75 ns, which reasonably matches the selected shaping time (nominally 1 μs) for the VA1 input preamplifier.

The internal structure of the FE controller[16] is relatively simple, since it must repeat exactly the same operating cycle for 1024 times (corresponding to the number of strips to be read-out) except for the special cycles 1, 257, 513, 769, for which the output of a new pair[17] of VA1 chips is selected as input for the ADC, and the last (1025^{th} cycle) that brings the controller in its quiescent (*idle*) state again. These special cycles are activated by specific bit sequences in the data/strobe stream sent by the DSP controller; for example, for most of the command sequence the data bit is 1, corresponding to a standard cycle, while a sequence of two 0 bits signals the need to switch to a new VA1 pair during the following cycle.

The answer data stream is formed by 3 12-bit words per cycle, encoded in data/strobe protocol, and is simultaneously sent by the 3 FE controllers to the DSP controller, that sequentially stores it in a dedicated area of the DSP memory. Each of the 3 data/strobe words is decoded and loaded in a parallel 12-bit buffer by an independent logical machine working with the clock reconstructed from the data/strobe signals coming from the corresponding FE subsection.

A dedicated section of the DSP controller (*FE answer controller*), synchronous with *DSP_clock* and operating according to the control sequence on the *FE_answ_contr* input, transfers the 3 parallel 12-bit words in a further buffer stage, where they remain available for a full cycle (40 *DSP_clock* cycles) for the transfer in the DSP memory by means of 3 successive IDMA write cycles. This mechanism for data read-out does not depend on the presence of activity on each of the 3 answer links and therefore can work also in case of failure of any of the 3 subsections.

[16] The detailed structure of the FE controller FPGA is discussed in my Physics Degree thesis [15].
[17] As explained in sect. 1.8, the outputs of the 8 VA1 chips are short-circuited in pairs.

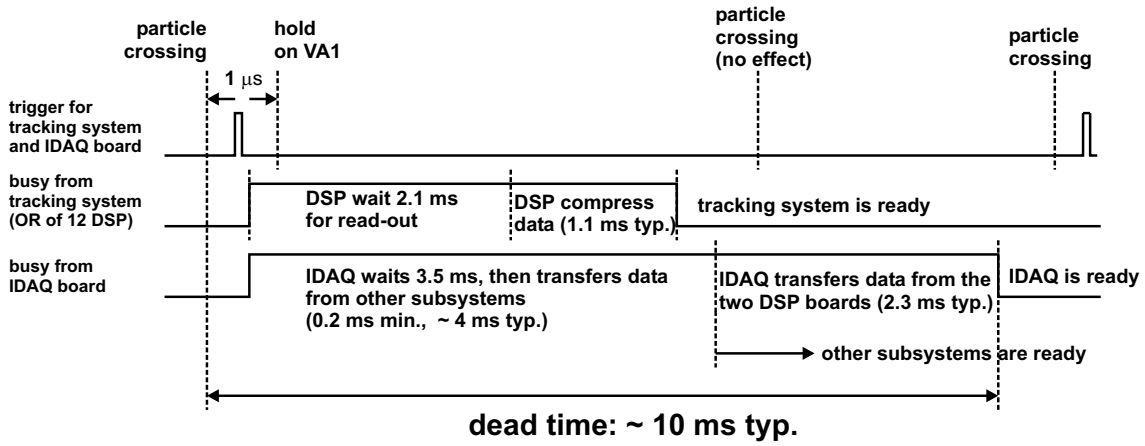

Figure 3.27: basic scheme of the acquisition process for the PAMELA apparatus. Actually a separate *busy* signal is generated by each of the two DSP boards of the tracking system.

3.7.1 Dead time of the PAMELA apparatus

The tracking system contributes to most of the dead time of the PAMELA apparatus, because of the large amount of data produced by its detectors (about 90% of the total data) that have to be read-out and processed.

The IDAQ board, managing the general acquisition procedure, sends the commands to transfer the processed data from the DSP memories of the tracking system to the PSCU after having waited for a predefined time interval (chosen accordingly to the operation of the electronics of the tracking system) and after having managed all the other subsystems.

As mentioned in the discussion of the trigger board in sect. 1.6, the IDAQ board asserts its own *busy* line in response to a trigger pulse, and negates it only after completing the data transfer from all the DSP boards; only at this point the whole apparatus is ready for a new trigger (all *busy* are 0). Thus the dead time of the apparatus is equal to the duration of the *busy* pulse from the IDAQ board, which is typically of the order of 10 ms. The various contributions to the dead time of PAMELA are discussed in what follows (see fig. 3.27 for reference).

On each FE board of the tracking system, the FE controller FPGA works exploiting both edges of the clock signal reconstructed from the data/strobe inputs. Each operating cycle of the FE controller lasts 40 clock transitions, corresponding to $40 \cdot 50$ ns $= 2$ μs at the 20 Mpbs bit rate of the data/strobe command link. The total time necessary for the read-out of the 1024 strips is ≈ 2.05 ms; this is also the read-out time for the whole tracking system, because the 36 FE sections are operated in parallel by the 12 DSP controller FPGA.

Each of the 12 DSP of the tracking system receives the trigger pulse on an edge-sensitive interrupt input (*irqe*); it immediately sets to 1 an output flag (*fl0*) that is reset to 0 only when the DSP has finished the compression of the read-out data and is ready for a new trigger pulse. The OR of the 6 *fl0* signals on each DSP board forms the *busy* output which is sent to the trigger board. The DSP waits a fixed number of *clkin* pulses starting from the *irqe* edge, corresponding to a time interval of 2.1 ms, that is sufficient to let the read-out of the strips be completed, before starting processing the data.

The maximum time employed by a DSP of the tracking system to perform the data compression is 1.6 ms; this means that the compression ends after at most 3.7 ms from the trigger. Therefore the wait time for the IDAQ board has been fixed in such a way that the transfer of processed data from the tracking system to the PSCU does not start before this time interval has elapsed; taking into account also the time necessary to transfer the data from the DSP boards of the other subsystems of PAMELA, the IDAQ board typically begins the read-out of the tracking system after ~ 7.5 ms from the trigger.

The time necessary to transfer the data from the two DSP boards of the tracking system to the PSCU depends mainly on the compression factor achieved by each DSP for the current event, which determines the amount of data to be transferred (at a bit rate of 10 Mpbs); with a compression factor of 20, and considering the whole sequence of commands and answers, this time is 2.3 ms, for a typical overall dead time of ~ 10 ms, as mentioned at the beginning of this discussion.

3.8 Development and qualification of the DSP boards

A first prototype of the DSP board and its control logics was developed and tested during the first half of 2003. After this prototyping phase, a new version with improvements and corrections, both on the board layout and on the logical design, was completed on August 2003; this version has been used for data acquisition during the beam test of the PAMELA apparatus held on September 2003 at the SPS accelerator (CERN [43]) and subsequently integrated within the *technological model* of PAMELA, employed for tests of electric and functional compatibility with the Resurs-DK1 satellite.

Finally a further version of the DSP boards, carrying minimal variations with respect to the technological model, has been developed and then integrated within the final ("flight") model of the PAMELA apparatus, starting from July 2004. In parallel with this last stage of development of the DSP boards, the relays boards and their control logics were designed, tested and qualified.

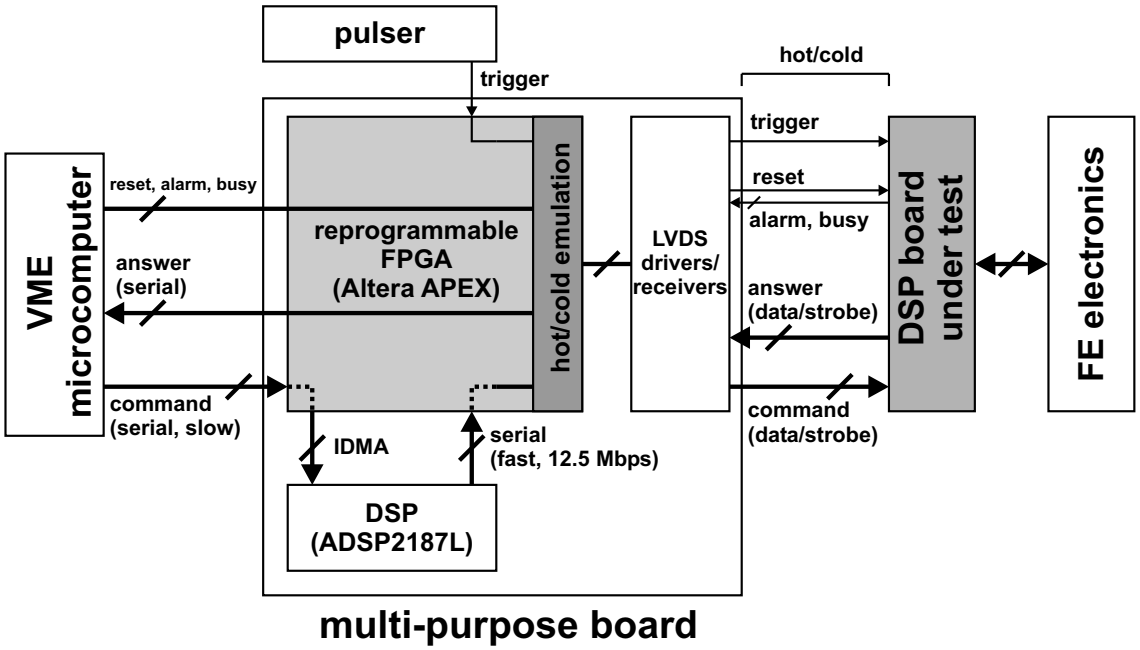

Figure 3.28: basic scheme of the system used to test the DSP boards.

Together with the development of the boards, specific test equipments have been prepared and control programs have been created. The test apparatus for the DSP boards is shown in fig. 3.28: a multi-purpose board, controlled by a VME computer (MVME2431 [44]), reproduces the behaviour of the external devices (IDAQ, IHK and trigger boards).

Command sequences (up to many thousands of commands) are generated in the computer through specifically developed programs and then transmitted to the multi-purpose board, equipped with a DSP microcomputer (ADSP2187L) and a reprogrammable FPGA (Altera APEX20KE [45]) suitably configured. The FPGA chip transfers each command in the DSP memory through the IDMA port, then activates the DSP through a control flag; the DSP sends the command to the FPGA, that performs the data/strobe encoding and selects the *hot* or *cold* link according to its internal configuration. The intermediate DSP stage is necessary because the bit rates available on the serial output of the VME computer are too slow compared with the actual rate achieved by the IDAQ board (16 Mbps).

The answers generated by the DSP board under test are directly passed to the VME computer where they can be analyzed. The test system can also send *trigger* and *reset* signals and monitor the state of *busy* and *alarm* lines; the *trigger* is generated by a pulser with adjustable frequency.

Tests have been performed starting with a preliminary phase of electric checks on

the DSP board, to verify the correctness of the layout and assembling, the voltage levels generated by the linear regulator circuits and the operation of the current limiting stages. Then a first group of functional tests has been done, sending single commands or triggers, verifying the presence of the correct sequences of digital signals on the external pins of the chips and on the connectors (by means of oscilloscope or logical-state analyzer), and checking the content of the answer registered by the VME computer.

Other functional tests have been done through automatic procedures, managed by a test program executed by the VME computer, and reproducing the basic functionalities of the PSCU of PAMELA. The VME computer sends previously prepared sequences of commands, combining different operations on FPGA, DSP and flash devices of the DSP board, and immediately checks the integrity of the answer data by means of the CRC code; besides, the value of the *alarm* line is continuously monitored. The accumulated answer data are then checked for consistency with other specific analysis programs.

Continuous transfers of several hundreds of Mbyte have been achieved with this method, without detecting any anomalous condition, thus qualifying the DSP board against the presence of critical signal paths in the control logics, possibly not identified and removed during the design and simulation phases, and that may sporadically introduce logical errors (upsets) in the signal chains, as seen in sect. 3.2.

With these automatic tests also the complete reliability of the internal memories and of the IDMA ports of the employed DSP chips has been verified; this was necessary because during the prototyping phase a DSP was found showing an uncorrect behaviour of its internal memory: specifically it was impossible to write the value 1 on a certain set of locations. To detect the presence of damaged memory locations and verify the proper functionality of the IDMA port (addressing, writing, reading) the following test procedure has been used:

- set all the bits of all the DSP memory locations to 0, then read and check;
- test the possibility of changing the value of each bit from 0 to 1 and to correctly read this value: write 1 on all the bits of all the locations, then read and check;
- repeat the previous step for the changing of the bit value from 1 to 0;
- test that each single location is correctly addressed when writing or reading: fill the memory with a series of increasing values, then read and check.

Following the usual practice for qualification of electronic systems, the DSP boards have been subjected, normally powered, to several thermal cycles (8 hours period, for a

total of about 50 hours) between $-20°$ C and $+70°$ C, which represent the limits for admitted operating temperatures common to all the employed devices. The stresses caused on the microscopic structures (transistors, resistances etc.) are much greater than the ones expected during the 3 years of operation in space[18]. Components surviving this *burn-in* procedure are safely expected to guarantee a correct operation for the whole duration of the mission. After the thermal cycles, the electric and functional tests have been repeated to verify the absence of damaged components.

The qualification tests have been continued and concluded during the integration of the DSP boards within the PAMELA apparatus, until the complete definition and check of the whole acquisition system of PAMELA and of the procedures managed by the PSCU have been achieved.

3.9 Relays boards

The present section describes the structure and operation of the power supply architecture for the front-end electronics of the tracking system and in particular of the two *relays boards*.

Before entering the detailed description of the relays board, a brief introduction is necessary on the general structure of the power supply system of the PAMELA apparatus. This is organized as a tree structure to distribute the power coming from the satellite to the about 100 independent sections, each equipped with a high-stability voltage regulator and current limiting circuit.

The first stage of the supply system is composed of 3 *intermediate power modules* (IPM), doubled for redundancy, converting the incoming +27 V voltage into +17 V, required for the operation of the DC/DC converters forming the second stage; a total of 32 DC/DC converters (each doubled for redundancy) are present, generating voltages that range from the 3.6 V required by the LVTTL digital boards to the -900 V to operate the PMT of the scintillator detectors.

A further stage, for the OR-ing of the *hot* and *cold* voltages, is inserted between the DC/DC converters and their loads; a typical configuration, in the case of redundancy of the loads, has been discussed in sect. 2.5.

The general control of the IPM and DC/DC converter stages is performed directly by the PSCU, that can independently select the active IPM of each *hot/cold* pair. A

[18] As mentioned in sect. 2.1, the temperature range inside the Resurs satellite is limited between 5° C and 40° C.

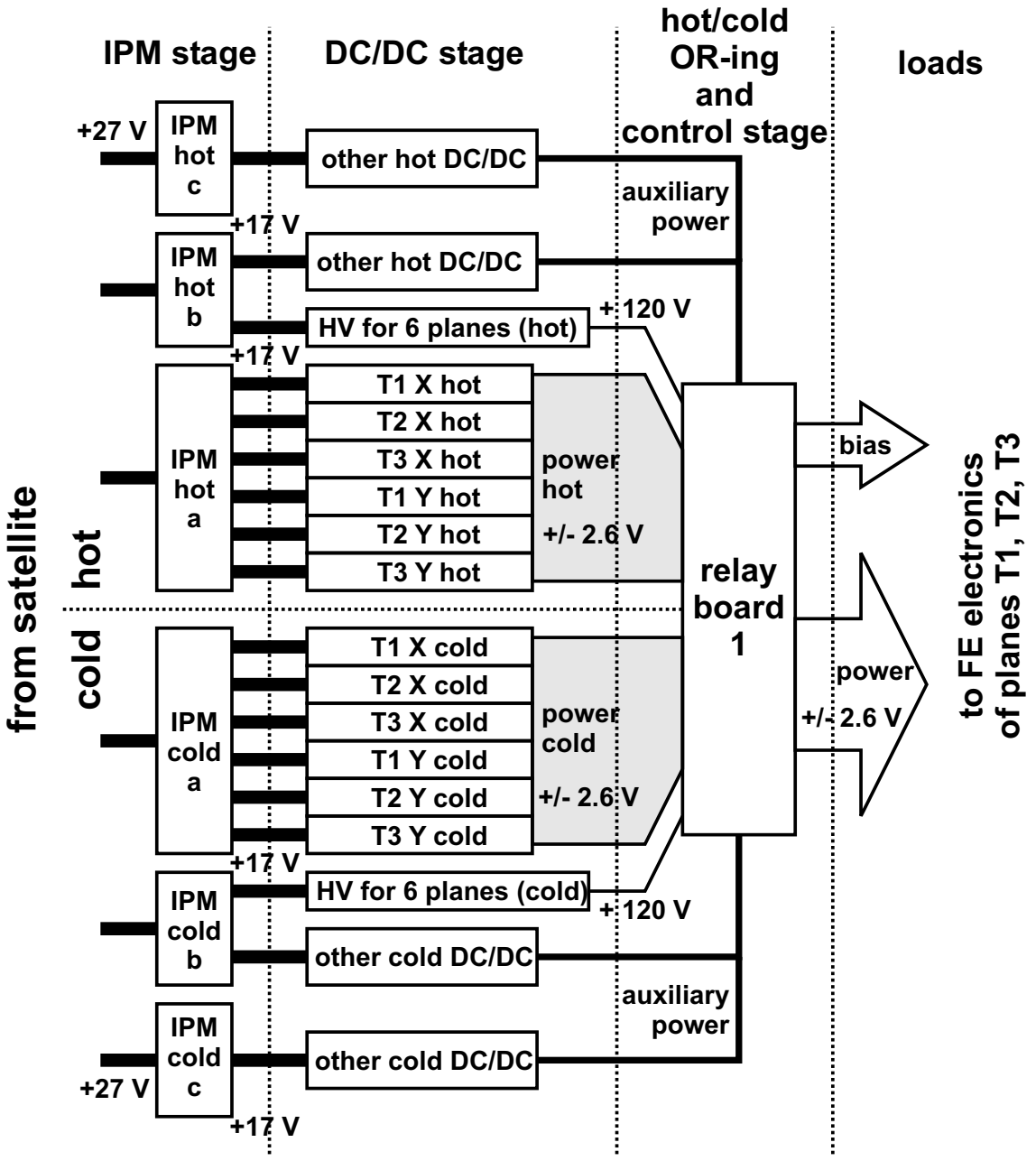

Figure 3.29: scheme of the power supply architecture for half tracking system (planes T1 to T3).

dedicated power control board (also doubled) activates single loads or groups of loads by acting on the OR-ing stages and also on single DC/DC converters.

The power supply for the FE electronics of the tracking system has an analogous structure (fig. 3.29), with the difference that the OR-ing stages are housed on the two dedicated relays boards together with their control system. This choice is due to the relatively high number of independent sections to manage (36) and to the peculiar complications in the design of the OR-ing stage, related to the fact that the power lines for

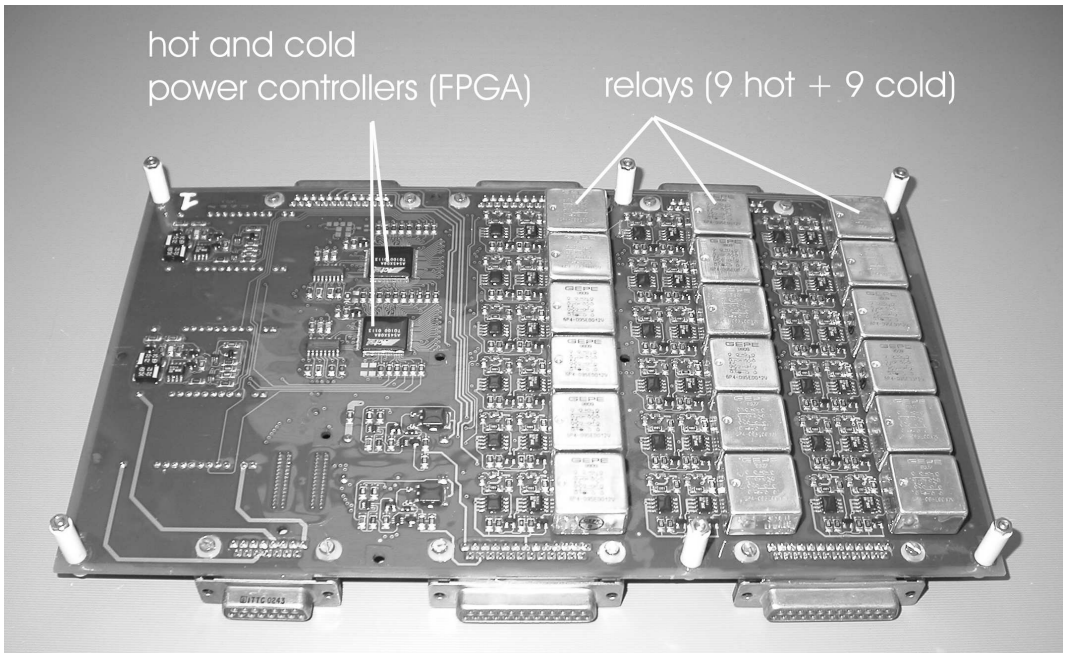

Figure 3.30: one side of a relays board.

the Y views are referred to the bias voltage applied to the corresponding n^+ strips (+80 V above the ground of PAMELA).

Fig. 3.30 shows a picture of a relays board. Each of the two boards is dedicated to half tracking system (3 planes, corresponding to 6 views or equivalently 18 FE sections). Many power lines enter a relays board:

- the low voltages (± 2.6 V with respect to the ground of the X or Y view) for the operation of the FE sections, from 6 (doubled) DC/DC converters, one for each of the views;
- the high voltage (+120 V) for the generation of the bias for 3 planes of the tracking system and also for two sections of the calorimeter (corresponding to half instrument); note that a unique (doubled) DC/DC converter generates the high voltage for the bias of both the tracking system and the calorimeter;
- a set of auxiliary power lines for the internal operation of the board.

As already said, the basic task of the relays board is to perform the OR-ing of the *hot/cold* power and bias lines. Besides, since the unique DC/DC converter for the bias has 10 independent loads (corresponding to the 6 planes of the tracking system plus the 4 sections of the calorimeter), the board is capable of excluding any of these loads in case of permanent failure (e.g. short-circuit).

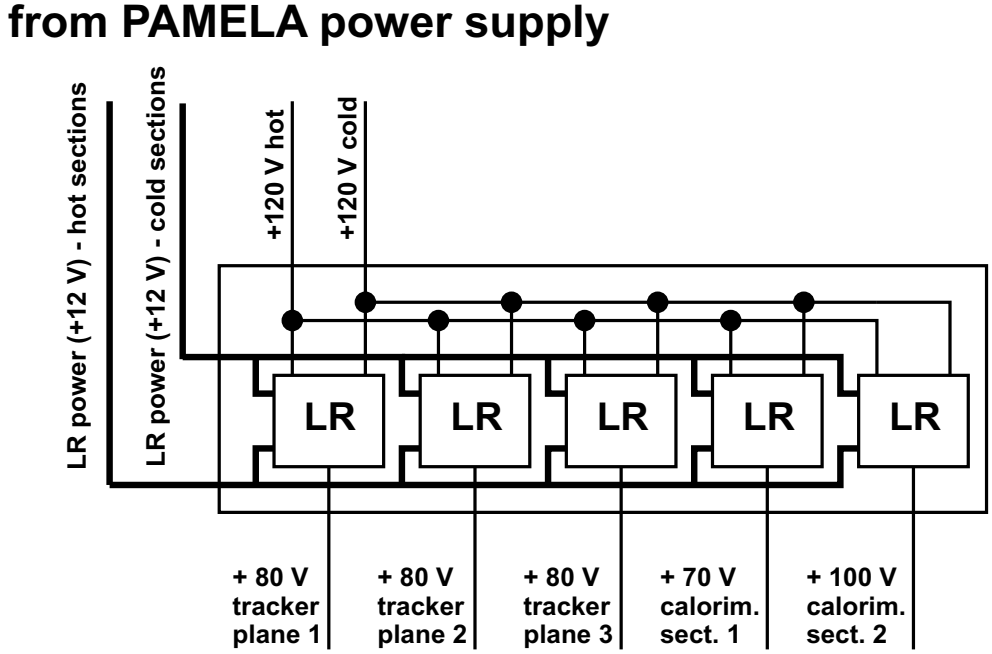

Figure 3.31: relays board: scheme of the bias control section with the S9002 linear regulator (LR) circuits.

The required functionalities are provided by the on-board control logics, composed of two FPGA chips (A54SX08A, [32]), which are managed by the PSCU through the intermediate housekeeping (IHK) board.

The general problem when doing the OR-ing operation on the *hot* and *cold* power lines is to guarantee the independence of the two power supplies, avoiding the presence of a direct current path between their outputs. The typical solution is to employ an active circuit (commonly referred to as a *protection stage*), exploiting the switching characteristics of diodes or transistors, which in its simplest form is exemplified by the diode stage discussed in sect. 3.3; in the present case three different methods have been used, as described below.

The bias voltages are generated by dedicated linear regulator circuits (S9002 [39], see fig. 3.31); an S9002 circuit is split into two symmetric halves, each producing a regulated output which is internally OR-ed through a diode protection stage. The on/off state of the *hot* or *cold* half of the S9002 depends on the level of an input LVTTL signal directly driven by the control logics.

The situation is more complex for the power lines to the FE sections (fig. 3.32). The lines for the FE electronics of X sides are referred to the voltage applied to the corresponding p^+ strips, coinciding with the ground of PAMELA; these lines will be indicated as *low voltage* (LV) power. On the other hand, as already mentioned, the lines

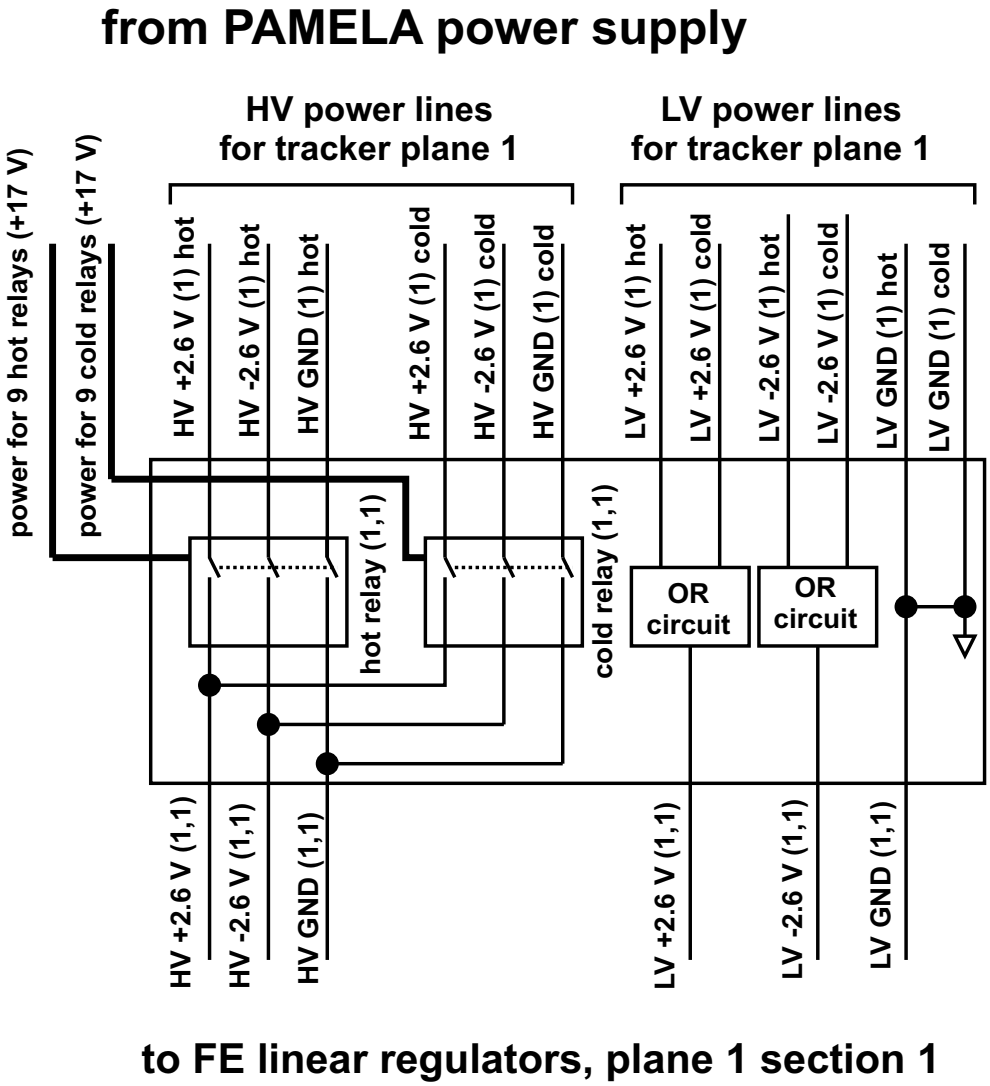

Figure 3.32: relays board: scheme of the power control for one section of FE electronics.

for the Y-side electronics are referred to +80 V with respect to the ground of PAMELA and will be indicated as *high voltage* (HV) power.

For the LV power, the OR-ing circuit exploits two MOS switches, inserted on the *hot* and *cold* lines respectively, that are automatically driven open when the corresponding input voltage is not applied (the DC/DC converter is off).

For the HV power, the OR-ing must also avoid a direct current path between the HV lines and the rest of the board, which is referred to the ground of PAMELA. The problem has been solved by using a *latching relay* (GP4-095E00 [46]) for each triplets of HV power lines +2.6 V, −2.6 V, GND (all referred to +80 V).

This latching relay is an electro-mechanical switch formed by two metallic coils inductively coupled to a ferromagnetic armature which carries 4 independent electrical contacts;

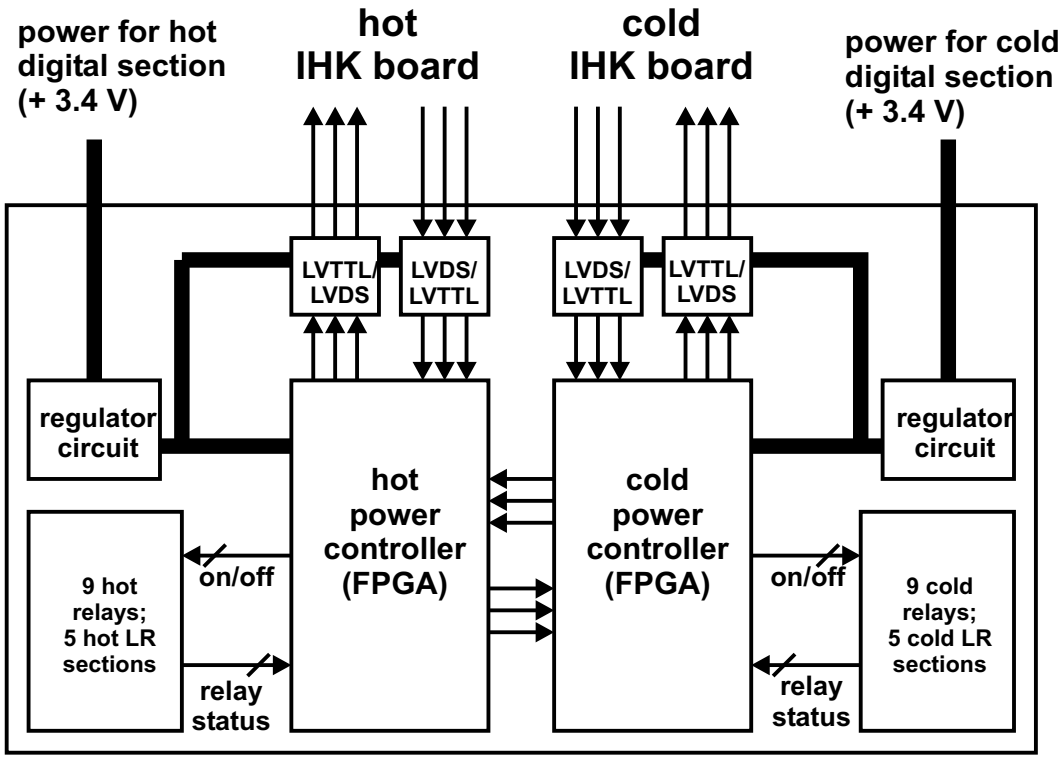

Figure 3.33: relays board: scheme of the digital section.

coils and armature are electrically insulated from each other with layers of material of high dielectric strength. When a current pulse of the proper sign flows through one of the coils, the induced magnetic field causes the displacement of the armature, that remains in this position until a current pulse is injected in the other coil; the two independent coils allow to switch between two different positions of the armature, corresponding to the closing or opening of all the contacts. The current pulse on each of the two coils is generated by a dedicated circuit in response to an LVTTL signal from the control logic.

For each FE section two relays operate on the *hot* and *cold* triplets respectively; after the relay stage the power lines are short-circuited in pairs. An advantage of this solution is that each relay maintains the selected state (on/off) until a new current pulse flows through one of the coils, with a significant power saving.

The setting and monitoring of the on/off state of relays and bias regulators is managed by a *power controller* FPGA chip, doubled for redundancy (see fig. 3.33). The *hot* FPGA directly operates all the *hot* power and bias sections, formed by 9 relays and the *hot* halves of the 5 linear regulators. This controller can also act on the *cold* sections through an internal bidirectional serial link with the *cold* FPGA.

The controller receives commands from the IHK board and, if requested, can send

back data on the status of the relays and bias sections, through LVDS links with a serial 3-lines protocol analogous to that shown in fig. 3.18. A *write* command contains the pattern of on/off states to be implemented on the 3 relays and regulators for one plane of the tracking system, or on the two regulators for the calorimeter. A *read* command implies an answer containing the actual current pattern of on/off states. The command code contains the indication about the addressed FPGA; in the case of a command for the other FPGA, the controller simply activates the internal bidirectional serial link and waits for the end of operation by the other FPGA before returning to its *idle* state.

Chapter 4

The geometric factor of PAMELA

The present chapter is dedicated to the description of a method that I have developed for the determination of the geometric factor of the PAMELA apparatus, conceived to obtain a sufficiently small calculation error within a reasonably short computation time.

A precise knowledge of the geometric factor and of the other terms that contribute to the detection efficiency of the PAMELA apparatus is of fundamental importance for the measurement of the absolute intensities of primary cosmic rays on the basis of the actually detected intensities; this general framework is discussed in sect. 4.1.

In sect. 4.2 and 4.3 the characteristics of the calculation method are explained and justified, with particular attention for the optimizations employed to minimize the computation error. Results and comments are in sect. 4.4.

4.1 Measurement of absolute differential intensities of cosmic rays with PAMELA

The fundamental purpose of the PAMELA experiment is to study the energy spectra of the antiproton $\overline{\mathrm{p}}$ and positron e^+ components in the primary cosmic rays outside the atmosphere, by means of the identification of the incoming particle and by the precise determination of its direction of incidence and momentum modulus $|\vec{p}|$, through the direct measurement of the magnetic rigidity $\rho = |\vec{p}|/\mathrm{z}\, e$[1] of the particle. With this same method the apparatus will also be able to study the spectra of protons, electrons and light nuclei.

The quantity that characterizes the flux of cosmic rays in a region of space is the *absolute differential intensity*, usually defined, for a given species, as the number of particles incident at a given position and time, with rigidity ρ and direction of incidence

[1]Note that ρ can assume both positive and negative values, depending on the charge z of the particle (expressed in electronic charges e).

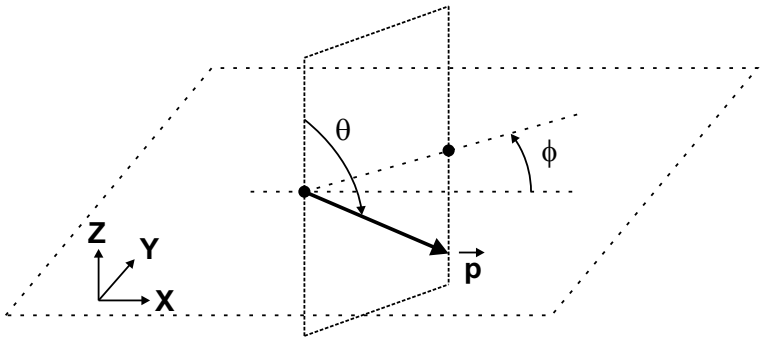

Figure 4.1: definition of the polar angles θ and ϕ characterizing the direction of incidence of a particle with momentum $\vec{p}$.

specified by angles θ[2] and ϕ (see fig. 4.1), per unit rigidity, solid angle, area normal to the direction of incidence and time:

$$\mathrm{J} = \frac{\mathrm{dN}}{\mathrm{d}|\rho|\ \mathrm{d}\Omega\ \mathrm{dS}_{\perp}\ \mathrm{dt}} \tag{4.1}$$

which is commonly expressed in particles per $(\mathrm{GV/c}) \cdot \mathrm{sr} \cdot \mathrm{cm}^2 \cdot \mathrm{s}$.

For a specific experimental apparatus, such as PAMELA, the quantity that is actually measured, in an observation time interval $(0, \Delta\mathrm{t})$ and for a range of rigidities, characterized by an average value ρ, is the number of particles of a given species that are detected, $\mathrm{N}_{\mathrm{det}}(\rho)$, where *detected* here means identified and completely characterized by the apparatus.

It can be easily shown[3] that, in the hypothesis of uniform and isotropic flux[4], the average value of J in the observation time and in the rigidity range considered is related to $\mathrm{N}_{\mathrm{det}}(\rho)$ by the formula:

$$< \mathrm{J} > = < \frac{\mathrm{dN}}{\mathrm{d}|\rho|\ \mathrm{d}\Omega\ \mathrm{dS}_{\perp}\ \mathrm{dt}} > = \frac{\mathrm{N}_{\mathrm{det}}(\rho)}{\mathrm{G}(\rho) \cdot \epsilon \cdot \Delta\mathrm{t} \cdot \Delta|\rho|} \tag{4.2}$$

where:

- $\mathrm{G}(\rho)$ is the *geometric factor* of the apparatus, usually expressed in $\mathrm{cm}^2 \cdot \mathrm{sr}$ and defined as the 4-dimensional integral:

$$\mathrm{G}(\rho) = \int_{\mathrm{S}_0,\Omega} \mathrm{d}\Omega\ \mathrm{dS}\ |\cos\theta|\ \mathrm{f}\,(\mathrm{x}, \mathrm{y}, \theta, \phi; \rho) \tag{4.3}$$

[2]Throughout this work, the zenith angle θ characterizing the direction of the particle momentum $\vec{p}$ is used, instead of the more common angle $\theta' = \pi - \theta$.

[3]For a detailed treatment of the subject see for example the classical paper of Sullivan [47].

[4]The flux of cosmic rays entering the PAMELA apparatus during its orbital motion can clearly be considered uniform over the small volume of the apparatus; the flux entering the acceptance window from the zenith part of the sky is also expected to be substantially isotropic. On the other hand, a non-negligible dependence of the intensity on the time is foreseen, due to the effect of the geomagnetic field (that changes along the orbit) on the particles of lower energies.

where:

- S_0 is a reference plane of the apparatus, orthogonal to the Z axis and positioned at height $z = z_0$[5];
- Ω is the total solid angle;
- the weight function f is 1 or 0 whether or not the particle crossing the reference plane at coordinates (x,y), with direction of incidence (θ, ϕ) and rigidity ρ, satisfies the acceptance requirements of the apparatus;

- ϵ represents the various contributions to the *detection efficiency* of the apparatus, and depends in general both on ρ and the particle species.

These two terms characterize the detection capabilities of PAMELA and are discussed in what follows.

Note that the definition of $G(\rho)$ reported in (4.3), including an explicit dependence on the rigidity ρ of the incident particle, is an extension of the usual "simple" definition of geometric factor:

$$G_{geom} = \int_{S_0,\Omega} d\Omega \; dS \; |\cos\theta| \; f\,(x, y, \theta, \phi) \tag{4.4}$$

as a constant depending only on the geometric characteristics of the apparatus. This extension is necessary to properly take into account the fact that in presence of a magnetic field, like in the case of PAMELA, the acceptance window varies with the rigidity of the particle. In fact particles with smaller $|\rho|$ are more deflected by the field toward the lateral walls of the magnetic cavity, where they are lost before reaching the lower planes of the tracking system (see fig. 4.2). As a consequence, it is expected that the effective acceptance window (and hence the geometric factor) decreases for smaller $|\rho|$ and that $G(\rho) = G_{geom}$ only in the limit case of very high $|\rho|$ (practically rectilinear trajectories).

It can be observed here that the PAMELA apparatus has been designed to maximize the product $G \cdot \epsilon$, within the limits of the various constraints (mass, volume, cost, power etc.) imposed by the characteristics of the mission, in such a way to increase the number of detected events.

[5]For the calculation of G discussed in the present chapter, the PAMELA coordinate system has been adopted, as introduced in sect. 1.2, with its origin at the centre of the magnetic cavity of the spectrometer and axes oriented as shown in fig. 1.3. In particular the reference plane S_0 has been fixed in correspondence of the upper face of the magnetic cavity of spectrometer, at a nominal height $z_0 = 21.83$ cm. It must be also noted that for a fixed experimental configuration the value of the geometric factor does not depend on the choice of S_0 [47].

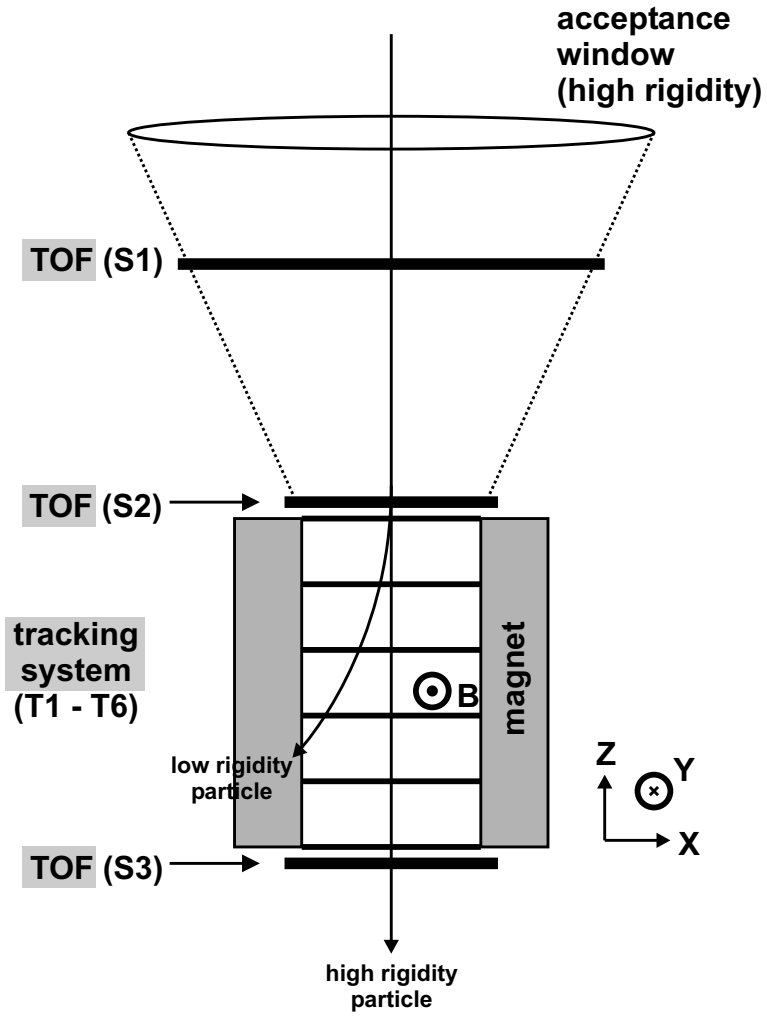

Figure 4.2: the motion of a particle with high positive rigidity ρ traversing the PAMELA apparatus is not greatly affected by the magnetic field within the spectrometer cavity; on the other hand, a particle with significantly smaller rigidity is deflected toward the wall of the magnet. As a consequence, the effective acceptance window tends to decrease for smaller values of ρ.

The precise knowledge of both the geometric factor and the detection efficiency of PAMELA is critical for the determination of the absolute differential flux on the basis of the measured number of detected particles.

4.1.1 Geometric factor

The geometric acceptance requirements of PAMELA, corresponding to events that the apparatus can detect (with detection efficiency ϵ, as discussed in sect. 4.1.2), can be specified as follows:

- the pattern of TOF scintillator layers crossed by the particle must be compatible with the trigger configuration selected; as mentioned in sect. 1.6, the trigger pattern

normally employed requires an event signal from at least one of the two layers of scintillator detectors forming each of the 3 TOF planes S1, S2, S3:
(S11 OR S12) AND (S21 OR S22) AND (S31 OR S32);

- the particle must enter the upper face of the magnetic cavity and exit from its lower face, without crossing the lateral walls of the cavity and hence compromising the rigidity measurement operated by the spectrometer;

- the particle must cross all the 6 detectors of the tracking system (T1 to T6) within their sensitive areas, defined for each plane by the total area of the array of the 6 Si sensors.

The situation is exemplified in fig. 4.3, where for clarity only the constraints posed by the magnetic cavity of the spectrometer are considered, in the limit of high rigidities, for which the trajectories are not affected by the magnetic field and are well described as straight lines: a particle (case 1 in figure) that enters the cavity from the upper face and exits on the lower face satisfies the acceptance requirements; another particle (case 2) that instead exits from a lateral wall of the cavity does not satisfy the requirements. In general the presence of the magnetic field both inside and outside the cavity introduces a dependence of the geometric factor (eq. 4.3) on the initial rigidity of the particle, as mentioned when discussing the definition of $G(\rho)$.

It can be also observed that the geometric factor for a given rigidity ρ, taking into account only the effects of the geometry and of the magnetic field of PAMELA, does not depend on the particle species. In fact the trajectory of a charged particle in an arbitrary magnetic field $\vec{B}(\vec{r})$ is uniquely determined once specified, in the reference plane S_0, the initial rigidity ρ, the crossing coordinates (x,y) and the direction of incidence (θ, ϕ); this can be seen by considering the expression of the curvature k of the trajectory at a point $\vec{r}$, in the plane orthogonal to the direction of $\vec{B}$ (already given in sect. 1.7, eq. (1.3)):

$$k(\vec{r}) \approx 0.3 \cdot |\vec{B}(\vec{r})| \cdot \frac{z}{|\vec{p}(\vec{r})|} \propto \frac{|\vec{B}(\vec{r})|}{\rho(\vec{r})} \tag{4.5}$$

and observing that k at any point depends only on the local value of $\vec{B}$ and ρ.

Note also that $|\vec{p}|$, and hence $\rho = |\vec{p}|/z\,e$, is a constant of the motion since the magnetic force, orthogonal to $\vec{p}$, does not perform any work on the particle. Clearly this is exactly true only neglecting the energy losses of the particle in the thin detector layers of the TOF and tracking system; the correction coefficient to be applied to $G(\rho)$ to take into account these energy losses can be included in the general efficiency term ϵ, as discussed below.

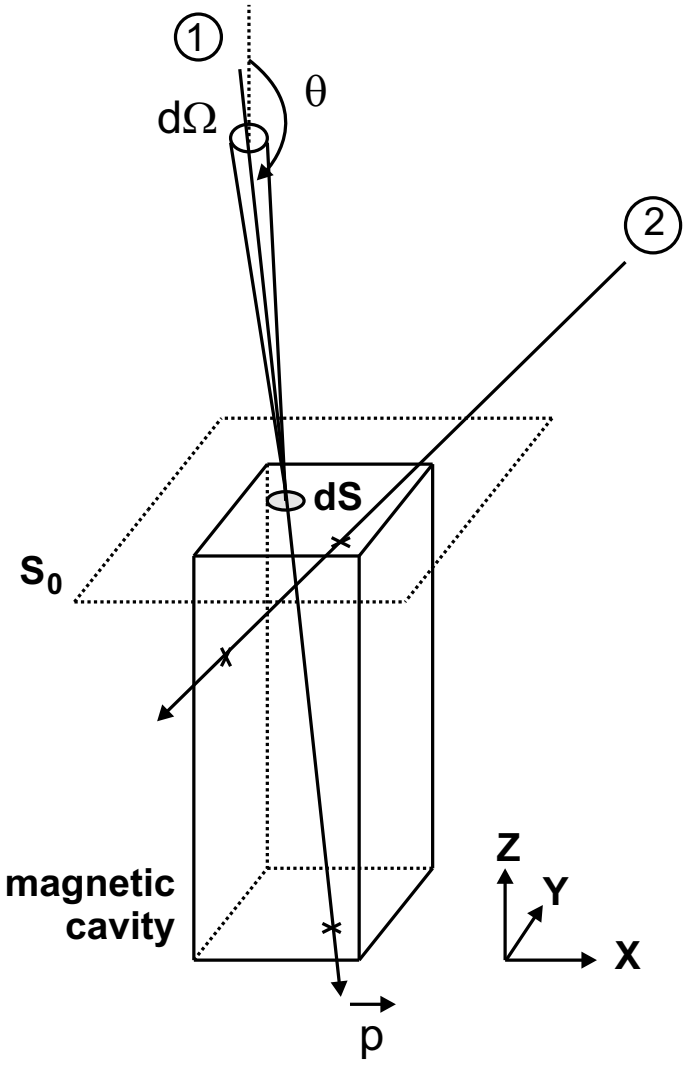

Figure 4.3: geometric acceptance of the magnetic cavity of PAMELA, in the case of high rigidities (rectilinear trajectories): particles crossing the reference plane S_0 (upper face of the cavity) in the element dS and within the solid angle $d\Omega$ (case 1) contribute to the geometric factor integral with a term $d\Omega\ dS\ |\cos\theta|$; particles exiting from the lateral walls of the cavity (case 2) are outside the acceptance domain and do not contribute to the geometric factor.

4.1.2 Detection efficiency

As mentioned in the derivation of equation (4.1), for a given species and rigidity only a fraction ϵ of the incident particles, satisfying the geometric acceptance requirements of PAMELA, are actually identified and fully characterized (*detected*) by the apparatus. The other particles can be lost for several reasons:

- they interact within the traversed thin detector layers (TOF and tracking system) where they can loose part of their initial energy, be deflected or even absorbed; the most significant processes are multiple scattering, energy loss by ionization, energy

radiation by bremsstrahlung for $e^{\pm}$, annihilation for antiprotons;

- the apparatus is not ready to accept a new event because it is performing the read-out of the previous one (the apparatus is *dead*);
- the particle is not completely identified and characterized, even if the apparatus is ready (*live*), because of the intrinsic detection inefficiencies of the various detectors.

Assuming that these contributions do not depend from each other, they can be summarized in the expression:

$$\epsilon = \epsilon_{\text{int}} \cdot \epsilon_{\text{live}} \cdot \epsilon_{\text{eff}} \tag{4.6}$$

where ϵ_{int}, ϵ_{live}, ϵ_{eff} are defined as the fraction of particles, entering the geometric acceptance of PAMELA, that respectively are not lost for the effects of the interactions with the crossed materials, that are incident during the *live time* of the apparatus, that are correctly detected by the apparatus.[6]

The coefficient ϵ_{eff}, representing the intrinsic detection efficiency of PAMELA, is determined by the performances of the various detectors, in particular by the efficiency of:

- the TOF system for the generation of the trigger signal (ϵ_{trigg}) and for the identification of downward-going particles (ϵ_{down});
- the tracking system in the determination of the rigidity and of the sign of the charge for the discrimination between antiparticle and corresponding particle (ϵ_{track});
- the calorimeter in the discrimination of $\overline{\text{p}}$ from electrons and e^+ from protons (ϵ_{calo});

or synthetically, assuming also in this case that the single contributions are all independent: $\epsilon_{\text{eff}} = \epsilon_{\text{trigg}} \cdot \epsilon_{\text{down}} \cdot \epsilon_{\text{track}} \cdot \epsilon_{\text{calo}}$.

All the coefficients defined above will be separately determined in the course of the analysis work which is foreseen on the data collected by the PAMELA apparatus during its flight, integrating the information given by all the various detectors and developing algorithms to proper handle these data. For the present work, which is more oriented to study the experimental characteristics of the apparatus from a "hardware" point of view, the attention is focused on the terms ϵ_{trigg} and ϵ_{live}. A method for the measurement of the term ϵ_{trigg} is described in chapter 5, while the term ϵ_{live} is briefly discussed in what follows.

[6] Note that in general ϵ_{int}, ϵ_{live}, ϵ_{eff} can depend on both the particle species and rigidity ρ; here the corresponding indexes are omitted for simplicity.

The coefficient ϵ_{live} will be obtained by means of automatic measurements performed by the trigger board (see also sect. 1.6 and 3.7.1). The board, with dedicated counters, measures the dead and live times associated with each single acquisition: for the event j, the board registers the dead time τ_{j} elapsed between the instant of the internal generation of the trigger signal and the instant when all the *busy* inputs from the various subsystems are released, meaning that the apparatus is ready for a new acquisition cycle. The board also registers the subsequent live-time interval (until a new trigger signal is generated) and the total number of triggers since the last initialization of the system.

Assume for example that during the observation time Δt, a number N of particles of given species and rigidity enters the apparatus within the acceptance window, and that M_{tr} is the total number of triggers registered during Δt, including also triggers generated by other particles. In the hypothesis of a constant probability per unit time that a particle of the considered sample crosses PAMELA, the average number of particles, that are lost from the original sample N because the apparatus is dead, is proportional to the ratio between the dead time and the observation time:

$$\text{N}_{\text{lost}} = \text{N} \cdot \frac{1}{\Delta\text{t}} \cdot \sum_{\text{j}=1}^{\text{M}_{\text{tr}}} \tau_{\text{j}} \tag{4.7}$$

On the other hand, N_{lost} is the difference between N and the mean number N_{live} of particles incident during the live time of the apparatus: $\text{N}_{\text{lost}} = \text{N} - \text{N}_{\text{live}}$. Solving for N gives:

$$\text{N} = \frac{\text{N}_{\text{live}}}{1 - \frac{1}{\Delta\text{t}} \cdot \sum_{\text{j}=1}^{\text{M}_{\text{tr}}} \tau_{\text{j}}} \tag{4.8}$$

and the ratio:

$$\frac{\text{N}_{\text{live}}}{\text{N}} = 1 - \frac{1}{\Delta\text{t}} \cdot \sum_{\text{j}=1}^{\text{M}_{\text{tr}}} \tau_{\text{j}} \tag{4.9}$$

gives the average value of ϵ_{live}.

The dead time varies in general from event to event, mainly because the different characteristics of each incident particle imply different signal patterns in the various detectors and a varying amount of processed data that have to be transferred from the DSP boards of the detectors to the central control unit of PAMELA (PSCU); besides, the parameters of the data processing algorithms can be varied during the operation of the apparatus.

Given the typical value for the dead time of $\tau = 10$ ms and the expected mean trigger rate f = 12 Hz, it can be concluded that the typical average value for ϵ_{live} is:

$$1 - \frac{1}{\Delta\text{t}} \cdot \sum_{\text{j}=1}^{\text{f}\cdot\Delta\text{t}} \tau = 1 - \text{f} \cdot \tau = 88\% \tag{4.10}$$

4.2 Geometric factor: calculation of the integrand function

The geometric factor as a function of the rigidity ρ of the incident particle, already introduced in equation (4.3), is given by the 4-dimensional integral:

$$\mathrm{G}(\rho) = \int_{\mathrm{S_0}} \mathrm{dx\ dy} \int_0^{2\pi} \mathrm{d}\phi \int_{\frac{\pi}{2}}^{\pi} \mathrm{d}\theta\ \sin\theta\,|\cos\theta|\,\mathrm{f}\,(\mathrm{x},\mathrm{y},\theta,\phi;\rho) \tag{4.11}$$

with:

- the reference plane $\mathrm{S_0}$ chosen in correspondence of the upper face of the magnetic cavity;
- the angular domain of integration limited to the downward hemisphere, characterized by $\pi/2 < \theta < \pi$, since the events of interest for PAMELA are only the downward-going particles coming from the external space;
- the weight function f equal to 1 or 0 depending whether the trajectory, uniquely defined by the set $(\mathrm{x},\mathrm{y},\theta,\phi;\rho)$, satisfies the acceptance requirements.

In the calculation discussed here these acceptance requirements are represented by the selection cuts described below and illustrated in fig. 4.4 (see also table 4.1 for a detailed list of all the geometric quantities involved in the calculation):

- for each pair of layers of the TOF system: the trajectory must cross at least one of the two rectangles defining the layers;[7]
- the trajectory must traverse the parallelepiped defining the magnetic cavity of the spectrometer by crossing its upper and lower faces (rectangles), respectively defined by the maximum and minimum heights of the magnetic tower[8], without escaping from the lateral walls of the cavity[9]; this automatically assures that the 4 intermediate planes (T2 to T5) of the tracking system are crossed by the particle, because the array of 6 sensors forming a complete plane is larger than the cross section of the cavity;
- the trajectory crosses the rectangles defining the detectors T1 and T6. Note that these two planes are positioned slightly outside the magnetic cavity (about 0.4 cm apart along Z).

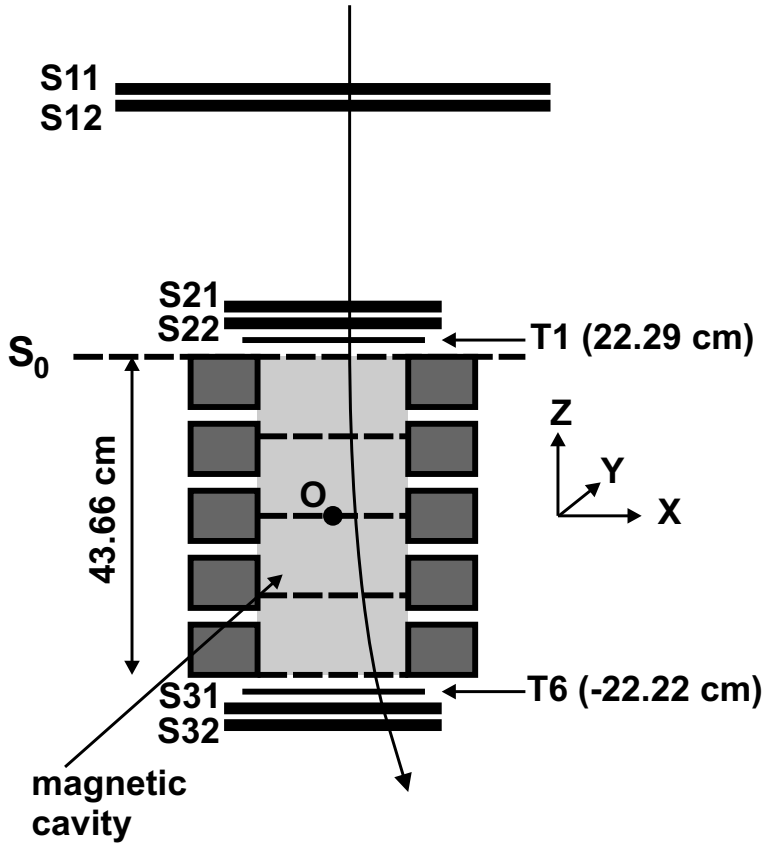

Figure 4.4: scheme of the elements considered for the definition of the acceptance requirements of PAMELA in the calculation of the geometric factor. The 5 dashed lines represent the heights at which the verification of the containment of the trajectory inside the magnetic cavity is performed. The coordinate system of PAMELA, with origin O in the centre of the cavity, is also indicated.

element	height (z), cm	x min., cm	x max., cm	y min., cm	y max., cm
S11 c.p.	+53.74	-20.4	+20.4	-16.5	+16.5
S12 c.p.	+53.04	-20.4	+20.4	-16.5	+16.5
S21 c.p.	+23.94	-9.0	+9.0	-7.5	+7.5
S22 c.p.	+23.44	-9.0	+9.0	-7.5	+7.5
T1 c.p.	+22.29	-8.07	+8.07	-7.25	+6.75
cavity upper face	+21.83	-8.07	+8.07	-6.57	+6.57
cavity lower face	-21.83	-8.07	+8.07	-6.57	+6.57
T6 c.p.	-22.22	-8.07	+8.07	-6.75	+7.25
S31 c.p.	-23.49	-9.0	+9.0	-7.5	+7.5
S32 c.p.	-24.34	-9.0	+9.0	-7.5	+7.5

Table 4.1: nominal geometry (in the PAMELA frame) of the elements considered in the calculation of the geometric factor. Elements are listed for decreasing z; c. p. = central plane.

[7]The various thin detector layers of the TOF and tracking system have been approximated as rectangles

These selection cuts correspond to 13 different heights z_i for which the coordinates (x_i, y_i) of the trajectory must be contained within a specific rectangle; the value of f associated with the trajectory is 1 or 0 according to whether or not all the 13 requirements are met.

For a given set $(x, y, \theta, \phi; \rho)$ characterizing the incidence of the particle on the reference plane S_0, the determination of the trajectory in the actual magnetic field of PAMELA[10] has been performed by adopting the same tracking routine employed in the reconstruction of real events acquired by the spectrometer, derived by the GRKUTA routine included in the GEANT software package[11]. This routine is based on the interpolation method of Runge-Kutta for solving the equation of motion of a charged particle in an inhomogeneous magnetic field.

4.3 Geometric factor: integration method

The calculation of the geometric factor can practically be done by analytical integration only for values of $|\rho|$ sufficiently high to neglect the effect of the magnetic field (trajectories approximated as straight lines) and for relatively simple geometries defining the acceptance domain; examples of such analytical calculations can be found in [47]. For more complex geometries, and in particular when the dependence of $G(\rho)$ on the magnetic field must be studied, like in the case of the PAMELA apparatus, a numerical approximation of the integral must be calculated and the approximation error estimated.

A generic numerical calculation of the integral (4.11) is represented by a weighted sum of the values of the integrand function, evaluated for a certain number of points in the 4-dimensional space (x, y, θ, ϕ).

In the present work the choice of the integration method has been done with the main purpose of maximizing the achievable accuracy for a given computation time T_{comp} or equivalently for a given number of points, since in practice, once chosen the machine performing the calculation, T_{comp} is proportional to the number of times that the integrand is evaluated.

As a starting point, the 4-dimensional integration domain is divided into N^4 equal

positioned at the average height of the corresponding detector.

[8]Considering also the thin (0.3 mm) Al protective layer covering the free surfaces of the modules.

[9]This check is done at three intermediate planes inside the cavity volume, as shown in fig. 4.4.

[10]Note that the magnetic field of PAMELA has been mapped over the whole volume of interest for the geometric acceptance, both inside and outside the magnetic cavity.

[11]GEANT [48] is a simulation tool developed at CERN for the description of the tracking and interactions of particles inside detectors.

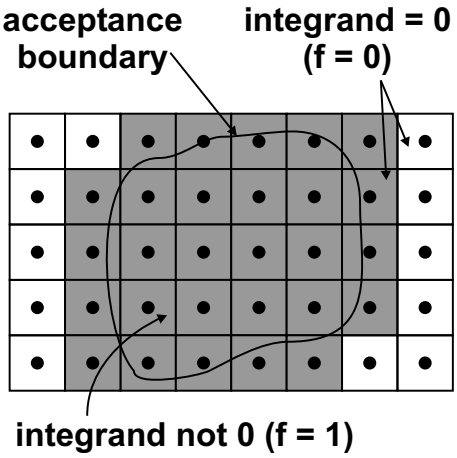

Figure 4.5: calculation of the geometric factor (in 2 dimensions for simplicity): the step-like integrand $\sin\theta\,|\cos\theta|\,\mathrm{f}$, evaluated at the centre of each cell, varies in general over the volume of a cell; therefore all the cells (in grey), either completely within the acceptance boundary or crossed by it, contribute to the calculation error.

cells, in such a way that the integration interval for each dimension is covered by N cells. $\mathrm{G}(\rho)$ can then be expressed as the sum of the integrals over the single cells:

$$\mathrm{G}(\rho) = \sum_{1}^{\mathrm{N}^4} \int_{\mathrm{cell}} \mathrm{dx\ dy\ d}\phi\ \mathrm{d}\theta\ \ \sin\theta\,|\cos\theta|\,\mathrm{f}\,(\mathrm{x},\mathrm{y},\theta,\phi;\rho) \tag{4.12}$$

and calculated by approximating the integral over each cell with the integrand value, computed at the central point, multiplied by the volume of the cell, $\mathrm{V_{cell}}$:

$$\mathrm{G}(\rho) \approx \sum_{1}^{\mathrm{N}^4} \mathrm{V_{cell}} \cdot [\sin\theta\,|\cos\theta|\,\mathrm{f}\,(\mathrm{x},\mathrm{y},\theta,\phi;\rho)]_{\mathrm{central\ point}} \tag{4.13}$$

It is clear that only the cells whose centre is inside the acceptance domain contribute to the sum (see fig. 4.5).

In order to reduce the calculation error, we can observe that, if the integrand is made constant within the acceptance boundary, the cells completely within the boundary (internal cells) do not contribute to the error, since for these cells the value of the integrand at the central point, multiplied by $\mathrm{V_{cell}}$, is exactly equal to the integral over the cell (see fig. 4.6).

This condition can be satisfied in the present case with a suitable change of variable; observing that $\cos\theta \leq 0$ in the integration domain and defining:

$$\beta = \frac{1}{2}\cos^2\theta \tag{4.14}$$

the integral takes the form:

$$\mathrm{G}(\rho) = \int_{\mathrm{S_0}} \mathrm{dx\ dy} \int_0^{2\pi} \mathrm{d}\phi \int_0^{\frac{1}{2}} \mathrm{d}\beta\ \hat{\mathrm{f}}\,(\mathrm{x},\mathrm{y},\beta,\phi;\rho) \tag{4.15}$$

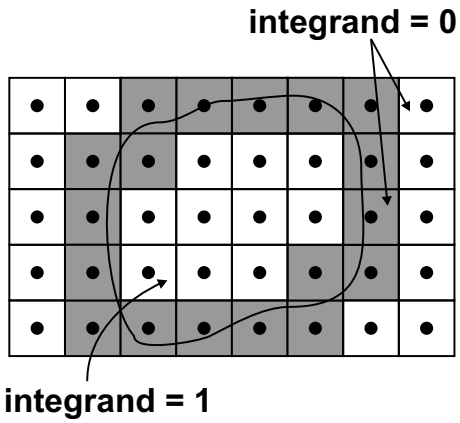

Figure 4.6: if the integrand is constant within the acceptance domain, only the cells (in grey) traversed by the acceptance boundary contribute to the calculation error.

and its approximation becomes:

$$\mathrm{G}(\rho) \approx \sum_{1}^{\mathrm{N}^4} \mathrm{V}_{\mathrm{cell}} \cdot [\hat{\mathrm{f}}\,(\mathrm{x}, \mathrm{y}, \beta, \phi; \rho)]_{\mathrm{central\ point}} \tag{4.16}$$

with $\hat{\mathrm{f}}$ equal to 1 or 0 (depending whether the acceptance conditions are satisfied or not) and constant within the acceptance domain.

With this solution one can roughly expect a dependence $\delta \sim 1/\mathrm{N}$ of the calculation error δ on the number N of points per dimension. This expectation can be justified by observing that the cells traversed by the 4-dimensional surface defining the acceptance boundary are of the order of N^3, while their average contribution to the total error δ is of the order of their volume ($\propto 1/\mathrm{N}^4$).

From a review of more refined numerical integration methods, using non-uniform distributions of points in the integration domain (Gaussian quadrature, Romberg method; see for example [49]), it can be concluded that a dependence better than $\delta \sim 1/\mathrm{N}$ cannot be achieved for the specific integral $\mathrm{G}(\rho)$; in fact these methods assume the continuity of the integrand and its derivatives, condition that is not satisfied by the step-like function f.

Other methods for the calculation of geometric factors use Monte Carlo algorithms, based on a random or quasi-random distribution of points; they are usually preferred because it is possible to include in the computation also the effects of several kinds of interactions of the incident particle with the traversed materials. In the present approach I have chosen to treat separately the calculation of the geometric properties of the apparatus with the purpose of achieving the best possible accuracy. A comparison between the method discussed here and a standard Monte Carlo calculation [50] of the geometric

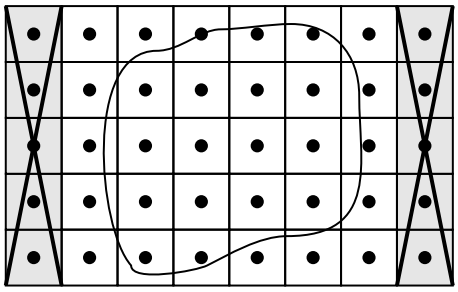

Figure 4.7: with a reduced integration domain, obtained by excluding cells for which the integrand is surely 0, a given approximation error can be reached with a smaller number of evaluations of the integrand.

factor of PAMELA will be done at the end of this chapter.

It can also be observed that the angular integration domain, defined as the whole downward hemisphere, is surely greater than the actual acceptance; with a reduction of this domain it is possible to reach the same accuracy with a significantly smaller number of evaluations, avoiding to compute the integrand function for cells that are surely outside the acceptance and do not contribute to the overall sum (see fig. 4.7). In doing this optimization it must be paid attention not to wrongly exclude also parts of the acceptance domain, thus introducing an underestimation of the true geometric factor.

In particular the integration interval in the angular variable β can be reduced by deriving, with analytical calculations, a conservative lower integration limit $\beta_{\min}(\rho)$ or a corresponding $\theta_{\min}(\rho)$.[12]. This has been achieved by constructing a conservative simplified model of the acceptance interval for β, taking into account only the constraints on the motion of the particle inside the magnetic cavity (without considering the further constraints imposed by the presence of the external detector planes: S1, S2, T1, T6, S3) and assuming the field $\vec{B}$ inside the cavity to be exactly uniform and directed along the Y axis, with modulus $|\vec{B}\,| = 0.5$ T larger than the actual maximum (0.48 T); this is a good approximation of the true spatial dependence of $\vec{B}$, characterized by B_y at least an order of magnitude greater than B_x and B_z and rather uniform over the cavity volume (see also sect. 1.7.1).

The limits $\theta_x^{\min}$ and $\theta_y^{\min}$[13], given by the projection of the motion on the XZ and YZ planes respectively, have been separately calculated. Under these hypotheses the projection on the XZ plane (orthogonal to $\vec{B}$) of the trajectory inside the cavity is simply

[12]Note that β is a monotonically increasing function of θ in the interval $\frac{\pi}{2} < \theta < \pi$; $\beta = 0$ represents horizontally incoming particles, $\beta = \frac{1}{2}$ corresponds to vertical incidence.

[13]θ_x is defined as the polar angle between the Z axis and the projection on the XZ plane of the momentum of the particle entering the magnetic cavity; see also fig. 4.10 and 4.11; θ_y is similarly defined.

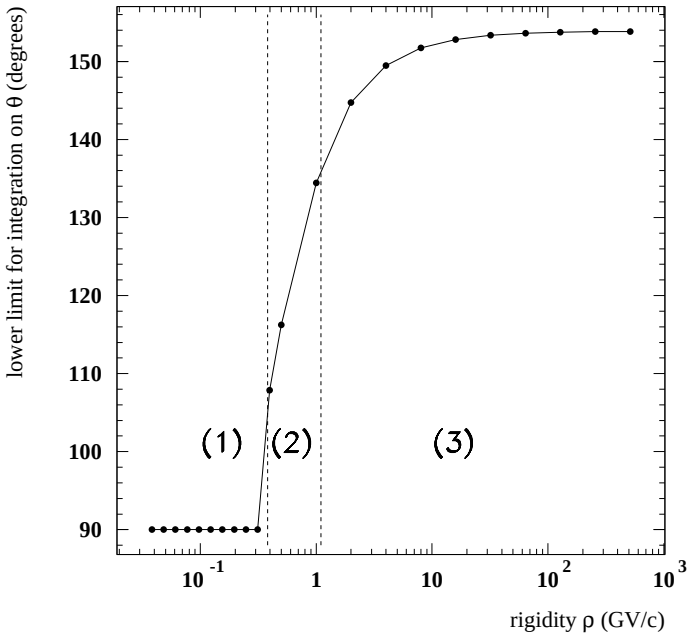

Figure 4.8: results of the calculation of $\theta_{\min}$ for the considered values of $|\rho|$; $\theta_{\max} = 180°$ (vertical incidence).

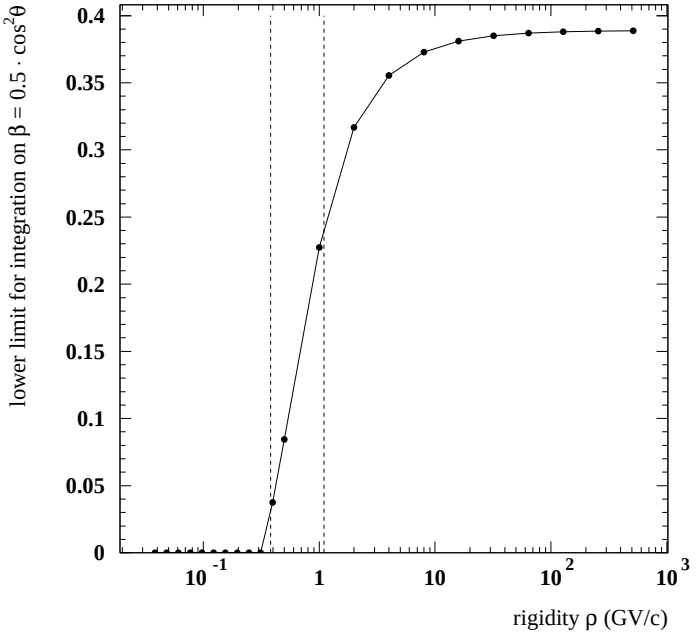

Figure 4.9: results of the calculation of $\beta_{\min}$ for the considered values of $|\rho|$; $\beta_{\max} = 0.5$ (vertical incidence).

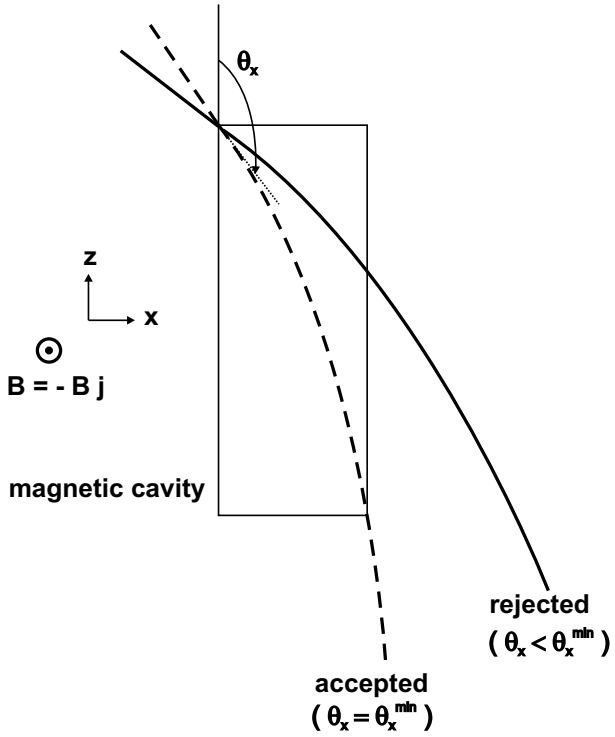

Figure 4.10: determination of $\theta_{\rm x}^{\rm min}$ for a value of $|\rho|$ within the range (3) of fig. 4.8.

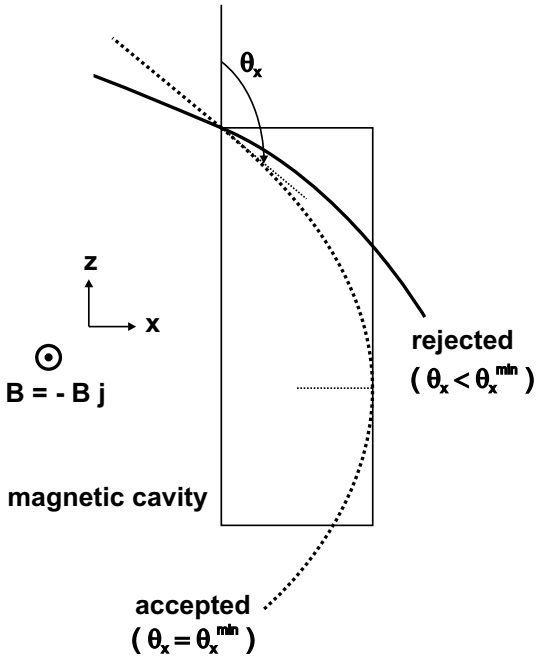

Figure 4.11: determination of $\theta_{\rm x}^{\rm min}$ for a value of $|\rho|$ within the range (2) of fig. 4.8.

described by an arc of circumference, while in the YZ plane the trajectory does not differ much from a straight line; besides, because of the symmetry of the system, the resulting $\beta_{\min}(\rho)$ does not depend on the sign of the rigidity.

Finally the angle $\theta_{\min}$ (between 90° and 180°) is uniquely determined by:

$$(\tan\theta_{\min})^2 = (\tan\theta_{\mathrm{x}}^{\min})^2 + (\tan\theta_{\mathrm{y}}^{\min})^2 \tag{4.17}$$

The results of the calculation for $\theta_{\min}$ are shown in fig. 4.8 and the corresponding values of $\beta_{\min}$ are in fig. 4.9; the values of $|\rho|$ have been chosen to cover the whole range of interest ($\approx 0.04 - 500$ GV/c) for the calculation of the geometric factor, as discussed in the next section.

Some comments on these results are necessary.

As a first observation it must be noted that the major contribution to the variation of $\theta_{\min}$ with the rigidity comes from the motion in XZ plane, where the curvature of the circumferences covered by the particles significantly increases for lower values of $|\rho|$.

For the highest values of $|\rho|$ (range indicated as (3) in fig. 4.8) the limit angle in the XZ plane $\theta_{\mathrm{x}}^{\min}$ corresponds to a trajectory that enters the magnetic cavity from an upper vertex and exits from the opposite one (see fig. 4.10). For increasing values of $|\rho|$ this limit trajectory tends to become a straight line and $\theta_{\mathrm{x}}^{\min}$ approaches the purely geometric constraint set by the dimensions of the cavity, which amounts to 159.7°, corresponding to an aperture of 20.3° of the acceptance window (the associated value of the 3-dimensional angle $\theta_{\min}$ is 154°); in this limit the integration interval in β is reduced to about 0.11, with a factor 4.5 of gain respect to the integration over the whole downward hemisphere ($0 < \beta < 0.5$).

For intermediate values of $|\rho|$ (range indicated as (2) in fig. 4.8), because of the increased curvature, the particle can invert its velocity along the X axis inside the magnetic cavity and then reach the bottom of the cavity (see fig. 4.11), resulting in smaller values of $\theta_{\min}$.

Finally, for the lowest values of $|\rho|$ (range (1) in fig. 4.8), the search for an analytical limit $\beta_{\min}$ becomes quite complicated, while at the same time it can be expected that in this range $\beta_{\min}$, continuing to decrease monotonically with $|\rho|$, is < 0.04 (minimum value in range (2)), with a corresponding gain of less than ≈ 1.1 in the reduction of the integration domain. Therefore in range (1) $\beta_{\min}$ has been simply set to 0; the effect of this choice is discussed in the next section.

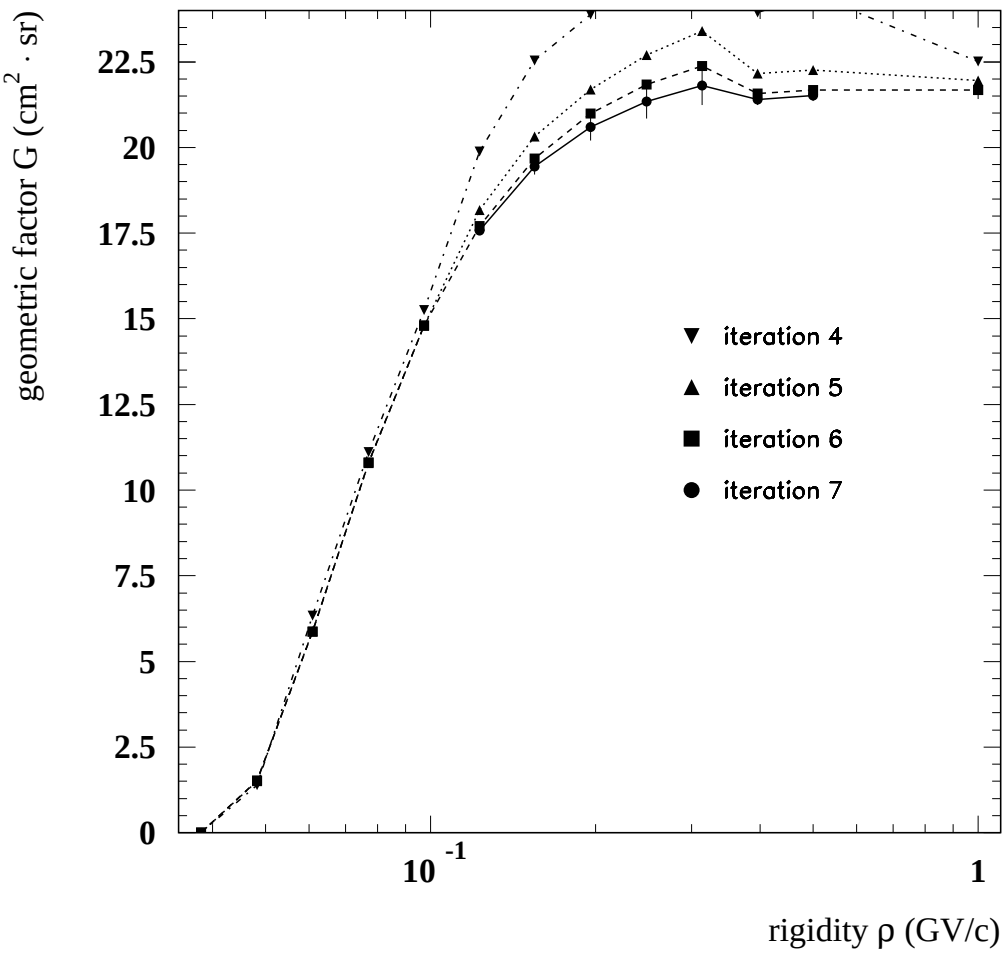

Figure 4.12: results of the calculation of G(ρ) in the low-rigidity range.

4.4 Geometric factor: results

The calculation of the geometric factor G(ρ) of PAMELA has been done for positive and negative rigidity values, chosen to cover the interval of interest $0 < |\rho| < 500$ GV/c [14].

The results of the calculation, shown in the two graphs 4.12 and 4.13 and summarized in table 4.2, are discussed in what follows.

The sequence of values of $|\rho|$, for which the calculation of G has been done, was determined starting from the maximum value, $|\rho| = 512$ GV/c, and taking half the current $|\rho|$ as the next step. For values of $|\rho|$ below 0.5 GV/c the geometric factor begins to

[14]The maximum detectable rigidity for the spectrometer of PAMELA is ≈ 700 GV/c. Besides, the geometric factor is not expected to show any significant dependence on the rigidity of the particle in the high-$|\rho|$ limit, as confirmed by the calculation (see below).

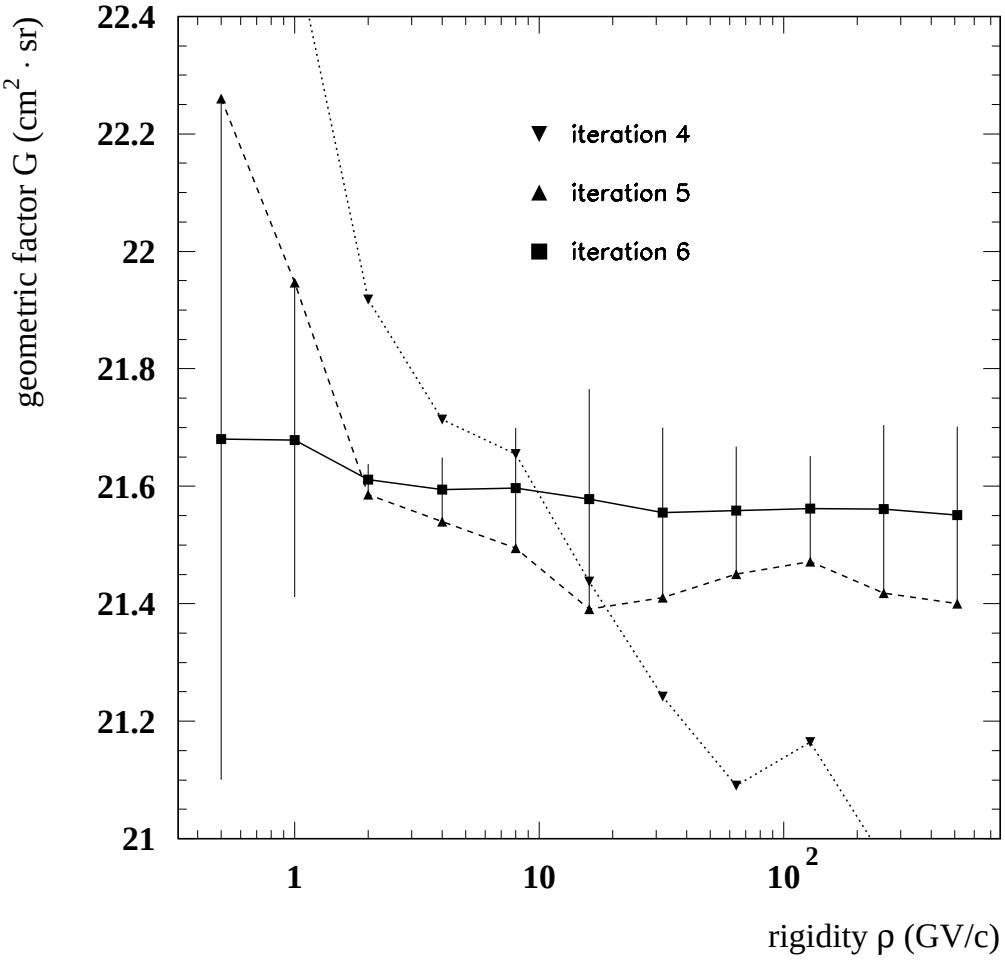

Figure 4.13: results of the calculation of G(ρ) in the high-rigidity range.

significantly differ from the high-rigidity limit; hence the density of $|\rho|$ values has been increased below 0.5 GV/c, with a smaller proportionality factor (set to 1.264) in such a way to obtain, in this rigidity range, 10 values of $|\rho|$ with a corresponding G greater than 0. With this system a total of 22 rigidity values has been determined, as reported in table 4.2, from 512 GV/c to 0.0381 GV/c for which the geometric factor is exactly 0.

The results shown here are only those for positive values of rigidity, since the calculated values of G(ρ) do not show a significant dependence on the sign of ρ, within the uncertainty of the calculation.

The determination of G for each value of ρ has been done according to the method

rigidity ρ, GV/c	geometric factor G, $cm^2 \cdot sr$	calc. error ΔG, $cm^2 \cdot sr$	relative error	range for β_{min}	number s of iterations
0.0381	0	0	//	1	6
0.0482	1.519	0.019	0.013	1	6
0.0609	5.866	0.016	0.0029	1	6
0.077	10.804	0.014	0.0013	1	6
0.0973	14.8005	0.0055	0.00038	1	6
0.123	17.58	0.13	0.0075	1	7
0.155	19.43	0.23	0.012	1	7
0.196	20.59	0.39	0.019	1	7
0.248	21.33	0.49	0.023	1	7
0.313	21.80	0.56	0.026	1	7
0.396	21.40	0.16	0.0076	2	7
0.5	21.51	0.16	0.0077	2	7
1	21.67	0.26	0.012	2	6
2	21.611	0.026	0.0012	3	6
4	21.594	0.054	0.0025	3	6
8	21.59	0.10	0.0047	3	6
16	21.57	0.18	0.0087	3	6
32	21.55	0.14	0.0067	3	6
64	21.55	0.10	0.0050	3	6
128	21.562	0.089	0.0042	3	6
256	21.56	0.14	0.0066	3	6
512	21.55	0.15	0.0070	3	6

Table 4.2: table summarizing the values of geometric factor and their errors, calculated for different rigidity values in the range $\approx 0.04 - 500$ GV/c.

described in the previous sections:

$$G(\rho) \approx \sum_{1}^{N^4} V_{cell} \cdot [\hat{f}\,(x, y, \beta, \phi; \rho)]_{central\ point} \tag{4.18}$$

starting with one single cell corresponding to the chosen 4-dimensional integration domain, then iterating the procedure, with the length of the cells along each dimension reduced to half (and their total number multiplied by $2^4 = 16$) at each new iteration. Iteration 6 corresponds to a total of $\approx 1.05 \cdot 10^6$ cells, iteration 7 to $\approx 16.8 \cdot 10^6$ cells.

The time required by a computer with good performances to calculate a value $G(\rho)$, for a given iteration level s, increases for lower rigidities, since the stronger curvatures of the trajectories in the magnetic field imply a greater number of interpolation points along the path of the particle to be considered by the routine GRKUTA. For iteration 6 this computation time increases from 1 to 6 hours, for iteration 7 it ranges from 15 to 90 hours (note the approximate proportionality of the calculation time to the total number

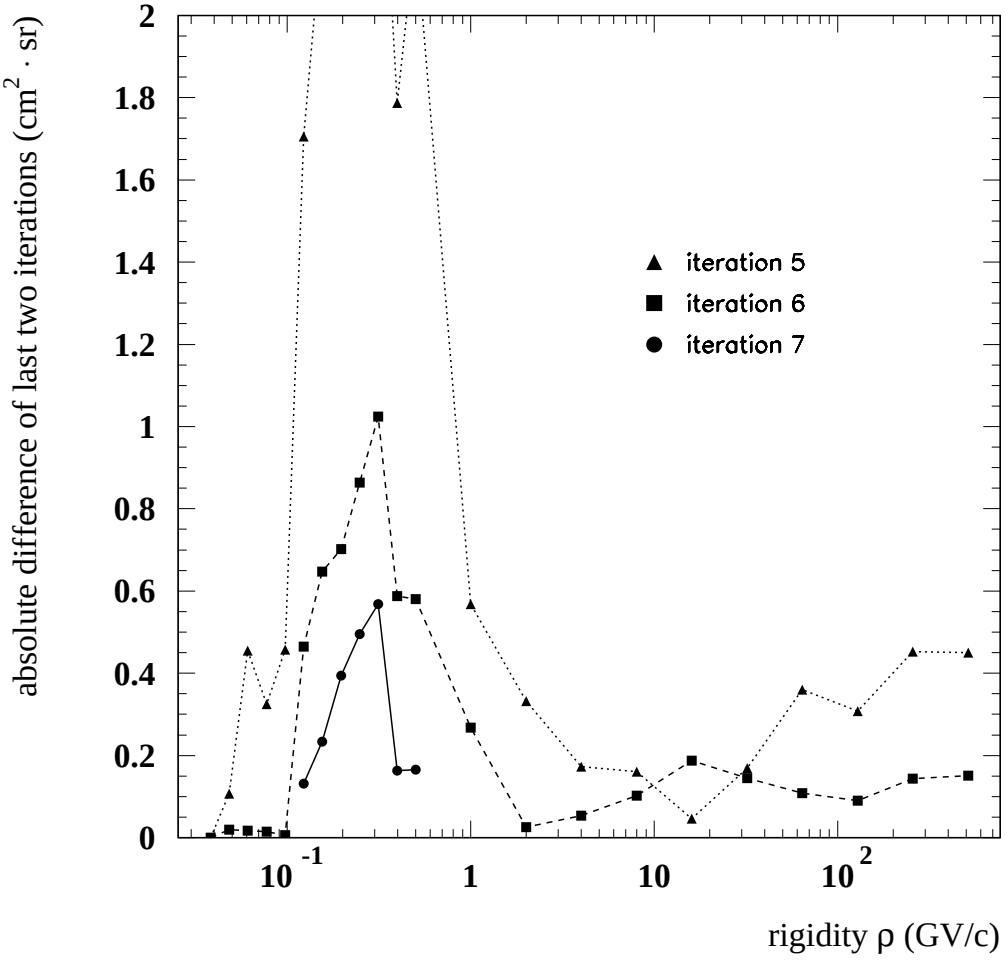

Figure 4.14: calculated errors $\Delta G(\rho)$ for various iterations. $\Delta G(\rho)$ for step s is given by the absolute value of the difference between the values of $G(\rho)$ for iteration s and s-1.

of cells, increasing by a factor 16 at each iteration).

In the two graphs 4.12 and 4.13 the systematic computation errors $\pm\Delta G(\rho)$ for the last iteration are also reported, with $\Delta G(\rho)$ for level s estimated as the absolute difference between the value of G obtained at step s and that given by the previous iteration:

$$\Delta G_s(\rho) = |G_s(\rho) - G_{s-1}(\rho)| \tag{4.19}$$

The errors calculated for different iterations over the whole range of rigidities are also reported in fig. 4.14. It can be observed that, for a given step s, in the intermediate range $\rho \approx 0.15 - 0.5$ GV/c the error $\Delta G(\rho)$ is significantly greater than for other values of ρ; this can be explained with the fact that for this range the integration domain has not been sufficiently reduced, around the acceptance boundary, with the simple method exposed

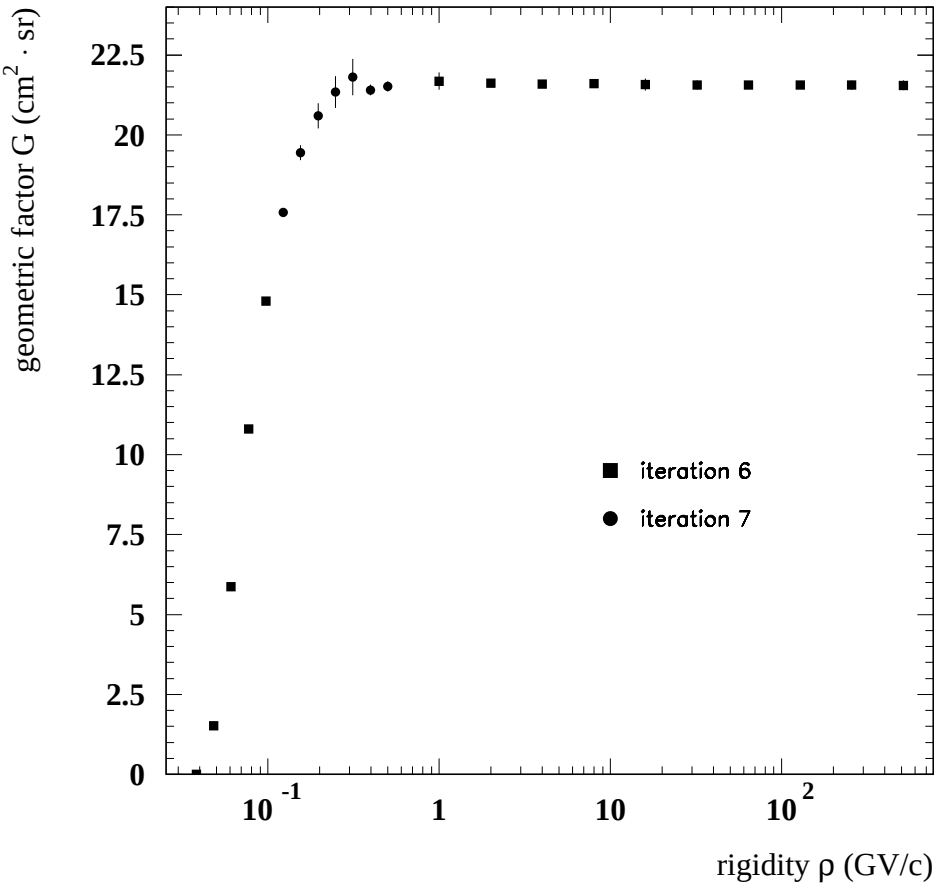

Figure 4.15: the values of G(ρ) with the corresponding calculation errors, resulting from the last iteration for each value of rigidity.

in sect. 4.3. By applying a further iteration (s = 7) to these rigidity values a reduction of the error by roughly a factor 2 have been achieved, as expected from the dependence $\Delta G(\rho) \sim 1/N$ on the number of cells per dimension already discussed, but at the expense of a T_{comp} 16 times greater.

The effect of the correct choice of the integration boundaries on the result of the calculation is evident in fig. 4.15, reporting G(ρ) (and the associated error) corresponding to the last iteration for each value of rigidity (see also table 4.2). An abrupt change in the value of G can be observed in correspondence of the transition between the range (2) and (1) for the determination of β_{min}; for $\rho = 0.313$ GV/c the value of the geometric factor is clearly higher than the one expected on the basis of a continuous dependence of G on the rigidity, meaning that a major contribution to the calculation error is introduced, for this value of ρ, by the poor matching between the integration domain, set to the whole

downward hemisphere, and the real acceptance limit in β.

On the other hand we can also observe that the adopted method for the estimation of the calculation error describes correctly the actual level of approximation; in fact the two nearby values of G (for $\rho = 0.248$ GV/c and $\rho = 0.396$ GV/c) differ from G($\rho = 0.313$ GV/c) by a quantity that is smaller than the associated uncertainty.

It can be seen that for rigidities greater than ≈ 0.3 GV/c there is no dependence of the geometric factor on ρ (within the uncertainties), because the effect of the magnetic field on the motion of the particles does not change significantly the effective acceptance domain for different rigidities; in this regime the trajectories are well approximated as straight lines and the value of G is determined only by the geometric characteristics of the system.

A linear fit performed over the 10 highest rigidity values gives:

$$G_{hr} = (21.60 \pm 0.13)\ cm^2 \cdot sr \tag{4.20}$$

where the systematic computation error ΔG_{hr} has been estimated by averaging over the 10 values of $\Delta G(\rho)$; the relative uncertainty is 0.6%.

For lower values of ρ, the effect of the magnetic field on the trajectories becomes increasingly important: incident particles within the combined geometric acceptance of the cavity and of the external detector planes tend to be deflected toward the lateral walls of the cavity before they can escape from its lower face; correspondingly the geometric factor decreases until it becomes 0 for ρ between 0.04 and 0.05 GV/c, meaning that no particle can satisfy the acceptance requirements of the apparatus for smaller values of ρ.

A consistency check of the computation algorithm has been done, by applying it to simplified models of the PAMELA apparatus, for which an analytical calculation is also possible; specifically the two following models have been considered:

- **model 1**: only the magnetic cavity of the spectrometer, in the hypothesis $\vec{B} = 0$, without the TOF system and the planes T1 and T6 of the tracking system; in this model, characterized by the absence of magnetic field, G does not depend on ρ. Applying the present method for $\rho = 1000$ GV/c, the value $G = 22.0 \pm 0.1\ cm^2 \cdot sr$ is obtained. Besides, the simplified geometry makes it possible to do an analytical calculation[15] giving $G \approx 22.0\ cm^2 \cdot sr$, in agreement with the previous value, within the computation uncertainty;

[15]The calculation has been done following [47] with a correction on the formula that is wrongly reported there.

- **model 2**: same as model 1, but with the true magnetic field of PAMELA; the calculation for $\rho = 1000$ GV/c gives $G = 22.0 \pm 0.1$ cm$^2 \cdot$ sr, consistent with the value obtained in absence of magnetic field (model 1), as expected in the high-rigidity limit.

It can be concluded that the present method for the calculation of the geometric factor, with the relatively simple optimizations discussed in sect. 4.3, gives a good accuracy within a reasonably small computation time; considering in particular the high-rigidity range ($|\rho| \geq 1$ GV/c), 1-2 hours per rigidity value are needed by a single computer to obtain a relative uncertainty not greater than 1%.

As a term of comparison, we can estimate the computation time (for a single machine) required to reach this same degree of accuracy with the standard Monte Carlo method used for the simulation of the PAMELA apparatus; this is implemented in the program GPAMELA, developed by the collaboration to simulate and study the motion and interactions of charged particles incident on the instrument and based on the GEANT software package (CERN).

The uncertainty in the GPAMELA calculation of the geometric factor is expected to have a dependence $\sim 1/\sqrt{N_p}$ on the total number N_p of particles simulated. The application of this method [50] has reported, in the high-rigidity range, a relative uncertainty of about 2.5% for $N_p = 10^6$ particles, with a computation time per rigidity value of the order of 12 hours. To obtain the desired relative uncertainty of 1%, we would have to increase N_p by a factor $(2.5)^2 = 6.25$. The computation time T_{comp} would also be increased by the same factor, thus becoming about 75 hours (at least ≈ 40 times greater than for the present method).

Improvements of the dependence $\sim 1/\sqrt{N_p}$ of the uncertainty in the GPAMELA calculation are expected with the introduction of a non-uniform (or quasi-random) Monte Carlo algorithm; further improvements are expected with other optimizations, such as the ones discussed in sect. 4.3, on the choice of the integration domain.

Chapter 5

Characterization of the TOF detectors in the flight configuration

The present chapter contains the description of a study that I have done on the characteristics of the signals produced by the 6 scintillator layers of the TOF system of PAMELA[1], in particular for the determination of the attenuation length of scintillation light and of the trigger efficiencies for the various strips of each layer.

This analysis is based on cosmic rays collected by the apparatus, between March and May 2005, in the laboratories of Rome (Italy) and Samara (Russia), during the latest qualification phases of PAMELA in its final ("flight") configuration, with data acquisition sessions alternated to the program of mechanical and electrical works and tests on various parts of the system.

The TOF strips interested by each event have been identified by using the precise spatial information on the trajectory of the incident particles given by the tracking system. The possibility of performing a combined analysis of the data produced by the TOF and tracking system of PAMELA allowed, on one hand, to obtain a characterization of the response of each strip along its length and, on the other hand, to do useful cross-checks between the spatial information independently provided by the two systems.

During and after the integration of the three double planes (S1, S2, S3) of scintillators and of the related electronics with the rest of PAMELA, the TOF system has been mechanically and electrically tested several times. Besides, the whole apparatus has sustained strong mechanical stresses during the qualification tests at the IABG facilities (Germany [22]) that certified its ability to be launched on board of the Resurs-DK1 satellite, without apparently reporting significant damages. Therefore a final qualification of the TOF detectors has naturally become one of the main tasks to be performed, after all

[1]The characteristics of the TOF scintillator system of PAMELA and of the related electronics (including the trigger board) are discussed in sections from 1.4 to 1.6.

these various phases of work and test.

The first three sections of the present chapter are dedicated to the measurement of the characteristic attenuation lengths of scintillation light for the various strips of the TOF system. In sect. 5.1 the propagation mechanisms of the scintillation light along a strip, from the point of production by an ionizing particle to the collection at the interface with the light guide, are briefly discussed; sect. 5.2 contains the description of the method that I have developed for the analysis of the data collected by the PAMELA apparatus, and in particular for the determination of the coordinate of incidence of the ionizing particle along the strip, using the information provided by the tracking system; in sect. 5.3 the results of the analysis are presented.

The rest of the chapter describes the measurement of the trigger efficiency for the various strips and layers of the TOF system. Sect. 5.4 contains a general discussion on the adopted method; in sect. 5.5 the measured trigger efficiencies are reported and commented. Finally the overall efficiency of the TOF system is calculated in sect. 5.6.

5.1 Attenuation lengths of TOF scintillators

The structure and basic operation of the 3 double planes of scintillators strips, constituting the TOF system of PAMELA, has been introduced in sect. 1.4. Here the mechanisms of light propagation along a scintillator strip and of collection on the light guides positioned at both ends of the strip are discussed in further detail.

When an ionizing particle crosses one of the 24 strips of the TOF system, the fluorescence light generated in the interaction with the molecules of the material (BC-404 [6]) is guided toward the two read-out PMT, positioned at the opposite ends, mainly by means of total reflections on the lateral walls of the strip and of the plexiglas light guide connecting it to the input window of the PMT; the PMT converts the light incident on the photo-cathode into an electric current pulse at its anode. This pulse, on one hand, enters the fast discriminator of the time section of the FE electronics, starting the timing measurement and producing the *event signal* for the generation of the trigger; on the other hand, it reaches the charge section where it is integrated by a charge preamplifier and finally converted into a 12-bit digital value: the digital charge value, for each of the 48 electronic channels (PMT) of the TOF system, depends linearly on the intensity of light collected on the photo-cathode of the corresponding PMT.

During the propagation of the light along the strip a fraction of the initial intensity is inevitably lost mainly because of partial loss at the walls of the strip and because of

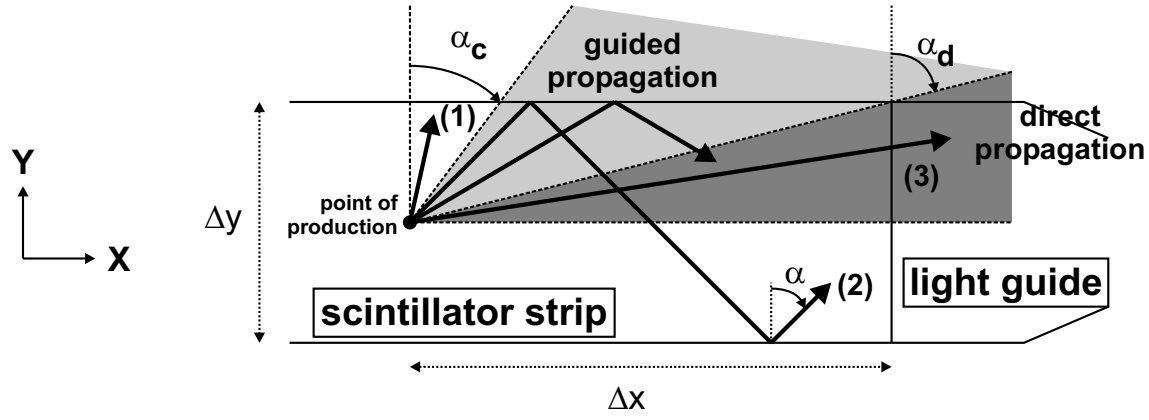

Figure 5.1: two-dimensional drawing illustrating the mechanisms of light collection on the light guide positioned at one end of a strip of the TOF system.

self-absorption in the bulk of the scintillator; this in general introduces a dependence of the height of the anode pulse on the coordinate of production of the light by the ionizing particle along the strip.

A simplified two-dimensional model of the mechanism of light propagation is illustrated in fig. 5.1 and explained in what follows.[2]

Consider a strip of the TOF detectors (of thickness 7 or 5 mm, depending on the TOF plane), directed along the X axis in the PAMELA reference system[3], which is crossed by a charged particle, moving along the Z axis, at its centre and at a distance Δx from the strip end.

The particle releases part of its energy within the scintillator material with the production of isotropically distributed fluorescence light. Consider for simplicity the light emitted in the XY plane, with angle of emission α (between 0 and 90°) respect to the Y axis, in the hypothesis of perfectly polished lateral walls and absence of diffusion within the bulk of the scintillator material.

If the angle of emission α is small (case (1) in fig. 5.1), the light, propagating along a straight line, reaches the lateral wall of the strip where it is partly reflected back and

[2]A complete and precise 3-dimensional treatment of the light propagation along the strip, to calculate the effective light collection efficiency as a function of the point of production within the strip volume, taking into account the geometric characteristics of the strip and light guide and the optical properties of the lateral walls of the strip, requires the use of a dedicated numerical simulation tool, like for example Guide 7 or GUIDEIT [51].

[3]The PAMELA reference frame has been defined in sect. 1.2.

partly transmitted across the boundary of the scintillator material; the fraction of light that is reflected back in the scintillator then continues propagating in straight line inside the strip and maintaining the same angle of incidence α with respect to the lateral wall, until another partial reflection happens at the opposite wall and so on. As mentioned in sect. 1.4, each scintillator strip has been enveloped in a thin aluminized mylar foil; with this system a fraction of the light escaping from the walls of the strip is reflected back by the Al film. Anyway it is expected that after several reflections most of this light will be lost out of the scintillator and will not contribute significantly to the intensity entering the light guide at the end of the strip.

For angles of emission α greater than α_c, *critical angle*[4] characteristic of the external boundary of the scintillator (case (2) in figure), the light incident on the wall is subjected to total internal reflection and hence propagates along the strip with practically no loss of intensity; this fraction of the light is usually indicated as *guided light*.

Finally for $\alpha > \alpha_d$, with α_d given by

$$\tan \alpha_d = \frac{\Delta y}{2 \cdot \Delta x} \tag{5.1}$$

(case (3) in figure), the light can directly enter the guide, without reaching the lateral wall of the strip. This part of the light is indicated as *direct light*.

The guided light component constitutes the main contribution to the overall intensity collected on the guide for points of incidence of the ionizing particle along most of the strip length; in fact the direct light contribution is practically important only when the light is produced in the vicinity of the strip end. In fact, given the isotropic angular distribution of the scintillation light, the fraction of direct light is proportional to the solid angle under which the end wall is seen from the point of production, and hence decreases as the square of the distance Δx.

The guided light travelling from the point of production toward the end of the strip is attenuated because of the self-absorption inside the scintillator volume (*bulk attenuation*) and, in practical applications, also because of a non-negligible partial loss at the lateral walls (*surface attenuation*). This surface attenuation is due to the fact that in a real strip the walls are not perfectly polished[5], but instead microscopic defects and impurities are

[4]The critical angle α_c for total internal reflection of light propagating in a material characterized by a refraction index n_1 and reaching the interface with another material of index $n_2 < n_1$ is given by $\sin \alpha_c = n_2/n_1$. In the present case the two materials are respectively the scintillator BC-404 ($n_{BC404} = 1.58$) and the thin layer of N_2 (or air during the tests on earth) which is expected to be present between the scintillator and its envelope of mylar; $\alpha_c = \arcsin(n_{air}/n_{BC404}) \approx 39°$.

[5]The surfaces of the scintillator are "optically polished" when the dimensions of the irregularities left by the machining process are much smaller than the wavelength of the incident light (in the present case ≈ 400 nm).

in general present; as a consequence the critical angle can locally be greater than α_c thus causing a partial reflection even in the case that $\alpha > \alpha_c$.

The derivation of the effective dependence of the light intensity collected at the end of the strip, I_f, on the intensity I_i of the initially produced light at a distance Δx, taking into account the combined effect of the bulk and surface attenuations, the spatial geometry of the strip and the optical properties of the scintillator walls and of the external Al reflecting layer, is quite complicated and can be done with the help of purposely written computer programs (see for example [51]). Anyway for practical applications a good phenomenological approximation of this dependence is given by the exponential law:

$$I_f = I_i \cdot \exp\left(-\frac{\Delta x}{\lambda_{att}}\right) \tag{5.2}$$

where λ_{att} is the *effective attenuation length* of scintillation light in the strip. The study described in the next sections uses this simple phenomenological model for the measurement of the effective attenuation length λ_{att} for the various strips of the TOF system, comparing it to the known bulk attenuation length characteristic of the BC-404 (140 cm [6]).

5.2 Method for attenuation lengths measurement

The method described here, for the measurement of the effective attenuation length λ_{att} of the TOF scintillator strips, after the integration of the TOF system within PAMELA, uses the cosmic-ray events acquired during the final qualification tests of the apparatus and is based on the combined analysis of the charge signals collected by the TOF electronics for each channel (PMT) of the system and of the spatial information given by the tracking system on the position along the scintillator strip where the incident particle has initially produced the scintillation light.

The data produced by the TOF and tracking system can be processed by custom designed *event reconstruction* software programs, developed within the PAMELA collaboration for the automatic extraction of the significant physical information from the bulk of data produced by the apparatus at each event; in particular for the tracking system this means the automatic determination of the trajectory of the incident particle on the basis of the pattern of signals on the microstrips of the 6 detector planes.

At the time when the present analysis has been done, the available preliminary versions of these programs were designed to produce an output compatible with the PAW [52] physics analysis software. Therefore I have worked in the PAW environment with the creation of automatic procedures and specific FORTRAN routines to perform the analysis,

starting from the data produced by the event reconstruction software of PAMELA.

The method employed to determine $\lambda_{\rm att}$ considers minimum ionizing particles (MIP) within the sample of collected cosmic-ray events. It can be observed that at low altitudes (below 1 km on the sea level) the cosmic-ray flux is mostly composed of muons ($\mu^{\pm}$), whose intensity is about a factor of the order of 100 greater than for the other species of charged particles. These muons can be considered MIP for rigidities $|\rho|$ greater than ≈ 0.5 GV/c, as can be derived from the dependence of the linear energy transfer (LET) of an ionizing particle on its velocity, given by the Bethe-Bloch formula.[6]

A MIP traversing a strip of the TOF system is characterized by a distribution of the LET that is well represented by a Landau function, which describes the energy lost by relativistic particles across thin layers of material [3]. On the other hand, the intensity of the scintillation light produced by an ionizing particle in a thin layer of plastic scintillator is given by the semi-empirical Birks law [3], stating that for small LET the light intensity is proportional to the linear energy transfer.

As a conclusion, applying a selection cut to the collected sample of cosmic rays to reject particles with $|\rho|$ smaller than ≈ 0.5 GV/c, the distribution of the intensity of scintillation light is expected to be described by a Landau function.

By dividing each strip, along the longitudinal axis, into suitably small sections (*pads*) such that it is possible to neglect the effect of the light attenuation over the pad length, the distribution of the charge signal for a TOF channel (PMT), for MIP crossing the strip in correspondence of a specific pad, is expected not to significantly differ from the Landau distribution, characterized by a most probable (peak) value of the charge signal $\rm Q_{mp}$.

Finally $\lambda_{\rm att}$ can be determined by fitting an exponential function to the observed dependence of $\rm Q_{mp}$ on the central coordinate of the pad. Further details of the method will be given in the next section.

It must also be noted that the distribution of the incidence directions for the cosmic muons at earth is expected to cover sufficiently well all the acceptance window of PAMELA when the apparatus is pointing toward the zenith (maximum aperture $\approx 26^{\circ}$).

In the selection of the data sample to analyze we must also take into account that the proportionality factor between the intensity of light, entering the input window of the PMT, and the 12-bit digital output of the corresponding charge section, can be varied by

[6]The ionizing power of a particle species in a thin layer of material is commonly represented by the *linear energy transfer* (LET), defined as the ratio between the energy released by ionization and the length of the path inside the traversed material. Usually this quantity is normalized to the density of the material and typically measured in $\rm MeV \cdot cm^2/mg$. The LET of a charged particle decreases for higher velocities until it reaches a *minimum ionization plateau* [3].

adjusting the gain of the PMT.[7] Clearly this makes it necessary to separately consider, for each PMT, data samples collected with different gain settings, because the associated digitally converted Landau distributions for a given pad would be significantly shifted from each other.

The separate adjustment of the PMT gains is of fundamental importance because, together with the variation of the discrimination thresholds for the time sections of the TOF FE electronics, it allows to set the system in an optimized configuration, taking into account the unavoidable differences, among the various electronics channels, in the proportionality factor between the intensity of light, entering the PMT photo-cathode after surviving the attenuation in the scintillator, and the charge producing the corresponding anode signal. This proportionality is affected by many parameters, such as the effective attenuation length of the scintillator strip, the transmission efficiencies of the optical couplings strip/light guide and guide/photo-cathode, and the characteristics of the PMT, in particular the quantum efficiency and the (adjustable) gain of the electron multiplier structure; these parameters generally show important differences from channel to channel and also variations with time.

During the final qualification sessions of the PAMELA apparatus, many different combinations of high voltage and discriminator thresholds have been tested to find the best configuration to be normally used in flight. The present analysis is focused on the statistically most significant sample, collected with the "flight" high-voltage configuration.

5.2.1 Determination of the position on the TOF layers

In the present method the determination of the coordinates of the crossing points of a particle in the 6 TOF layers, necessary for the identification of the traversed strips and the association of the corresponding charge signal to a specific pad, is done on the basis of the information on the particle trajectory given by the event reconstruction program of the tracking system (indicated from now on simply as *tracking program*).

The tracking program provides information on the rigidity ρ of the particle and on the crossing point (x,y) and direction of incidence (θ, ϕ) on a reference plane chosen at a height z = 23.5 cm, in correspondence of the S2 scintillator detector. From this information and from the knowledge of the map of the magnetic field, inside and outside the cavity of the spectrometer, the crossing points of the particle for the 6 layers of the TOF system have

[7]The PAMELA apparatus is equipped with the electronics for the adjustment of the gain for the 48 PMT of the TOF system; this is achieved by changing the high voltage applied between the anode and photo-cathode of the PMT, and exploiting the exponential increase of the gain with the high voltage. For the specific characteristics of the R5900 PMT employed in the TOF system see [7].

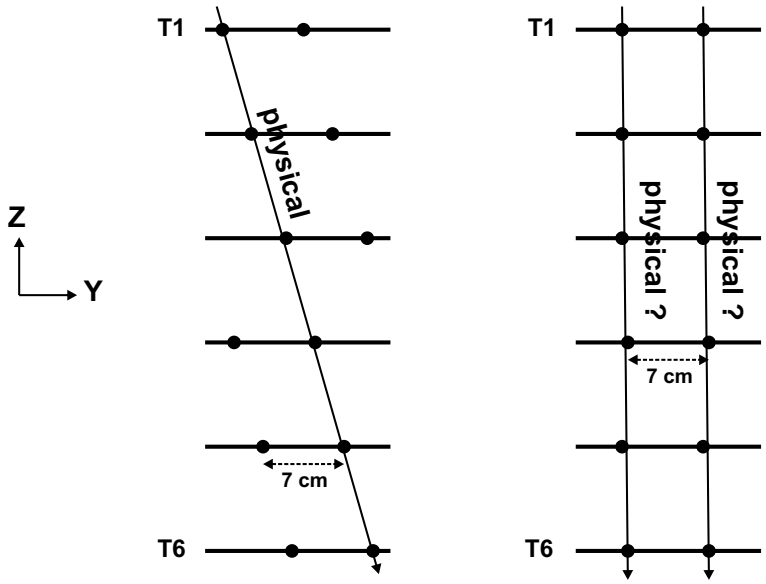

Figure 5.2: at left: the identification of the physical trajectory with only the information from the tracking system is relatively simple for inclined incidence; at right: the identification is much more difficult for vertical incidence.

been reconstructed, using the same method described in the discussion on the geometric factor and based on the GRKUTA routine (see end of sect. 4.2).

A complication in the use of the information given by the tracking program comes from the fact that the crossing point of the particle in a silicon plane of the tracking system is not uniquely identified along the Y axis, since each read-out channel of the Y view of a detector is associated with two microstrips separated by 7 cm (see sect. 1.7.2 and fig. 1.19), thus yielding two possible Y coordinates per plane for the same event. The tracking program, only on the basis of the pattern of microstrips that have reported a significant signal for each of the 6 Si detector planes, is able to identify the physical trajectory if this ambiguity can be clearly solved, like in the example at left in fig. 5.2; otherwise the program cannot do better than identifying two possible and mutually excluding trajectories, separated by 7 cm, like in the example at right side in fig. 5.2.

An efficient method to solve this ambiguity is to use the spatial information given by the TOF system itself, with the pattern of scintillator strips that have provided an event signal for the 3 layers (S12, S21 and S32) which are segmented along the Y axis (with strips directed along the X axis, see fig. 1.6 and table 5.1). The employment of this method in the data analysis is described in the next section.

layer	number of strips	Z (cm)	thickness (mm)	strip dim. X (cm)	strip dim. Y (cm)	strip cross-section (cm^2)	pads per strip
S11	8	+53.74	7	5.1	33.0	3.57	15
S12	6	+53.04	7	40.8	5.5	3.85	16
S21	2	+23.94	5	18.0	7.5	3.75	8
S22	2	+23.44	5	9.0	15.0	4.50	8
S31	3	-23.49	7	6.0	15.0	4.20	9
S32	3	-24.34	7	18.0	5.0	3.50	12

Table 5.1: geometrical characteristics of the TOF scintillator strips; the number of longitudinal pads chosen for the analysis of the attenuation lengths are also indicated.

5.3 Data analysis for attenuation lengths

The initial sample used in this analysis consists of 14620 events, collected with the apparatus, pointing toward the zenith, in the standard ("flight") configuration, for which the tracking program has reported either a unique physical trajectory or two mutually excluding trajectories (for the unsolved ambiguity in Y coordinate). The trigger pattern is the standard one: (S11 OR S12) AND (S21 OR S22) AND (S31 OR S32).

The applied cuts for data selection have been minimized in order to keep the fraction of rejected events low enough to be able to populate a sufficient number of pads for each scintillator strip. The number of longitudinal pads for each strip is indicated in table 5.1; the choice of the pad length ($\approx$ 2 cm) has been done by taking into account the request of a sufficiently high number of events per pad to correctly reconstruct the expected Landau distribution, at the same time avoiding the distortion of this distribution because of the light attenuation, in particular for the strips characterized by the shortest $\lambda_{\rm att}$.

The following two cuts have been applied, in order to determine the correct path of the particle in the apparatus for those events ($\approx$ 6900) for which the tracking program reconstructs two mutually excluding trajectories:

- **cut 1**: for each of the 3 TOF layers segmented along the Y axis, one and only one strip has generated an event signal for the trigger board on one or both the corresponding channels (PMT);

- **cut 2**: one and only one of the two reconstructed tracks crosses the 3 strips identified in the previous cut.

A further cut on the rigidity of the incident particle measured by the tracking system ($|\rho| > 0.5$ GV/c) assures that the considered particles are MIP, as discussed in the

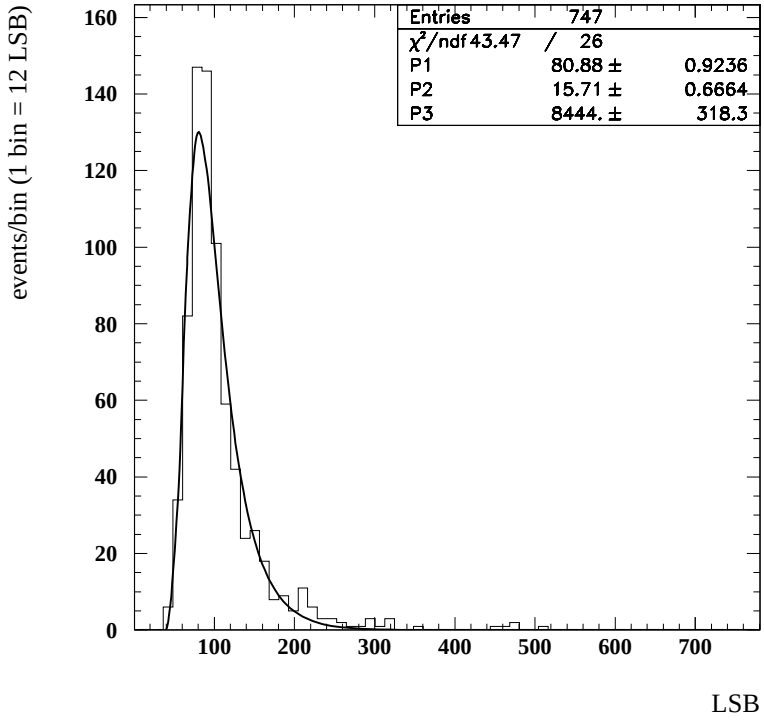

Figure 5.3: distribution of the charge signal (in LSB) for pad 3 of strip 1 of TOF layer S21, PMT 2; superimposed is the Landau fit. P_1, P_2 and P_3 respectively represent the parameters Q_{mp}, ξ and A of the formula (5.3).

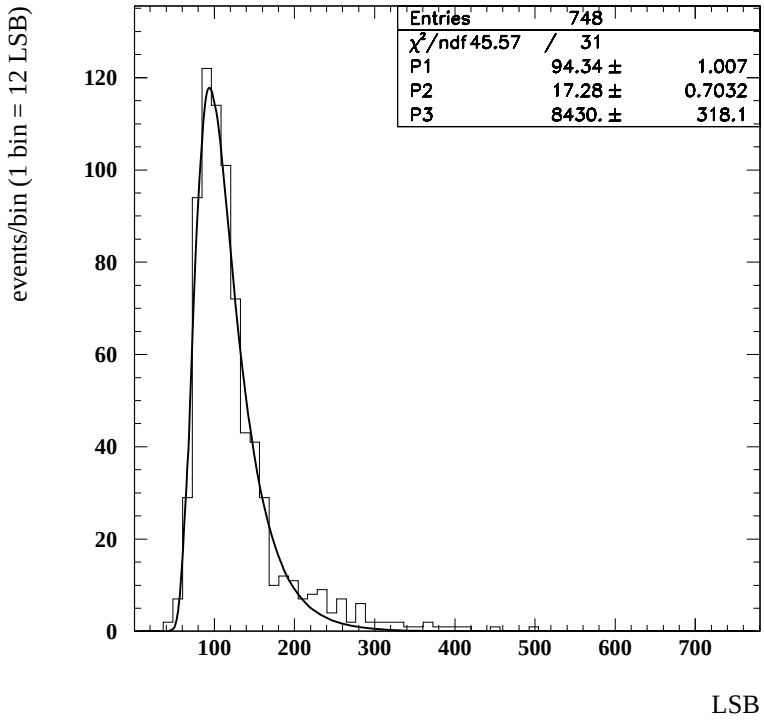

Figure 5.4: distribution of the charge signal (in LSB) for pad 6 of strip 1 of TOF layer S21, PMT 2; superimposed is the Landau fit. P_1, P_2 and P_3 respectively represent the parameters Q_{mp}, ξ and A of the formula (5.3).

previous section. With these cuts about 10% of the events are rejected.[8]

For each of the surviving events and for each TOF layer, the strip crossed by the incident particle has been determined on the basis of the tracking information and the corresponding charge signals for the two PMT has been associated with a specific pad of that strip.

The distribution, over the considered event sample, of the 12-bit charge signals Q for a given PMT and pad has been fitted with the Landau function:

$$f(Q) = \frac{A}{\sqrt{2\pi}\,\xi}\, e^{-\frac{1}{2}\left(\lambda + e^{-\lambda}\right)} \quad \text{with} \quad \lambda = \frac{Q - Q_{mp}}{\xi} \tag{5.3}$$

where Q_{mp} represents the most probable value of Q (peak value).

Examples of two typical distributions of charge signals Q (expressed in least significant bits, LSB, of the 12-bit value) for a fixed strip and PMT (layer S21, strip 1, PMT 2) and two different pads (pad 3 and 6, about 7 cm apart) are shown in fig. 5.3 and 5.4. It can be seen that the number of events contributing to the distribution is about 800 and the χ^2/d.o.f. of the fit is $1.5 - 1.7$ meaning that the Landau function describes sufficiently well the actual distribution of Q. The difference between the values of Q_{mp} (81 and 94 LSB respectively) means that the intensity of light collected in the considered PMT is smaller for the farther pad (pad 3).

To finally determine the attenuation length, the distribution of the most probable values Q_{mp} along the pads, for a given strip and PMT, has been fitted with an exponential function. Further selections have been necessary on the set of considered pads for each strip and PMT.

Since this study aims at measuring the attenuation length given by the combined effect of bulk and surface attenuation on the guided light component, the nearest pad to each PMT has been excluded from the fit; in fact for this pad the contribution of direct light, entering the guide without any reflection on the lateral walls of the strip, is expected to be of the same order of the intensity of guided light. For the other pads the intensity of direct light is expected to be much smaller, given its dependence on the inverse square of the distance between the point of production and the end wall of the strip.

The number of events per pad tends to decrease along the strip going from the center toward both ends; this clearly is an effect of the shape of the acceptance window of PAMELA, that excludes a greater domain of directions of incidence near its boundaries.

[8] It is clear that a more efficient analysis would also employ the spatial information given by the calorimeter to identify the physical track in the case of ambiguity; for the present purpose the simple cross-check between TOF and tracking system is quite sufficient.

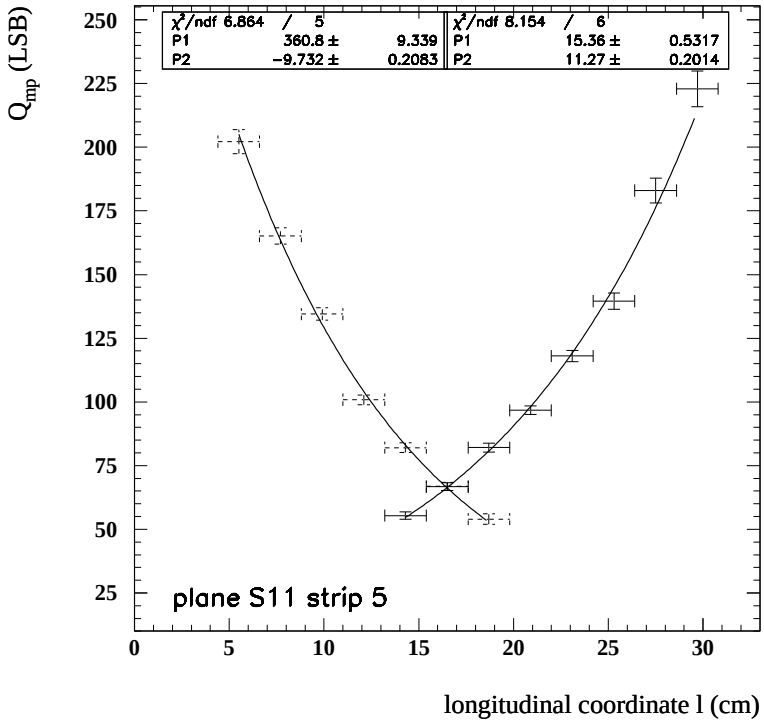

Figure 5.5: most probable charge signal Q_{mp} for the two PMT of strip 5, layer S11, as a function of the longitudinal coordinate l, for the selected pads; the two corresponding exponential fits are superimposed. Note that the selected pads are different for the two PMT, as explained in the text. The attenuation length is given by $\lambda_{att} = |P_2|$ (cm).

For this reason the pads have been considered in the fit only if the accumulated statistics was significant and the Landau distribution sufficiently well fitted: pads with less than ≈ 80 events or with a χ^2/d.o.f. of the fitted Landau greater than 3 have been rejected, thus excluding the most external pads on both ends of each strip. For the same reason the most external strips of layer S11 (strip 1 and 8) have not been considered in this analysis, since only a fraction of their area is within the acceptance window of the apparatus.[9]

The results of the exponential fit for several representative strips are shown in figures 5.5 to 5.9. Each figure refers to one strip and, for the two PMT, reports the dependence of the measured Q_{mp} on the pad centre coordinate l, for the pads considered according to the selection cuts mentioned above. The two PMT are respectively distinguished by the continuous and dashed lines used to draw the error bars (the horizontal error bar represents the length of the pad). The figure also reports the two exponential functions

[9]The plane S1 has been designed to cover a greater area than the one matching the purely geometric acceptance of the spectrometer, to be able to study also particles of lower energies that significantly bend within the volume of the magnetic cavity of PAMELA; see also the discussion, in sect. 4.3, on the limit zenith angle θ_{min} for the calculation of the geometric factor.

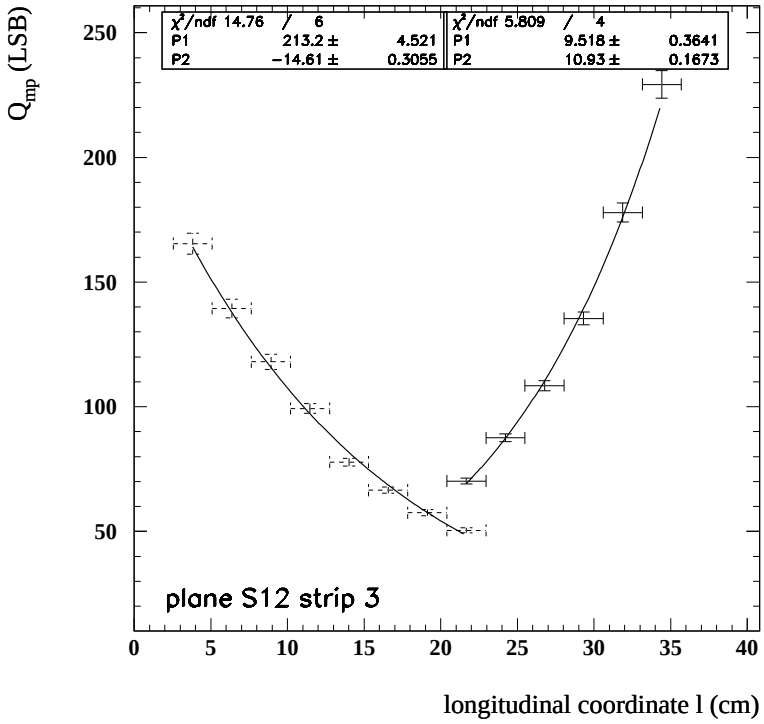

Figure 5.6: same as fig. 5.5 for strip 3 of layer S12.

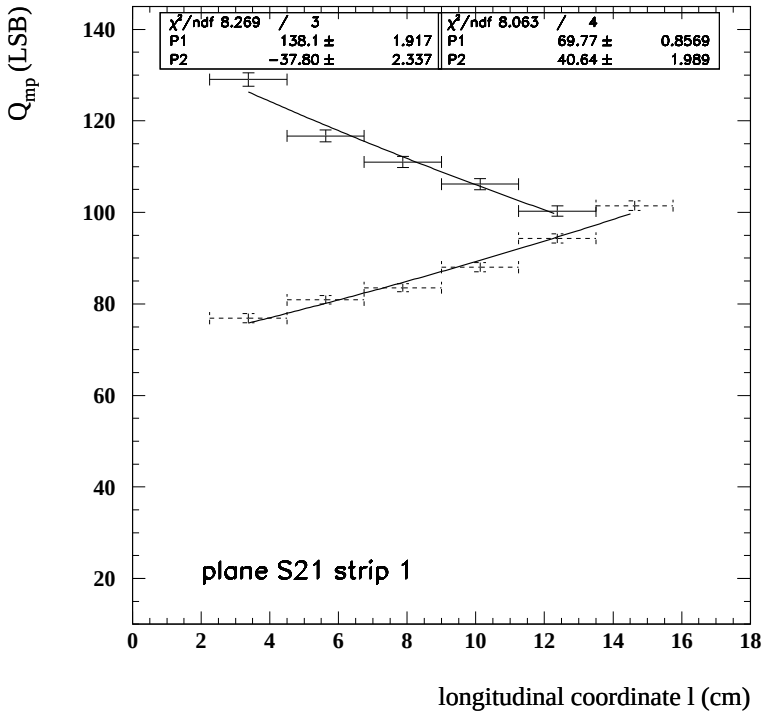

Figure 5.7: same as fig. 5.5 for strip 1 of layer S21.

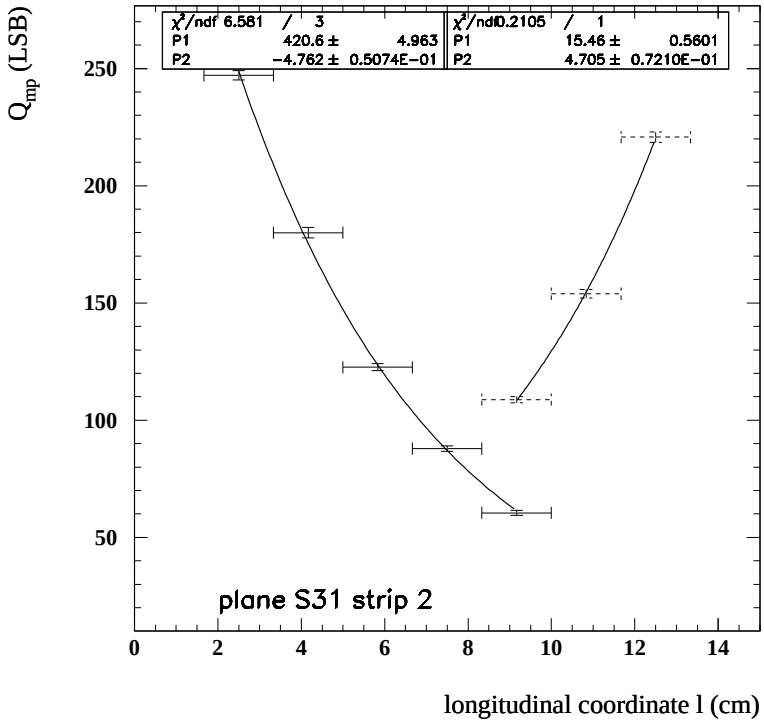

Figure 5.8: same as fig. 5.5 for strip 2 of layer S31.

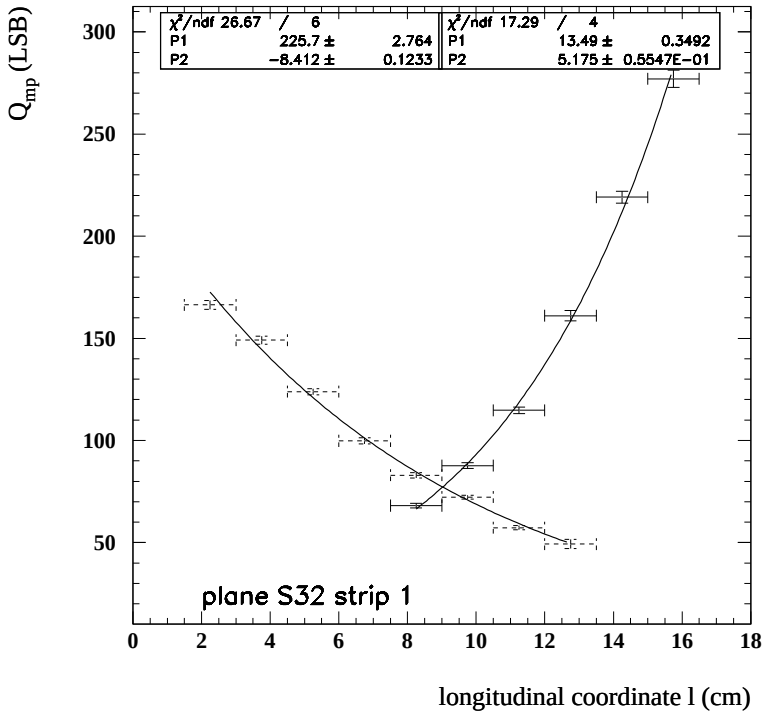

Figure 5.9: same as fig. 5.5 for strip 1 of layer S32.

fitted on the experimental points:

$$Q_{mp}(l) = P_1 \exp(\frac{l}{P_2}) \tag{5.4}$$

relating Q_{mp} (in LSB) to the longitudinal coordinate l along the strip (x or y in the reference of PAMELA, depending on the orientation of the strip) expressed in cm. The parameters of the fit, characteristic of the channel considered, are P_1 (LSB) and P_2 (cm), whose absolute value is the effective attenuation length ($\lambda_{att} = |P_2|$); note that P_2 is negative for the PMT positioned at the end of the strip corresponding to lower values of the longitudinal coordinate (PMT 1), positive for the PMT at the opposite end of the strip (PMT 2).

Note also the variation in the absolute response of the two PMT of a given strip, due to the intrinsic differences in the characteristics of the optical couplings and signal amplification chain, as already mentioned.

Tables 5.2, 5.3 and 5.4 report all the results for the considered strips and PMT.

Some comments should be made at this point.

As expected, the values of λ_{att} obtained with the independent fits for the two PMT of a given strip show differences of the order of 10% that are generally larger than the statistical uncertainties estimated from the fit; this can be explained with the fact that the hypothesis of exponential dependence of the intensity of guided light on the travelled distance is a good but not exact approximation of the true dependence. It can also be observed that for the same reason the values of χ^2/d.o.f. show quite significant variations.

Hence the simple average of the two measurements has been assumed as a more realistic estimate of λ_{att} (indicated as "mean" in the tables), with a systematic error, related to the exponential approximation, equal to the absolute difference between this average and each of the two values given by the fit. In the few cases for which this absolute difference is smaller than the statistical errors from the fit, the higher of the two statistical uncertainties has been conservatively taken as error in the λ_{att} measurement.

Comparing the measured λ_{att} for the various strips, it can be seen that strips of the same layer show quite similar values, while large differences are observed between different layers. In particular the central layers (S21 and S22) have the best values of attenuation length (greater than 30 cm), at least a factor ≈ 3 higher than for the S11 and S12 layers (about 10 cm) and for S31 and S32 (between 5 and 7 cm). On the other hand all these values are much smaller than the bulk attenuation length of the scintillator BC-404 (140 cm).

It can be concluded that the surface attenuation is the major contribution to λ_{att} for

layer	strip	PMT	att. length (cm)	error (cm)
S11	2	PMT 1	7.98	0.77
		PMT 2	10.62	0.36
		mean	9.3	1.3
	3	PMT 1	10.39	0.50
		PMT 2	9.21	0.20
		mean	9.80	0.59
	4	PMT 1	7.80	0.38
		PMT 2	8.64	0.11
		mean	8.22	0.42
	5	PMT 1	9.73	0.20
		PMT 2	11.27	0.20
		mean	10.50	0.76
	6	PMT 1	9.91	0.22
		PMT 2	8.54	0.26
		mean	9.23	0.68
	7	PMT 1	9.75	0.35
		PMT 2	8.44	0.28
		mean	9.10	0.65
S12	1	PMT 1	6.2	1.8
		PMT 2	8.80	0.90
		mean	7.5	1.3
	2	PMT 1	9.67	0.32
		PMT 2	7.32	0.46
		mean	8.5	1.1
	3	PMT 1	14.61	0.30
		PMT 2	10.93	0.16
		mean	12.8	1.8
	4	PMT 1	8.40	0.30
		PMT 2	12.20	0.21
		mean	10.3	1.9
	5	PMT 1	7.80	0.37
		PMT 2	7.46	0.15
		mean	7.63	0.37
	6	PMT 1	8.8	1.4
		PMT 2	8.4	3.4
		mean	8.6	3.4

Table 5.2: measured attenuation lengths for layers S11 and S12.

layer	strip	PMT	att. length λ (cm)	error (cm)
S21	1	PMT 1	37.8	2.3
		PMT 2	40.6	1.9
		mean	39.2	2.3
	2	PMT 1	-	-
		PMT 2	29.7	2.0
		mean	-	-
S22	1	PMT 1	-	-
		PMT 2	60	11
		mean	-	-
	2	PMT 1	36.2	2.1
		PMT 2	43.4	2.5
		mean	39.8	3.5

Table 5.3: measured attenuation lengths for layers S21 and S22. The exponential fit for two PMT has not been done because, with the available statistical sample, the corresponding number of pads, selected according to the cuts discussed in the text, was less than 3.

layer	strip	PMT	att. length λ (cm)	error (cm)
S31	1	PMT 1	5.506	0.050
		PMT 2	6.91	0.12
		mean	6.21	0.70
	2	PMT 1	4.762	0.050
		PMT 2	4.705	0.072
		mean	4.733	0.072
	3	PMT 1	4.710	0.089
		PMT 2	5.510	0.054
		mean	5.11	0.40
S32	1	PMT 1	8.41	0.12
		PMT 2	5.175	0.055
		mean	6.8	1.6
	2	PMT 1	5.438	0.087
		PMT 2	8.01	0.22
		mean	6.7	1.2
	3	PMT 1	8.73	0.44
		PMT 2	5.748	0.088
		mean	7.2	1.4

Table 5.4: measured attenuation lengths for layers S31 and S32.

the TOF strips of PAMELA, and that the optical properties of the strips differ significantly for the 3 double planes; in particular for the layers of S1 and S3 they are much worse than what would be expected on the basis of the standard processes that are usually employed in the fabrication of scintillator strips.

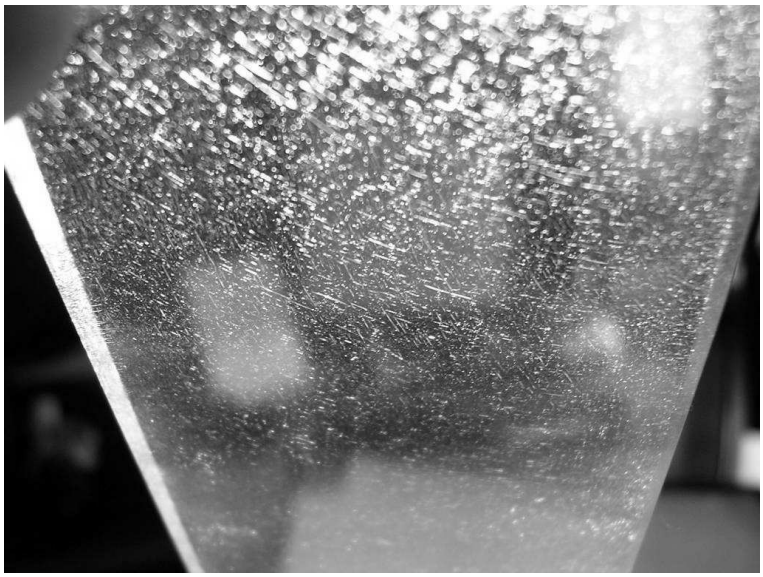

Figure 5.10: a detail of an "engineering model" strip of the TOF system with micro-cracks; the width of the strip (horizontal dimension in figure) is about 5 cm.

The reason for this unexpected worsening, that was pointed out by the present analysis, has been investigated by the PAMELA collaboration. The most probable explanation takes into account that an additional machining process was performed on the strips of S1 and S3 to reduce their thickness from the 10 mm standard to the desired 7 mm one; on the other hand, for the S2 strips the 5 mm thickness was already made available by the manufacturer (Bicron): the hypothesis is that during the additional processing a deterioration of the scintillator surfaces happened. Only with this dedicated analysis, during the final qualification tests, the effect was discovered; a subsequent study of the available "engineering model" spare strips, used for the first prototyping phase of the TOF system, has reported the presence of micro-cracks (fig. 5.10) on the strip surface, that are known to worsen the reflectivity properties; since these strips have been machined with the same procedure used for the "flight model" ones, it is reasonable to assume the presence of micro-cracks also on the "flight model" strips.

Anyway after having reported these not satisfying results on the attenuation lengths, a detailed analysis of all the data collected during the final qualification tests has been done, to verify if the characteristics of the TOF system, from the point of view of the trigger efficiency and of the time resolution, were still compatible with the requirements of the design. In particular the analysis of the trigger efficiencies for all the strips of the TOF system is described in the following sections.

5.4 Trigger efficiency

For the determination of the trigger efficiency of the strips and layers of the TOF system I have done an analysis similar to the one described in the previous sections, using cosmic-ray events collected by the PAMELA apparatus during the final qualification tests in Rome and Samara and exploiting the spatial information given by the tracking system of PAMELA.

The efficiency ϵ of a strip is defined as the probability that an incident particle, crossing the strip within the acceptance and live time of the PAMELA apparatus, is accompanied by the generation of an *event signal* toward the trigger board by *at least one* of the two PMT read-out channels associated with that strip.[10] An analogous definition holds for each of the 6 layers of the TOF system, considering the generation of an event signal from at least one of the corresponding PMT channels.

The overall efficiency of the TOF system for a specific trigger configuration[11] can then be derived from the efficiencies of the single scintillator layers, given that each layer and the corresponding electronics independently reacts to the passage of an ionizing particle.

The efficiencies of the various strips and layers can be determined with the analysis of the data collected by the PAMELA apparatus in the standard trigger configuration:

$$\text{(S11 OR S12) AND (S21 OR S22) AND (S31 OR S32)} \tag{5.5}$$

where the generation of a trigger happens in presence of an *event signal* in correspondence of at least one of the layers of each pair. In fact, with the redundant double-plane structure, characteristic of the TOF system, and with this specific trigger pattern, it is possible to detect and characterize also particles for which a given layer has not generated an event signal (because its trigger efficiency is less than 1); at the same time it is possible to use the information from these events and from those for which the same layer has produced an event signal, to measure the efficiency of the layer.

The trigger efficiency depends in particular on the distribution of the intensity of scintillation light produced by an ionizing particle crossing the TOF layers, varying in general with the particle species and rigidity, according to the Bethe-Bloch formula as mentioned in sect. 5.2; the smallest value of efficiency is reached for minimum ionizing particles (MIP), which on the other hand constitute the events of primary interest for the PAMELA mission (antiprotons with $|\rho|$ up to ≈ 190 GV/c, positrons with $|\rho|$ up to

[10] The operation of the TOF FE electronics for the generation of the event signals is described in sect. 1.5.1.

[11] The various available trigger patterns for PAMELA have been presented in sect. 1.6.

≈ 270 GV/c).

In order to understand the real performances of the trigger system, the efficiencies have been measured by using the cosmic-ray events acquired in Rome and Samara and applying a cut on the rigidity, as discussed in sect. 5.2, to assure that the selected events (substantially all muons) are MIP.

In general the measured efficiency of a scintillator strip or layer depends on the spatial and angular distribution of the incident particles used in the analysis (usually indicated as *conditions of illumination*), because of the not completely uniform sensitivity of the detector; therefore the efficiencies measured with the present method, using the cosmic-ray MIP at earth, contain a systematic discrepancy with respect to the efficiencies that will be measured during the extra-atmospheric flight of the apparatus and used for the determination of the absolute differential fluxes of particles. In any case it can be reasonably assumed that the efficiencies measured here are a good approximation of those expected in flight, since the angular dependence of the cosmic-ray flux at earth does not differ too much from that expected outside the atmosphere, which is practically isotropic.

It must also be noted that the trigger efficiency of a specific TOF PMT channel depends on the various parameters that affect the height of the anode pulse at the PMT output and on the low and high thresholds of the discriminator generating the corresponding event signal. In particular the height of the anode pulse depends on the characteristic attenuation length of the strip read-out by the PMT; this effect introduces a dependence of the efficiency of the PMT on the longitudinal coordinate of incidence along the strip.

When considering the combined efficiency of the two PMT of a strip, this dependence is greatly reduced, on one hand because of the complementary paths travelled by the scintillation light from the point of production to the two PMT at the opposite ends of the strip, on the other hand because the adjustable parameters for each PMT signal channel have been set to take into account the effect of the light attenuation. The adjustable parameters are, as already mentioned in other points of the discussion, the high voltage for the PMT, affecting the gain of the electron multiplier structure, and the low and high thresholds for the discrimination of the anode output of the PMT.

The time available during the final qualification tests of the apparatus has been divided in such a way that several different combinations of PMT high voltage, threshold settings and also trigger patterns have been tried, thus testing tens of possible configurations of the system.

The analysis has been focused on the most significant data samples, acquired with

the standard high voltage configuration and trigger pattern (5.5), and studying the dependence of the trigger efficiency on the threshold setting, in particular with all the low and high thresholds set at their minimum (10 mV for the low ones and 30 mV for the high ones, indicated as "configuration 1") and at significantly higher values ("configuration 2": all the low ones at 33 mV and the high ones at 100 mV); these two samples are actually a subset of the events considered in the analysis of the attenuation lengths. The "flight" threshold configuration was set at an intermediate point between these two ones, to obtain the best compromise between a high efficiency and a small number of spurious event signals due to noise: 23 mV and 70 mV for the low and high thresholds of most PMT except for a group of 4 particularly noisy PMT, with thresholds set at 33 mV and 100 mV respectively. Unfortunately, up to now, there was not the possibility of doing long term data acquisition with this configuration, so that in the following discussion only the analysis for the previously mentioned two configurations will be presented.

5.5 Data analysis for trigger efficiency

By operating with a method similar to the one discussed in sect. 5.3 for the study of the attenuation lengths, a set of general selection cuts has been applied to the initial samples of acquired events (≈ 2300 and ≈ 8860 for the threshold configurations 1 and 2 respectively) for which the tracking program has reconstructed either a unique physical trajectory or two mutually excluding trajectories, because of the Y coordinate ambiguity.

The analysis of the data has been repeated for each of the 6 TOF layers, to separately determine the trigger efficiencies for the single strips of the layer and for the whole layer.

Once the layer under test has been fixed, the following cuts have been applied to determine the physical trajectory:

- **cut 1**: for each of the TOF layers segmented along the Y axis (and different from the layer under test), one and only one strip has generated an event signal for the trigger board on one or both the corresponding channels (PMT);

- **cut 2**: one and only one of the two reconstructed tracks crosses the (2 or 3) strips identified in the previous cut.

For the selection of MIP muons the following cut has been applied:

- **cut 3**: the rigidity of the incident particle, measured by the tracking system, is such that $|\rho| > 0.5$ GV/c.

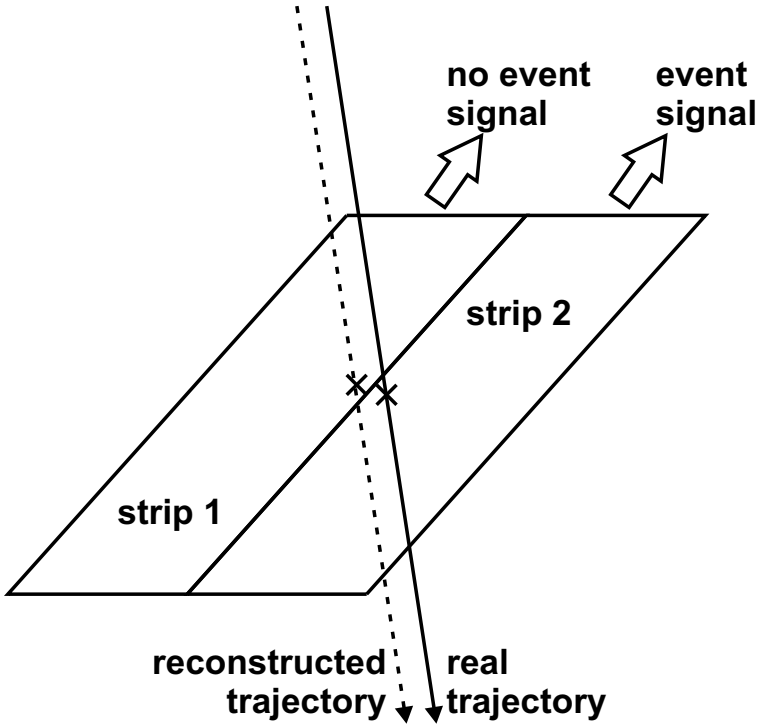

Figure 5.11: systematic error in the measurement of the efficiency of a TOF strip, caused by the discrepancy between the nominal and true position of the strip in the PAMELA reference system. The particle has crossed strip 2 which has produced an event signal on at least one of its two PMT channels, but the reconstructed trajectory is wrongly associated to the adjacent strip (strip 1) which has not generated any event signal; this causes an underestimation of the true efficiency for strip 1.

A further cut has been used to assure that the selected sample contains only events acquired without any dependence on whether or not the layer under test has produced an event signal:

- **cut 4**: given the layer under test (e.g. S11), an event signal has been generated by the other layer of the same double plane (e.g. S12).

This cut exploits the fact that, with the standard trigger configuration (5.5), the presence of an event signal for the other layer of the same double plane (e.g. S12) implies that the trigger has been generated independently from the behaviour of the layer under test (e.g. S11). If this cut were not applied, the sample would contain a certain number of events that have been acquired *because* the layer under test has produced an event signal (necessary to have the trigger in absence of event signal for the other layer of the same double plane), thus causing an overestimation of the true trigger efficiency.

The procedure for the measurement of the trigger efficiency for each single strip of the layer under test is described below; the determination of the efficiency of the whole layer

is analogous, with some differences that will be pointed out when discussing the results of the analysis.

For each strip under test two values are recorded in the analysis: the number N of events with a reconstructed physical trajectory crossing the strip and the fraction N_{ok} of the N events for which the strip has given an event signal. The best estimate of the corresponding trigger efficiency is then given by the ratio N_{ok}/N; the statistical uncertainty associated with this measurement will be separately discussed in app. A.

The measurement of efficiency for a TOF strip is potentially affected by greater systematic uncertainties, with respect to the ones present in the determination of the corresponding attenuation length. In fact, the identification of the strip that has been crossed by the incident cosmic ray is based on the trajectory reconstructed by the tracking program in the PAMELA reference system, which is operatively defined in a consistent way within the reconstruction procedure itself; the obtained trajectory is then compared with the nominal positions of the TOF strips in this reference system. These nominal positions are known with an accuracy of the order of the mm, which takes into account the combined contributions of the mechanical tolerances of the various elements connecting each strip to the magnetic tower of the spectrometer. If the discrepancy between the true position of the strip and the nominal one is such that for a specific event (see fig. 5.11) the reconstructed trajectory is not associated with the strip actually crossed (strip 2), but with an adjacent strip (strip 1) which correctly has not given any event signal, this implies an underestimation of the efficiency measured for strip 1. The resulting relative systematic uncertainty on the efficiency of strip 1 is of the order of the fraction of events with trajectories external to the strip and within a distance δ from the strip border, where δ (~ 1 mm) is the spatial displacement of the strip from its nominal position. This fraction of events, on the other hand, is of the order of the ratio between δ and the transverse width of a strip (between 5 and 9 cm, depending on the layer), thus being $1 - 2\%$, comparable with the statistical uncertainties obtained with the present measurement, as discussed below.

To avoid introducing this systematic error in the measurements, a fiducial cut has been applied on the nominal area of each considered strip, excluding events for which the trajectory reconstructed by the tracking program crosses the strip in the vicinity of its lateral walls (within a distance δ_{cut}). By studying the variation of the measured efficiencies with δ_{cut} it can be concluded that a value $\delta_{cut} = 2$ mm is sufficient to remove this systematic contribution for all the strips; this applied cut is consistent with the

previously estimated order of magnitude for the displacement δ.

An acceptable drawback of the application of this further selection cut is a reduction of the order of some percent of the available statistical sample. It must also be considered that the excluded external area of the strip is expected to be characterized by a smaller trigger efficiency with respect to the rest of the area, because particles with inclined trajectories that cross the strip near its external border can escape from the lateral wall before reaching the lower face of the strip, thus releasing a significantly smaller amount of energy within the strip volume.

A more refined analysis, avoiding the application of this fiducial cut on the area, will be done after the accumulation of a large statistical sample of events, which will make it possible to align the strip with respect to the PAMELA reference system, with the precise determination of its true position.

The efficiencies ϵ measured with the present method for the single strips of the 6 layers, in the two considered threshold configurations, are shown in figures from 5.12 to 5.17; the complete results of the analysis are summarized in tables 5.5 (configuration 1) and 5.6 (configuration 2).

It can be generally observed that the efficiencies, taking into account the statistical uncertainties, are compatible with values greater than 98% for almost all the strips and for both configurations, and for most strips are above 99%; in particular the application of higher thresholds (configuration 2) does not generally cause a significant decrease in the measured values of ϵ. Besides, strips with different attenuation lengths λ_{att} do not show a corresponding variation of ϵ, as expected on the basis on the considerations made in the previous section (compare for example ϵ for the strips of S2 and S3, characterized by λ_{att} of more than 30 cm and less than 7 cm respectively).

Note also that the statistical uncertainties are much greater for layers S11 and S12; this is clearly a consequence of the fact that these layers are divided into a larger number of strips (8 and 6 respectively) with respect to the other ones, and therefore the number of events collected on a single strip is correspondingly smaller.

As mentioned before, a similar analysis has been done also for the 6 TOF layers, each one considered as a whole. The same event selection cuts described before have been applied, with the only difference that in the definition of the fiducial area, only the external part of the layer has been excluded and not the region between adjacent strips within the layer; in fact this is sufficient to take into account the possible systematic error caused by the use of the nominal position of the layer in the PAMELA reference.

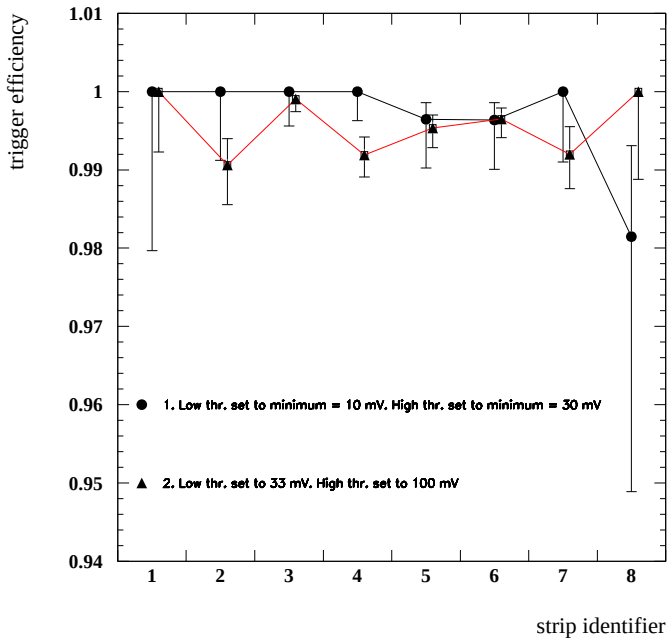

Figure 5.12: measured trigger efficiencies for the 8 strips of layer S11.

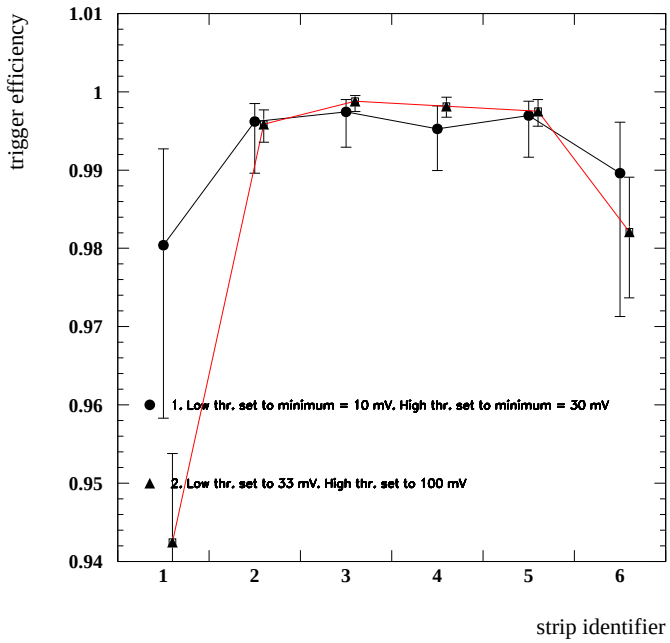

Figure 5.13: measured trigger efficiencies for the 6 strips of layer S12.

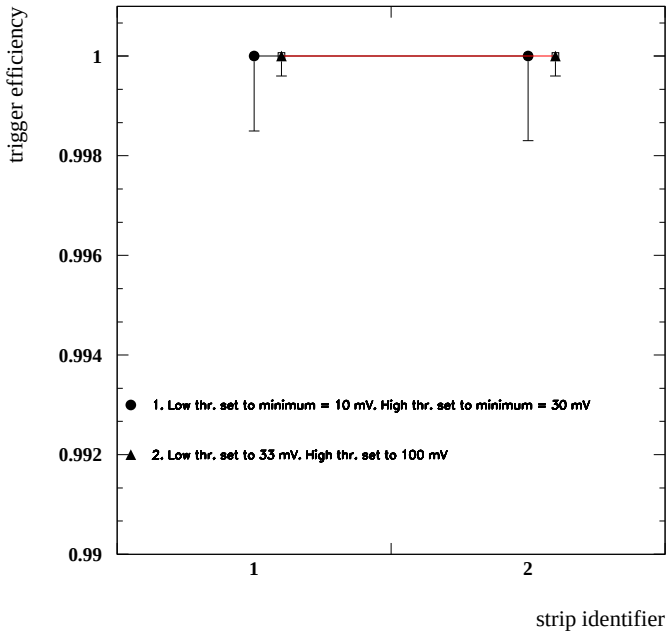

Figure 5.14: measured trigger efficiencies for the 2 strips of layer S21.

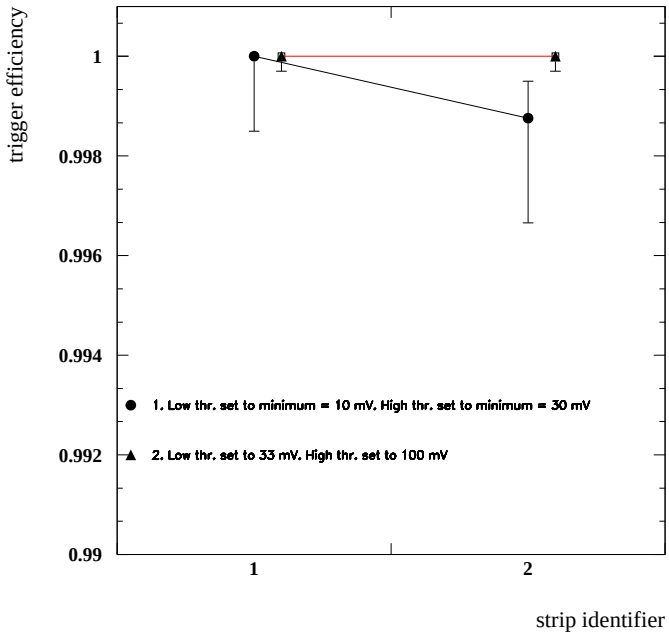

Figure 5.15: measured trigger efficiencies for the 2 strips of layer S22.

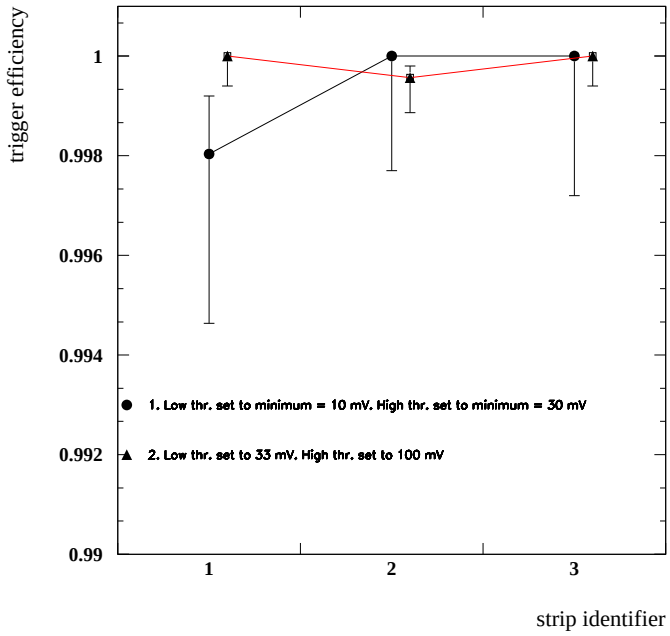

Figure 5.16: measured trigger efficiencies for the 3 strips of layer S31.

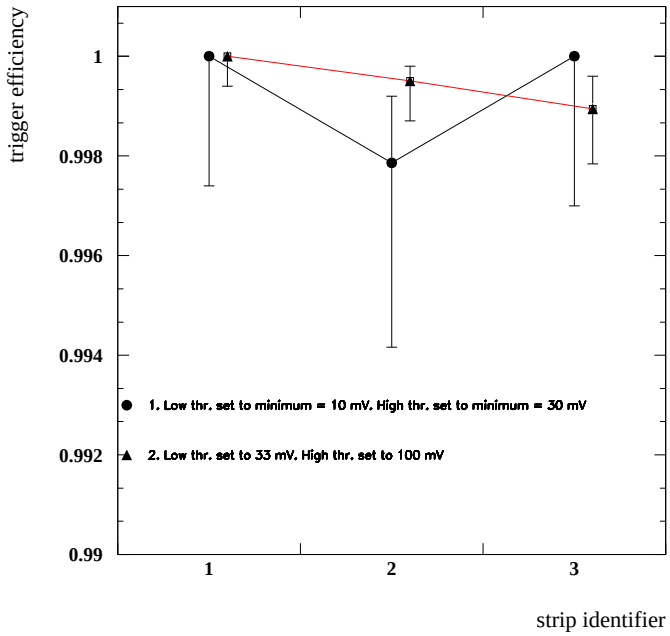

Figure 5.17: measured trigger efficiencies for the 3 strips of layer S32.

layer	strip	N_{ok}	N	ϵ	$\Delta\epsilon_-$	$\Delta\epsilon_+$
S11	1	62	62	1	0.020	0
	2	145	145	1	0.0088	0
	3	286	286	1	0.0044	0
	4	339	339	1	0.0037	0
	5	281	282	0.9964	0.0062	0.0021
	6	274	275	0.9963	0.0063	0.0022
	7	141	141	1	0.009	0
	8	53	54	0.981	0.032	0.011
S12	1	100	102	0.980	0.022	0.012
	2	261	262	0.9961	0.0065	0.0023
	3	387	388	0.9974	0.0045	0.0015
	4	419	421	0.9952	0.0053	0.0029
	5	328	329	0.9969	0.0052	0.0018
	6	95	96	0.9895	0.0183	0.0065
S21	1	821	821	1	0.0015	0
	2	730	730	1	0.0016	0
S22	1	853	853	1	0.0015	0
	2	802	803	0.9988	0.0021	0.0007
S31	1	508	509	0.9980	0.0033	0.0011
	2	541	541	1	0.0022	0
	3	453	453	1	0.0027	0
S32	1	495	495	1	0.0026	0
	2	467	468	0.9978	0.0037	0.0013
	3	421	421	1	0.0029	0

Table 5.5: measured efficiencies for the various TOF strips in the threshold configuration 1.

For a fixed layer under test the trigger efficiency has been estimated as the ratio N_{ok}/N, where N is the number of events with a reconstructed trajectory crossing the layer and N_{ok} is the fraction of these events for which *any* strip of the layer has given an event signal. The ϵ measured with this method, for the 6 layers, are summarized in table 5.7 (configuration 1) and 5.8 (configuration 2).

Note that the number of events N, accumulated for each layer, is greater than the sum of the events used for the measurement of ϵ for the single strips; this was expected since the fiducial area of the layer, considered in this analysis, is larger than the sum of the fiducial areas of the single strips forming the layer. Note also that the values of N are comparable for the various layers, since the initial sample is composed of events with a particle crossing all the three double planes (standard trigger pattern).

We can observe that the absence, in this case, of a fiducial cut on the area along the boundary between adjacent strips, allows the inclusion in the selected sample also

layer	strip	N_{ok}	N	ϵ	$\Delta\epsilon_-$	$\Delta\epsilon_+$
S11	1	165	165	1	0.0077	0
	2	635	641	0.9906	0.0050	0.0033
	3	1041	1042	0.9990	0.0016	0.0006
	4	1346	1357	0.9918	0.0028	0.0023
	5	1286	1292	0.9953	0.0025	0.0016
	6	1137	1141	0.9964	0.0024	0.0014
	7	620	625	0.9920	0.0044	0.0035
	8	113	113	1	0.011	0
S12	1	344	365	0.942	0.013	0.011
	2	1200	1205	0.9958	0.0022	0.0018
	3	1640	1642	0.9988	0.0013	0.0007
	4	1618	1621	0.9981	0.0013	0.0011
	5	1197	1200	0.9975	0.0018	0.0015
	6	384	391	0.9820	0.0083	0.0070
S21	1	3103	3103	1	0.0004	0
	2	3113	3113	1	0.0004	0
S22	1	3337	3337	1	0.0003	0
	2	3285	3285	1	0.0003	0
S31	1	2032	2032	1	0.0006	0
	2	2298	2299	0.9996	0.0007	0.0002
	3	2015	2015	1	0.0006	0
S32	1	1925	1925	1	0.0006	0
	2	2024	2025	0.9995	0.0008	0.0003
	3	1890	1892	0.9989	0.0011	0.0007

Table 5.6: measured efficiencies for the various TOF strips in the threshold configuration 2.

of events characterized by a particle passing in the vicinity of this boundary, and, if sufficiently inclined, sharing its path on both strips, with a smaller energy release for each; the inclusion of these events gives a more correct evaluation of the actual performance of the layer in the generation of the event signal.

On the other hand, it must be considered that the adopted definition of efficiency for the layer under test implies that also events with a signal, typically due to noise, associated to a strip that has *not* been crossed by the particle, contribute to the calculated N_{ok}; also in this case the definition takes into account the actual behaviour of the system, since the presence of a spurious event signal on a layer, with a reconstructed physical trajectory and with event signals on at least two other TOF double planes (in the standard trigger configuration) is not sufficient to conclude that the event must be considered as spurious.

The measured efficiency ϵ for a given layer is thus different from the value ϵ' expected considering the ratio between the sum of the N_{ok} values for the single strips of that layer

layer	N_{ok}	N	ϵ	$\Delta\epsilon_-$	$\Delta\epsilon_+$
S11	1732	1734	0.9988	0.0012	0.0007
S12	1710	1711	0.9994	0.0010	0.0003
S21	1597	1597	1	0.0008	0
S22	1702	1705	0.9982	0.0012	0.0010
S31	1579	1580	0.9994	0.0011	0.0003
S32	1476	1482	0.9959	0.0022	0.0014

Table 5.7: measured efficiencies for the 6 TOF layers in the threshold configuration 1.

layer	N_{ok}	N	ϵ	$\Delta\epsilon_-$	$\Delta\epsilon_+$
S11	6897	6913	0.9977	0.0006	0.0005
S12	6883	6923	0.9942	0.0009	0.0009
S21	6408	6413	0.9992	0.0004	0.0003
S22	6826	6838	0.9982	0.0006	0.0005
S31	6670	6675	0.9993	0.0004	0.0002
S32	6171	6181	0.9984	0.0006	0.0005

Table 5.8: measured efficiencies for the 6 TOF layers in the threshold configuration 2.

and the sum of the corresponding selected events N:

$$\epsilon' = \frac{\sum_{\text{strips}} N_{ok}^{\text{strip}}}{\sum_{\text{strips}} N^{\text{strip}}} \tag{5.6}$$

In fact the two effects (different fiducial cut and inclusion of "noisy" events) tends respectively to decrease and increase ϵ respect to ϵ'.

As a general conclusion of this analysis it can be stated that the high efficiencies measured for the 6 TOF layers are quite satisfying and do not justify the replacement of the detectors, in spite of their bad performances in terms of attenuation lengths; similar conclusions have been drawn from an independent analysis on the time resolution of the system [5].

For this reason, during the final qualification tests of PAMELA, taking also into account the strict limits in the available time before starting the integration of the apparatus within the Resurs-DK1 satellite, the collaboration decided not to replace the TOF detectors.

5.6 Trigger efficiency of the TOF system in the "flight" configuration

After having directly measured, with the method described above, the efficiencies of the single layers of the TOF system, it is possible to calculate the trigger efficiency of the single double planes (ϵ_{S1}, ϵ_{S2}, ϵ_{S3}) and the overall efficiency of the TOF system (ϵ_{trigg}) in the "flight" trigger configuration:

(S11 OR S12) AND (S21 OR S22) AND (S31 OR S32)

Note that a direct measurement of the trigger efficiency of the single double planes is only possible by acquiring events with an active trigger configuration that excludes the double plane under test[12], while a direct measurement of ϵ_{trigg} in the "flight" configuration is clearly impossible and its calculation, starting from the ϵ of the single TOF detectors, is unavoidable.

The dependence of ϵ_{S1}, ϵ_{S2}, ϵ_{S3} and ϵ_{trigg} on the measured ϵ for the single layers can be derived as follows. Given the efficiencies of the two layers of a double plane, generically indicated here as $\epsilon_A = P(A)$ and $\epsilon_B = P(B)$, the combined efficiency of the double plane is defined as $\epsilon_{OR} = P(A \text{ OR } B)$; taking into account the independence of the two probabilities ϵ_A and ϵ_B, the following formula can be derived:

$$\epsilon_{OR} = 1 - (1 - \epsilon_A) \cdot (1 - \epsilon_B) = \epsilon_A + \epsilon_B - \epsilon_A \cdot \epsilon_B \tag{5.7}$$

For plane S1, composed of layers S11 and S12, we can write:

$$\epsilon_{S1} = \epsilon_{S11} + \epsilon_{S12} - \epsilon_{S11} \cdot \epsilon_{S12} \tag{5.8}$$

Similar relations hold for the other two planes S2 and S3.

Finally the total trigger efficiency of the TOF system in the standard trigger configuration is obtained, given the independence of the probabilities ϵ_{S1}, ϵ_{S2} and ϵ_{S3}, as:

$$\epsilon_{trigg} = P(\text{S1 AND S2 AND S3}) = \epsilon_{S1} \cdot \epsilon_{S2} \cdot \epsilon_{S3} \tag{5.9}$$

and substituting ϵ_{S1} with (5.8) etc..

Before specifying how the uncertainties for ϵ_{S1}, ϵ_{S2}, ϵ_{S3} and ϵ_{trigg} have been determined, it can be noted that the values of ϵ_{S1}, ϵ_{S2}, ϵ_{S3} and ϵ_{trigg} obtained with this calculation, starting from the measured ϵ in tables 5.7 and 5.8, differ from 1 by a quantity $\delta\epsilon$ that is at

[12]The only long term acquisitions during the final tests in Rome and Samara were done with the standard trigger configuration.

plane	ϵ	$\Delta\epsilon_-$
S1	1	0.0007
S2	1	0.0008
S3	1	0.0008
TOF (standard)	1	0.0008

plane	lower limit (68% c. l.)
S1	$\epsilon > 0.9993$
S2	$\epsilon > 0.9992$
S3	$\epsilon > 0.9992$
TOF (standard)	$\epsilon > 0.9992$

Table 5.9: calculated efficiencies for the 3 double planes and the overall TOF system with the standard trigger pattern, in the threshold configuration 1.

plane	ϵ	$\Delta\epsilon_-$
S1	1	0.0001
S2	1	0.0002
S3	1	0.0002
TOF (standard)	1	0.0002

plane	lower limit (68% c. l.)
S1	$\epsilon > 0.9999$
S2	$\epsilon > 0.9998$
S3	$\epsilon > 0.9998$
TOF (standard)	$\epsilon > 0.9998$

Table 5.10: same as table 5.9 in the threshold configuration 2.

least one order of magnitude smaller than the uncertainty $\Delta\epsilon_-$ in the direct measurements of ϵ for the single layers ($\delta\epsilon < 2 \cdot 10^{-5}$ in all cases and for both configurations). Therefore we can reasonably assume that also the uncertainty associated with the calculated values ϵ_{S1}, ϵ_{S2}, ϵ_{S3} and ϵ_{trigg} is much larger than $\delta\epsilon$ and hence approximate the calculated efficiencies with 1 and indicate a conservative confidence interval $\Delta\epsilon_-$ (at the 68% level). Alternatively a conservative lower limit ϵ_{low} at the 68% confidence level can be defined, given by $\epsilon_{low} = 1 - \Delta\epsilon_-$.

To estimate this confidence interval (or limit) it can be observed that in a hypothetical direct measurement of one among ϵ_{S1}, ϵ_{S2}, ϵ_{S3} and ϵ_{trigg}, with a sample of N events, the fraction of events with a positive outcome, N_{ok}, would follow a binomial distribution, as in the case of the fractions of positive events for the single layers. Then we can assume, conservatively, that the calculation performed here of, for example, ϵ_{trigg}, is equivalent to a direct measurement $N_{ok}^{meas} = N$ (giving as best estimate for ϵ_{trigg} the value 1 obtained from the calculation), with N given by the *lowest* number among the 6 values of N associated to the direct measurements used in the calculation of ϵ_{trigg}; analogous assumptions can be made for ϵ_{S1} etc. , considering the values of N associated to the 2 corresponding layers.

With this approximation, and applying the method for the estimation of the uncertainties already used for the single layers and described in app. A, the values shown in tables 5.9 and 5.10 are obtained for $\Delta\epsilon_-$ (and for the lower limit ϵ_{low}). We can observe that the uncertainties $\Delta\epsilon_-$ are all greater than $1 \cdot 10^{-4}$ and are consistent with the initial hypothesis that $\delta\epsilon << \Delta\epsilon_-$.

Appendix A

Statistical uncertainty in the efficiency measurement

The present appendix contains a discussion of the method adopted in chapter 5 for the determination of the confidence intervals $\Delta\epsilon_-$ and $\Delta\epsilon_+$ in the measurement of the trigger efficiencies for the TOF system of PAMELA.

Considering for example a scintillator strip of the TOF system and a sample of N events, characterized by the presence of a charged particle crossing the strip, the fraction N_{ok}, for which the strip has produced an event signal, is expected to follow a binomial distribution:

$$B(N_{ok}, \epsilon) = \frac{N!}{N_{ok}! \cdot (N - N_{ok})!} \cdot \epsilon^{N_{ok}} \cdot (1 - \epsilon)^{(N - N_{ok})} \tag{A.1}$$

where the efficiency ϵ ($0 < \epsilon < 1$) is defined as the probability that the strip, crossed by an ionizing particle, generates an event signal. The binomial distribution is discrete, limited ($0 \leq N_{ok} \leq N$) and asymmetric with respect to the most probable value of N_{ok}, equal to the integer part of $\epsilon \cdot (N + 1)$.[1]

For a measured value N_{ok}^{meas}, the best estimate of the efficiency is given by $\epsilon_{meas} = N_{ok}^{meas}/N$, for which N_{ok}^{meas} is the most probable value of the corresponding distribution $B(N_{ok}, \epsilon_{meas})$.

The statistical uncertainty associated with the efficiency measurement is represented as a *confidence interval* at a specified confidence level (for example at the 68%), according to a general procedure first introduced by Neyman [53] [3]. The confidence interval includes all the physically allowed values ϵ, on both sides of the best estimate ϵ_{meas}, that satisfy the following condition:

- given the distribution of probability for N_{ok} with parameter ϵ, $B(N_{ok}, \epsilon)$, the mea-

[1] For $\epsilon = 1$, a degenerate distribution $B(N_{ok}, 1)$ is expected, which is 0 for $0 \leq N_{ok} < N$ and 1 for $N_{ok} = N$. The most probable value for N_{ok} is N in this degenerate case.

sured N_{ok}^{meas} is contained in the interval of values contributing to 68% of the total probability; this interval of values of N_{ok} is known as the *acceptance interval* for the considered ϵ.

Note that according to this definition the acceptance interval is not unique; a *coverage condition* must also be specified for its construction.

A usual practice for the determination of the confidence interval is to approximate the binomial distribution with a Poisson distribution (discrete and asymmetric, but with no upper limit) or more commonly with a Gaussian one (continuous, symmetric and unlimited):

$$G(N_{ok}, \epsilon) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp[-\frac{(N_{ok} - N_{ok}^{mp}(\epsilon))^2}{2 \cdot \sigma^2}] \tag{A.2}$$

where the parameters of the Gaussian distribution are given by[2]:

$$N_{ok}^{mp}(\epsilon) = \epsilon \cdot N \qquad \sigma(\epsilon) = \sqrt{N \cdot \epsilon \cdot (1 - \epsilon)} \tag{A.3}$$

With this approximation and with a symmetric coverage condition for the construction of the acceptance interval around $N_{ok}^{mp}(\epsilon)$, the obtained confidence interval (at the 68% level) is *centered* on ϵ_{meas} with length given by twice the quantity[3] (see also fig. A.1):

$$\Delta\epsilon = \frac{\sigma(\epsilon_{meas})}{N} = \sqrt{\frac{1}{N} \cdot \frac{N_{ok}^{meas}}{N} \cdot \left(1 - \frac{N_{ok}^{meas}}{N}\right)} \tag{A.4}$$

The advantage of the Gaussian approximation is the possibility of using this analytical relation to determine the confidence interval, avoiding a numerical calculation. On the other hand this simplified approach is physically consistent only when the set of values of N_{ok} characterized by a non-negligible probability, according to the Gaussian distribution $G(N_{ok}, \epsilon_{meas})$, is contained in the allowed interval $0 \leq N_{ok} \leq N$ (in a typical efficiency measurement the critical condition to be satisfied is $N_{ok} \leq N$, since $\epsilon_{meas} \approx 1$). This can be obtained by requiring, for example, that only 2.3% or less of the Gaussian-distributed values of N_{ok} are greater than N, or equivalently, for the properties of the Gaussian function, that $N_{ok}^{meas} + 2 \cdot \sigma(\epsilon_{meas}) \leq N$, which for $\epsilon_{meas} < 1$ implies:

$$N > 4 \cdot \frac{\epsilon_{meas}}{1 - \epsilon_{meas}} \tag{A.5}$$

while for $\epsilon_{meas} = 1$ the condition can never be satisfied.

[2]This approximation is discussed in standard data analysis textbooks; see for example [55]. Note that the parameter $N_{ok}^{mp}(\epsilon)$ is real.

[3]It is implicitly assumed that it is possible to neglect the variation of σ with ϵ within the confidence interval; this is correct in practical cases, for which $\Delta\epsilon$ is much smaller that ϵ_{meas}.

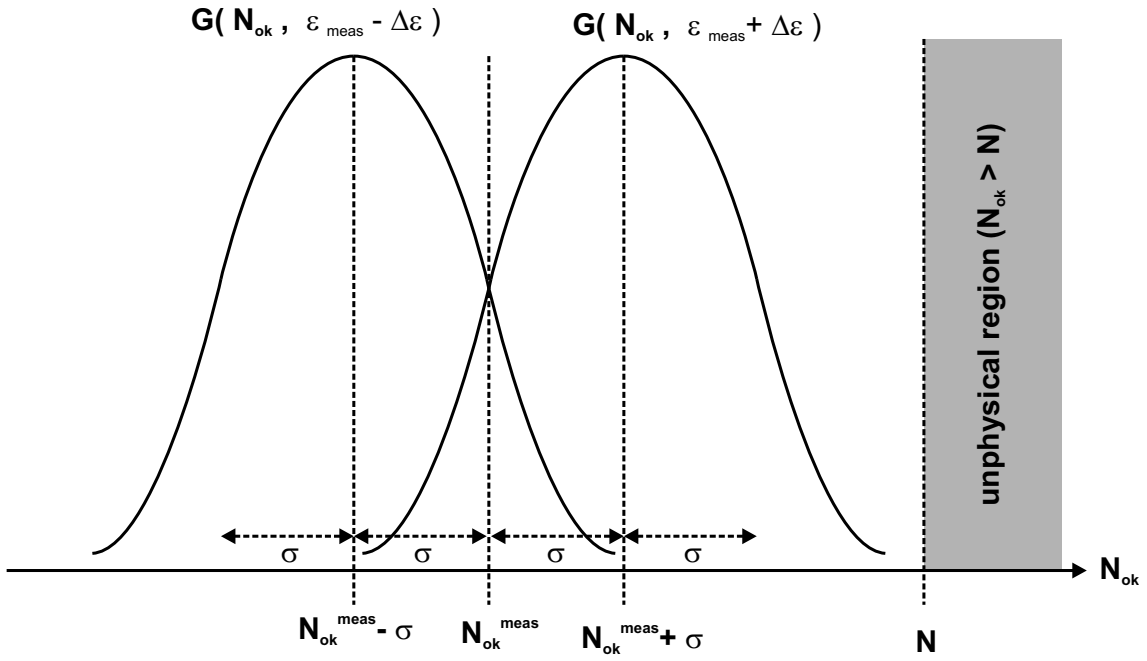

Figure A.1: determination of the 68% confidence interval for the Gaussian probability distribution (A.2). The distribution $G(N_{ok}, \epsilon_{meas} + \Delta\epsilon)$ (at right), with the corresponding acceptance interval (at the 68% level) centered in $N_{ok}^{meas} + \sigma$ and of amplitude twice the value of σ, characterizes the upper limit of the confidence interval, $\epsilon_{meas} + \Delta\epsilon$; a similar consideration holds for the lower limit $\epsilon_{meas} - \Delta\epsilon$.

In the analysis of trigger efficiencies discussed in chapter 5, there are many cases where $\epsilon_{meas} = 1$; in many other cases[4] $\epsilon_{meas} < 1$ but N is too small to comply with (A.5).

From the above observations it can be concluded that in this situation, with high measured efficiencies and small statistical samples, the approximation with an unlimited (Poisson or Gaussian) distribution does not guarantee a correct evaluation of the confidence interval and that a numerical calculation, starting from the binomial distribution, must be performed; in particular the obtained confidence interval is expected to be in general *not centered* around the best estimate of ϵ, because of the asymmetry of the binomial distribution, that is more pronounced for smaller values of N.

In the present work the general method of *unified intervals* outlined by Feldman and Cousins [54] (originally applied to a Gaussian and a Poisson distribution) has been used in the construction of the binomial acceptance interval for a given ϵ. According to this method the allowed values of N_{ok} ($0 \leq N_{ok} \leq N$) are included in the acceptance interval in decreasing order of the likelihood ratio:

$$R(N_{ok}, \epsilon) = \frac{B(N_{ok}, \epsilon)}{B(N_{ok}, \epsilon_{max}(N_{ok}))} \tag{A.6}$$

[4]For example strip 2 of layer S22 in the threshold configuration 1, with $\epsilon_{meas} = 0.9988$ and $N = 803$ (see table 5.5).

where $\epsilon_{max}(N_{ok}) = N_{ok}/N$ is the value of ϵ maximizing the binomial probability $B(N_{ok}, \epsilon)$ for a given N_{ok}. For a specified confidence level CL, the acceptance interval is completely filled when the sum of the probabilities over all the included values becomes $\geq$ CL.[5]

I have developed a dedicated FORTRAN routine that uses this construction criterion. The required inputs are:

- the numbers N and N_{ok}^{meas}, characterizing the efficiency measurement performed;
- the desired resolution $\delta\epsilon_{res}$ for the calculation of the upper and lower branch of the confidence interval, $\Delta\epsilon_{\pm}$;
- a range of values around ϵ_{meas} to be considered $(\epsilon_{min}, \epsilon_{max})$.[6]

The routine begins to scan the specified range starting from $\epsilon = \epsilon_{max}$ and decreasing by steps equal to $\delta\epsilon_{res}$ until an ϵ is found that belongs to the confidence interval (at the 68% level); then $\Delta\epsilon_+$ is given by $\epsilon - \epsilon_{meas}$. $\Delta\epsilon_-$ is calculated with an analogous procedure, starting from ϵ_{min}.

To determine, at a given step in the scan, whether the current ϵ belongs to the confidence interval, the corresponding acceptance interval is constructed according to the rule of Feldman and Cousins, using as ordering criterion the likelihood ratio

$$R(N_{ok}, \epsilon) = \frac{B(N_{ok}, \epsilon)}{B(N_{ok}, \frac{N_{ok}}{N})} = \frac{\epsilon^{N_{ok}} \cdot (1-\epsilon)^{(N-N_{ok})}}{\left(\frac{N_{ok}}{N}\right)^{N_{ok}} \cdot \left(1 - \frac{N_{ok}}{N}\right)^{(N-N_{ok})}} \tag{A.7}$$

and analogous expressions in the degenerate cases $\epsilon = 1$ or $N_{ok} = N$.

As a practical illustration of the method, consider a certain $\epsilon < 1$; it can be easily demonstrated analytically that the ratio (A.7) has a unique maximum in correspondence of one of the two integer numbers N_A and $N_B = N_A + 1$ such that $N_A \leq \epsilon \cdot N < N_B$, and that $R(N_{ok}, \epsilon)$ decreases monotonically moving away from the maximum both toward higher and lower values of N_{ok}.

The routine exploits this property for the determination of the sequence of values to be included in the acceptance interval: it starts by considering the two integers N_A and N_B, identifying which of the two has a higher likelihood ratio and hence is the first value of the sequence, $N_{ok}^{(1)}$; then the pair $N_{ok}^{(1)} - 1$, $N_{ok}^{(1)} + 1$ is considered and the second value,

[5] Since the binomial distribution is discrete, it is in general not possible that the cumulative probability exactly matches the predefined confidence level; hence the present conservative definition is adopted.

[6] Clearly $(\epsilon_{min}, \epsilon_{max})$ must be larger than the confidence interval $(\epsilon_{meas} - \Delta\epsilon_-, \epsilon_{meas} + \Delta\epsilon_+)$, to avoid an underestimation of $\Delta\epsilon_{\pm}$.

$N_{ok}^{(2)}$, determined; if for example $N_{ok}^{(2)} > N_{ok}^{(1)}$, as a third step the routine takes $N_{ok}^{(3)}$ from the pair $N_{ok}^{(1)} - 1$, $N_{ok}^{(2)} + 1$ and so on.[7]

For each new value $N_{ok}^{(n)}$, the sum of the probabilities $B(N_{ok}, \epsilon)$ over all the values $N_{ok}^{(1)}..N_{ok}^{(n)}$ is calculated; the construction ends either when this sum is $\geq 68\%$ and N_{ok}^{meas} has not yet been included in the acceptance interval (meaning that ϵ is not part of the confidence interval) or when $N_{ok}^{(n)} = N_{ok}^{meas}$ and the previously calculated sum over $N_{ok}^{(1)}..N_{ok}^{(n-1)}$ is $< 68\%$ (meaning that ϵ is part of the confidence interval).

To calculate $\Delta\epsilon_{\pm}$ for the efficiency measurements shown in the tables of sect. 5.5 and 5.6, the resolution $\delta\epsilon_{res}$ and the range $(\epsilon_{min}, \epsilon_{max})$ have been adjusted to limit the number of necessary steps. The confidence intervals have been calculated within few minutes for the available small statistical samples, with $\delta\epsilon_{res}$ set to 0.0001 for all the cases.

[7] If, for a certain step n, $N_{ok}^{(n)} = N$, the following values of the sequence are uniquely defined, because the acceptance interval is limited to the physically allowed values $N_{ok} \leq N$.

Bibliography

[1] M. Boezio *et al.*, The PAMELA Space Experiment, Proc. of the 29$^{\text{th}}$ International Cosmic Ray Conference (2005), submitted for publication; see also http://icrc2005.tifr.res.in

[2] PAMELA official web page: http://wizard.roma2.infn.it/pamela

[3] Particle Data Group, Review of particle physics, **Physics Letters B 592** (2004) and references therein.

[4] P. Spillantini, Ballon experiments, **Frascati Physics Series Vol. XXIV**, 149 (2002)

[5] G. Osteria *et al.*, The TOF and Trigger systems of the PAMELA experiment: Performances of the Flight Model, Proc. of the 29$^{\text{th}}$ International Cosmic Ray Conference (2005), submitted for publication; see also http://icrc2005.tifr.res.in

[6] Saint-Gobain Crystals: http://www.detectors.saint-gobain.com

[7] Hamamatsu: http://sales.hamamatsu.com

[8] M. Aguilar *et al.*, The Alpha Magnetic Spectrometer (AMS) on the International Space Station: part I - results from the test flight on the space shuttle, **Phys. Rep. 366**, 331 (2002)

[9] O. Adriani *et al.*, The magnetic spectrometer of the PAMELA satellite experiment, **Nucl. Instr. and Meth. in Phys. Res. A 511**, 72 (2003)

[10] L. Bonechi *et al.*, A powerful tracking detector for cosmic rays: the magnetic spectrometer of the PAMELA satellite experiment, **Nuclear Physics B (Proc. Suppl. 125**, 308 (2003)

[11] F. Taccetti *et al.*, Pamela tracking system: status report, **Nucl. Instr. and Meth. in Phys. Res. A 485**, 78 (2002)

[12] PAMELA Florence group: http://hep.fi.infn.it/PAMELA

[13] Vacuumschmelze: http://www.vacuumschmelze.de

[14] Sypris: http://www.sypris.com

[15] S.B. Ricciarini, Sviluppo dell'elettronica digitale per il sistema tracciante dello spettrometro PAMELA, Physics Degree Thesis, Florence University (2001), http://hep.fi.infn.it/PAMELA/tesi/tesi.html [in Italian]

[16] S. Straulino, Simulation of Silicon Microstrip Detectors for a Cosmic Ray Experiment, Ph.D. Thesis, Bologna University (2003), http://hep.fi.infn.it/PAMELA/tesi/tesi.html

[17] S. Straulino *et al.*, Spatial Resolution of Double-Sided Silicon Microstrip Detectors for the PAMELA apparatus, accepted for publication in **Nuclear Instr. and Methods in Phys. Res. A**, also available at http://arxiv.org/abs/hep-ex/0510012

[18] Ideas: http://www.ideas.no

[19] Analog Device: http://www.analog.com

[20] M. Boezio *et al.*, The magnetic spectrometer of the PAMELA satellite experiment, **Nucl. Instr. and Meth. in Phys. Res. A 511**, 72 (2003)

[21] F. Giambi, Calibrazione e allineamento dell'esperimento PAMELA su satellite, Physics Degree Thesis, Florence University (2003), http://hep.fi.infn.it/PAMELA/tesi/tesi.html [in Italian]

[22] IABG: http://www.iabg.de

[23] J. R. Wertz and W. J. Larson, Space mission analysis and design, Space Technology Library (1999)

[24] F. B. McLean, H.E. Boesch and T.R Oldham, "Electron-Hole Generation, Transport, and Trapping in SiO_2", Ionizing Radiation Effects in MOS Devices and Circuits, T. P. Ma and P. V. Dressendorfer ed., Wiley (1989)

[25] G. J. Deboo, Integrated circuits and semiconductor devices: theory and application, McGraw-Hill (1997)

[26] M. Boscherini *et al.*, Radiation damage of electronic components in space environment, **Nucl. Instr. and Meth. in Phys. Res. A 514**, 112 (2003)

[27] https://creme96.nrl.navy.mil

[28] A. S. Tanenbaum, Computer Networks, Prentice Hall (1981)

[29] Institute of Electrical and Electronics Engineers: http://www.ieee.org

[30] E. Petersen, Cross Section Measurements and Upset Rate Calculations, **IEEE Trans. on Nucl. Sc., vol. 43, n. 6** (1996)

[31] Gesellschaft für Schwerionenforschung: http://www.gsi.de

[32] Actel: http://www.actel.com

[33] Quicklogic: http://www.quicklogic.com

[34] Z. Navabi, VHDL. Analysis and Modeling of Digital Systems, McGraw-Hill (1993)

[35] S.Rajan, Essential VHDL. RTL Synthesis Done Right (1998)

[36] D. Naylor and S. Jones, VHDL: A Logic Synthesis Approach, Chapman & Hall (1997)

[37] K. Skahill, VHDL for Programmable Logic, Cypress Semiconductor (1995)

[38] Actel HDL Coding. Style Guide, Actel (2000), http://www.actel.com

[39] CAEN Aerospace: http://www.caen.it/aerospace

[40] Statek: http://www.statek.com

[41] National Semiconductor: http://www.national.com

[42] AMD: http://www.amd.com/us-en

[43] CERN: http://user.web.cern.ch/user/Welcome.html

[44] Motorola: http://www.motorola.com

[45] Altera: http://www.altera.com

[46] Leach International: http://www.leachintl.com

[47] J. D. Sullivan, Geometrical factor and directional response of single and multi-element particle telescopes, **Nuclear Instr. and Methods 95**, 5 (1971)

[48] GEANT: http://wwwasd.web.cern.ch/wwwasd/geant/index.html

[49] W. H. Press *et al.*, Numerical Recipes in Fortran 77, Cambridge University Press (1992), http://www.library.cornell.edu/nr/cbookfpdf.html

[50] J. Lund, Antiparticle identification Studies for the PAMELA Satellite Experiment, Doctoral Thesis, Royal Institute of Technology, Stockholm (2004)

[51] GUIDEIT v.1.1: www.npl.uiuc.edu/ftp/guideit/manual.ps

[52] PAW: http://wwwasd.web.cern.ch/wwwasd/paw

[53] J. Neyman, **Phil. Trans. Royal Soc. London A 236**, 333 (1937)

[54] G. J. Feldman and R. D. Cousins, Unified approach to the classical statistical analysis of small signals, **Phys. Rev. D 57, 7**, 3873 (1997)

[55] S. L. Meyer, Data Analysis for Scientists and Engineers, Wiley (1975)