ROMA TRE
UNIVERSITÀ DEGLI STUDI

SCUOLA DOTTORALE IN SCIENZE MATEMATICHE E FISICHE
DOTTORATO DI RICERCA IN FISICA

XXIV CICLO

The cosmic ray spectrum measured by the ARGO–YBJ experiment in the 1–1000 TeV energy range.

Paolo Montini

Coordinatore:
Prof. Guido Altarelli
Prof. Orlando Ragnisco

Docente Guida:
Prof. Stefano Maria Mari

Contents

Introduction

The cosmic rays spectrum spans a huge energy interval up to 10^{20} eV or more. Since the discovery of the cosmic radiation in 1912, many experimental efforts have been devoted to the study of cosmic ray properties. The study of the energy spectrum and mass composition in the energy range $1 - 10000$ TeV has been an active field for many years, with various methods and techniques applied. The study of the properties of Cosmic Rays offers an unique opportunity to investigate several non–thermal phenomena, obtaining information about the early evolution of the Universe. Due to the extreme variability of the flux intensity two broad classes of experimental techniques have been developed: direct and indirect measurements. Direct measurements can access only the low energy region because of the restricted dimension and the limited exposure of the detector. Otherwise they can directly measure the energy and the composition of the cosmic radiation. The indirect technique based on the measurement of the extensive air showers with surface arrays allows the detection of events produced by primaries with energy up to 10^{20} eV. This technique relies on Monte Carlo simulations in order to unfold the relevant information about the primary particle from the collected data sample. Unfortunately, despite a great experimental effort and a huge amount of data collected, several questions concerning the acceleration and propagation mechanisms, the origin of the knee and the composition at the knee are still under discussion and investigation. In the energy region $(1 \div 100)$ TeV, in particular, direct measurements of the cosmic ray flux made with emulsion chamber experiments (RUNJOB and JACEE)

show large discrepancies and uncertainties. Recent measurements made by new generation balloon–borne experiment (CREAM) show that the proton and helium spectra in the range $(2 \div 250)$ TeV are harder than in the low energy region. The CREAM data, however, present large uncertainties at the highest energies. Since proton and helium nuclei are the bulk of cosmic ray particles at energies below the knee ($\sim 3. \times 10^{15}$ eV) the study of the spectrum at these energies is of primary importance. A better understanding of these topics can be obtained by extending the indirect measurements in the low energy region covered by balloons or satellite. The study of the composition around the knee plays an important role in the understanding of the origin of the knee itself. Since the direct measurements cannot be performed at energies $E > 100$ TeV the composition at the knee must be studied by extensive air shower experiments. Showers produced by different primaries present several differences that are covered by large variations of the primary energy, uncertainties in the hadronic interaction models and by shower fluctuations. A detailed measurement of the lateral particle density distribution can be used in order to discriminate showers induced by different primaries.

The ARGO–YBJ experiment is a full–coverage extensive air shower array operating at the Yangbajing international cosmic ray observatory in Tibet (P.R. China) at about 4300 m. The detector is a full–coverage array of RPC chambers covering an area of about 10^4 m^2 and providing a high resolution image of the shower front. Due to its characteristics (full–coverage, high altitude, high segmentation) the ARGO–YBJ experiment can detect extensive air showers with a low energy threshold. The detector was designed and built in order to explore the cosmic ray spectrum in the energy range $(1 \div 10^4)$ TeV and is able to overlap the direct measurements in a wide energy range below 100 TeV, an energy region not accessible by other EAS experiments. The ARGO–YBJ observable is the multiplicity of the shower, namely the space time distribution of the charged particles in the shower front. Due to shower fluctuations the energy distribution of the primary cosmic rays cannot be related to the observed multiplicity distribution and must be evaluated by means of an unfolding procedure. This is a classical unfolding problem that can be dealt with the bayesian technique.

In this work the measurement of the cosmic ray spectrum in the energy range $(1 \div 10^3)$ TeV is presented by using a bayesian approach. In the first chapter the general properties of the cosmic ray energy spectrum and composition

are summarized and the open questions are discussed. In chapter 2 and 3 the characteristics and performances of the ARGO–YBJ detector are described. In chapter 4 the bayesian unfolding method, used in this analysis, is described. Chapter 5 is devoted to the data analysis and the results.

CHAPTER 1

Spectrum and composition of cosmic rays

The primary cosmic radiation is made of particles reaching the Earth's atmosphere from outside the Solar System. They are essentially composed by protons, helium nuclei, a small fraction of heavier nuclei, electrons and photons. Although already discovered in 1912, cosmic rays sources and propagation mechanisms are still subject of intense research. Understanding cosmic rays origin and propagation through the interstellar medium is a fundamental problem which has several implications on models of the structure of the universe. During the last decades several progresses have been made and the picture of cosmic rays observations is now evolving. One of the difficulties related to the study of the cosmic rays properties is that cosmic rays are essentially composed by charged particles that are being scattered by the galactic magnetic field. During the propagation in the interstellar medium cosmic rays lose their original direction and the flux at Earth is essentially isotropic. The only observable quantities which can give us some informations are the energy distribution and the elemental composition of primary cosmic rays. At highest energies, above 10^{18} eV there is a chance of locating emission sources by searching for anisotropies. The observation techniques can be grouped into two broad classes: direct and indirect measurements. At energies below 10^{14} eV the flux of cosmic rays is sufficiently large that the individual nuclei can be studied with flying detectors like balloon or satellites. Direct measurements with satellite or balloon–borne detectors can determine

the energy and the mass of the primary particle only in the low energy region (up to a few hundreds of TeVs) due to the strong decrease of the cosmic ray flux. From this kind of experiments we know that the majority of particles are nuclei of elements from hydrogen to iron. Magnetic spectrometers can provide information about the cosmic ray flux and composition at very low energies. Ground based experiments are characterized by a long exposure time and a large collecting area. These features are needed to investigate the rapidly decreasing cosmic ray flux with increasing energy. These kind of experiments, however, cannot identify the mass of the primary particle that produced the shower. In the high energy region we rely only on observations extensive air showers (EAS) providing indirect information about the energy and the mass of cosmic rays particles. The information needed is inferred from the particle cascades initiated by the interaction of cosmic rays in the Earth's atmosphere. This technique requires a very good knowledge of the EAS development in the atmosphere and of the interaction mechanisms of high energy particles with air. Monte Carlo simulations are used as reference patterns. Despite the large amount of data collected by air shower experiments the determination of the spectrum and of the composition presents several open questions.

1.1 Energy Spectrum

The all–particle spectrum is the energy spectrum of all primary cosmic rays and does not take into account the charge or the mass, so it comprises all the species of elements and all of isotopes. This spectrum spans a very wide energy range from less than 1 GeV up to 10^{11} GeV. Apart from particles produced during solar flares the primary cosmic rays come from outside the solar system. At very low energies the flux of galactic cosmic rays is modulated by the solar activity. There's a significant anticorrelation between solar activity ($T \sim 11$ yr) and the intensity of the flux at energies below 10 GeV. The flux of low energy particles decreases during periods of high solar activity and reaches the maximum during the low activity phase. The flux at very low energies is also influenced by the Earth's magnetic field and the intensity of the cosmic radiation in the GeV range depends on the time and the location. Above the energies of 10 GeV, where the flux is no longer influenced by Earth's magnetic field modulated by solar activity, the cosmic ray spectrum is well described by a power law $dN/dE \propto E^{-\gamma}$. As can be inferred

from figure 1.1 there is a prominent change of the spectral index at energies around 10^{15} eV, generally referred as the *knee* [1]. At energies around 10^{18} eV another change of the spectral index occurs, known as the *ankle*, and the spectrum becomes harder again. Values of the spectral index are $\gamma \approx 2.7$ below $E \approx 4 \cdot 10^{15}$ eV, above this energy the spectral index changes to $\gamma \approx 3.1$. Many experiments have performed and published measurements of the cosmic ray spectrum in the knee region and up to the highest energies, as reported in figure 1.1. The individual measurements are in agreement and shows a change of the spectral index at $E \approx 4 \cdot 10^{15}$ eV. The origin of the knee is still an unresolved question in cosmic rays research. Nearly the totally of the cosmic rays particles is thought to be of galactic origin. The ankle can be interpreted as a transition between the galactic and extra–galactic component. At the far end of the spectrum, above $\sim 5 \cdot 10^{19}$ eV, the cosmic ray flux is expected to be strongly suppressed due to the interaction with the cosmic microwave background (CMB). This feature is known as Greisen–Zatsepin–Kuzmin (GZK) cutoff [2, 3]. Since heavier nuclei are destroyed by photo–disintegration the flux at the highest energies is expected to be dominated by protons. Protons lose energy while propagating through the CMB field: the most important energy loss mechanism is the pion photo production with an energy threshold of about $6 \cdot 10^{19}$ eV. The suppression of the flux above this energy was observed by the HiRes [4] and Auger [5] experiments. Above the GZK cutoff energy cosmic rays cannot travel more than ~ 20 Mpc and the observation of a possible extension of the cosmic ray spectrum would indicated a local origin for these particles.

1.2 Composition

All the elements of the periodic table have been found in cosmic rays. For elements lighter than nickel the individual energy spectrum has been measured. The relative abundance is shown in figure 1.2, compared with the abundance of elements in the solar system. The two distributions look very similar, however there are some differences that reveal information about the acceleration and the propagation of cosmic rays. Lighter elements like lithium, beryllium and boron are more abundant in cosmic rays rather than in the solar system. Since these nuclei are not present in the final products of stellar nucleosynthesis, they are assumed to be produced in spallation processes of the more abundant carbon, nitrogen, oxygen and iron nuclei during their

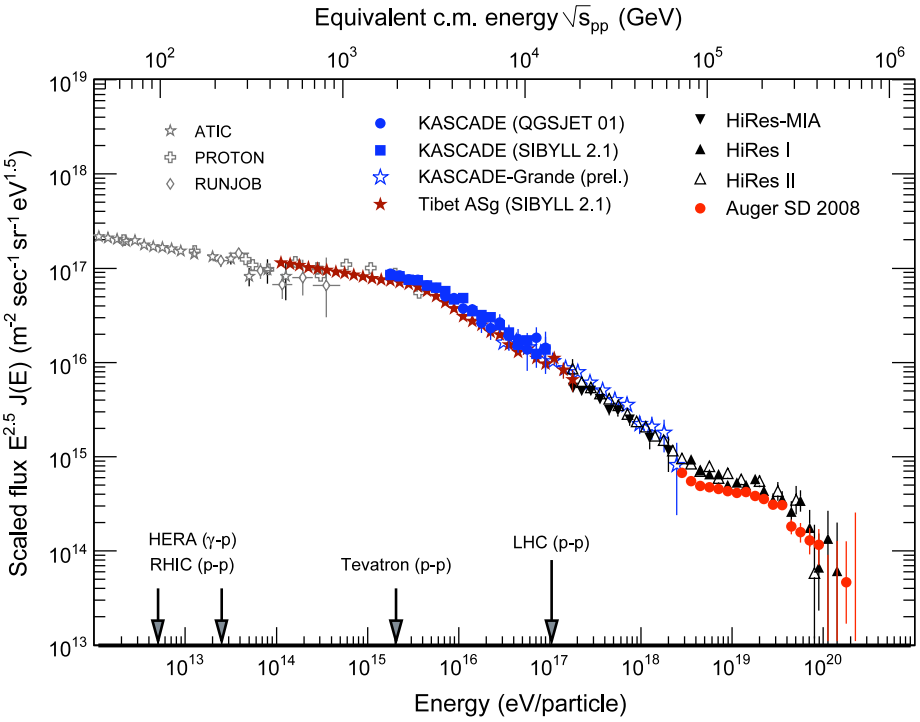


Figure 1.1: All-particle spectrum of cosmic rays measured with direct detectors and air shower experiments For a better visualization of the knee the flux was multiplied by $E^{2.5}$. See [6] for references.

journey through the Galaxy. The cosmic rays accelerated in their sources propagate towards the Earth through the interstellar medium. Heavier nuclei can interact with the ISM and fragment into lighter elements. For this reason they are usually called secondary cosmic rays. The cosmic rays as they are observed near Earth consist therefore in a mixture of a primary and a secondary component. Since the spallation cross section is known at GeV energies from the accelerator data, the ratio between secondary and primary cosmic rays can be used as an estimator of the propagation path length and of the residence time of cosmic rays in the Galaxy. Also the measurements of the isotopic composition can provide information about the cosmic ray transport in the galaxy. As an example the measure of the boron to carbon ratio as a function of energy can be used to estimate the propagation path length. Recent measurements of the B/C ratio have been performed by the TRACER experiment [7]. The PAMELA collaboration has also presented

the measurements of the fluxes of boron and carbon [8]. Usually lighter elements (p, He) are referred as the *light component*, while the heavier ones (CNO, NeMgSi, Fe) are referred as the *heavy component*.

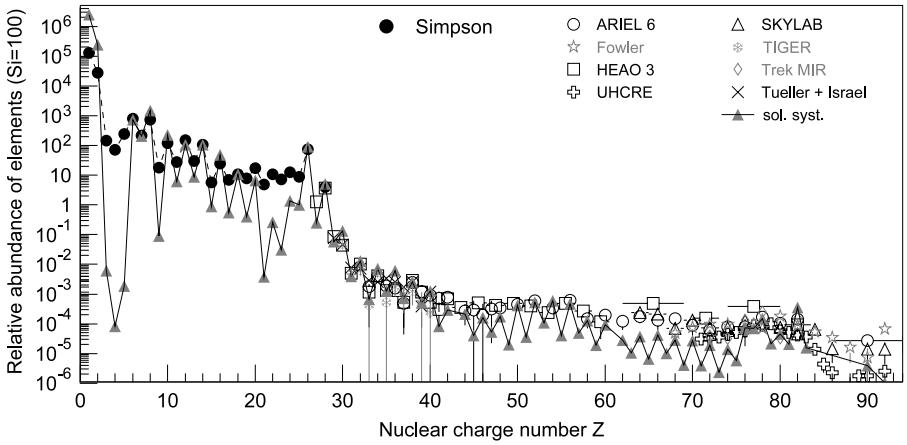


Figure 1.2: Relative abundance of elements found in cosmic rays as a function of the nuclear charge Z (referred to Silicon). The abundance of elements in the solar system is also shown.

1.3 Production and acceleration

After almost 100 years from the first observation of the cosmic radiation, the origin of cosmic rays is not yet fully known. Since the abundance of elements in cosmic rays is very similar to the abundance of elements in the solar system it indicates that cosmic rays are regular matter accelerated to very high energies. The bulk of cosmic rays is assumed to be accelerated in the shock waves of supernova remnants. This feature was discovered by Baade and Zwicky [9] in 1934. Supernova remnants were proposed as cosmic rays sources due to energy balance considerations [10]. The power necessary to produce the galactic cosmic rays can be estimated as

$$L_{CR} = \frac{v_D \rho_E}{\tau_r} \simeq 5 \cdot 10^{40}\, erg/\mathrm{s} \tag{1.1}$$

where v_D is the volume of the galactic disk, τ_r is the residence time in the disk and ρ_E is the energy density of cosmic rays (about $1\,\mathrm{eV/cm^3}$). Comparing the results in eq. 1.1 with the typical power released in a supernova

explosion it results evident that the shock waves of supernova remnants can be one of the possible acceleration site for galactic cosmic rays. The acceleration mechanism has remained unclear until Fermi proposed a model of acceleration which involves the interaction of particles in galactic magnetic field [11]. This lead to the so–called first order Fermi mechanism in which the responsible for the acceleration of cosmic ray is the shock front produced in a supernova explosion. At present it is believed that cosmic rays are accelerated in a process called diffusive shock acceleration (DSA) [12, 13]. The basic idea is that a charged particle can gain an amount of energy $\Delta E \propto E$ every time the particle passes from the upstream unshocked region to the downstream shocked region and back. At each passage between the upstream and downstream region there is a probability that the particle can escape the shock and propagate through the Galaxy. This process leads naturally to a power law spectrum like $N(E) \propto E^{-\gamma}$. The highest energy particles are those that have remained longer in the shock region. After a time T the maximum energy can be written as

$$E_{max} \sim Ze\beta_s \cdot B \cdot TV_s \tag{1.2}$$

in which $\beta_s = V_s/c$ is the velocity of the shock. This leads to an upper limit assuming that the minimum diffusion length is equal to the Larmor radius of a particle of charge Ze in a magnetic field B. Using typical values for supernovae $E_{max} \approx Z \times 10^{14}\,\text{eV}$ [14]. More recent estimates give a value for the maximum energy up to one order of magnitude larger [15]. Highest energies can be reached by introducing the reaction of accelerated particles onto the accelerator. Cosmic rays generate the magnetic structure in which they are scattered, therefore the acceleration time can be reduced and the maximum energy can reach larger values. Evidences of particle acceleration in supernova remnants come from the observations of non–thermal radio, X–ray and gamma–ray radiation. The observation of non–thermal X–rays in young SNRs indicates the presence of electrons in the $10 \div 100\,\text{TeV}$ energy range that emit synchrotron radiation [16]. From the data of the H.E.S.S. experiment it is established that there are high energy cosmic ray particles in the shell of a supernova remnant. The measured γ–ray energies imply the efficient acceleration of particles up to energies greater than 100 TeV, moreover the spectral index of the observed radiation is compatible with the predictions of shock acceleration in SNRs [17]. The flux of cosmic rays is nearly isotropic up to the highest energies. The isotropy is due to the

interaction with the cosmic rays with the galactic magnetic field. Since the Larmor radius of a 1 PeV proton in the galactic magnetic field ($B \approx 3\mu G$) is about 0.4 pc there's no possibility to find any point sources of galactic cosmic rays. The current paradigm for the cosmic ray acceleration is based on a continuous distribution of sources that leads to a simple power–law spectrum for all the elements. A variation from the simple power law spectrum could indicate a non–uniform distribution of sources. For energy above the knee there are no obvious candidates, although several objects like Gamma Ray Bursts or active galactic nuclei can be listed as possible sources.

1.4 Recent observations

Many experimental efforts have been devoted to the study of the cosmic ray properties. The indirect measurements in the region of the hundreds of TeVs show an overall agreement within a factor of two [18], while the direct measurements have to cope with the systematic uncertainties due to analysis of data taken during different flights. Direct measurements have provided a measurement of the flux at energies up to $\sim 10^{11}$ eV/nucleon. In figure 1.3 the measurement of the proton and helium spectra performed by several balloon–borne experiments is reported. At higher energies the measurements made by the balloon–borne emulsion chambers RUNJOB [19, 20] and JACEE [21] show large discrepancies and uncertainties. A New generation balloon–borne active experiments can perform longer flight and increase higher statistics at the highest energies. Recent more precise measurements of the cosmic ray flux have been carried out by the balloon-borne CREAM experiment [22, 23] during a long duration flight. These measurements show that the proton and helium spectra in the range $(2.5 \div 250)$ TeV are both flatter than in the lower energy measurements.

1.5 Open questions

1.5.1 Proton and helium spectra

Despite the large amount of data collected by several experiments there are several open questions concerning the energy spectrum and the composition of cosmic rays. The evolution of the proton and helium spectra can be an

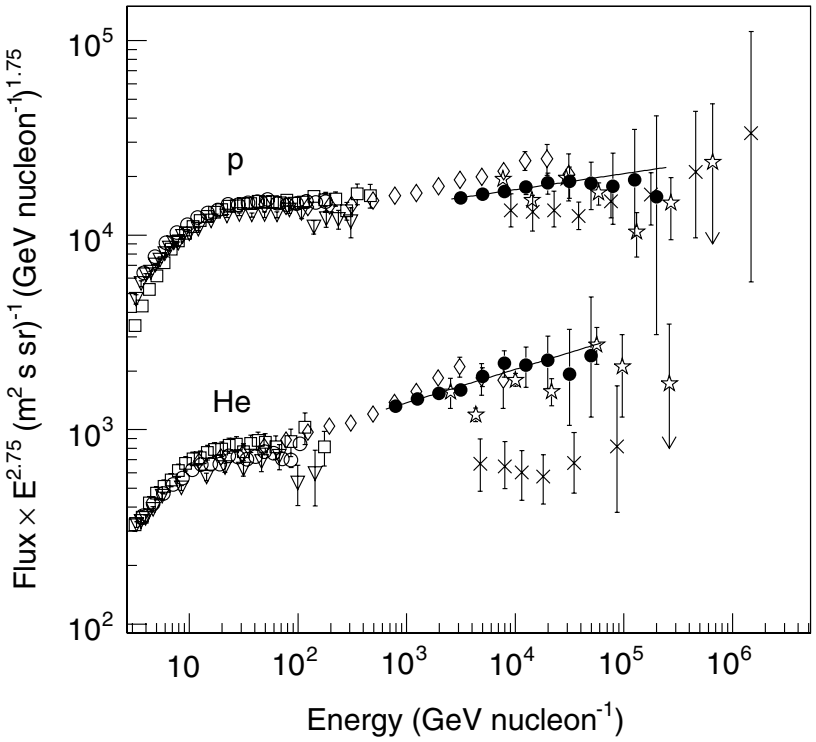


Figure 1.3: Measurements of the proton and helium spectra at the top of the atmosphere from data of several baloon-borne experiment. Results from CREAM [23] (filled circles) , BESS (squares), CAPRICE98 [24] (downward triangles), ATIC–2 [25] (diamonds), AMS [26] (open circles), RUNJOB (crosses) and JACEE (stars) are reported.

indication of different population of cosmic ray sources or acceleration sites. According to the DSA theory the spectral index does not depend on the mass of the elements. A possible explanation the difference between proton and helium spectra is that protons and helium nuclei are coming from different sources or acceleration sites. Since a single class of sources cannot provide different spectral indexes for different masses, at least two kind of sources are needed [27]. Each class of sources can produce a power law spectrum with its specific spectral index and maximum energy, therefore the galactic cosmic rays are essentially a mixture of particles accelerated by different sources. In many models [28, 29] the responsible for the acceleration of cosmic rays up to the knee are the explosions of normal supernovae directly into the interstellar medium and the explosion of massive stars into their former stellar wind, like

Wolf–Rayet stars. If elements heavier than hydrogen are accelerated in the latter sources their spectra would be harder than the spectrum of protons. The explosion of normal supernovae into a homogeneous stellar medium leads to the acceleration of protons up to $\sim 10^5$ GeV. The spectral hardening can be connected also to the anisotropy of cosmic rays at energies around 10 TeV observed by the Milagro experiment, which observed two regions of excess with high significance [30]. These two regions are inconsistent with pure gamma–ray emission. Moreover the energy spectrum of one of these two regions is different from the spectrum of the isotropic cosmic rays and can be described as a harder proton spectrum with a cutoff. The same regions have been observed also by the ARGO–YBJ experiment [31]. Another possible explanation is that most massive stars are born in association and they explode as supernovae near their parent molecular clouds. The effect of supernova explosions in a small region of the Galaxy is the formation of large bubbles (superbubbles) of very hot material. Superbubbles are one of the greatest injector of energy in the Galaxy. Most massive stars are in dense clusters called OB associations, each containing several hundreds of stars. OB associations have been proposed as a candidate source of galactic cosmic rays due to energy balance considerations [32]. This hypothesis is also supported by the observation of the isotopic abundance of heavy nuclei [33]. These sources can produce different power–law spectrum for protons and helium nuclei because in the stellar wind and in supernova explosion the hydrogen has a lower density and a higher velocity rather than helium, therefore the fraction of helium nuclei in the center of the superbubble is higher than in the outer region [34]. The energy spectrum at the source can also be harder than previous hypothesis based on the low–energy data. The hardening can be related to the predicted concavity of the spectrum before the knee region [35]. In diffusive shock acceleration the accelerated particles can amplify the magnetic field close to the shock surface. As a consequence higher energy particles can gain energy faster. This feature could lead to a hardening of the spectrum as the energy increases and deviations from a single power law [36].

1.5.2 The origin of the knee

Despite the great experimental and phenomenological effort made to investigate the knee region, the origin of the knee is still an unresolved question in

cosmic ray physics. Several hypotheses about the origin of the knee have been formulated [37]. Some hypotheses are related to the acceleration mechanisms and the knee is due to the fact that the spectrum at the sources exhibits a break. A special case is the single source model in which the structures in the knee region are due to a single SNR from a near recent explosion [38]. Another possible explanation is that the knee is due to propagation effects. As the energy increases the probability that a cosmic ray particle could escape the galactic region increases. The Larmor radius of a proton in the galactic magnetic field

$$r_L = 1.08\text{pc}\frac{E/\,\text{PeV}}{Z \cdot B/\mu\text{G}}$$

can become greater than the size of the galactic disk as the energy increases. Above a certain energy the cosmic rays are not more confined into the Galaxy. Other models state that the origin of the knee can be related to the interaction of high energy protons with massive neutrinos or the photodisintegration in dense photon field. Experimental observations rule out these models with a high confidence level. The origin of the knee can also be related to new physics in the elementary particle interactions in the atmosphere. A certain amount of the total energy may be transferred into non observed channels. Actually the measurements indicate that the knee is probably due to a decrease of the flux of light elements. If the knee is due to a break in the source spectrum or to the leakage of high energy particles from the Galaxy, the energy spectra of elements of charge Z should exhibit a cutoff at energy $E_c^Z = Z \cdot E_c^p$, where E_c^p is the cutoff energy for protons. The knee in the all–particle spectrum is due to the cutoff of protons and the spectrum at highest energies becomes steeper due to the cutoff of elements with increasing atomic number. In order to solve the knee "puzzle"a detailed measurement of the energy spectrum of the single component or of the light and heavy group is needed.

1.6 The role of the ARGO–YBJ experiment

In order to understand the behavior of the spectrum at TeV energies an to investigate the composition at the knee an EAS experiment able to cover a wide energy range is needed. The ARGO–YBJ experiment, described in chapter 2, is a full–coverage array which is in full and stable data taking at the Yangbajing International Cosmic Ray Observatory. The full–coverage

technique and the high space granularity allow the measurement the distribution of the charged particles in the shower front with unprecedented detail, leading to the possibility of a detailed study of the characteristics of the showers. The ARGO–YBJ data can be used to investigate the shower space time structure in order to obtain information about the processes that regulate the development of the shower in the atmosphere. Exotic phenomena like multicore events can also be investigated. The detector is characterized by a low energy threshold and high duty cycle. The high altitude allows the detection of very small size showers before they are completely absorbed in the atmosphere. Due to these characteristics a very wide energy range below 100 TeV can be explored, overlapping the direct measurements. Moreover the detector can provide a high space–time resolution image of the front of showers produced by cosmic rays with energies up to 10^4 TeV, allowing the measurements of the energy spectrum and composition up to the knee region. The position of the maximum of a shower produced by a primary with energy in the knee region occurs at the atmospheric depth of the ARGO–YBJ experimental site, therefore it's possible to investigate the characteristics of showers in the knee region with very high detail.

CHAPTER 2

The ARGO-YBJ experiment

The ARGO-YBJ experiment (Astrophysical Radiation with Ground-based Observatory at YangBaJing) is a full coverage air shower detector designed and optimized for the detection of cosmic rays with an energy threshold of a few hundreds GeV. The detector consists of a single layer of resistive plate chambers (RPCs) covering an area of about $10^4\,\mathrm{m}^2$ and providing a high resolution image of atmospheric showers initiated by primaries of energies up to $10^3\,\mathrm{TeV}$. The shower array is able to monitor the northern emisphere in the declination range $-10^\circ < \delta < 70^\circ$ with high duty cycle. The experiment has been designed to cover a wide range of issues in γ-ray astronomy and in cosmic rays physics.

2.1 Overview

The Argo-YBJ experiment [39, 40, 41] was designed and developed for the detection of extensive air showers produced by cosmic rays interacting with Earth's atmosphere. The detector is located in the Yangbajing International Cosmic Ray Observatory in the Tibet region (P. R. China) at an altitude of about 4300 m a.s.l. (see figure 2.2). The site chosen for the experiment is located in the village of Yangbajing (latitude $[30^\circ\,60'\,38''\,N]$, longitude $[90^\circ\,31'\,50''\,E]$), about 90 km far from Lhasa.

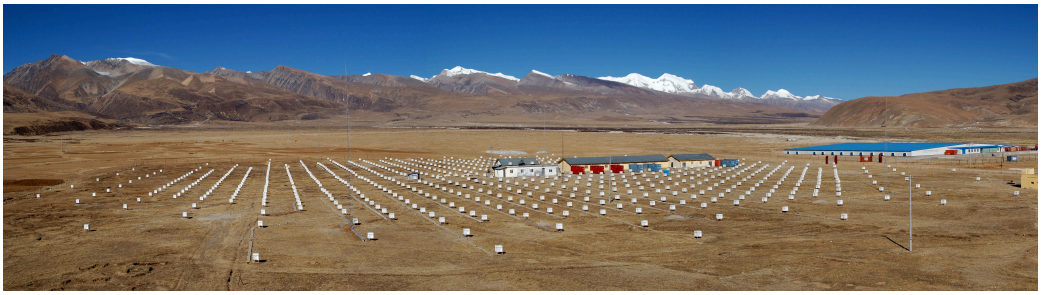


Figure 2.1: Landscape of the Tibet plateau and of the Yangbajing International Cosmic Ray Observatory.

The experiment has been designed to cover a wide range of physical issues both in γ-ray astronomy and in cosmic ray physics, among which are:

- γ -ray astronomy at a threshold of a few hundreds GeV
- Primary cosmic ray spectrum and composition in the $(1 \div 10000)$ TeV energy range
- γ-ray burst at energies above 100 GeV
- Proton and anti-proton ratio at TeV energies
- Solar and heliosphere physics
- Extensive air shower properties.

The experiment was designed to operate in two independent acquisition modes characterized by different energy threshold. This feature was designed in order to extend the range of physical phenomena accessible by the detector.

2.2 The ARGO-YBJ detector layout

The ARGO-YBJ detector consist on a single layer of 1836 resistive plate chambers [40] (RPCs) covering an area of about 6700 m^2. The RPCs are arranged in a central *full coverage* carpet of about 5700 m^2 surrounded by an external ring ($\sim$ 1000 m^2) partially instrumented with additional 276 RPCs in order to improve the reconstruction of showers with external cores. The

Figure 2.2: Satellite view of the ARGO-YBJ experiment and the Yangbajing village.

external ring also improves the angular resolution of the whole detector. The detector layout is shown in figure 2.3. The detector has a modular structure, the basic module is a cluster ($5.72 \times 7.64\,m^2$) consisting in 12 RPCs ($2.850 \times 1.254\,m^2$) each. Each RPC is read–out by 80 copper strips ($61.8 \times 6.75\,cm^2$) logically arranged in 10 pads ($55.6 \times 61.8\,cm^2$). Each RPC is also equipped with two large size electrodes called big pads ($1.23 \times 1.39\,m^2$). The whole carpet is made of 153 clusters, 130 on the central full coverage carpet and 23 on the guard ring. The full coverage carpet has an active area of about 93%. All the environmental and detector parameters like atmospheric pressure, external air temperature, humidity, high voltage and current drawn by each RPC are continuously monitored.

2.2.1 Resistive Plate Chambers

Resistive Plate Phambers (RPCs) are gas detectors able to detect the passage of a charged particle by ionization and multiplication processes that occur in the gas mixture inside the chambers. The RPCs are frequently used in high energy physics experiments thanks to a high detection efficiency ($\sim 98\%$) and relatively low production cost. The use of RPCs in a surface set-up allows many practical advantages like easy mounting without mechanical support, robust assembling and easy access to any part of the detector. The RPCs are low noise detector usually operated for triggering and tracking purposes,

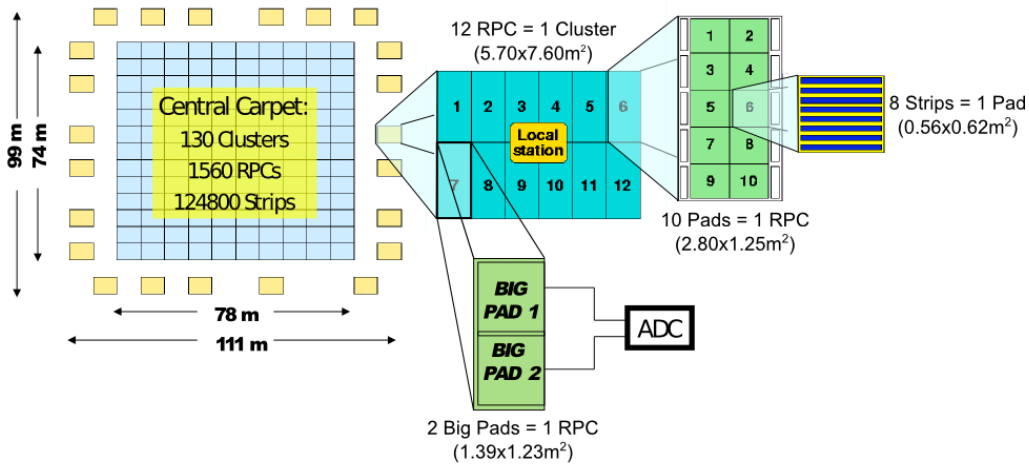


Figure 2.3: Layout of the ARGO-YBJ detector.

however the layout of the chambers built for the ARGO-YBJ experiment has been optimized for the detection of EAS secondaries. Each chamber consists on two Bakelite foils, a polymer with a resistivity $\rho \sim 5 \cdot 10^{11}\,\Omega\,\mathrm{cm}$, assembled to form a gas gap 2 mm wide. A grid of plastic cylindrical spacers (10 cm pitch) is used in order to maintain the gas volume plane and flat. A schematic layout of a RPC is reported in figure 2.4. In order to have a homogeneous distribution of the high voltage to the Bakelite electrodes, a layer of conductive material (graphite) is laid on the electrode surface. The choice of a high-resistivity material for the electrodes is related to the need of decreasing the dead time of the detector. In this way the signal formation on the electrodes is a localized phenomena and the regions of the chamber which are not interested in the ionization process remain sensitive.
The RPCs used in the ARGO-YBJ experiment are operated in streamer mode at a voltage of about 7400 V with a gas mixture of Tetrafluoroethane R134A, Isobutane and Argon in the proportion 75:10:15%. Argon is the active component of the gas mixture while the other two components are needed to quench the discharge by absorbing the ultraviolet photons and secondary electrons emitted in recombination processes. The signal is picked up inductively by using a strip panel consisting on a copper foil 17 µm thick cut into 80 strips ($6.75 \times 61.80\,\mathrm{cm}^2$ each) glued on a 0.2 mm thick film of plastic material (PET). The time resolution of the order of $\sim$ 1ns and the high space–time granularity allow the sampling of the shower front with

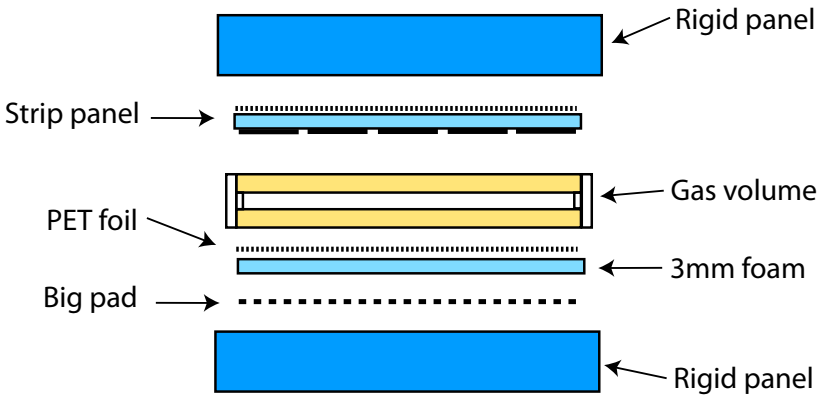


Figure 2.4: Schematic cross–section of a resistive plate chamber used in the ARGO-YBJ experiment

unprecedented detail. In order to extend the measurable energy range each RPC is equipped with two large size electrodes ($1.23 \times 1.39\,\mathrm{m}^2$) which provide a signal whose amplitude is expected to be proportional to the number of charged particles detected. The characteristics of the two readout system will be described in the next chapter.

2.3 Trigger Logic

The ARGO-YBJ experiment was designed to study a wide range of phenomena both in γ–ray astronomy and in cosmic ray physics. The observable quantity of the apparatus is the space–time distribution of the charged particles in the shower front. Each shower is therefore characterized by a particular space–time distribution of fired strips on the detector. In order to discriminate events different both by nature and energy a trigger system has to be implemented. The trigger system must distinguish between showers produced by primaries of a few GeV, which are expected to produce only a few hits on the detector surface, and showers with thousands of hits on the detector. The trigger system of the ARGO-YBJ detector is based on three subsystems: Low Multiplicity Trigger (LMT), High Multiplicity Trigger (HMT) and Fast Trigger (FT) [42]. All these trigger algorithms select an event on the basis of the time distribution of the fired pads and their multiplicity on the carpet. The low multiplicity trigger is based on a four level

hierarchic architecture. The first level, called Level–0, analyzes the multiplicity in groups of four contiguous clusters in a time window of about 150 ns. The second level (Level–1) processes the signals coming from Level–0 and search for a coincidence in groups of 12 clusters in a time window of 200 ns. The third level (Level–2) and the fourth level (Level–3) search for a coincidence in 65 clusters in 360 ns and in 130 clusters in 400 ns respectively. The threshold for the LM trigger is set to 20 pads, corresponding to a maximum event rate of about 3.5 kHz (see figure 2.5) with a dead time less than 4%.

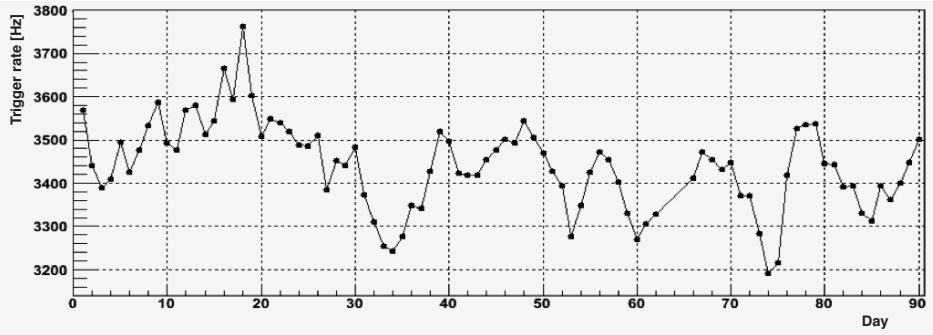


Figure 2.5: Values of the trigger rate for each day of the period January–March 2011.

2.4 Observation techniques

The ARGO–YBJ detector was designed to operate in two independent operation modes: the shower mode and the scaler mode.

2.4.1 Shower mode

The shower mode is based on the requirements that a minimum number of pads must be fired in the central carpet with a defined space–time pattern. For these events the position and time of any fired pads is recorded in order to reconstruct the shower parameters like the core position, the shower size and the arrival direction. For each fired pad the time and the mask of the fired strips within the pad are recorded. Data recorded in shower mode will be used in γ–ray astronomy and in cosmic ray physics. Moreover due to the low energy threshold, the high angular resolution and the high detector

stability the ARGO–YBJ experiment can measure the antiproton/proton flux ratio via the observation of the Moon shadowing effect [43]. The results of the analysis based on the data collected during the period between July 2006 and December 2009 are reported in figure 2.6. Data taken since november 2007

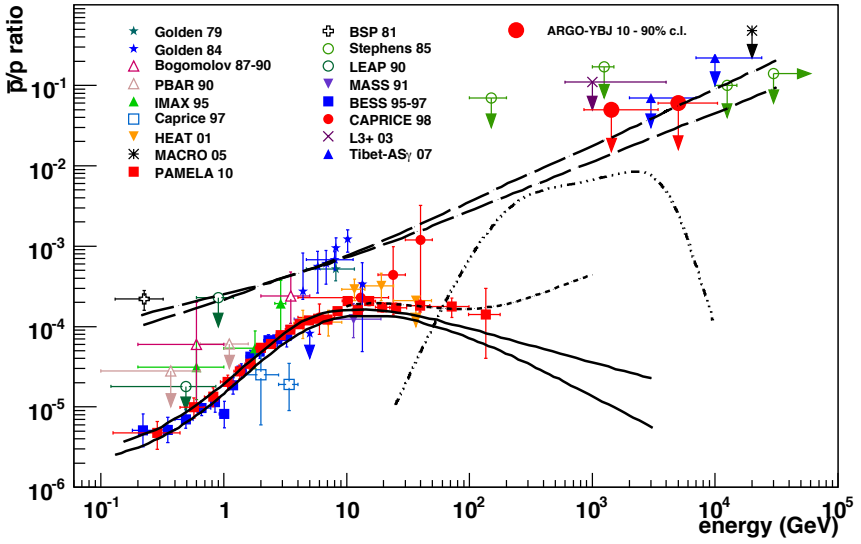


Figure 2.6: The $\bar{p}/p$ flux measured by the ARGO–YBJ experiment compared with the results of other experiments.

have been analyzed in order to look for a few–degree anisotropies in the arrival directions of primary cosmic rays. Several regions with significant eccess of events have been found, with a maximum energy of about 10 TeV [31]. In figure 2.7 the ARGO–YBJ sky map is reported, showing the significance of the observations.

2.4.2 Scaler mode

The lower energy limit of the detector ($\approx$ 1 GeV) is reached by using the scaler mode technique, in which the total counting rates of each cluster are recorded at fixed time (500 ms) with no information on the arrival direction and spatial distribution [44]. The counts from different pads of the same cluster are put in coincidence in narrow time window of 150 ns. Four low multiplicity channels are implemented for event multiplicities from $\geqslant 1$ to $\geqslant 4$ with a mean measured counting rates of about $\approx$ 40 kHz, $\approx$ 2 kHz,

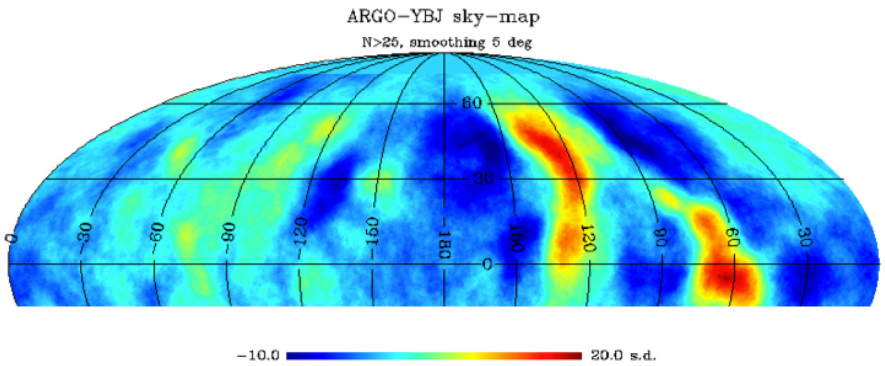


Figure 2.7: ARGO–YBJ sky map in celestial coordinates showing the significance of the observations.

$\approx$ 300 Hz and $\approx$ 120 Hz respectively. The main component of the DAQ system is composed of a series of scaler boards. Each board can manage signals coming from 19 clusters and is connected to a master GPS clock in order to receive a 10 MHz reference frequency. The counting rate of each cluster and multiplicity channel is recorded in a fixed an programmable time interval (i.e. 500 ms). Since The cosmic ray spectrum decreases very steeply as the energy increases, most of the events detected with this operation mode are related to particles coming from small air showers produced by primaries in the $1 - 100$ GeV energy range. The number of particles is too small to reconstruct the shower parameters and is not possible to evaluate the energy or the direction of the primaries. However it is possible to monitor the counting rate of each cluster as a function of time in search of any possible variation that can be related to cosmic phenomena like Gamma Ray Bursts. The search for a GRB signal has been carried out in coincidence of the prompt emission detected by satellites. In the period between December 2004 and April 2011 131 GRBs occurred in the ARGO–YBJ field of view. For 110 of these the search of a high energy emission was performed, showing no statistically significant signal [45]. In figure 2.8 the fluence upper limits in the $1 - 100$ GeV energy range as a function of the zenith angle are reported. The data taken in scaler mode can also be used to monitor the overall detector stability.

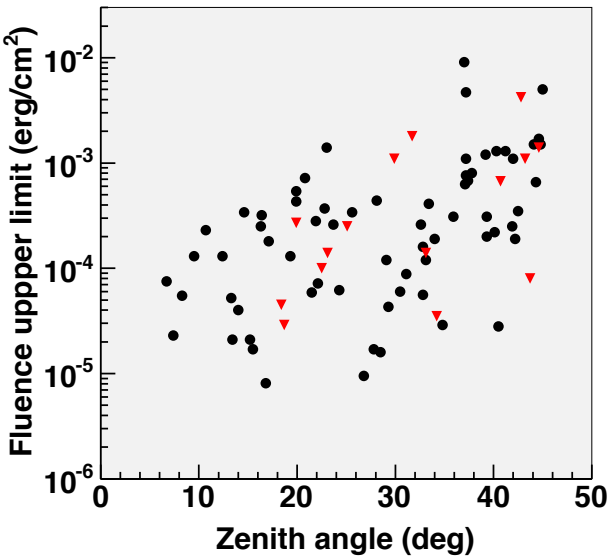


Figure 2.8: Fluence upper limits in the $1 - 100$ GeV energy range as a function of the zenith angle obtained by extrapolating the measured keV $-$ MeV spectra. Triangles represents GRBs with measured redshift, circles GRBs assumed at $z = 1$.

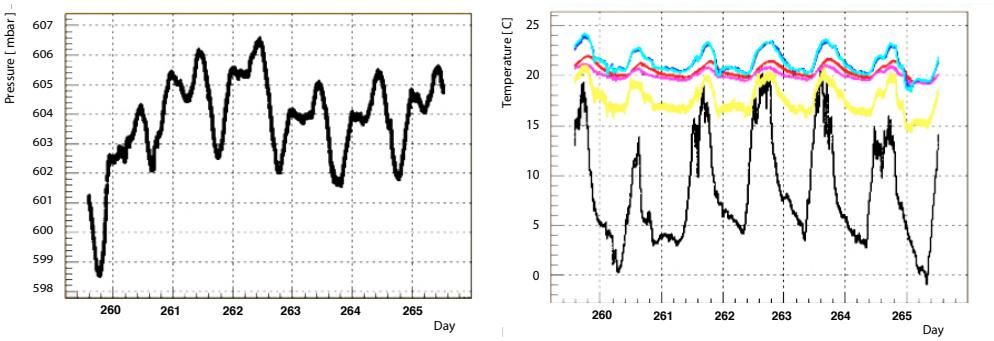
2.5 Detector monitoring and data quality

The ARGO–YBJ experiment produces a large amount of data. One of the major problem is to have fast and efficient tools to check the quality of the data and to select them for analysis purposes. The ARGO–YBJ experiment collects about 200 TB of data each year. A series of detector monitoring and data quality check tools has been implemented by the collaboration. The first step consists in the online monitoring of the detector operation. The detector control system (DCS) has been designed to monitor the following quantities:

- gap currents for each RPC, voltage of the RPC front–end electronics and receiver board, the local temperature and the barometric pressure in the experimental hall
- the applied voltage and the absorption current in each high–voltage channel
- the gas pressure.

In figure 2.9a and 2.9b the daily trends for pressure and temperature are reported. These data are crucial for many detector check and allow the correla-

tion of the RPC gap current with the environmental conditions. The second

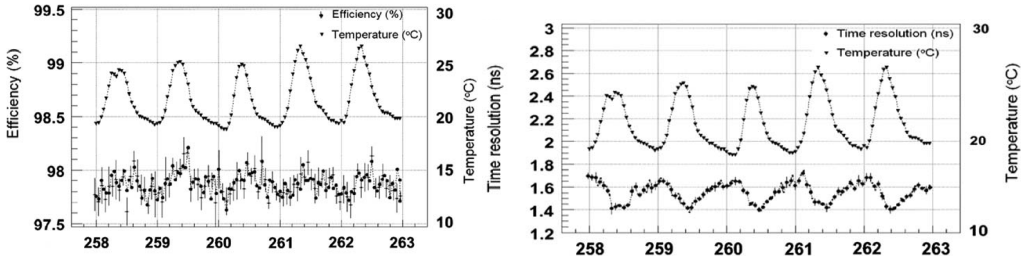


(a) Pressure **(b)** Temperature

Figure 2.9: Monitoring of the environmental condition of the ARGO–YBJ detector site. Values of pressure are reported in figure (a). Values for temperature outside the building (black), inside (yellow), on the ground (purple) and on the RPC carpet (red) are shown. Also the temperature of the gas at the beginning (blue) and at the end (cyan) of the gas line are reported. Data were taken in September 2011.

step consists on an online monitoring of the whole data taking process and of the overall detector performance. These two steps are performed at Yangbajing laboratory. The data check manager (DCM) performs an offline analysis of the detector operation and data consistency before the offline reconstruction of measured quantities. The monitoring of quantities like trigger rate or azimuth and zenith distributions are a powerful method to check the quality of the data taking. All the offline analysis are performed on dedicated computer farms. Since a selection procedure of good data taking periods based on the data quality and on the detector performance is needed for any kind of analysis, a set of easy tools has been implemented [46]. These tools provide a list of good runs based on the parameters recorded by both the DCS and the DCM and on the reconstructed parameters. Due to its geographical position and the high altitude the working condition of the YBJ laboratory are quite distinctive. The experimental site is characterized by a very low average atmospheric pressure (about 600 mb) and the temperature in winter can reach very low values. In order to study the detector performance in these extreme condition a multi–layer RPC telescope has been installed near the ARGO-YBJ detector. This telescope continuously monitors the detector

efficiency and time resolution by using atmospheric muons [47]. In figure 2.10a and 2.10b a plot of the efficiency and time resolution monitoring with respect to the room temperature is reported.



(a) efficiency **(b)** time resolution

Figure 2.10: Monitoring of the efficiency (a) and of the time resolution (b) of the ARGO–YBJ RPCs respect to the temperature inside the detector building as a function of the day number of 2008.

2.6 Event reconstruction

All the data collected by the experiment are processed by an object–oriented code (medea++ [48]) which provides root ntuples with the quantities used for high–level analysis. The same code is used in order to reconstruct the Monte Carlo events. The code performs three levels of reconstruction. The first level, called RecLevelO, decodes the DAQ file and convert all the information in a format which is then elaborated in the next reconstruction level. Monte Carlo events do not need the decoding and this level is skipped. The second level (RecLevel1) connect each hit to the corresponding pad and for each detector element the corresponding hits can be retrieved. All the hits located in disconnected detector elements are discarded. The third level (RecLevel2) performs the event reconstruction and is essentially organized in three steps: noise filter, core reconstruction and direction reconstruction. The noise filter allow to reject the hits due to background or noise. All the hits selected are fitted according to a plane. The core position is determined by using the Likelihood method based on the Nishimura Kamata Greisen function [49, 50, 51]. The core position is estimated and a conical fit is performed to the hits in order to reconstruct the shower direction.

CHAPTER 3

ARGO-YBJ detector performance

As described in the previous chapter, the ARGO–YBJ experiment consists in a full–coverage detector made by a single layer of RPCs covering an area of about $10^4\,\mathrm{m}^2$. The detector provides a high–resolution space–time image of the shower front. The detector was designed to investigate a very wide energy range, from a few TeVs up to the knee region. For these reason two different readout systems were implemented. The digital readout system allow the detection of showers with a particle density around the core up to $\sim 10\,\mathrm{particles/m}^2$ and can be used to measure the primary spectrum up to about 100 TeVs. Above these energies the response of the digital readout system saturates. In order to extend the measurable energy range up to the PeV region, where particle densities around the core are of the order of $10^3\,\mathrm{particles/m}^2$ or more, each RPC has been equipped with two large size electrodes that provide a signal whose amplitude is proportional to the number of charged particles detected. Special care was devoted to the detector calibration, for both digital and readout system. In this chapter the two readout system and the respective calibration procedures are described.

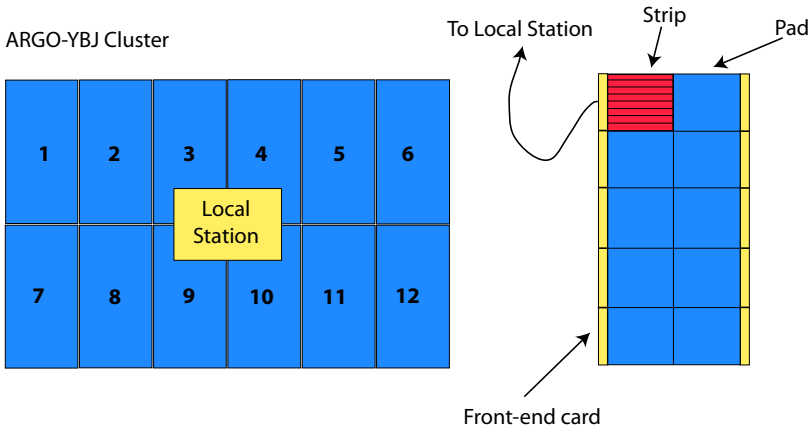


Figure 3.1: Schematic view of a cluster, made of 12 RPCs. Each RPC is made of 10 pads composed of 8 strips each.

3.1 Digital readout system

Each chamber is equipped with 80 copper strips which represent the space granularity of the detector. The fast-OR of 8 contiguous strips signal defines a logic unit called pad ($55.6 \times 61.8\,\mathrm{cm}^2$). The pad signals are used for triggering purposes and provide the time pattern of the shower front. A scheme of the read-out system is reported in figure 3.1. The time resolution of about 1 ns and the space granularity of the detector allow the imaging the shower front with extreme detail. An example of high accuracy sampling capability of the ARGO-YBJ detector is shown in figure 3.2. Signals coming from the strips readout are converted into digital signals by the front end boards mounted directly on the strip plane and sent to a special board called Local Station (LS). The local station is the basic unit of the data acquisition system and provides control and read–out of an entire cluster. The local station is essentially composed of 12 input boards which receive signals from each RPC in one cluster, an input/output board which is devoted to the communication with the Central Station and 4 TDCs (Time to Digital Converter) for time measurements. The Central Station is the unit in charge of memorizing the event. Each input board receives the output digital signal coming from the front end boards installed on a chamber and produces a FAST–OR output for each one. Each FAST–OR signal is sent to the TDCs and to the trigger system. When the trigger conditions are satisfied a *common stop* signal is

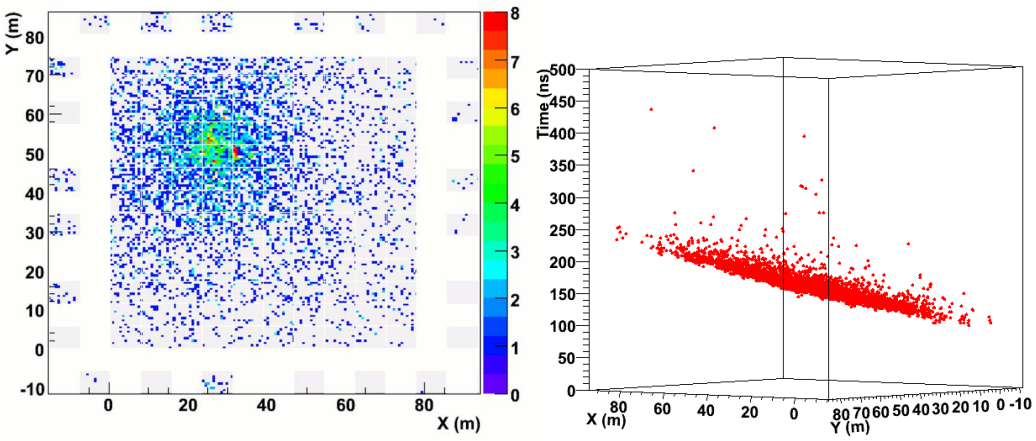


Figure 3.2: A typical event triggered by the ARGO-YBJ detector. The space hit density is obtained by the pattern in x–y projection, whereas the arrival direction is inferred from the time distribution, shown on the vertical coordinate.

sent to the TDCs and the local station assembles the data frame containing the addresses of fired strips and all the timing information from the TDCs. These information are transferred to the Central Station for event building and data storage. The Central Station contains the modules that receive the pad signals and produce the trigger signal. The digital read out system has a density of 23 strips/m^2 and can be used to investigate the primary spectrum up to energies of a few hundred TeVs. At greater energies the response starts to saturate, as shown in [52]. In figure 3.3 the average strip and pad size are compared with the size and the truncated size of induced proton showers at Yangbajing altitude. The figure shows that the strip size start to saturate at a few hundred TeV. In order to extend the dynamic range of the detector an analog readout system of the RPC detectors was implemented. This system, described in the next section, allow the detection of showers with a particle density around the core of about 10^4 particles/ m^2 and therefore allow the investigation of the PeV energy range.

3.2 Analog readout system

In order to extend the detector energy range up to the PeV region an analog charge readout system was implemented. Each RPC has been equipped with two large size copper electrodes (Big Pads) facing the lower side of the RPC

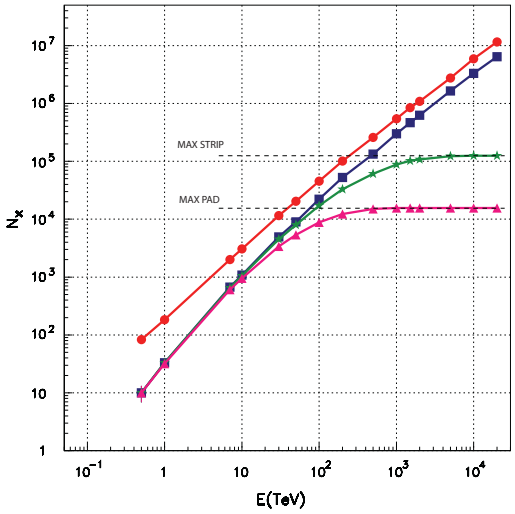


Figure 3.3: Shower size (red circles) and truncated size (blue squares) for proton induced showers at Yangbajing. The values of the average pad size (purple triangles) and strip size (green stars) are reported.

gas volume. In figure 3.4 the layout of the big pads in a cluster is reported. Each big pad provides a signal proportional to the number of charged par-

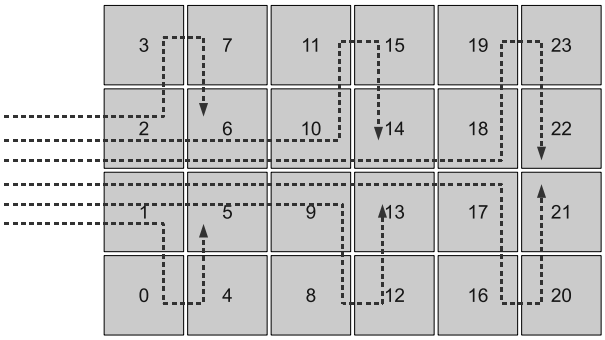


Figure 3.4: Scheme of the big pads in one cluster. The arrows shows the direction of the gas flow.

ticles reaching the detector surface. At present the analog system has been fully implemented on the 130 central clusters. The analog readout has been fully integrated in the DAQ system. Data taking started in December 2009. The amplitude of the big pad signal ranges from a few millivolts up to ten

volts. Each cluster is equipped with a custom crate containing three ADC boards and one control module that manages the analog data. The system is able to operate with different full–scale values, namely 0.33, 0.66, 1.3, 2.5, 5, 10, 20 and 40 V [53, 54]. Operating the system at different scales allows an overlap between digital and analog data, necessary for calibration purposes. The ADC digitalization and data collection in each cluster starts when the local number of hits is greater than a programmable threshold. Four different threshold are implemented: $\geqslant 16$, $\geqslant 32$, $\geqslant 64$ and $\geqslant 73$ hits. The data acquisition start when the local trigger is confirmed by the main detector trigger. Due to the local trigger the highest trigger rate of the analog readout is about 130 Hz. Since the trigger rate of the experiment is about 3.5 kHz, the analog readout does not introduce additional dead time. For each event satisfying the local trigger condition both digital and analog information are recorded. In figure 3.5 a typical event recorded by the two readout systems is shown. While the digital readout response saturates in the core region, the analog readout allow the measurement of the particle distribution in the shower front with high detail. In the period between December 2009 and June 2010 the detector was operated at 330 mV full scale, in which the particle distribution measured with the analog system overlaps the data taken with the digital readout. These data have been used to study the detector behavior and for calibration purposes. From July to middle August 2010 the system was operated at an intermediate full scale which corresponds to about 2.5 V. Since middle August 2010 the full scale was set to 20 V. The number of big–pads in the central carpet is 3120, apart from dead channels or channels with some problems which are of the order of 3%. By using the analog readout system the saturation of the digital readout can be avoided, especially in the core region, allowing the sampling of showers with particle density up to about 10^4 part/m^2, as shown in figure 3.6.

3.3 Detector calibration

3.3.1 Software timing calibration

The detector measures the arrival time of the charged particles in the shower front and the primary direction is reconstructed by means of a high resolution space–time picture of the shower. A timing resolution is necessary in order to achieve a high angular resolution and pointing accuracy. The calibration

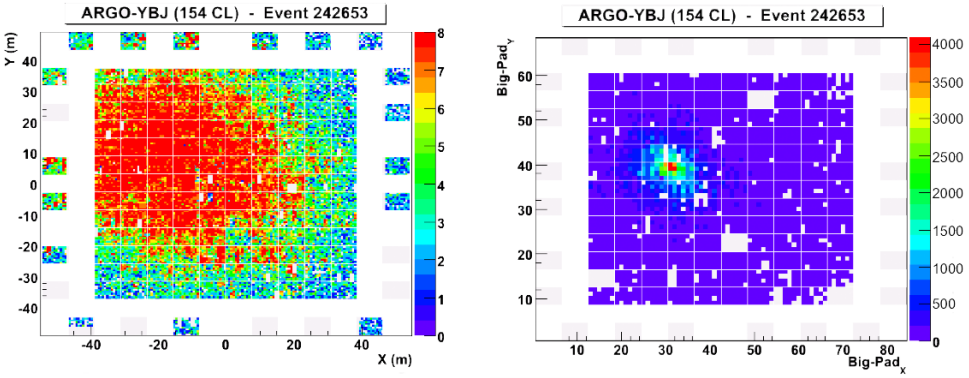


Figure 3.5: An event triggered by the two different readout system: digital (left) and analog (right). Is clearly visible the saturation of the digital readout.

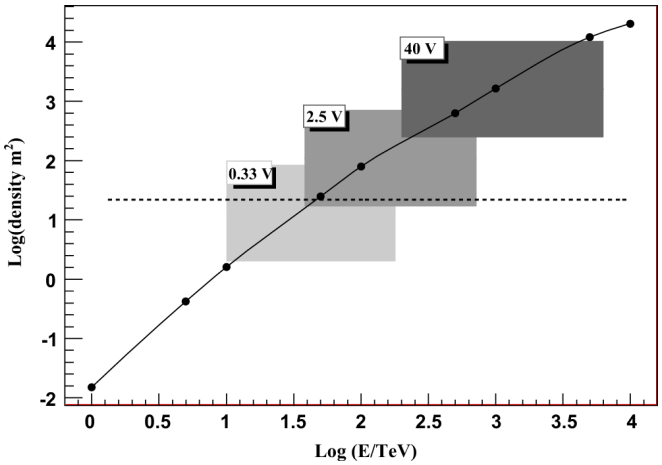


Figure 3.6: Charged particle density around the shower core as a function of energy, according to simulations of proton–induced air showers at Yangbajing altitude. The dotted line represents the maximum particle density that can be measured with the digital readout ($23/\,\mathrm{m}^2$). The boxes represent the different full–scales of the analog readout. The data at the highest energies are affected by a limitation on the maximum number of detectable particles introduced in order to avoid the instability of the software simulation of the detector.

removes systematical time offsets due to differences in the length of the cables, in the electronic boards and so on. Since the number of pixels is very large, a software time calibration is the fastest solution. The method takes the

secondary particles in a shower as a test beam because the shower axis is parallel to the primary direction. If the primary direction is known the detector can be calibrated with a set of shower events. Due to detector time offset there is a difference between the reconstructed direction and the true one. This difference corresponds to the slope of a characteristic plane defined by the time offset of detector units fired by the events [55]. Events firing the same units are characterized by the same characteristic plane. The direction cosines of each plane are the average of the direction cosines of the whole event set if the shower azimuth is uniformly distributed. The characteristic plane is estimated by the average of the whole event set, the reconstructed directions are corrected accordingly to the characteristic plane and used to evaluate the time offset of each detector unit [56]. Since the detector has a modular structure, the timing calibration was performed also during the construction phase, taking into account data coming from the active clusters. Data used for the calibration procedure are the standard data, only requiring the stability of the detector performance. Only events with core reconstructed in the central carpet and with more than 1500 fired pads are take into account for the calibration procedure. An example of the results of the calibration procedure are shown in figure 3.7. After the calibration the azimuth distribution becomes almost flat as expected for an isotropic cosmic ray flux. The small modulation observed after the calibration procedure is due to geomagnetic effects. Usually the calibration is valid up to 30–40 days.

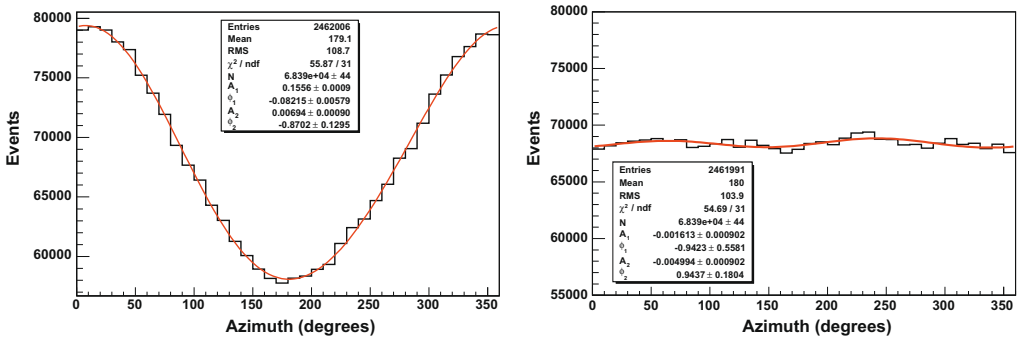


Figure 3.7: Azimuth distribution before (left) and after (right) the calibration procedure

3.3.2 Charge readout calibration

This is the first time that the RPC detectors are used with an analog readout system in extensive air shower experiments. The behavior of the detector at high particle densities cannot be predicted therefore an accurate calibration procedure is needed [57]. Since the ARGO–YBJ RPCs are operating at high altitude an on–site calibration is necessary. The main problem is to determine the exact number of charged particles impinging on the RPC to be calibrated. In order to perform a calibration of the analog readout of the RPC a telescope (see fig. 3.8) was setup with two scintillation detectors and five RPCs. The scintillators are used to measure the number

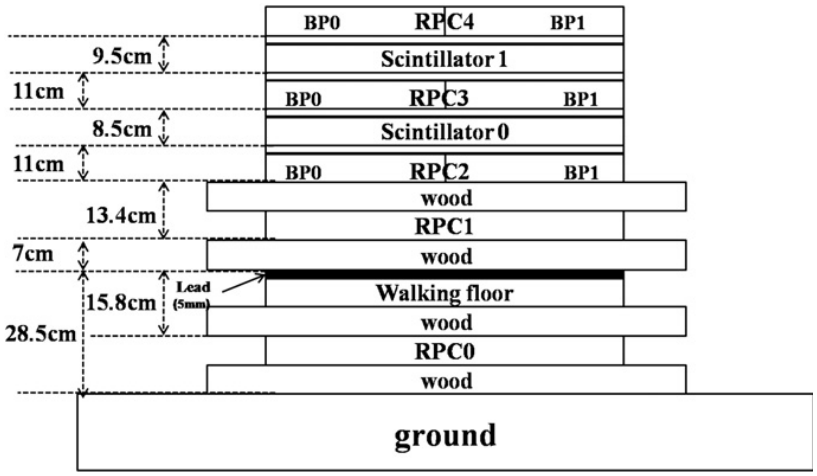


Figure 3.8: Sketch of the calibration telescope

of particles impinging on them, the RPC3 act as the one to be calibrated, RPC2 and RPC4 provide pad and strip signals that can be used to check the calibration. RPC0 and RPC1 allow to pick up coincident events from the ARGO–YBJ carpet. Charged particles in extensive air showers are used as calibration beam. The scintillation detector ($275 \times 125\,\mathrm{cm}^2$) consists on 11 tiles covering the RPC. Each scintillation detector is read–out by one photomultiplier tube. The telescope is plugged in the ARGO–YBJ DAQ system acting as an additional cluster, however a dedicated DAQ was implemented for calibration purposes. In order to convert the ADC counts to particles two calibration steps are needed: first converting the ADC counts to amplitude and then amplitude to particles. The calibration was performed by using

the data taken at 330 mV full scale, where the particle density measured by the analog readout is comparable to the one measured by the digital system. After the electronic calibration the resulting amplitude was compared with the number of fired strips. In figure 3.9 the measured amplitude as a function of the number of fired strips is reported. The fit was performed in the range between 8 and 15 strips. Above this threshold the density effect starts to be relevant, below 8 strips the electronic noise dominates. A linearity between the number of fired strips and the big–pad signal amplitude is quite evident. The second step of the calibration procedure allow the conversion of the values of the amplitude into the number of detected particles. In figure 3.10

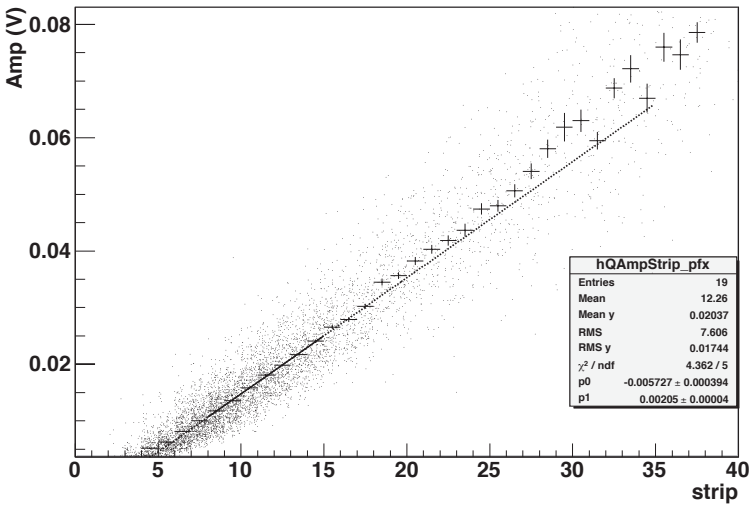


Figure 3.9: Single big pad amplitude versus number of fired strips. The line represent the fit performed in the range $8 - 15$ strips.

the relation between the number of particles and the RPC charge output is shown. The values of the residual nonlinearity as a function of the number of particles crossing the RPC is also shown. Data used for detector calibration were collected between the 80^{th} day and 160^{th} of 2010. The gain values were monitored over the whole period and the dependance on the atmospheric pressure and temperature was studied. These parameters are measured with high accuracy, namely ± 0.25 K and ± 0.5 mbar by the detector control system. The gain shows a daily modulation and correlation with pressure and temperature [53]. The gain of each channel can be corrected for pressure and

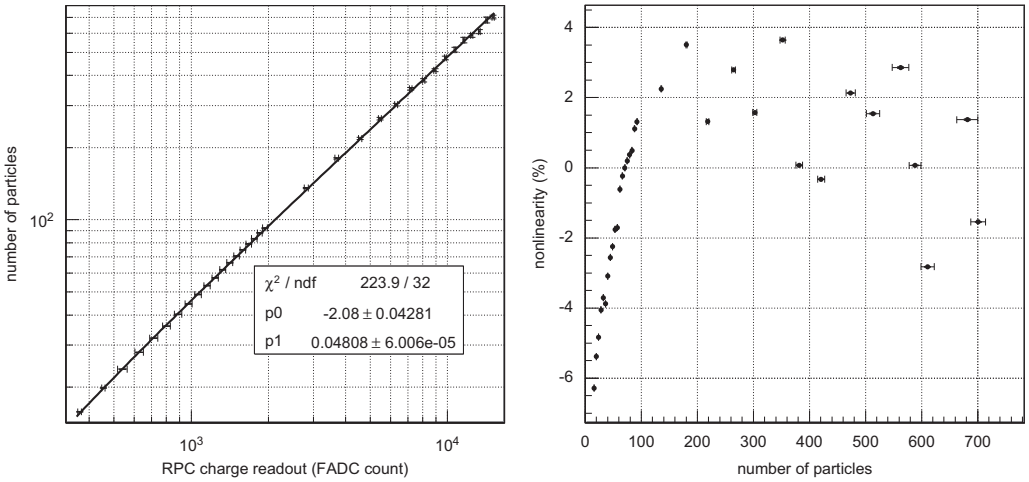


Figure 3.10: Number of charged particles as a function of the RPC charge readout (left) and nonlinearity of the calibration.

temperature dependency according to the relation

$$BP(i) = BP_0(i)\frac{P_0}{P}\frac{T}{T_0} \tag{3.1}$$

in which P_0 and T_0 are chosen as reference values and $BP_0(i)$ is the gain of the i–th big pad at pressure P_0 and temperature T_0.

3.4 Performance of the data taking

The ARGO-YBJ detector is characterized by high space–time resolution and high granularity. These characteristics allow a very detailed measurement of the shower profile. The high time resolution of the order of ~ 1 ns allow an accurate reconstruction of the shower direction. Since the shower front is sampled with high detail, it is possible to obtain very detailed informations about the lateral distribution of the particles in the shower. For events with the core localized inside the full–coverage area of the detector the reconstructed core position and the lateral density profile can be determined with a very high precision. The angular resolution has been measured with different methods and is better than 0.6° for events with more than 200 fired pads. Experimental results show that the angular resolution ψ_{72}, defined as the angle containing $71,5\%$ of the reconstructed events coming from a fixed

direction is about 0.4° for events with pad multiplicity of about 500 [58]. The values of the measured angular resolution as a function of the shower

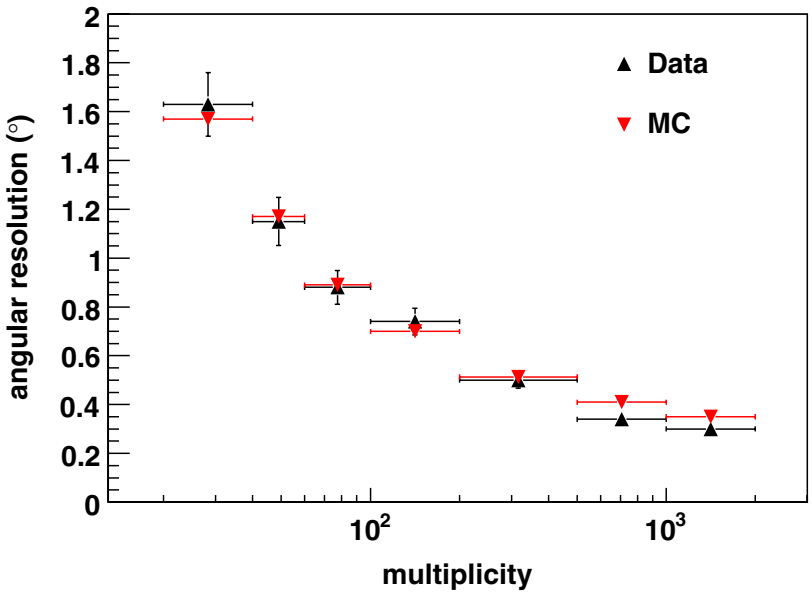


Figure 3.11: Measured angular resolution of the ARGO–YBJ detector (black triangles) compared to the expectations from the Monte Carlo simulation (downward red triangles) as a function of the shower multiplicity. The horizontal bars represent the multiplicity bins.

multiplicity are shown in figure 3.11. The accuracy of the reconstructed core position was estimated to be about 5 m from Monte Carlo events with the core localized in the full–coverage area of the detector. The analog readout system allows the extension of the operating energy range up to several PeVs. The installation of the central carpet was completed in June 2006. The guard ring was completed in spring of 2007 and was plugged in the DAQ system in November 2007. The detector is in operation since July 2006. After a period dedicated to commissioning and tests the detector is in full and stable data taking since November 2007 with a mean duty cycle of about 85% and with a mean dead time of about 4%. All data are recorded in a temporary buffer installed in the experiment building then they are transferred to IHEP (Beijing, China) and CNAF (Bologna, Italy) data centers for event reconstruction and analysis. The data transfer rate is about 220 GB/day. At present time more than $3.5 \cdot 10^{11}$ events have been collected.

CHAPTER 4

Energy spectrum deconvolution methods

In many experiments the distribution of the measured observables generally differs from the expected distribution of the "true" quantities. Since we are not working with ideal detectors, the distribution of the observables will be distorted by several effect (background, detector efficiency and so on). The true distribution must be inferred from the observed one. Generally one assume a mathematical parametric function that describes the true distribution of the observed quantity and that includes the noise and the detector effects. The problem is reduced in the estimation of the values of the free parameters of the function. This is the so–called *parametric inference* because all the information that can be inferred from the observed events is contained into the model parameters. This method, however, imply an interpretation of the observed data. If we want to evaluate the distribution of the true quantities apart from a strong interpretation of the observed data but taking into account the distortions of the spectrum due to physics and detector effects we are dealing with an *unfolding* problem. When dealing with only one variable the method used to unfold the true distribution from the observed data is the *bin to bin* correction. A generalized efficiency is evaluated from the ratio between the observed events falling in a certain bin of the reconstructed variable and the Monte Carlo events falling in the same bin. The value of this efficiency is therefore used to the estimate the number of true events in a certain bin starting from the observed events. This method, however,

does not take into account the migration of the events from one bin to one other and the correlation between adjacent bins. An attempt to solve the migration problem consists in building a matrix which connects the events observed in one particular bin to the events observed in the other bins. This matrix can usually be inverted and applied to the measured data. A singularity in this matrix, however, lead to inversion problems. The bayesian inference, on the other hand, is a way to learn information about physical quantities from the experimental data by using probability theory [59, 60, 61].

As it is well known the development of an atmospheric shower presents several fluctuations. Since the values these fluctuations can be very large the correlation between the primary energy and the space–time distribution of the shower particles cannot be obtained on an event by event basis. The energy distribution of the incoming particles must be evaluated by means of an unfolding procedure. The space–time distribution of the particles of the shower front is called multiplicity of the shower. The correlation between observed multiplicity and the primary energy is a cause–effect problem that can be dealt with the bayesian technique [62]. Moreover in the bayesian approach no assumption on the shape of the spectrum are needed.

4.1 Bayesian inference and bayesian unfolding

The purpose of the unfolding is to find the true number of events $x(C_i)$ related to each cause C_i given the observed spectrum of effects E_j and assuming some knowledge about the quantities that allow the connection between causes and effects. Since the connection between causes ad effects are of probabilistic nature, also the link between the connection between effects and causes will be probabilistic (see figure 4.1). The quantity that can be determined is the probability that a spectrum of causes might have produced the observed effects, namely

$$P(\mathbf{x}_C|\mathbf{x}_E, \Lambda, I), \tag{4.1}$$

where $\mathbf{x}_C = \{x(C_1), x(C_2), \ \ldots, \ x(C_{n_C})\}$ is the number of events in each bin of the true distribution, $\mathbf{x}_E = \{x(E_1), x(E_2), \ \ldots, \ x(E_{n_E})\}$ is the observed spectrum of effects, I contains the information under which the analysis is performed and Λ is the smearing matrix. The smearing matrix Λ is defined

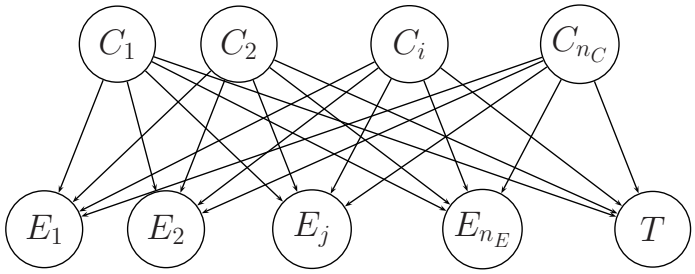


Figure 4.1: Schematic view of the probabilistic links between causes and effects. The thrash bin (T) is used to describe the inefficiencies.

in probabilistic terms as

$$\lambda_{ji} = P(E_j|C_i, I) \tag{4.2}$$

which values can be determined by means of a Monte Carlo simulation. The quantity $P(\mathbf{x}_C|\mathbf{x}_E, \Lambda, I)$ is related to the smearing matrix Λ by the Bayes' theorem which states that:

$$P(\mathbf{x}_C|\mathbf{x}_E, \Lambda, I) = \frac{P(\mathbf{x}_E|\mathbf{x}_C, \Lambda, I) \cdot P(\mathbf{x}_C|I)}{\sum_{\mathbf{x}_C} P(\mathbf{x}_E|\mathbf{x}_C, \Lambda, I) \cdot P(\mathbf{x}_C|I)} \tag{4.3}$$

The denominator of eq. 4.3 is only a normalization factor, so we can rewrite the Bayes' formula as:

$$P(\mathbf{x}_C|\mathbf{x}_E, \Lambda, I) \propto P(\mathbf{x}_E|\mathbf{x}_C, \Lambda, I) \cdot P(\mathbf{x}_C|I), \tag{4.4}$$

where $P(\mathbf{x}_E|\mathbf{x}_C, \Lambda, I)$ is called *likelihood*, $P(\mathbf{x}_C|I)$ *prior* and $P(\mathbf{x}_C|\mathbf{x}_E, \Lambda, I)$ *posterior*. The fact that the posterior has a dependence on the prior can look trivial, however, in terms of bayesian analysis, the prior is necessary in order to evaluate the posterior starting from the likelihood. The values of the prior can be chosen according to all relevant information that might come from previous experimental data or from theoretical models. The prior, in this case, assigns different probabilities to the different possible spectra. The values of the prior usually have a weak influence on the values of the posterior and the inference procedure is often dominated by the values of the likelihood. If we choose a vague prior like $P(\mathbf{x}_C|I) = const$ we assume that all the possible spectra have the same probability to occur and the equation 4.4 becomes:

$$P(\mathbf{x}_C|\mathbf{x}_E, \Lambda, I) \propto P(\mathbf{x}_E|\mathbf{x}_C, \Lambda, I). \tag{4.5}$$

It follows that the most probable spectrum of true events $\mathbf{x}_C$ is the one that maximizes the likelihood. The problem rely in finding an expression for the likelihood. If we consider a certain number of events in one cause bin $x(C_i)$, the number of events in the all effect bins $x(E_j),\ j = \{1, \ldots,\ n_E\}$ including the thrash bin for the inefficiencies can be described by a multinomial distribution:

$$\mathbf{x}_E|_{x(C_i)} = \frac{x(C_i)!}{\prod_j^{n_E+1} x(E_j)!} \prod_j^{n_E+1} \lambda_{ji}^{x(E_j)} \tag{4.6}$$

with $\lambda_{ji} = \{\lambda_{1,i},\ \lambda_{2,i},\ \ldots,\ \lambda_{n_E+1,i}\} = \{P(E_1|C_i, I),\ \ldots,\ P(E_{n_E+1}|C_i, I)\}$. The likelihood $P(\mathbf{x}_C|\mathbf{x}_E, \Lambda, I)$ will therefore be a sum of independent multinomial distributions, which doesn't have a closed expression. For this reason is not possible to apply the bayesian unfolding method to the true and observed spectrum. It's possible, however, to apply the bayes theorem to causes and effects, as described in the next section.

4.1.1 Unfolding a sample of data

Consider a set of independent causes C_i $(i = 1, \ldots, n_C)$ which can produce an effect E_j $(j = 1, \ldots, n_E)$. These quantities are connected by a probability that one of the effect E_j comes from at least one of the causes C_i (see figure 4.1). Let us assume that we know the initial probability of each cause $P(C_i)$ and the conditioned probability $P(E|C_i)$ that the effect E occurs given the *i-th* cause C_i. The Bayes formula states that:

$$P(C_i|E_j, I) = \frac{P(E_j|C_i, I)P(C_i|I)}{\sum_i P(E_j|C_i, I) \cdot P(C_i|I)} \tag{4.7}$$

or

$$\theta_{ij} = \frac{\lambda_{ji} P(C_i|I)}{\sum_i \lambda_{ji} \cdot P(C_i|I)} \tag{4.8}$$

where $\theta_{ij} = P(C_j|E_i, I)$ and $\lambda_{ji} = P(E_j|C_i, I)$. The conditioned probability λ_{ij} must be evaluated by means of a full Monte Carlo simulation as $\lambda_{ij} = N(E_j)^{MC}/N(C_i)^{MC}$. Having evaluated $P(C_i|E_j, I)$ we can use it to distribute the events observed in each effect bin among all the cause bins (see figure 4.2). The true spectrum can be estimated by repeating this sharing for all the effect bins and taking into account the inefficiency. The number of counts in C_i due to the observation in E_j will be

$$N(C_i)|_{n(E_j)} \propto P(C_i|E_j, I) \cdot N(E_j) = \theta_{ij} \cdot N(E_j), \tag{4.9}$$

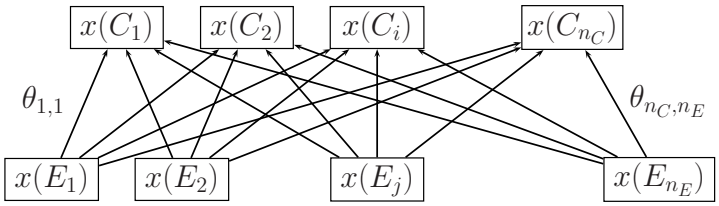


Figure 4.2: Sharing of the counts between effects bins and the cause bins

therefore the number of events in C_i due to all the observations can be estimated as

$$N(C_i) \propto \sum_{j=1}^{n_E} P(C_i|E_j, I) \cdot N(E_j) = \sum_{j=1}^{n_E} \theta_{ij} \cdot N(E_j). \tag{4.10}$$

Taking into account the efficiencies the eq. 4.10 becomes:

$$N(C_i) \propto \frac{1}{\varepsilon_i} \sum_{j=1}^{n_E} \theta_{ij} \cdot N(E_j), \quad \varepsilon_i \neq 0 \tag{4.11}$$

where

$$\varepsilon_i = \sum_{j=1}^{n_E} P(E_j|C_i, I) = \sum_{j=1}^{n_E} \lambda_{ji}. \tag{4.12}$$

Notice that if $\varepsilon_i = 0$ also $N(C_i) = 0$. From the number of unfolded events evaluated in eq. 4.11 it is possible to estimate the total number of unfolded events, the probability of each cause C_i and the overall efficiency. The prior used in equation 4.7 assigns a probability to each cause. A flat prior means that all the causes has the same probability. This is a strong assumption that can be recovered by using an iterative procedure. The values of $P(C_i|E_j, I)$ are computed by the equation 4.7 assuming a starting value of the prior $P(C_i|I)$. Using the equation 4.10 a more accurate value of $N(C_i)$ is derived, which can be used to estimate an updated value of $P(C_i)$ and to iterate. The iterative procedure ends when further variations on the values of $P(C_i)$ can be considered negligible. Usually a small number of iterations are enough to ensure a stable convergence of the procedure. If the number of iterations becomes to high the fluctuations can be amplified and the resulting spectrum may not reproduce the true one [59]. Monte Carlo simulations can be used to determine the optimum value for the number of iterations. In order to ensure

a stable and fast convergence an intermediate smoothing can be applied to regularize the posterior before using as new prior.

4.2 Unfolding of the cosmic ray spectrum

A typical air shower experiment like ARGO–YBJ detects the charged particles produced by the interaction of the primary nuclei with the atmosphere. The space–time distribution of the particles of the shower front is called multiplicity of the shower. The observable measured by the ARGO-YBJ experiment is the multiplicity distribution $N(M)$, which represents the number of events with a given multiplicity M collected in a fixed period Δt and within a solid angle Ω. The rate of observed events can be obtained by integrating the primary cosmic rays differential intensity $N(E)$ on the energy range of the primary particles and on the overall acceptance of the detector:

$$N(M) = \int_{E_1}^{E_2} \int_{\Omega} A_{eff}(E', M, \vartheta')\, N(E')\, dE'\, d(\cos\vartheta')\, d\varphi', \tag{4.13}$$

where the function $A_{eff}(E, M, \vartheta)$ is the effective area of the detector and is a function of the energy E, the multiplicity M and the zenith angle ϑ. Since the effective area does not depend on the polar angle φ and for nearly vertical events depend only weakly on the zenith angle ϑ, the mean value within the observation solid angle Ω it is often considered:

$$\bar{A}_{eff}(E, M) = \frac{1}{\Omega} \int_{\Omega} A_{eff}(E, M, \vartheta') d(\cos\vartheta')\, d\varphi'. \tag{4.14}$$

Combining equation 4.13 with equation 4.14 the rate of measured events is given by:

$$N(M) = \Omega \int_{E_1}^{E_2} A_{eff}(E', M)\, N(E')\, dE'. \tag{4.15}$$

This equation represents the connection between the cosmic ray flux intensity $N(E)$ and the observable quantity. The quantity $N(M)$ in equation 4.15 can be determined from the ARGO–YBJ data after applying a set of selection criteria on the measured data. The classical approach is based on a fitting procedure which requires a hypotesis on the shape of the spectra of different primaries. The function $A_{eff}(E', M)$ can be determined by means of a Monte Carlo simulation of the shower production and the detector response.
Extracting the cosmic ray flux from the equation 4.15 is a classical unfolding

problem and can be dealt with the bayesian technique [63]. The Bayesian approach does not require any assumption about the shape of the primary spectra.

4.2.1 Unfolding of the cosmic ray spectrum by using the bayesian method

Extracting the cosmic ray energy distribution from the observed multiplicity distribution is a problem that can be easily dealt with the bayesian technique. A primary cosmic ray of energy E_i (cause) produce an atmospheric shower that can be detected with a multiplicity M_j (effect). All the quantities described in the previous section can be redefined in terms of energy and measured multiplicity:

- $P(M_j|E_i)$: probability that a shower produced by a primary of energy E_i is detected with a multiplicity M_j

- $P(M_j)$: probability of detecting an event with multiplicity M_j, the quantity measured by the ARGO–YBJ detector

- $P(E_i)$: probability of observing a shower produced by a primary of energy E_i, the quantity to be obtained

These probabilities are connected by the Bayes theorem, which in the case of n_E independent energies (causes) each responsible for n_M multiplicity M_j states that:

$$P(E_i) \propto \sum_{j=1}^{n_M} P(E_i|M_j) \cdot P(M_j) \tag{4.16}$$

$$P(E_i|M_j) = \frac{P(M_j|E_i) \cdot P(E_i)}{\sum_{l=1}^{n_E} P(M_j|E_l)P(E_l)} \tag{4.17}$$

where $P(E_i|M_j)$ is the probability that a shower detected with a multiplicity M_j has be produced by a primary of energy E_i. In the bayesian scheme $P(M_j|E_i)$ must be evaluated by means of a Monte Carlo simulation, the values of $P(M_j)$ are calculated on the basis of the experimental data and the values of $P(E_i)$ are related to cosmic rays differential intensity. The bayesian unfolding is performed according to the iterative procedure described in the previous section: the values of $P(E_i|M_j)$ are computed by using the equation

4.17 assuming a starting value of the energy distribution. By using the equation 4.16 a more accurate value of $P(E_i)$ can be evaluated and is used to iterate. The iterative procedure continues until there are no further variations on $P(E_i)$. The relevant quantities are defined as follows:

$$P(E_i) = \frac{N(E_i)}{\sum_{E'} N(E')} \tag{4.18}$$

$$P(M_j) = \frac{N(M_j)}{N_{sel}} \tag{4.19}$$

where $N(E_i)$ is the number of events with energy E_i, N_{sel} is the total number of selected events and $N(M_j)$ is the number of detected events with multiplicity M_j.

4.2.2 Application of the bayesian unfolding method to the ARGO–YBJ data

In order to check the performance of the bayesian unfolding applied to the ARGO–YBJ data several consistency checks were performed. Two independent Monte Carlo data samples S_1 and S_2 were produced. The multiplicity distribution extracted from the sample S_1 was unfolded by using the conditioned probabilities evaluated from the sample S_2, starting from a flat prior. The unfolded spectrum was compared with the true spectrum of the sample S_1. The convergence of the unfolding procedure and the number of iterations was also checked. In figure 4.3 the results of these tests are shown. The unfolded distribution matches the true spectrum of the sample S_1 after only three iterations. In order to investigate eventual bias effects due to the characteristics of the Monte Carlo simulations used to evaluate the conditioned probabilities the following test were performed. Additional two independent Monte Carlo data sample S_3 and S_4 were generated with different energy distributions. These two samples were used to evaluate two conditional probabilities which in turn were used to unfold the multiplicity distribution extracted from the sample S_1. Both the unfolded spectrum has been compared with the true spectrum of the sample S_1, showing no significative differences. The results are shown in figure 4.4.

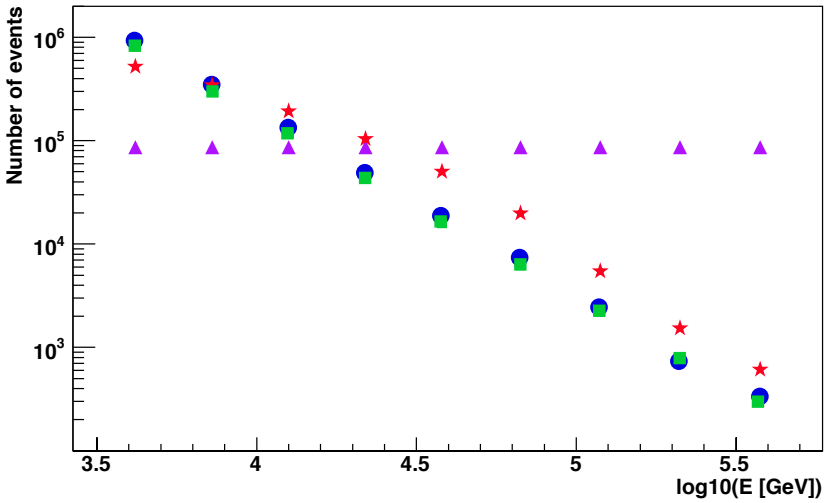


Figure 4.3: Check of the performance of the bayesian unfolding applied to a simulated data sample of the ARGO–YBJ experiment. Values for the prior (purple triangles), true distribution of the sample S_1 (blue circles) and unfolded distribution after one (red stars) and three iterations (green squares) are shown.

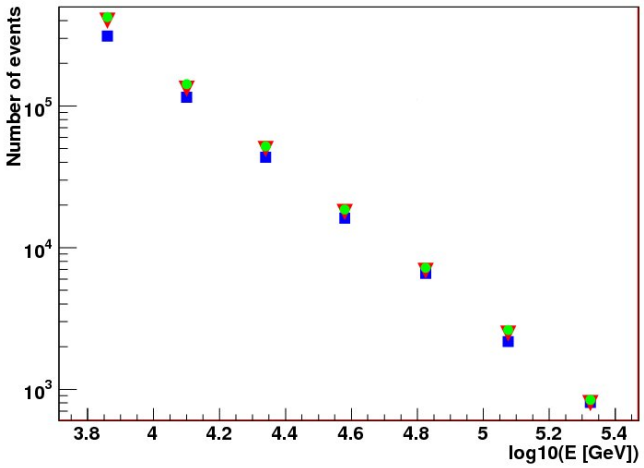


Figure 4.4: Energy distribution of the sample S_1 unfolded by using the conditioned probabilities evaluated from the sample S_3 (green circles) and S_4 (red triangles). The true energy distribution of the sample S_1 is also shown (blue squares).

CHAPTER 5

The cosmic ray spectrum

As described in the previous sections the indirect measurement of the cosmic ray energy spectrum and composition rely on the Monte Carlo simulation of the development of the showers and of the detector response. The ARGO–YBJ experiment observable is the shower multiplicity, namely the space–time distribution of the charged particles in the shower front. The energy distribution of the primaries cannot be determined by an event–by–event basis but must be evaluated by means of an unfolding procedure. As described in the previous chapter the unfolding of the cosmic ray spectrum from the observed data can be dealt with the bayesian method. In the bayesian unfolding scheme the necessary conditioned probabilities must be evaluated from Monte Carlo simulations.

5.1 The Monte Carlo data sample

In order to evaluate the relevant quantities used in the bayesian unfolding procedure a full Monte Carlo simulation is needed. Showers were produced by using the CORSIKA code (ver 6.90) [64] which provides a complete simulation of the shower development in the Earth's atmosphere. The electromagnetic component of the shower simulation is implemented by means of the EGS4 routines [65, 66], while for the hadronic component several options

are available. The Monte Carlo events used in this analysis were generated by using the QGSJET-II [67, 68] interaction model for the high energy hadronic interactions and FLUKA [69] for the low energy hadronic interactions. The data were generated in the energy range $(0.1 \div 10^4)$ TeV with energy distribution given by $dN(E)/dE = N_0 \cdot E^{-\gamma}$. Showers were sampled at the Yangbajing altitude. In order to accurately reproduce the detector response a full detector simulation based on GEANT-III [70] was applied. The accidental background generated by each pad has been included and a full trigger simulation performed, taking into account the measured efficiency and the time resolution of the RPCs detectors. For each event the response of the digital and analog readout was simulated. Monte Carlo events were produced in the same format as data and they have been processed through the same reconstruction code. Showers induced by protons, helium nuclei, CNO group and iron nuclei were generated.

5.2 The light component spectrum of cosmic rays in the multi-TeV region.

The ARGO–YBJ experiment takes data by means of an inclusive trigger [71] requiring a number of fired pads $N_{pad} \geqslant N_{trig}$ in the central carpet within a time coincidence window of 420 ns. During the runs selected for this analysis the detector was in stable data taking and N_{trig} was set to 20, with a resulting trigger rate of about 3.5 kHz and with a dead time of $\sim 4\%$. As described in section 3.3.1, special care is devoted to the time calibration of the 18360 pads of the detector in order to achieve a high pointing accuracy and good angular resolution. The core position (X_{core}, Y_{core}) of each detected shower is estimated by using the Likelihood method based on the Nishimura–Kamata–Greisen function. The accuracy of the reconstructed core position is expected to be about 5 m for events with the core localized in the full coverage area of the detector. In order to perform an unbiased analysis a set of selection criteria has been adopted for both Monte Carlo and data.

5.2.1 Data selection

A first selection of the data has been based on the quality of the runs and of the reconstruction procedure. Additional cuts have been applied in or-

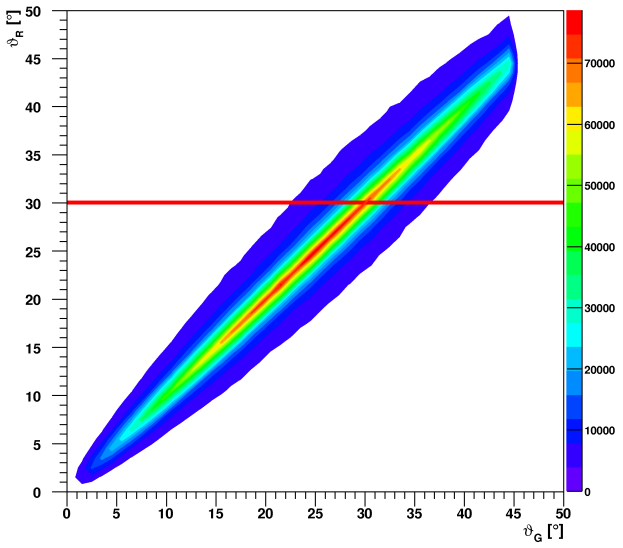
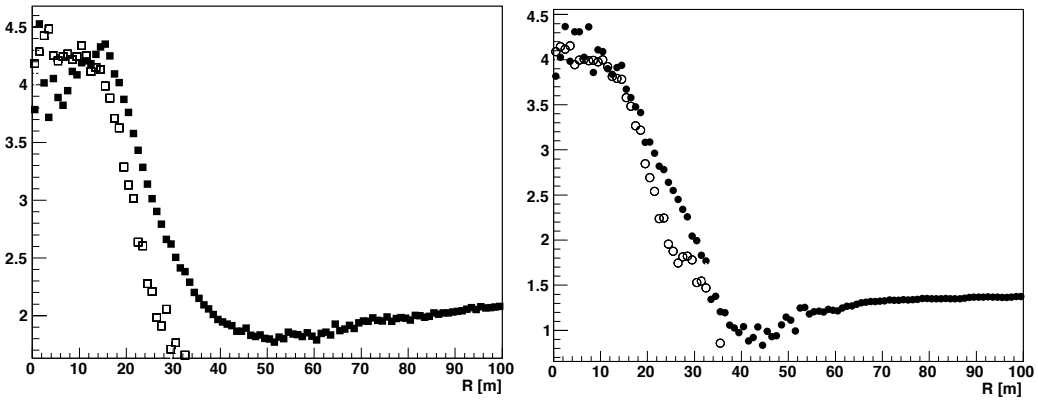


Figure 5.1: Distribution of the reconstructed zenith angle ϑ_R versus the generated zenith angle ϑ_G. The cut on $\vartheta_R \leqslant 30°$ is also shown.

der to estimate with good accuracy the probabilities used in the bayesian unfolding procedure and to make negligible the contamination of external events (like showers with the core position outside the detector area but mis–reconstructed inside, or events with mis–reconstructed zenith angle). The following selection criteria have been adopted for both Monte Carlo and data:

- The Monte Carlo events used in the analysis have been generated with the zenith angle (ϑ_G) in the range $0° \div 45°$. In order to avoid bias effects in estimating the quantities $P(M_j|E_i)$ and to improve the quality of the reconstruction, data and Monte Carlo events have been selected requiring the reconstructed zenith angle (ϑ_R) in the range $0° \div 30°$. This cut sets the solid angle Ω to about 0.842 sr. In figure 5.1 the distribution of the reconstructed zenith angle ϑ_R versus the generated zenith angle ϑ_G is reported, and the cut on ϑ_R is also shown.

- The Monte Carlo events have been generated in the energy range ($0.1 \div 10000$) TeV. In order to select well reconstructed events and reduce bias effects in the estimate of the Bayesian probabilities, mainly located at the edges of the energy range used in this analysis, the events have been selected



(a) Monte Carlo **(b)** Data

Figure 5.2: Ratio between the particle density measured in the innermost clusters (ρ_{in}) and in the outermost ones (ρ_{out}) as a function of the radius R from the detector center for Monte Carlo (filled squares) and data (filled dots). The values of ρ_{in}/ρ_{out} for events surviving the selection cuts used in this analysis are also reported for Monte Carlo (empty squares) and data (empty circles)

requiring the strip multiplicity M, measured for each event, in the range $500 \leq M \leq 50000$. This cut selects events well within the energy range used for the event simulation.

- the rejection of events falling far from the detector is achieved by comparing the average particle density measured by the innermost 20 clusters (ρ_{in}) to that measured by the outermost 42 ones (ρ_{out}). In figure 5.2a and 5.2b the values of the ratio ρ_{in}/ρ_{out} as a function of the distance from the detector center are shown. By requiring $\rho_{in} > 2.25 \times \rho_{out}$, the core position turns out to be inside an area of radius of about 28 m centered on the detector. No events generated outside the detector have been selected by this procedure.

The fraction of events passing the cuts used in this analysis has been checked to be consisted with Monte Carlo predictions, being about 24% both in data and simulations. As a consequence of the selection cuts used in this analysis only about 2% of the showers induced by the CNO group is included in the selected data sample. The cut based on particle density allows the selection of showers with well–shaped core, discarding events produced by heavier primaries. In figure 5.3 the fraction of the Monte Carlo events

selected by the described cuts is shown as a function of energy for Protons, Helium and CNO nuclei. The figure shows clearly that the events outside the energy range $(10^3 \div 10^7)$ GeV give a negligible contribution to the estimation of the conditioned probabilities $P(M_j|E_i)$. Furthermore the fraction of the CNO induced showers selected by the cuts is reduced by about ten times compared with the fraction of the protons and helium nuclei. In figures 5.4

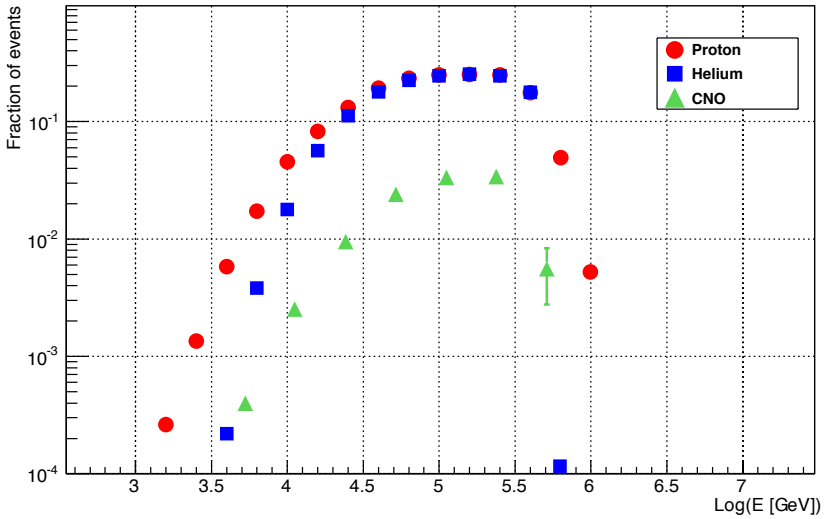


Figure 5.3: The fraction of the simulated events selected by the cuts used in this analysis is shown as a function of the energy for proton induced events (red circles), helium induced events (blue squares) and CNO induced events (green triangles).

and 5.5 the distribution of the distance of the core position from the centre of the ARGO-YBJ detector is shown both for Monte Carlo events and data. The figure shows clearly that the events selected by the cuts are located in an area of radius of about 28 m centered on the detector.

5.2.2 The unfolding procedure

In the bayesian method the energy bins are independent, without constraints among each other [61]. In order to measure the light–component spectrum the Monte Carlo events have been sorted in seven energy bins and twelve multiplicity bins. The same multiplicity bins were used to analyze a sample of data consisting on about 7.5×10^7 events collected in the first 2011 months.

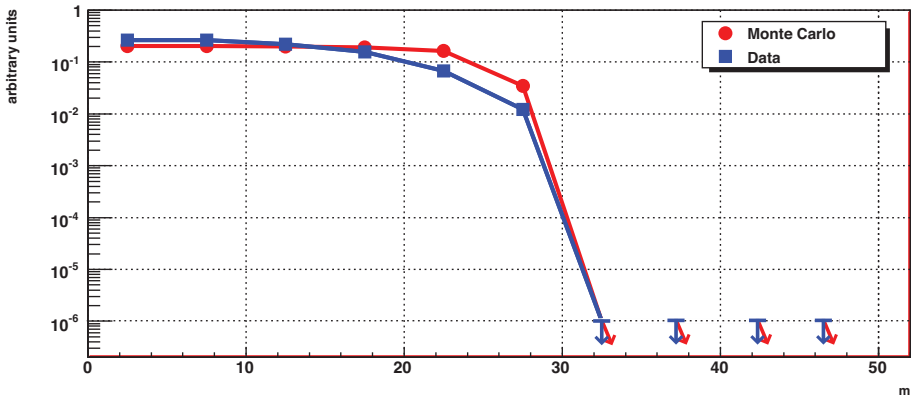
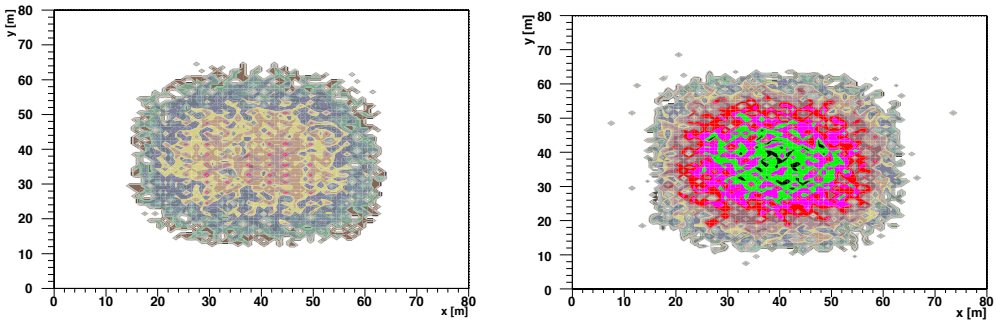


Figure 5.4: Distribution of the distance of the core position from the centre of the ARGO-YBJ detector for Monte Carlo events (red circles) and for data events (blue squares) selected by the criteria used in this analysis.



(a) Data **(b)** Monte Carlo

Figure 5.5: Distribution of the reconstructed core position for the selected data (a) and Monte Carlo (b) samples

Data were collected by the detector in its full configuration requiring an inclusive trigger of 20 fired pads in a time window of 420 ns. The bayesian unfolding has been performed by using a flat distribution as the initial value of the $P(E_i)$. The effect of using different prior distributions $P(E_i)$ has been evaluated as negligible. A soft smoothing [63] has been applied to the n–th value of $P(E_i)$ during the recursive procedure in order to ensure a stable convergence. A dedicated procedure has been used in order to evaluate the effective alive time of the detector for the runs used in this analysis and it

results to be about 96.2%.

5.2.3 Evaluation of systematical uncertainties

The measurements of the intensity of the spectrum are affected by a statistical error of the order of ±1%. A study of the possible sources of systematic effects has been performed. The sources of systematic effects that have been considered in this work are:

- effects related to the selection cuts on the measured quantities

- effects related to the reliability of the simulation of the detector response

- effects related to the variation of the fraction of the helium component

The quantities measured in this analysis are: the multiplicity of the shower (M), the reconstructed zenith angle ϑ and the particle density ρ. The contribution to the total uncertainty due the selection cuts on these quantities can be estimated by applying large variations to the selection cuts. The estimated uncertainty is about ±5%.
The relation between the measured strip multiplicity and the primary energy distribution has been studied in [72] and has been found in good agreement with Monte Carlo predictions. The reliability of the simulation procedure has been checked by comparing the distribution of several variables obtained applying the same selection cuts to data and to Monte Carlo events. As an example the distribution of the core position reported in figure 5.4 shows that the same fraction of events in the data and Monte Carlo samples has been selected in the same detector area. The multiplicity distribution of the Monte Carlo events has been compared to the experimental one by successfully applying the Kolmogorov–Smirnov test with $K = 0.081$ and 100 d.o.f. (see figure 5.6). A conservative estimate of the uncertainty due to the effect of different run condition (stability of the detector and environmental parameters) has been obtained by analyzing three different runs taken in different periods. The contribution to the total uncertainty turns out to be about ±4%, slightly larger in the edge bins.
The last contribution to the total uncertainty is related to the fraction of the helium component used to evaluate the conditioned probabilities $P(M_j|E_i)$. The variation of the helium component in the range $(7 \div 11)\%$ gives a negligible effect. In order to evaluate the effect of a large variation of the helium

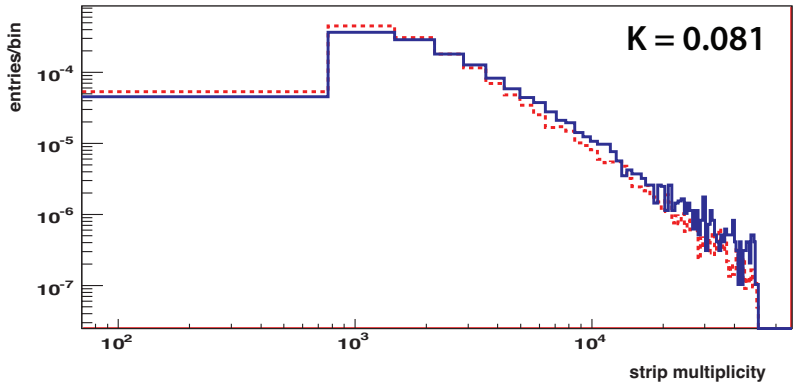


Figure 5.6: Normalized strip multiplicity distribution for Monte Carlo (dotted red line) and data (blue solid line), the Kolmogorov–Smirnov variable (K) is also shown.

component the unfolding procedure has been performed again by using a Monte Carlo data sample composed by 50% protons and 50% helium nuclei. The values of the spectrum obtained by using this sample are consistent, within a few percent, with the values obtained with the sample containing 9% of helium.
Therefore we estimate that the total uncertainty affecting this result is not exceeding 10%. In order to have an evaluation of the systematics due to the use of different hadronic interaction models, the conditioned probabilities $P(M|E)$ have been computed for three energy values namely: 10 TeV, 30 TeV, 120 TeV by using also the SIBYLL model [73, 74]. The values obtained are reported in figure 5.7 as a function of the multiplicity M of the detected showers and compared to the same results obtained by using QGSJET-II. The plots show that the probabilities provided by the two interaction models are very similar suggesting therefore a very small contribution to the systematic uncertainty.

5.2.4 The light component spectrum in the energy range $5 - 250$ TeV.

The values of the energy spectrum obtained by applying the Bayesian unfolding procedure to all data collected in the three runs are shown in figure 5.8.

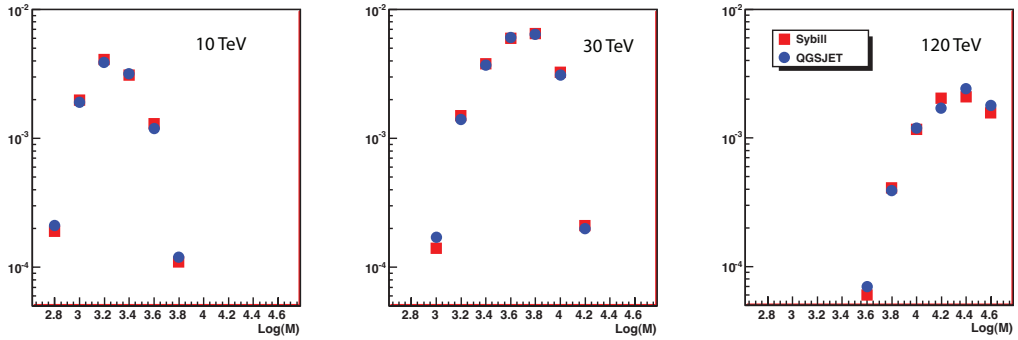


Figure 5.7: Conditioned probability $P(M|E)$ evaluated in three different energy bins 10 TeV, 30 TeV, 120 TeV) by using QGSJET-II (blue circles) and SYBILL (red squares) hadronic interaction models.

The spectrum covers the energy region from 5 TeV up to 200 TeV. The measured intensities are reported with the total uncertainty obtained by adding the systematic uncertainties, as estimated in the previous section, to the statistical errors. It can be noted that the total uncertainty on each experimental point doesnt exceed 10%. As shown in section 5.2.1, the contribution to the energy spectrum of elements heavier than helium nuclei is negligible. The data are compared with the recent results of the CREAM experiment and with the best fit provided by Hörandel to proton and helium experimental fluxes [18]. The point at 80 TeV represents the p + He intensity measured by EAS-TOP and MACRO experiments [75] at the Gran Sasso Laboratory combining the simultaneous detection of the electromagnetic and Cherenkov light components of atmospheric showers and of high energy muons. The ARGO-YBJ data agree remarkably well (within about 15 %) with the values obtained by adding up the proton and helium fluxes measured by CREAM both concerning the total intensities as well as the spectrum slope. The value of the spectral index of the power–law fit representing the ARGO–YBJ data is -2.62 ± 0.04, which should be compared to $\gamma_p = -2.66 \pm 0.02$ and $\gamma_{He} = -2.58 \pm 0.02$ obtained by CREAM. We note that the ARGO-YBJ data are mainly induced by protons since the average energy of helium primaries contributing to events with a given multiplicity M is about 1.5 times greater than the average proton energy.

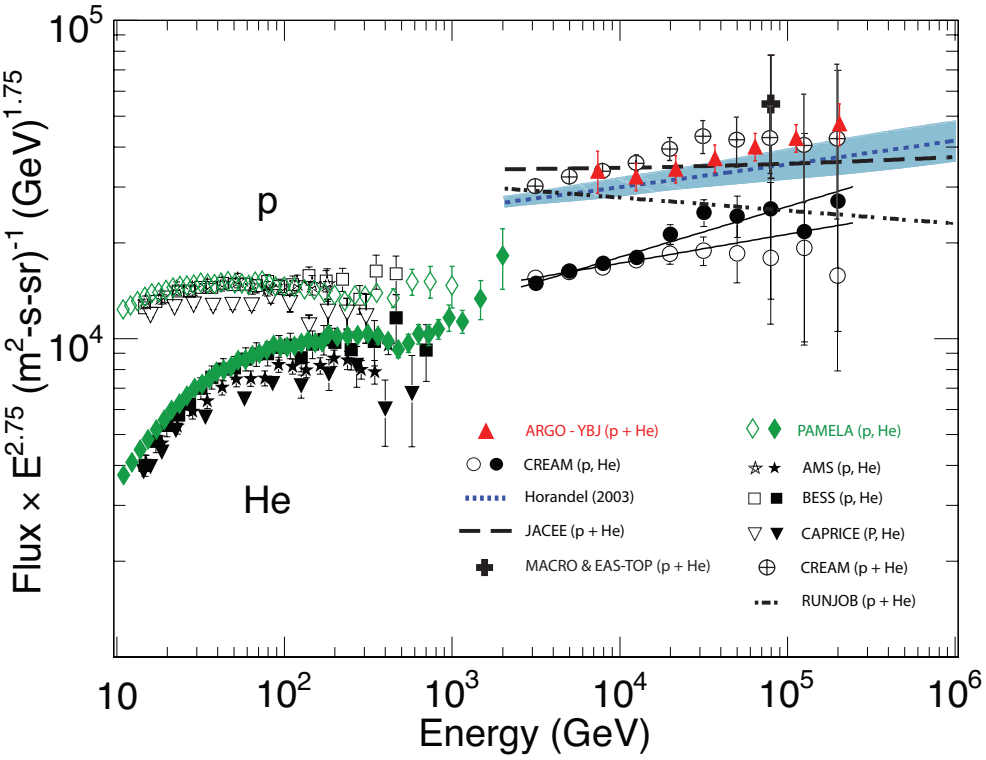


Figure 5.8: The differential energy spectrum of the light-component (proton and helium) measured by ARGO-YBJ (filled red triangles) superimposed to the proton spectrum (open circles) and helium spectrum (filled circles) measured by the CREAM experiment [22]. The crossed circles represent our sum of the proton and helium data measured by CREAM. The blue dotted line represents the best fit to proton and helium data performed by Hörandel [18]. The black cross represents the proton and helium flux measured by EAS-TOP and MACRO collaborations [75]. The spectra measured by PAMELA [76] (green diamonds), AMS [77] (stars), BESS [78] (squares), CAPRICE [24] (inverted triangles) are shown. Results from RUNJOB [19] (dashed–dotted line) and JACEE [21] (dashed line) are also reported.

5.3 The cosmic ray spectrum at higher energies

The determination of the elemental composition of the primary cosmic ray around the knee plays an important role in the understanding of the origin of the knee itself. As described in chapter 2 the ARGO–YBJ experiment

measures the space–time distribution by means of a full–coverage carpet of RPC detectors. The digital readout can measure showers with energies up to a few hundred TeVs. At greater energies the saturation effects start to become relevant. The maximum particle density that can be measured with the digital readout system is of the order of $10\,\text{particles}/\,\text{m}^2$. The analog readout system allow the detection of showers up to $10^4\,\text{particles}/\,\text{m}^2$ in the core region, allowing the extension of the detector operating range up to the PeV region. Since the ARGO–YBJ experiment is characterized by high granularity and space–time resolution the distribution of particles in the shower front can be measured with high precision. Showers produced by different primaries presents several differences in their lateral particle distribution. The analysis of the particle density as a function of the distance from the core of the shower can be used to discriminate the composition of cosmic rays.

5.3.1 Study of the chemical composition

Although the determination of the primary energy and direction is relatively easy, the estimation of the mass is a very difficult task. Showers produced by different primaries present several difference that are covered by the variation of the primary energy, by the uncertainties in the hadronic interaction models and by fluctuations of the shower itself. A mass composition discrimination can be based on a measurement of the charged particle distribution in EAS [79]. A detailed study of the charged particle distribution was performed on the Monte Carlo data sample described in section 5.1. Showers produced by protons, helium nuclei, CNO group and iron nuclei in the energy range $(31.6 \div 10^4)\,\text{TeV}$ were taken into account. A full detector simulation was applied, including the response of the analog readout system. The lateral particle density distribution of simulated showers was studied and a preliminary set of mass discrimination parameters was determined.

5.3.2 Effective particle density distribution

Since the charged particles are distributed around the core of the shower, the determination of the core position must be very accurate. In order to determine the core position for events detected by the analog readout system of the ARGO–YBJ detector a dedicated reconstruction algorithm based on

the determination of the center of mass around the big pad with the maximum signal was developed. The resolution of the core reconstruction was estimated by analyzing the Monte Carlo events and turns out to be of the order of 0.5 m, as reported in figure 5.9. In order to measure the lateral

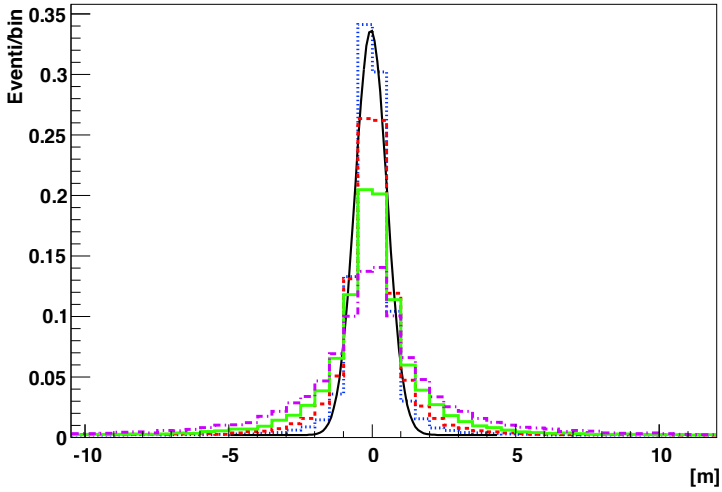


Figure 5.9: Resolution of the core position reconstruction for a sample of events with $E \approx 100\,\text{TeV}$. Data for protons (blue), helium (red), CNO group (green) and iron (purple) are reported. The black curve is the best fit with a Gauss function.

particle density distribution with good accuracy the events were selected by the following criteria:

- reconstructed zenith angle is in the range $(0° \div 30°)$
- reconstructed core position located inside a radius of 30 m centered in the detector center
- multiplicity of the shower N_{part}, measured for each event must be in the range $2000 \leqslant N_{part} \leqslant 3 \cdot 10^7$.

As described in the previous section, the lateral density distribution of the charged particles in the shower front presents a large number of fluctuations, in particular at distances far from the core position. Showers produced by heavier nuclei often present a multi–core structure, with several clusters of particles far from the core position. In order to emphasize this characteristic the effective particle density was taken into account. The particle density was

computed on a series of circular rings centered on the shower core position. Only the big–pads with $N_{part} > 0$ have been taken into account. Events were sorted in four multiplicity bins, reported in table 5.1. For each multiplicity bin the charged particle density was studied and a set of composition–sensitive parameters was determined.

Multiplicity Range	Mean energy [TeV]			
	p	He	CNO	Fe
M1 $2.0 \cdot 10^3 - 2.72 \cdot 10^4$	44.9	46.9	49.8	55.3
M2 $2.72 \cdot 10^4 - 2.22 \cdot 10^5$	89.9.0	127.8	152.5	218.4
M3 $2.22 \cdot 10^5 - 2.59 \cdot 10^6$	617.4	713.2	762.5	965.7
M4 $2.59 \cdot 10^6 - 3.0 \cdot 10^7$	8013	6514	7018	7465

Table 5.1: The multiplicity bins used in this analysis and the corresponding mean energy for each primary type.

5.3.3 Composition sensitive parameters

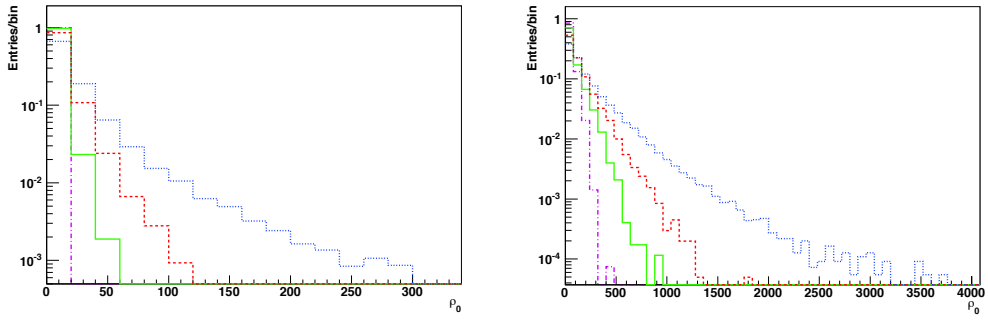
The following composition–sensitive parameters have been determined by analyzing the lateral particle density distributions:

- ρ_0: particle density measured in an area of about $4\,\mathrm{m}^2$ around the core and centered in the core position

- $\beta = \rho_{36}/\rho_0$: ratio between the particle density measured at 36 m and the particle density around the core

- $B = \sum_{i=0}^{36} \rho_i/\rho_0$: sum of the values of the ratio ρ_i/ρ_0 up to 36 m from the core position.

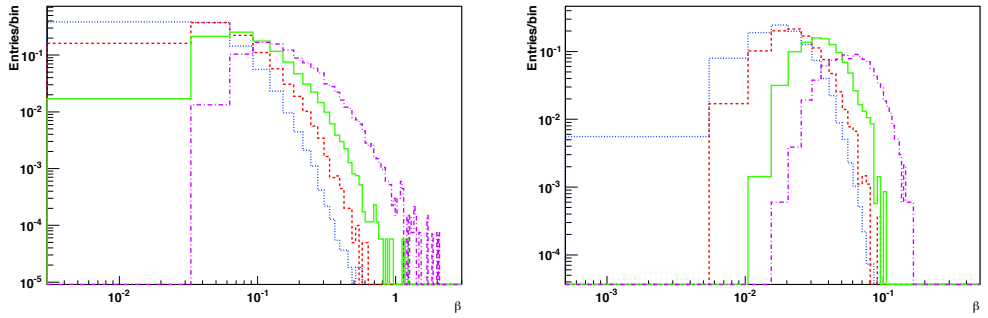
The parameter ρ_0 measures the charged particle density around the shower core. The distribution of the ρ_0 parameter in the multiplicity range M_2 and M_3 is reported in figure 5.10. The parameters β and B has been introduced to describe the shape of the shower front. The first is related to the extension of the core region, which is narrower for showers produced by light nuclei (protons and helium) and broader for the heaviest ones (CNO, iron). The parameter B also take into account the different space–time pattern of the charged particle distribution as a function of the primary mass. The



(a) M_1 **(b)** M_2

Figure 5.10: Distribution of the ρ_0 parameter in the multiplicity ranges M_1 and M_2. Values for protons (blue), helium nuclei (red), CNO group (green) and iron nuclei (purple) are reported.

distribution of β in the multiplicity range M_3 and M_4 is reported in figure 5.11. By choosing different values of the parameter β it is possible to select showers mainly induced by the light component (proton and helium nuclei) or by the heavy one (CNO and iron nuclei).



(a) M_2 **(b)** M_3

Figure 5.11: Distribution of the β parameter in the multiplicity ranges M_2 and M_3. Values for protons (blue), helium nuclei (red), CNO group (green) and iron nuclei (purple) are reported.

5.3.4 Light/heavy component discrimination

The distribution of the parameters described in the previous section has been used to set up a selection function

$$S_{L/H}(\rho_0, \beta, B) \tag{5.1}$$

that allow the discrimination of showers produced by light or heavy primaries. In table 5.2 the values of the discrimination efficiency and contamination for the light component (LH) and the heavy component (HC) are reported. The values of the efficiencies and contaminations are referred to the Monte Carlo data sample used in this analysis. At present time only a few months of data

	ρ_0, β		ρ_0, B	
	M_3	M_4	M_3	M_4
LC	0.18 (0.03)	0.55 (0.15)	0.14 (0.01)	0.63 (0.05)
HC	0.38 (0.35)	0.60 (0.15)	0.64 (0.18)	0.43 (0.08)

Table 5.2: Values of the efficiency (contamination) due to a selection of the light and heavy component based on the combination of the parameters ρ_0, β and B.

have been completely calibrated and reconstructed. This reduced amount of data does not allow the measurement of the composition, that it will be addressed when more data will become available.

5.4 The all–particle spectrum in the energy range $200 - 800\,\mathrm{TeV}$

The analog readout system has been completely installed on the 130 central clusters of the ARGO–YBJ detector and it is taking data since December 2009. Data used in this analysis was taken during the period between July 2010 and August 2010, in which the analog system was operated at 2.5 V full–scale, corresponding to an energy range between $(10^2 \div 10^3)\,\mathrm{TeV}$. Although the data sample is not sufficiently large to allow the measurement of the composition, it can be used to obtain a preliminary measurement of the all–particle spectrum. By using these data the measurement of the cosmic ray spectrum can be extended from 200 TeV up to 800 TeV.

5.4.1 Data analysis

A first selection of the data based on the reconstruction quality an detector stability was applied. The selected data sample consists on about 10^6 events containing both analog and digital information. In order to perform an unbiased analysis both data and Monte Carlo events were selected by requiring the following criteria:

- reconstructed zenith angle ϑ_R in the range $(0^\circ \div 30^\circ)$
- number of particles N_{part} in the range $(2 \cdot 10^3 \div 2.5 \cdot 10^6)$
- reconstructed core position (X_{core}, Y_{core}) located in a radius of 30 m from the detector center
- the cluster with the highest particle multiplicity must be located in the innermost 20 clusters of the full–coverage carpet

The Monte Carlo events were sorted in three energy bins and eight multiplicity bins. The same multiplicity bins were used to analyze the sample of data. The Monte Carlo data sample was used to evaluate the conditioned probabilities $P_{ALL}(N_{part}|E)$ used in the bayesian analysis. The bayesian unfolding procedure was applied to the selected data sample by using a flat distribution for the initial values of $P(E_i)$. A soft smoothing was applied in order to ensure a stable convergence.

5.4.2 Evaluation of the uncertainties

Due to the limited size of the data sample the statistical uncertainty turns out to be $\pm 3\%$. Systematical uncertainties are due to effects related to the selection cuts used in this analysis and to the variation of the fractions of light and heavy component used to build the Monte Carlo data sample. This effect was studied by building different Monte Carlo data samples assuming different values of the composition. The uncertainty in the determination of the all–particle spectrum is related to the uncertainties in the determination of the conditioned probabilities used in the bayesian unfolding technique. Large variation of the fractions of the light and heavy component used in the Monte Carlo data sample lead to a variation of the conditioned probability of about 20%. This variation is reflected in the determination of the all–particle spectrum.

5.4.3 The all–particle spectrum

In figure 5.12 the results of the bayesian unfolding procedure applied to the data collected in the period July–August 2010 are reported. These data was collected with the analog readout system at 2.5 V full-scale, corresponding to an energy range between $(10^2 \div 10^3)$ TeV. The measured spectrum covers the energy region from 200 TeV up to 800 TeV. The yellow band represent the range of variation of the measurements of the spectrum due to the variations on the values of the conditioned probabilities.

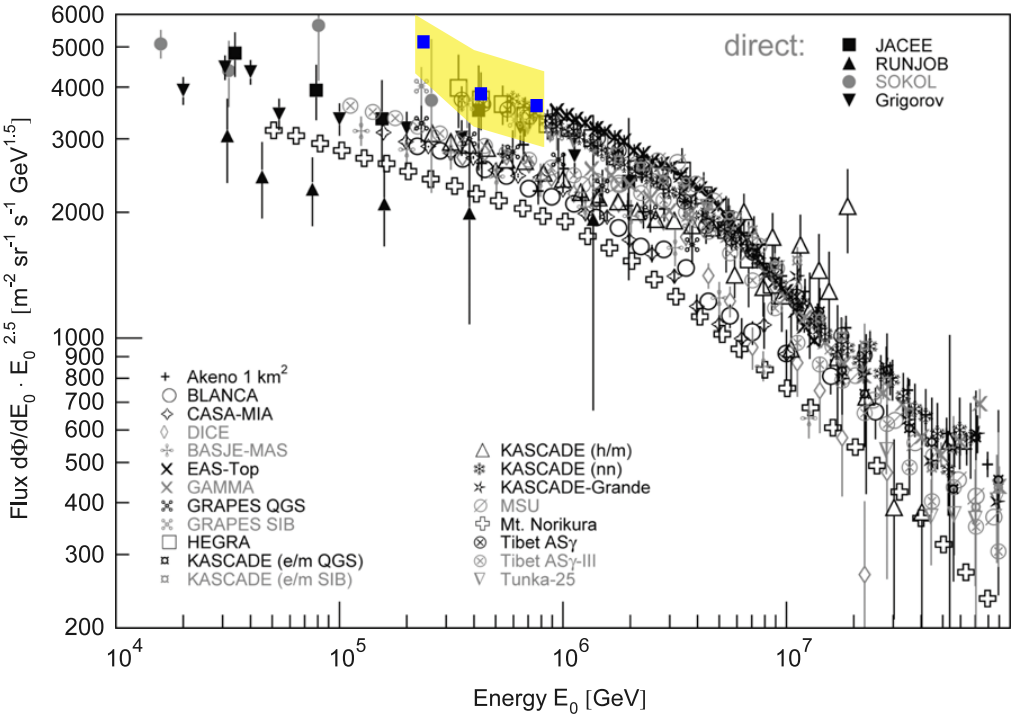


Figure 5.12: Preliminary measurement of the differential all–particle spectrum by the ARGO–YBJ experiment with the analog readout system (blue squares) compared with results from other experiments [6]. The yellow band represent the range of variation of the measurements of the spectrum due to the variations on the values of the conditioned probabilities.

5.5 Discussion

The energy spectrum of the light component of the primary cosmic rays from 10^{12} to 10^{15} eV plays an important role in understanding the mechanisms of acceleration, propagation and galactic confinement. Spectral differences between the proton and helium components might be related to different types of sources and acceleration environments [28, 32]. The present analysis does not allow the determination of the individual proton and helium contribution to the measured flux, however it shows a discrepancy with the spectra obtained at lower energies from direct measurements with the passive balloon-borne experiment JACEE and RUNJOB. The disagreement with the spectrum of the light component obtained adding the proton and helium spectra quoted by RUNJOB [19] is remarkable. This discrepancy takes place also at energies $E \sim 100$ TeV where the JACEE and RUNJOB experiments report a proton spectral index of about -2.8 [19, 21]. Deriving the primary energy spectrum from ground based EAS measurements introduces uncertainties related to the hadronic interaction model underlying the analysis. The strip multiplicity spectrum measured by ARGO-YBJ is mainly due to the electromagnetic component of the shower, the fraction of muons and hadrons being less than 10%. This component is sensitive to parameters governing the longitudinal development of the shower, like the proton–air inelastic cross section and the energy transferred to leading particle (elasticity). These processes influence the number of particles observed at ground level. A different longitudinal development of the shower could stretch or compress the energy scale in such a way affecting the spectral shape. In this context it is worthwhile to note that the proton-air inelastic cross section measured by ARGO-YBJ experiment in the energy range $(1 \div 100)$ TeV [72] has been found in good agreement with the values set in the CORSIKA/QGSJET code. According to the results shown in figure 6, the QGSJET and SYBILL models provide the same description of the longitudinal development of the shower in the energy range of interest. Moreover dedicated calculations find that the influence on the shower size of using different low-energy models (GHEISHA and FLUKA) is negligible [80]. These results give support to the present interpretation of the ARGO-YBJ experimental data.
The knee in the cosmic ray energy spectrum can be related to a change of the elemental composition. A measurement of the composition at the knee plays a fundamental role in the understanding of the mechanisms of produc-

tion and acceleration of particles at these energies. A detailed measurement of the effective lateral particle density distributions can be used in order to discriminate showers produced by different primaries.
A preliminary measurement of the all–particle spectrum in the energy range $(200 \div 800)$ TeV is presented. The all–particle spectrum presents large uncertainties due to the variations of the conditioned probabilities used in the bayesian unfolding technique that are connected to large variations on the fraction of the light and heavy component used to build the Monte Carlo data sample.

Conclusions

Despite a great experimental and theoretical effort made in order to understand the properties of the primary cosmic rays, several questions are still under discussion and investigation. Recent observations of the spectrum of the light component (protons and helium nuclei) performed by the CREAM experiment, a new–generation balloon–borne detector, present several discrepancies with data collected by past experiment. Moreover the origin of the knee is still an unresolved question in cosmic ray research. Several hypotheses have been presented, actually is believed that the origin of the knee is related to a change of the composition of the primary cosmic rays. A measurement of the composition in the knee region is crucial for the understanding of the origin of the knee.
The peculiar characteristics of the ARGO–YBJ experiment like high segmentation coupled to a digital readout, full coverage and high altitude location, allow the detection of small showers produced by primaries with energies of the order of 1 TeV. Moreover the analog readout can extend the detector operating range up to 10^4 TeV allowing the detection of showers in the knee region with very high detail.
The ARGO–YBJ observable is the shower multiplicity, namely the space–time distribution of the particles in the shower front. The energy and mass of the primary that initiated the shower cannot be evaluated on an event–by–event basis but must be determined by means of an unfolding procedure. The classical procedure is based on a fitting procedure that requires to formulate hypothesis about the shape of the energy spectrum. In this work an unfold-

ing procedure based on the Bayes theorem was used. The bayesian inference doesn't require any strong assumption about the shape of the spectrum and relies only on the knowledge of the conditioned probabilities that relate the observed shower multiplicity with the primary energy. These quantities can be evaluated by means of a Monte Carlo simulation of the development of the shower in the Earth's atmosphere and of the detector response. In this work a Monte Carlo data sample of EASs induced by protons, helium nuclei, CNO group and iron nuclei was produced and a full detector simulation was applied, including trigger and RPC's efficiency. Showers were generated in the energy range $(0.1 \div 10^4)$ TeV. A set of selection criteria based on the shower multiplicity, the reconstructed zenith angle and on the particle density on the detector surface was applied. These cuts allow the selection of showers mainly induced by protons and helium nuclei with core localized inside a fiducial area with radius ~ 28 m. The cut on the particle density allows the selection of showers with well–shaped core, discarding the events produced by heavier primaries. The fraction of CNO induced showers is reduced of about ten times compared with the fraction of protons and helium nuclei. A first selection of the data has been based on the reconstruction quality. A sample of 75×10^6 events was selected. The selection criteria described above have been adopted for both data and Monte Carlo events. The energy spectrum in the range $(5 \div 250)$ TeV was obtained by applying the bayesian unfolding procedure to the strip distribution obtained from the data sample. Statistical uncertainties are of the order of 1%. A study of the possible sources of systematical uncertainties has been performed. The sources of systematic effects are essentially effects related to the selection cuts used in data analysis, effects due to the reliability of simulation of the detector response and effects due to the variation of the fraction of the helium component. All of these effects have been taken into account and the overall uncertainty does not exceed 10%. The ARGO–YBJ data are fairly consistent with recent measurements of the light component spectrum performed by the CREAM collaboration, while they are in considerable disagreement with measurements reported by RUNJOB. The ARGO–YBJ measurements cover a wide energy range and bridges the energy gap between the low energy direct observations and the ground–based EAS experiments.

Showers produced by different primaries present several difference that are covered by fluctuations. A mass composition discrimination can be based on the lateral particle density distribution in EASs. In order to achieve a dis-

crimination of showers produced by primaries of different masses in the knee region a study of the lateral density distribution was performed on Monte Carlo events. A dedicated core reconstruction algorithm was developed. The study of the lateral particle density distribution allowed the determination of a set of discrimination parameters that can be used in order to discriminate showers produced by protons and light nuclei and showers produced by heavy nuclei. A preliminary measurement of the all–particle spectrum in the energy range $(200 \div 800)$ TeV has been performed by using the data taken with the analog readout system. Statistical uncertainties are of the order of 3%. The values of the spectrum are affected by large uncertainties due to the variations of the values of the conditioned probabilities used in the bayesian unfolding technique that are related to large variations of the fraction of the light and heavy component used to build the Monte Carlo data sample.

Bibliography

[1] G. V. Kulikov, G.B. Kristiansen *J. Exp. Theor. Phys.* **35** (1958) 635.

[2] K. Greisen *Phys. Rev. Lett.* **16** (1966) 748–750.

[3] G. T. Zatsepin and V. A. Kuzmin *JETP Lett.* **4** (1966) 78–80.

[4] R. U. Abbasi *et al. Phys. Rev. Lett.* **100** (2008) 101101.

[5] J. Abraham *et al. Phys. Rev. Lett.* **101** (2008) 061101.

[6] J. Blümer *et al. Progress in Particle and Nuclear Physics* **63** (2009) 293–338. and references therein.

[7] A. Obermeier *et al.* in *Proc.* 32^{nd} *ICRC*, no. 0675. Beijing (P.R. China), August, 2011.

[8] C. D. Santis in *Proc.* 32^{nd} *ICRC*, no. 1079. Beijing (P.R. China), August, 2011.

[9] W. Baade and F. Zwicky *Phys. Rev* **46** (1934) 76.

[10] V. L. Ginzburg and S. I. Syrovatskii, *The Origin of Cosmic Rays*. Pergamon Press. Classic Monograph, Jan. 1964.

[11] E. Fermi *Phys. Rev.* **75** (1949) 1169.

[12] A. R. Bell *Royal Astronomcal Society* **182** (1972) 147.

[13] R. Blandford and D. Eichler *Physics Reports* **154** (1987) 1.

[14] P. O. Lagage and C. J. Cesarsky *Astronomy and astrophysics* **125** (1983) 249.

[15] E. Berezhko *Astropart. Phys.* **5** (1996) 367.

[16] A. Bamba *et al. Astrophysical Journal* **589** (2003) 827.

[17] F. Aharonian *et al. Astronomy and astrophysics* **464** (2007) 253.

[18] J.R. Hörandel *Astropart. Phys.* **19** (2003) 193–220.

[19] A. V. Apanasenko et al. *Astropart. Phys.* **16** (2001) 13–46.

[20] V.A. Derbina et al. *The Astrophysical Journal* **628** (2005) L41–L44.

[21] K. Asakimori et al. *The Astrophysical Journal* **502** (1998) 278–283.

[22] H. S. Ahn et al. *The Astrophysical Journal Letter* **714** (2010) L89–L93.

[23] Y. S. Yoon *et al. The Astrophysical Journal* **728** no. 2, (2011) 122. `http://stacks.iop.org/0004-637X/728/i=2/a=122`.

[24] M. Boezio *et al. Astropart. Phys.* **19** (2003) 193.

[25] A. D. Panov *et al. Bull. Russ. Acad. Sci. Phys.* **73** (2009) 564.

[26] M. Aguilar *Phys. Rep.* **366** (2002) 331.

[27] V. I. Zatsepin and N. V. Sokolskaya *Astronomy and astrophysics* **458** (2006) 1–5.

[28] P. L. Biermann *Astronomy and astrophysics* **271** (1993) 649.

[29] T. Stanev *Astronomy and astrophysics* **274** (1993) 902.

[30] A. A. Abdo *et al. Phys. Rev. Lett.* **101** (2008) 221101.

[31] G. D. Sciascio *et al.* in *Proc.* 32^{nd} *ICRC*, no. 0507. Beijing (P.R. China), 2011.

[32] Y. Butt and A. M. Bykov *Ap. J.* **677** (2008) L21.

[33] W. R. Binns *et al. Space Sci. Rev.* **130** (2007) 439.

[34] Y. Ohira and K. Ioka *Ap. J. Letters* **729** (2011) L13.

[35] A. M. Hillas *J. Phys. G: Nucl. Part. Phys* **31** (2005) R95.

[36] D. C. Ellison *et al. Ap. J.* **540** (2000) 292.

[37] J. R. Hörandel *Astropart. Phys.* **21** (2004) 241–265.

[38] A. D. Erlykin and A. W. Wolfendale *J. Phys. G: Nucl. Part. Phys* **27** (2001) 1005.

[39] C. Bacci et al. *Nucl. Phys. Proc. Suppl.* **78** (1999) 38–43.

[40] G. Aielli et al. *Nucl. Instrum. Meth.* **A562** (2006) 92–96.

[41] C. Bacci et al. *Astropart. Phys.* **17** (2002) 151–165.

[42] S. Mastroianni et al., "The ARGO–YBJ inclusive trigger," in *Proc.* 29^{th} *ICRC*. 2005.

[43] G. D. Sciascio *et al.* in *Proc.* $32^{nd} ICRC$, no. 0225. Beijing (P. R. China), August, 2011.

[44] G. Aielli *et al. Astropart. Phys.* **30** (2008) 85–95.

[45] T. D. Girolamo *et al.* in *Proc.* 32^{nd} *ICRC*, no. 0574. Beijing (P. R. China), August, 2011.

[46] `http://www.fisica.unisalento.it/~demitri/idas/`.

[47] G. Aielli et al. *Nucl. Instrum. Meth.* **A608** (2009) 246–250.

[48] C. Bleve, S. Bussino, S. Catalanotti, I. D. Mitri, S. Mari, G. Marsella, and D. Martello, *How to with medea++*. ArgoYBJ-Note 002/03.

[49] J. Nishimura *Handbuch der Physik* **46** (1965) 1.

[50] K. Kamata and J. Nishimura *Progr. Theoret. Phys* **6** (1958) 93.

[51] K. Greisen, "Cosmic ray showers," *Ann. Rev. Nucl. Part. Sci.* **10** (1960) 63–108.

[52] L. Saggese, T. Di Girolamo, M. Iacovacci, and S. Mastroianni *Nucl.Instrum.Meth.* **A533** (2004) 55–59.

[53] M. Iacovacci and S. Mastroianni, "Stability and calibration of the analog RPC readout in ARGO–YBJ," in *Proc.* 32^{nd} *ICRC*. 2011.

[54] S. Mastroianni *et al. IEEE Transactions on Nuclear Science* **58** no. 4, (August, 2011) .

[55] H. H. He *et al. Astropart. Phys.* **27** (2008) 528.

[56] G. Aielli et al. *Astropart. Phys.* **30** (2009) 287–292.

[57] G. Aielli *et al. Nucl. Instrum. Meth. A* (2010) .

[58] B. Bartoli *et al. Phys. Rev. D* **84** (2011) 022003.

[59] G. D'Agostini *Nucl. Instrum. Meth.* **A362** (1995) 487–498.

[60] G. D'Agostini, *Bayesian reasoning in data analysis*. World Scientific, 2003.

[61] G. D'Agostini, Oct, 2010. `http://arxiv.org/abs/1010.0632`.

[62] V. Blobel report, 1984 CERN School of Computing, Aiuguablava, Catalonia, Spain, 1984.

[63] S.Bussino, E. De Marinis and S.M.Mari *Astropart. Phys.* **22** (2004) 81–94.

[64] D. Heck, J. Knapp, J. Capdevielle, G. Schatz, and T. Thouw Report **FZKA** 6019, Forschungszentrum Karlsruhe, 1998. `http://www-ik.fzk.de/corsika/physics_description/corsika_phys.html`.

[65] W. R. Nelson et al. SLAC Report **265**, 1985.

[66] A. F. Belayev et al. SLAC Report **6499**, 1994.

[67] S. Ostapchenko *Phys. Rev. D* **74** (2006) 014026.

[68] S. Ostapchenko *Nucl. Phys. B (Proc. Suppl.)* **151** (2006) 143 and 147.

[69] A. Fassò, A. Ferrari, J. Ranft, and P.R. Sala, "Fluka: a multi-particle transport code," tech. rep., CERN-2005-10 (2005), INFN/TC-05/11, SLAC-R-773.

[70] Cern Application Software Group, "GEANT - Detector description and simulation tool," tech. rep., CERN Program Library, long writeup W5013, 1993.

[71] A. Aloisio et al. *IEEE Transaction on Nuclear Science* **51** (2004) 1853.

[72] G. Aielli et al. *Physical Review D* **80** (2009) 092004.

[73] R. Engel, T. K. Gaisser, P. Lipari, and T. Stanev in *Proc.* 26^{th} *ICRC*. Salt Lake City (USA), 2006.

[74] E. J. Ahn, R. Engel, T. K. Gaisser, and T. Stanev *Phys. Rev. D* **80** (2009) 094003.

[75] M. Aglietta et al. *Astropart. Phys.* **21** (2004) 223–240.

[76] O. Adriani *et al. Science* (2011) science.1199172.

[77] J. Alcaraz *et al. Phys. Lett. B* **490** (2000) 27.

[78] S. Haino *et al. Phys. Lett* **594** (2004) 35.

[79] A. Cirillo, S. M. Mari, P. Montini in *Proc.* 32^{nd} *ICRC*, no. 0691. Beijing (P. R. China), August, 2011.

[80] H. Ulrich *et al. Nucl. Phys. B* **273** (2008) 175.