BERGISCHE
UNIVERSITÄT
WUPPERTAL

Fakultät für
Mathematik und Naturwissenschaften
Astroteilchenphysik

Muon density in extensive air showers measured with IceTop

Dissertation

zur Erlangung des Doktorgrades
doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt von

Daniel Bindig

am 27. Juli 2018

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20190128-120513-9
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20190128-120513-9]

Gutachter:
Prof. Dr. Klaus Helbing, Bergische Universität Wuppertal
Dr. Andreas Haungs, Karlsruher Institut für Technologie

Abstract

Cosmic Rays are particles that are accelerated from astrophysical sources and being propagated through the universe. When Cosmic Rays strike particles of the Earth's atmosphere, they undergo interactions which trigger a cascade of particles travelling to the Earth's surface. This air shower carries constituents of different types, such as hadrons, electrons, and muons. Measurements of these constituents can be compared with air shower simulations, in order to draw conclusions on the properties of the initial Cosmic Ray particle, such as nuclear mass, energy, and direction. In particular, the number of muons shows a dependance on the nuclear mass.

In this work, the detector signatures provided by the IceTop detector are analyzed. IceTop is the surface part of the IceCube detector situated at the South Pole. It consists of 81 stations ordered in a grid-like structure. Each station comprises two tanks filled with clear ice. Charged particles passing through ice produce Cherenkov light which is detected by photomultiplier tubes.

In this analysis a method is provided which allows the determination of muon number densities based on measured detector signatures. Initially, this method is developed on air showers simulated with the hadronic interaction model SIBYLL2.1, and having zenith angles between 0° and 36.9°. These air showers are subject to detector simulation and reconstruction, and undergo a selection of quality cuts. The selected set of simulated air showers is separated into bins in zenith angle and estimated energy. For each remaining air shower, the detector signatures are subclassified into signal and background, depending whether they contain muons or not.

The signal content of air showers is carved out by introducing energy and zenith angle dependent cuts on the tank charge and its distance to the shower axis. After the application of these cuts, two muon number estimators are defined basing on lateral charge and hit distributions. Systematic studies show that the latter is more robust against changes in simulation.

Using a conversion, the muon number estimator is transformed into a muon number density estimator, which is related to the true muon density known from air shower simulations. Finally, the muon number density in experimental data, recorded during the 2012/2013 season, is calculated by applying the conversion derived from simulations, on muon number density estimators derived from data. The result is the muon number density as function of reconstructed primary energy. Since the conversion is derived from air shower simulations, systematic uncertainties such as the hadronic interaction model or the primary mass propagate onto the final result.

Zusammenfassung

Die Astroteilchenphysik als relativ junger Forschungsbereich befasst sich mit den Eigenschaften der kosmischen Strahlung, die durch das Universum propagiert und zu einem geringen Bruchteil die Erde erreicht. Kosmische Strahlung setzt sich überwiegend aus Atomkernen verschiedenartiger chemischer Elemente, sowie weiterer Teilchen wie Elektronen, Positronen, Neutrinos und Photonen unterschiedlicher Wellenlängen zusammen. Die Entdeckung der kosmischen Strahlung durch Viktor Franz Hess im Jahre 1912 hat zahlreiche Fragen nach deren Quellen, Propagation, Energiespektrum und Zusammensetzung hervorgebracht, die auch heute noch Gegenstand aktueller Forschung sind.

Seit der Entdeckung kosmischer Strahlung wurden verschiedene Detektortypen zu ihrer Messung entwickelt. Beispielsweise messen auf Satelliten stationierte Detektoren kosmische Strahlung direkt, bevor Wechselwirkungen mit der Erdatmosphäre erfolgen. Da der differentielle Fluss kosmischer Strahlung einem Potenzgesetz folgend mit ungefähr $\propto E^{-2.7}$ abnimmt, und satelittengestützte Experimente nur sehr kleine effektive Flächen haben, kann der Nachweis kosmischer Strahlung höherer Energien nur durch erdgebundene Detektoren stattfinden. Dies geschieht indirekt über den Nachweis von Luftschauern. Dabei handelt es sich um Teilchenkaskaden, die entstehen, sobald Teilchen kosmischer Strahlung auf die Erdatmospäre treffen und eine Lawine von physikalischen Prozessen initiieren. Die in Luftschauern produzierten Teilchen sind elektromagnetischer, myonischer, und hadronischer Natur. Insbesondere die Anzahl der Myonen im Luftschauer weist eine starke Abhängigkeit von der Masse des Primärteilchens auf.

Basierend auf gemessenen Detektorsignaturen und Luftschauersimulationen lassen sich Rückschlüsse auf die Eigenschaften des ursprünglichen Teilchens der kosmischen Strahlung ziehen, wie beispielsweise Energie und Masse. Die Kenntnis des Energie- und Massenspektrums wiederum erlaubt das Bestätigen oder Ausschließen verschiedener Modelle, die die Produktion kosmischer Strahlung in Quellen und die Propagation durch das Universum beschreiben, und damit Vorhersagen für das Energie- bzw. Massenspektrum treffen.

Die in dieser Arbeit genutzten Daten stammen von IceTop, die Oberflächenkomponente des am Südpol stationierten IceCube-Detektors. IceTop besteht aus 81 Stationen mit jeweils zwei Tanks. Jeder Tank ist mit klarem Eis gefüllt, sowie mit zwei Photomultipliern ausgestattet, die Cherenkovlicht messen, das von geladenen Teilchen beim Durchqueren des Tanks erzeugt wird.

In dieser Arbeit wird die Myondichte auf IceTop-Niveau als Funktion der Energie des Primärteilchens bestimmt. Die dazu genutzte Methode wird zunächst mit Hilfe von Luftschauersimulationen basierend auf dem hadronischen Wechselwirkungsmodell SIBYLL2.1 entwickelt. Diese Simulationen werden genutzt, um in der IceTop-Detektorsimulation Signale zu generieren. Die Rekonstruktion des Luftschauers anhand dieser Signale liefert essentielle Charakteristika, wie zum Beispiel die ursprüngliche Richtung des einfallenden Teilchens, der Auftreffpunkt der Schauerachse im Detektor, und eine Schätzung der Energie des einfallenden Teilchens. Nach der Rekonstruktion wird die Anzahl der simulierten Schauer durch ausgewählte Qualitätsschnitte reduziert. Die verbleibenden Luftschauer werden gruppiert nach Bins in Zenithwinkel und Energieschätzer.

Um Myonen im Luftschauer zu isolieren, werden die Detektorsignaturen in IceTop mit Hilfe der Detektosimulation nach Signal und Untergrund aufgetrennt. Ist mindestens ein Myon für das Triggern eines Tanks verantwortlich, wird dieser Tank als Signal klassifiziert, andererseits als Untergrund. Die Signalkomponente wird weiter herausgearbeitet, indem Schnitte auf die gemessene Ladung in einem Tank, und dessen Abstand zur Schauerachse entwickelt werden. Es wird gezeigt, dass sich durch diese Schnitte der Anteil vom Signal an allen aufgezeichneten Ereignissen erhöht.

Basierend auf den Schnitten werden zwei Schätzer für die wahre Anzahl der Myonen im Luftschauer angegeben. Diese beziehen sich auf die laterale Verteilung der Tankladung und die laterale Verteilung der Tankanzahl, d.h. Tanks die eine Ladung registrieren. Gemessen an den systematischen Unsicherheiten ist die laterale Verteilung der Tankanzahl robuster als die Tankladung; deshalb wird nur diese im weiteren Verlauf verwendet.

Das angewandte Verfahren beinhaltet zunächst die Umwandlung des Schätzer für die Myonzahl in einen Schätzer für die Myondichte. Anschließend werden Umrechnungsfaktoren berechnet, die es ermöglichen, den Schätzer für die Myondichte auf die wahre Myondichte zu skalieren, die aus den Luftschauersimulationen bekannt ist. Analog werden aus den gemessenen Daten der Saison 2012 / 2013 basierend auf dem mit der Simulation entwickelten Verfahren ebenfalls Schätzer für die Myondichte bestimmt, die schließlich mit Hilfe der Kalibrationsfaktoren auf die Myondichte in den Daten skaliert werden. Das finale Ergebnis ist die Myondichte als Funktion der rekonstruierten Energie, d.h. der Energie des Primärteilchens.
Da das angewandte Verfahren von den Luftschauersimulationen abhängt, ergeben sich beispielsweise systematische Unsicherheiten aufgrund der hadronischen Wechselwirkungsmodelle, der Masse des simulierten Primärteilchens, oder von Schnee auf dem Detektor.

Contents

Abstract I

Zusammenfassung II

List of Figures IX

List of Tables XII

1 Introduction 1

2 Cosmic Rays 5
- 2.1 Sources and acceleration . . . 5
- 2.2 Extensive air showers . . . 8
- 2.3 Ground based measurements . . . 11
  - 2.3.1 Energy spectrum . . . 11
  - 2.3.2 Elemental composition . . . 13

3 The IceTop detector 15
- 3.1 Particles passing through matter . . . 15
- 3.2 InIce and IceTop . . . 15
- 3.3 Experimental data . . . 20
- 3.4 Local coincidences . . . 21
- 3.5 IceTop Trigger and Filter . . . 23
- 3.6 Charge calibration . . . 23
- 3.7 Snow on IceTop tanks . . . 25
- 3.8 Atmospheric conditions . . . 28

4 Simulation 29
- 4.1 Air shower simulation . . . 29
- 4.2 Cosmic Ray flux model . . . 32
- 4.3 Detector simulation . . . 33
  - 4.3.1 Shower core placement . . . 34
  - 4.3.2 Particle tracks . . . 36
  - 4.3.3 Cherenkov photons . . . 36
  - 4.3.4 Noise . . . 37

5 Reconstruction and event selection 39
5.1 Preparations of the used signals 39
5.2 Reconstruction procedure 40
  5.2.1 First guess reconstructions 41
  5.2.2 Functions used for reconstruction 42
  5.2.3 Shower size correction due to snow 44
  5.2.4 Overview of the reconstruction 44
5.3 Event selection 45
5.4 Comparison between simulation and data 50
  5.4.1 Passing rates 51
  5.4.2 Observables 51
5.5 Enhancement of events 51
  5.5.1 Charge recalibration 51
  5.5.2 Time residuals 53

6 Muon number estimation 57
6.1 Motivation 57
6.2 Separation of signal 60
  6.2.1 Properties of signal and background 61
  6.2.2 Cut on the hit charge 64
  6.2.3 Cut on the hit distance 64
  6.2.4 Discussion 66
  6.2.5 Consistency check 68
6.3 Muon number estimators 69
  6.3.1 Discussion 69
  6.3.2 The true muon number $\mu^{\mathrm{true}}(R_{ij})$ 70
  6.3.3 Consistency check 71
6.4 Summary 73

7 Calculation of muon number densities 75
7.1 The true muon number density $\rho_{\mathrm{true}}(R)$ 75
7.2 Discussion 76
7.3 Energy conversion 77
7.4 Development on simulation 77
7.5 Application on experimental data 78
7.6 Systematic uncertainties 79
7.7 Consistency check 81
7.8 Results 81
  7.8.1 Muon density as function of distance R 82
  7.8.2 Muon density as function of E_{rec} 82
7.9 Summary 83

8 Discussion and outlook 87

Appendix 89

Bibliography **95**

Glossary **XIII**

List of Figures

1.1 Hess' measurements during balloon flights 2

2.1 Hillas plot 7
2.2 Schematic development of an extensive air shower 9
2.3 The differential atmospheric muon flux separated into contributions from different particle types. 10
2.4 The differential energy spectrum of Cosmic Rays 12
2.5 Mean logarithmic mass as a function of the primary energy for various experiments 13

3.1 Schematic of Cherenkov light production 16
3.2 Schematic view of the IceCube detector 17
3.3 Digital Optical Modules (DOMs) 17
3.4 Different views of IceTop 18
3.5 Layout of an IceTop tank 18
3.6 A waveform recorded by a particular DOM in IceTop 20
3.7 Cabling between the DOMs within an IceTop station 22
3.8 Exemplary distributions of HLC and SLC hits 22
3.9 Charge distribution measured by a particular HG DOM, serving as example for the procedure of VEM calibration 24
3.10 S_μ/B_{em} as a function of the snow heights above IceTop tanks directly measured during three different months. 26
3.11 Snow height distribution on IceTop during July 2012. 27
3.12 Temporal atmospheric conditions at the South Pole 28

4.1 Number of muons per primary energy as a function of the primary energy for IceTop and Auger. 31
4.2 Radii in which shower cores are distributed on the IceTop array . . . 34
4.3 The observation level of the SIBYLL2.1 simulation sets compared to tank specific heights 35
4.4 Acceptance of an IceTop DOM as a function of the incident wavelength 36

5.1 Perpendicular distance to the shower axis $R_{T,p}$ 41
5.2 Visualization of air shower reconstruction in IceTop 45
5.3 Relation between true primary energy and the shower size S_{125} for vertical showers ($\cos\theta \geq 0.95$) initiated by proton primaries. 47

5.4 Quality of the reconstructed shower core and direction as function of the primary energy, shown for proton and iron initiated air showers, and for two zenith angle bins. 48
5.5 Visualization of the containment cut 49
5.6 Reduction of shower core migration. 50
5.7 Comparison of reconstructed observables between simulation and experimental data 52
5.8 Time Residuals in experimental data, shown for HLC and SLC hits. . 54
5.9 Comparison of time residuals for all hits within experimental data and simulation 55

6.1 Average light yield of electrons and muons vertically traversing an IceTop tank as function of their kinetic energy 58
6.2 Average kinetic particle energy for the muonic and electromagnetic shower components at observation level, as function of the perpendicular distance to the shower axis, shown for both primaries, and two bins in zenith angle and shower size S_{125}. 59
6.3 Visualization of electromagnetic and muonic distributions in a 2-dim contour plot, for vertical, iron induced air showers with $1.0 \leq S_{125} \leq 1.1$ 59
6.4 Particle combinations which cause hits in IceTop, shown for vertical, iron induced air showers with $1.0 \leq \log_{10}(S_{125}) \leq 1.1$. 60
6.5 Sketch of two exemplary rings in which the average charge is calculated 61
6.6 Total weighted hit charge, total weighted number of hit tanks, and the average charge as a function of the perpendicular distance to the shower axis for both signal and background, for both HLC and SLC hits, detected in vertical air showers initiated by iron primaries, shown for two bins in shower size S_{125}. 63
6.7 Determination of Q_{min} for vertical iron showers, and $1.0 \leq \log_{10}(S_{125}) \leq 1.1$. 65
6.8 Example of $R_{\text{min}}\,(S_{125})$, and an appropriate linear fit for vertical air showers induced by iron primaries. 65
6.9 Percentaged contribution ($1 \equiv 100\%$) of signal to all hits, shown for vertical air showers induced by proton and iron primaries, and two bins in shower size S_{125}, before and after applying cuts on hit charge and distance. 67
6.10 Charge distributions for four different thresholds for the charge expectation S_{exp} 68
6.11 Muon number estimators $\mu^{\text{est}}_{\text{ch, true}}\,(R)$, and $\mu^{\text{est}}_{\text{h, true}}\,(R)$, compared to the simulated muon number $\mu^{\text{true}}\,(R)$, shown for vertical proton an iron showers, and two bins in shower size S_{125}. 70
6.12 Comparisons between the muon number estimators $\mu^{\text{est}}_{\text{ch, h, true}}\,(R)$, and $\mu^{\text{est}}_{\text{ch, h, Data}}\,(R)$, divided by all tanks, for two bins in shower size S_{125} . . 72

7.1 Visualization of an ice block and an incoming muon. 76

7.2 The muon number density $\rho_{\text{true}}(R)$ calculated for vertical proton and iron showers, and three bins in true energy E_{true}. 76
7.3 Uncertainties on a conversion distribution $C_{\text{h}}(R)$, for vertical air showers induced by iron primaries. 80
7.4 Total uncertainties on the conversion distributions $C_{\text{ch, h}}(R)$ 80
7.5 Simulation-like and data-like consistency check 81
7.6 Muon number density $\rho_{\text{data}}(R)$ for air showers with zenith angles θ within $0.95 < \cos\theta \leq 1.0$. The derivation is based on the SIBYLL2.1 interaction model. 82
7.7 Muon number densities at 400 m, and 600 m perpendicular distance to the shower axis, for vertical air showers as a function of E_{rec}. The derivation is based on the SIBYLL2.1 interaction model. 83
7.8 Muon number densities for vertical air showers as a function of E_{rec}, and for three interaction models. 84

1 Best cuts on hit distance to the shower axis 89
2 Best cuts on hit charge 90
3 Muon number density $\rho_{\text{data}}(R)$ for air showers with zenith angles $0.80 < \cos\theta \leq 0.85$. The derivation is based on the SIBYLL2.1 interaction model. 91
4 Muon number density $\rho_{\mu}(E_{\text{rec}})$ for perpendicular distances of 400 m and 600 m to the shower axis. The derivation is based on the SIBYLL2.1 interaction model. 91
5 Muon number densities for inclined air showers as a function of E_{rec}, and for three interaction models. 92
6 Results produced with and without the application of cuts. 93

List of Tables

4.1 Properties of the air shower simulations 30

5.1 Relative and cumulative passing rates of individual cuts calculated for simulation and experimental data 51

7.1 Parameters of the functions converting S_{125} to E_{rec} 77

7.2 Uncertainties on the data points labeled with “p-cuts” in Fig. 7.8 in the case of SIBYLL2.1. 85
7.3 Uncertainties on the data points labeled with “Fe-cuts” in Fig. 7.8 in the case of SIBYLL2.1. 86

Chapter 1

Introduction

Looking into the cloudless night sky gives us the feeling of being part of something greater. Our eyes are hit by light from a variety of bright astronomical objects such as stars, galaxies, or planets reflecting light from the Sun. If one resides in dark and rather agrarian surroundings, the stars of the Milky Way can be clearly seen. Our home galaxy comprises millions of stars and planets, including the Earth, while it is just one of billions of other galaxies.

Humanity has always been fascinated by space and its intrinsic objects. Early in our history a deep desire for an understanding of the nature of the universe inspired a lot of people, some of whom are famous scientists today, to start detailed investigations. Modern physics would be impossible without their achievements.

Current physical research topics would also be different if Victor Franz Hess would not have performed the balloon flights he is famous for. The initial purpose of the balloon flights, he performed in 1912 [1], was the investigation of the radiation causing the discharge of an electroscope in air. This radiation was assumed to originate from radioactive materials in the Earth's crust. Hess measured the ionization rate as a function of the altitude and discovered that it decreases up to a height of about 1 km, and increases afterwards (see Fig. 1.1).

Hess explained his observations by the radiation coming from space interacting with the Earth's atmosphere [1]. With this conclusion, Hess was the founder of Astroparticle Physics [2]. This radiation was later called the *Cosmic Rays*. For its discovery, Hess was awarded the Nobel Prize in 1936.

The properties of Cosmic Rays were further investigated by other scientists in subsequent years. Kolhörster [4], Auger [5], and their colleagues detected coincident signals in ground based counters situated hundreds of meters apart from each other. They concluded that Cosmic Rays entering the Earth's atmosphere initiate a cascade of particles. These secondary particles spread out within the atmosphere while travelling to the ground. Thus, the Earth's surface is showered by a huge amount of particles initially caused by one Cosmic Ray particle. Based on this, Kolhörster and Auger defined the expression *Extensive Air Shower* (see Sec. 2.2 for details).

The knowledge of Cosmic Rays has been improved up today. Nevertheless, there are still open questions such as:

- What are the astrophysical sources of Cosmic Rays and where are they located?
- What are the mechanisms which produce Cosmic Rays in astrophysical sources and how are Cosmic Rays accelerated up to the highest energies which are measured?
- How can the spectral features of the energy spectrum of Cosmic Rays be explained?
- What are the principles of high energy hadronic interactions in the atmosphere, caused by Cosmic Rays?
- What is the average mass of Cosmic Ray nuclei at particular energies, i.e. what is the elemental composition of Cosmic Rays?

A few of these questions are reviewed in Ch. 2. In particular, this work is motivated by the last two questions. High energy interactions of hadrons in the atmosphere are modeled by various *hadronic interaction models*. These models are used within air shower simulations and differ in their outcomes (see Ch. 4). Thus, hadronic interaction models contribute large systematic uncertainties to analyses which use air shower simulations.

Measurements of the elemental composition of Cosmic Rays are linked to a variety of physical problems such as the questions above and may help to answer them. For example, the spectral features of the energy spectrum of Cosmic Rays are not yet completely understood (explained and illustrated in Sec. 2.1). Various theories attempt to clarify the origin of these features by modeling the propagation of Cosmic Rays through the universe after they have been emitted and accelerated from particular sources. Information provided by measurements of the elemental composition of Cosmic Rays can help to constrain these theories [6, 7].

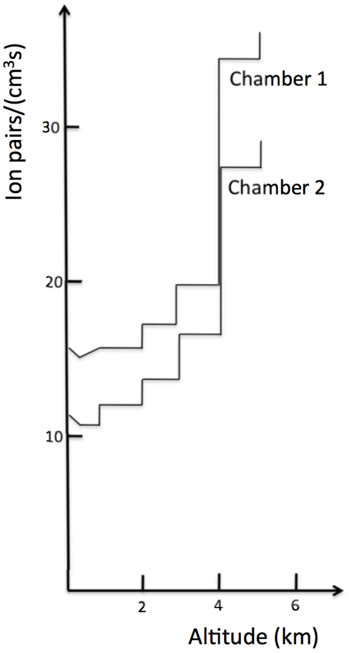

Figure 1.1: Hess' measurements during balloon flights [3]. They were performed using two Wulf electroscopes (*Chamber 1* and *Chamber 2*) [1, 3]. Hess discovered that the ionisation rate as a function of altitude increases for heights larger than $\sim 1\,\mathrm{km}$. He concluded that radiation coming from space is hitting the Earth's atmosphere.

The elemental composition of Cosmic Rays has been measured by several experiments over different energy ranges (see Sec. 2.3). However, further independent measurements are required in order to understand discrepancies between the results.

The goal of this work is a contribution to solve the problems described above. The experimental data used in this analysis is obtained by the IceTop air shower array which is located at the geographical South Pole. It is the surface part of the volume detector IceCube and comprises tanks situated in a grid-like structure, similar to the experimental setup developed by Kolhörster and Auger. The tanks are filled with ice and instrumented to detect Cherenkov light produced by particles passing through the tanks (see Ch. 3).

Together with air shower simulations, the experimental data is used to calculate a density of muons which are produced in extensive air showers (see Sec. 2.2). The density of muons calculated from experimental data can be used to address various problems:

- It can be compared to the density of muons predicted by hadronic interaction models. This procedure illustrates which hadronic interaction models describe experimental data badly.
- It can help to understand the *muon problem*, namely that the muon content in simulated air showers is lower than in experimental data, especially at high energies. For example, this has been reported by the Auger collaboration [15, 16] in 2015 [8].
- It can be used to study the composition of Cosmic Rays. As shown in Sec. 2.2, the number of muons in an air shower is related to the mass of the nucleus initiating the air shower. In turn, the knowledge of the Cosmic Ray mass composition can help to exclude or confirm astrophysical models explaining the production and propagation of Cosmic Rays.

This work is organized as follows:

An overview of potential acceleration mechanisms of Cosmic Rays is provided in **Ch. 2**. Furthermore, the development of extensive air showers is sketched and previous measurements of the energy spectrum and elemental composition of Cosmic Rays with ground based detectors are shown. The performance of IceTop is explained in **Ch. 3**. The creation of air shower and detector simulations are a crucial requirement to perform this analysis. The simulation sets and the process of the detector simulation are shown in **Ch. 4**. The reconstruction of air shower parameters and the event selection ensuring that high quality air showers are used for analysis are the topic of **Ch. 5**. An observable which is required for background reduction is defined in **Ch. 6**. Furthermore, it is shown that after introducing two simple cuts on observables related to IceTop tanks, a reliable and robust muon number estimator can be calculated. A method to derive muon number densities is presented in **Ch. 7**. Finally, the results and their significance for future analyses are discussed in **Ch. 8**.

Chapter 2

Cosmic Rays

The physical properties of Cosmic Rays have been investigated for roughly 100 years. In this period, our understanding of Cosmic Rays has been evolved: Cosmic Rays mostly comprise nuclei of different mass. Additionally, Cosmic Rays are enriched by particles such as electrons, positrons, antiprotons, photons, and neutrinos [29, 20]. Depending on the energy of Cosmic Rays, different types of sources can be assumed (see Sec. 2.1). While propagating through the universe, the charged component of Cosmic Rays is deflected by magnetic fields of variable scale. Thus, only neutral particles such as photons and neutrinos can point back to their source [7]. In this work, the expression "Cosmic Rays" refers to nuclei of different mass.

Cosmic Rays reaching the upper atmosphere of the Earth initiate the development of extensive air showers (see Sec. 2.2). The shower of particles propagates through the atmosphere while several physical processes change the number of particles in the shower. At the ground, the remaining particles can be detected. Physical experiments use the detector signatures to draw conclusions on energy and mass of the initial particle (see Sec. 2.3).

2.1 Sources and acceleration

One of the main objectives for constructing ground based air shower detectors is the discovery and understanding of Cosmic Ray sources. The origin of Cosmic Ray energies, spanning from a few hundred MeV to a few hundred EeV [20], can only be explained based on theoretical models.

One of the most famous models for the acceleration of Cosmic Rays is the *Fermi acceleration*. One differentiates between the Fermi acceleration of the *second* and the *first* order [21, 22, 23]:

- The second order Fermi acceleration describes charged particles accelerated in the *interstellar clouds* acting as magnetic mirrors. These are regions in the universe where gas, dust, and plasma are accumulated. Assuming the interstellar cloud moves with $v \ll c$, it can be shown that after scattering at an interstellar cloud the particle has an average relative energy gain given by

$$\left\langle \frac{\Delta E}{E} \right\rangle \sim \left(\frac{v}{c}\right)^2 \tag{2.1}$$

Furthermore the resulting differential energy spectrum can be calculated:

$$F(E)\,dE \sim E^{-1+\alpha}\,dE \tag{2.2}$$

where $F(E)$ is the rate of particles with energy E. The index α depends on the average energy gain after one collision and the probability that the particle remains in the area where it was accelerated.

- In the first order Fermi acceleration the Cosmic Ray particles are supposed to gain energy due to their collision with shock waves. In this case, the average relative energy gain is given by

  $$\left\langle \frac{\Delta E}{E} \right\rangle \sim \left(\frac{v}{c}\right) \tag{2.3}$$

  The energy spectrum is

  $$F(E)\,dE \sim E^{-2}\,dE \tag{2.4}$$

Since the second order Fermi acceleration is proportional to the square of v/c, the first order Fermi acceleration is more effective in accelerating Cosmic Rays to high energies.

Cosmic Rays with energies larger than $\sim 10^{18}\,\mathrm{eV}$ are called the *Ultra High Energy Cosmic Rays* (UHECRs) [25]. An overview of sources and in which fields the particles are accelerated to these energies was developed by Hillas which studied the relation between the magnetic field strength and size of the acceleration regions, and showed that the product of both represents an upper limit on particle acceleration [24],

$$B\,[\mu\mathrm{G}] \cdot L\,[\mathrm{pc}] > 2 \cdot E\,[\mathrm{PeV}]/Z \cdot \beta_\mathrm{s} \tag{2.5}$$

where B is the magnetic field component of the region perpendicular to the velocity of the particle, L is the size of the acceleration region, E is the energy of the particle, Z is the charge of the particle, and β_s is the velocity of the shock. Equation 2.5 is also known as the *Hillas criterion*. It can be written as [30],

$$E_\mathrm{max} \simeq 10^{18}\,\mathrm{eV} \cdot Z \cdot \beta_\mathrm{s} \cdot B\,[\mu\mathrm{G}] \cdot L\,[\mathrm{kpc}] \tag{2.6}$$

where E_max is the maximum energy a particle can gain in the corresponding acceleration region. In Fig. 2.1, the magnetic field strength B of a potential source candidate is shown as a function of its size L. According to Eq. 2.6, the magnetic field B as a function of size L can be described by a diagonal line as shown for protons with $E_\mathrm{max} = 10^{20}\,\mathrm{eV}$. The line represents an upper limit on potential sources for combinations of B and L, i.e. sources on this line accelerate Cosmic Rays up to $E_\mathrm{max} = 10^{20}\,\mathrm{eV}$. As can be seen in Fig. 2.1, the choice of appropriate source candidates in order to describe UHECRs is difficult.

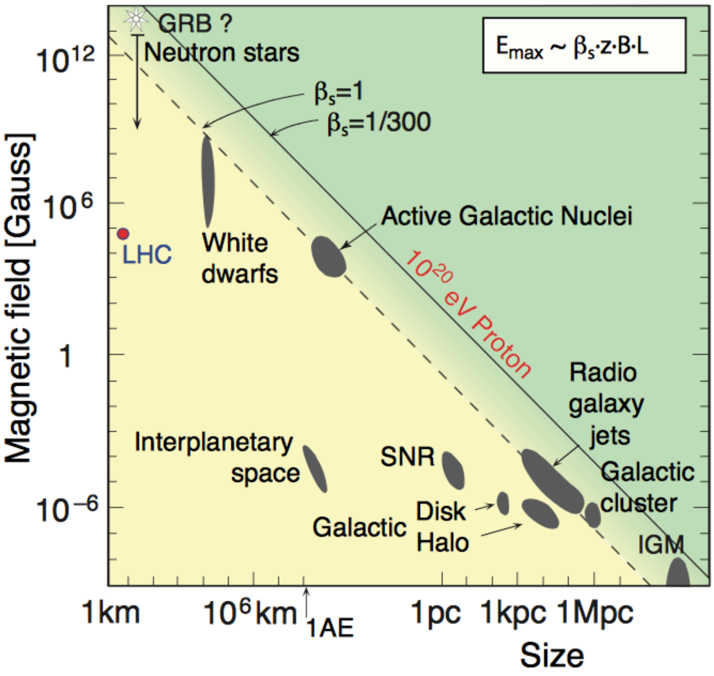

Figure 2.1: Hillas plot [30]. For various potential Cosmic Ray sources the relation between magnetic field strength B and size L is depicted. The diagonal lines represent upper limits on B and L for the acceleration of a proton up to 10^{20} eV for two shock velocities β_s. An overview of various sources can be found in Ref. [32]. The LHC is depicted in order to clarify that the energy of particles accelerated in it is far below energies of UHECRs.

2.2 Extensive air showers

Cosmic Rays striking the Earth's atmosphere interact with air molecules after traveling a particular distance which is called the *interaction length* λ. This length is shorter for nuclei with mass A than for a proton [7],

$$\lambda_A < \lambda_{\mathrm{p}} \,. \tag{2.7}$$

A Cosmic Ray particle, which induces an interaction in the atmosphere, is referred to as the *primary particle*. The counterparts or targets of the primaries are air molecules such as nitrogen (N_2 with 78.1 %), oxygen (O_2 with 21.0 %), and argon (Ar with 0.9 %) [34]. Particles which are produced in interactions or decay processes after the primary interaction, are called the *secondary particles*. The series of physical processes changing the amount of particles is called an *extensive air shower*[*] since the particles travel until a few of them reach the ground. The air shower develops along the initial momentum vector of the primary particle. The extrapolated track of the initial direction of the primary particle is called the *shower axis*. Due to decay, scattering, and interaction processes during air shower development the particles can have varying transverse momenta leading to a lateral spread of particles in the shower. This leads to the formation of a *shower front* which propagates to the ground. The point of maximum lateral particle density in an air shower is called the *shower core* which is reconstructed once particles are measured at the ground. Measurements of the lateral distributions of hadrons, electrons, and muons, as well as their analytical description are for example presented in Ref. [27].
Pions and kaons produced in primary interactions mostly decay in the following branches:

$$\mathrm{K}^{\pm} \rightarrow \mu^{\pm} + \nu_\mu(\overline{\nu}_\mu) \tag{2.8}$$
$$\pi^{\pm} \rightarrow \mu^{\pm} + \nu_\mu(\overline{\nu}_\mu) \tag{2.9}$$
$$\mu^{\pm} \rightarrow \mathrm{e}^{\pm} + \nu_\mathrm{e}(\overline{\nu}_\mathrm{e}) + \overline{\nu}_\mu(\nu_\mu) \tag{2.10}$$

$$\pi^0 \rightarrow \gamma + \gamma \tag{2.11}$$
$$\gamma \rightarrow \mathrm{e}^{+} + \mathrm{e}^{-} \tag{2.12}$$

According to these decay processes, a particular air shower can be subdivided into three components, as depicted in Fig. 2.2:

- The *hadronic air shower component* mostly comprises protons, neutrons, pions, and kaons. At high energies, knowledge about hadronic interactions is sparse. Thus, they have to be described by high energy hadronic interaction models (see Ch. 4).

---

[*] If not otherwise cited, the following overview about extensive air showers is based on Refs. [26, 28, 29].

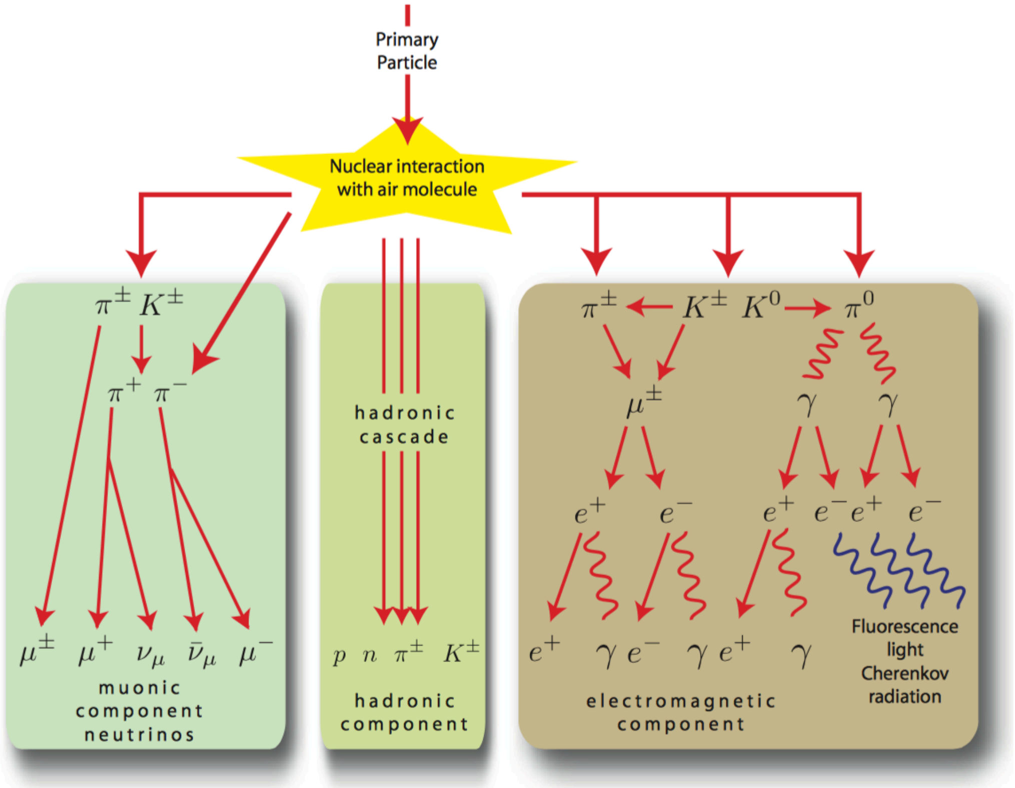
Primary Particle
Nuclear interaction with air molecule
$\pi^\pm$ $K^\pm$
π^+ π^-
$\mu^\pm$ μ^+ ν_μ $\bar{\nu}_\mu$ μ^-
muonic component neutrinos
hadronic cascade
p n $\pi^\pm$ $K^\pm$
hadronic component
$\pi^\pm$ $K^\pm$ K^0 π^0
$\mu^\pm$
γ γ
e^+ e^- e^+ e^- e^+ e^-
e^+ γ e^- γ e^+ γ
Fluorescence light Cherenkov radiation
electromagnetic component

Figure 2.2: Schematic development of an extensive air shower [35]. An air shower initiated by a hadron can be subdivided into three components.

- The *electromagnetic air shower component* comprises electrons, positrons, and photons. Electromagnetic *subshowers* containing these particles are dominated by radiative processes such as pair production and bremsstrahlung, as well as ionization caused by the electrons [30]. The electromagnetic air shower component can analytically be described in a simplified model called the *Heitler model* [31].

- The *muonic air shower component* comprises muons originating from different decay branches. One has to differentiate between the so-called *conventional muons* and the *prompt muons*. In Fig, 2.3 the contributions from different particles to the atmospheric muon flux is presented. Conventional muons are mostly produced by decays of pions and kaons. Up to a few PeV, the flux is dominated by conventional muons. At higher energies, the decay of pions and kaons into muons becomes rare and starts to compete with interaction processes [37]. In these interactions *charmed particles* such as mesons and baryons containing charm quarks are produced. These quickly decay into prompt muons which dominate the muon energy spectrum at higher energies [38, 39].

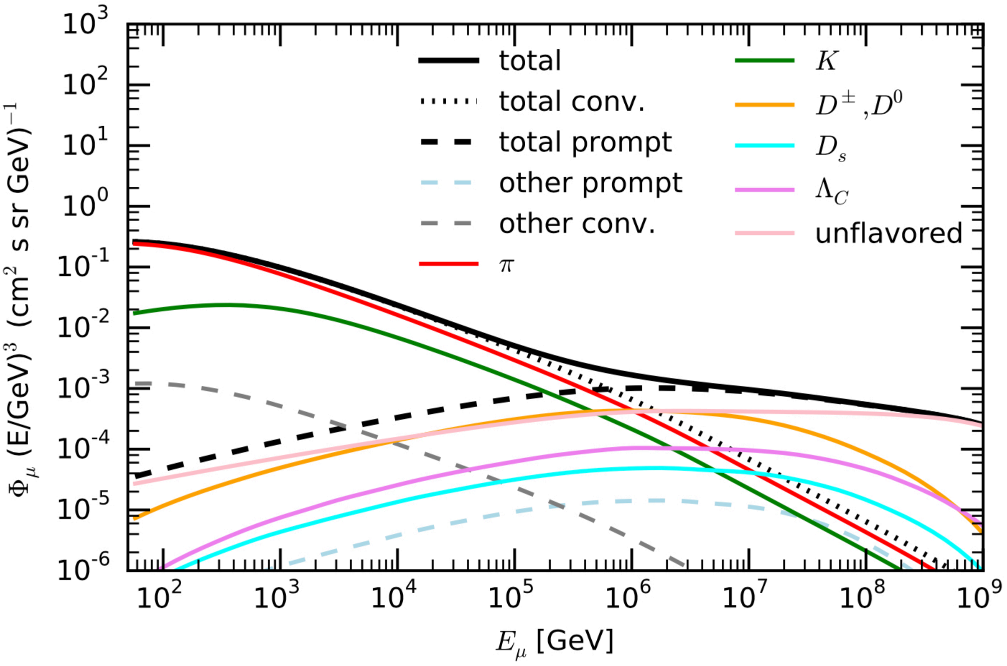

Figure 2.3: The atmospheric muon flux separated into contributions from different particle types [39].

In Ref. [28] it is shown that the relation between the mass A of a nucleus initiating an air shower, and the number of muons in it, can be modeled by

$$N_\mu^A = A^{1-\beta} N_\mu^{\mathrm{p}} \tag{2.13}$$

where N_μ^A is the number of muons in an air shower initiated by a nucleus with mass A, N_μ^{p} is the number of muons in an air shower induced by a proton, and β is a parameter related to the multiplicity of charged pions produced in secondary interactions. Equation 2.13 is valid for showers having equal primary energies. It shows that the number of muons is sensitive to the mass of the primary Cosmic Ray, i.e. the elemental composition of Cosmic Rays can be studied using the muon content of air showers.

2.3 Ground based measurements

Various experiments have measured the energy spectrum of Cosmic Rays and their elemental composition over different energy ranges. The properties of primary Cosmic Rays, such as mass and energy, determine the development of extensive air showers. Thus, air shower measurements can be used to infer information about the primary Cosmic Rays by reconstructing observables which are sensitive to primary energy and mass. The observables itself can be derived from air shower simulations (see Ch. 4). Results on the energy spectrum and elemental composition, in addition to those presented in this section, can be found in, for example, Refs. [13, 14, 17].

2.3.1 Energy spectrum

The differential all-particle energy spectrum of Cosmic Rays is illustrated in Fig. 2.4. The differential particle flux $F(E)$,

$$F(E) = \frac{d^4 N}{dE\, dA\, dt\, d\Omega} \tag{2.14}$$

denotes the number of Cosmic Ray particles N hitting the atmosphere per energy interval dE, area dA, time dt, and solid angle $d\Omega$. Up to primary energies of 100 TeV, Cosmic Rays can be measured directly [30]. This is done by detectors located on balloons or satellites in space. The restricted sensitive size of these detectors is not sufficient to detect Cosmic Rays of higher energies due to the steeply decreasing flux. Therefore, Cosmic Rays at higher energies are indirectly detected via extensive air showers at the ground where much larger detectors can be build. The event with the highest energy ever measured to date was detected by the Fly's Eye fluorescence detector (also called the *HiRes* - High Resolution Fly's Eye Cosmic Ray Detector) [18]. The reconstructed energy of the recorded event is $E = (3.2 \pm 0.9) \cdot 10^{20}$ eV [19]. The energy spectrum shown in Fig. 2.4 exhibits a few interesting features. It can be described by several power laws with different *spectral indices* γ, which denote the slopes and differ depending on the energy range,

$$F(E) = a \cdot E^{\gamma} \tag{2.15}$$

where a is a normalization constant. These are the slight changes of the spectral index at the *knee* at a few PeV, the *second knee* at roughly 100 PeV and the *ankle* at roughly 10 EeV. Understanding these features is a key goal in astroparticle physics because they imply information about Cosmic Rays at an early stage of their evolution. This information, for example, refers to changes of the origin, of acceleration mechanisms, and of propagation effects of Cosmic Rays.

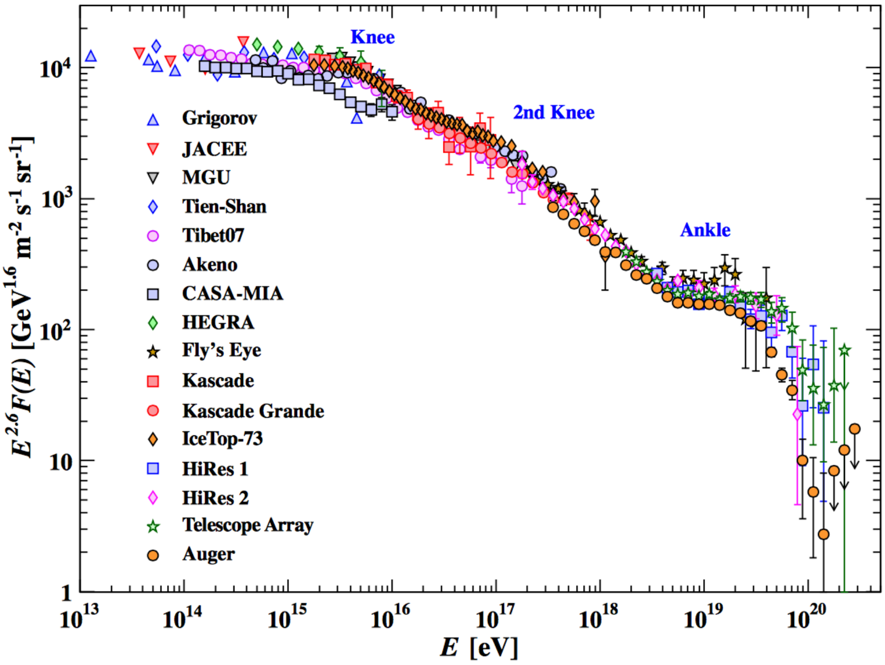

Figure 2.4: The differential energy spectrum of Cosmic Rays [33]. The particle flux as a function of the energy of the primary Cosmic Ray is shown. The spectrum is multiplied with $E^{2.6}$ in order to emphasize features in the structure.

At the knee of Cosmic Rays the fit spectral index changes from $\gamma_1 \approx -2.7$ to $\gamma_2 \approx -3.1$ [30]. A common explanation for the origin of the knee are the trajectories of nuclei in the galactic magnetic field. Since nuclei are charged, they follow a circular path. The radius of this path is called the *Larmor radius* and becomes larger with increasing energy of the nucleus [30]. The nucleus can leave the galaxy once its extension is exceeded by the Larmor radius. The radius depends on the charge of the nucleus Z, thus nuclei with higher mass require higher energies to escape.

The knee of Cosmic Rays can analytically be described using a phenomenological model, the *poly-gonato model* [73]. The quintessence of this model is that the all-particle energy spectrum as shown in Fig 2.4 can be parametrized by a superposition of energy spectra of individual nuclei. Each energy spectrum is given by a function which describes each particular knee as a smooth transition from the spectral index before the knee to the spectral index above the knee,

$$\frac{d\Phi_Z}{dE_0}(E_0) = \Phi_Z^0 \cdot E_0^{\gamma_1}\left(1 + \left(\frac{E_0}{E_Z}\right)^{\epsilon}\right)^{\frac{\gamma_2-\gamma_1}{\epsilon}} \tag{2.16}$$

where $d\Phi_Z$ is the differential energy spectrum at the energy value E_0, Φ_Z^0 is the absolute flux being a normalization constant, E_Z is the knee of the energy spectrum

of the nucleus with charge Z, γ_1 is the spectral index before the knee, γ_2 is the spectral index after the knee, and ϵ is a parameter defining the curvature of the function in the energy range of the knee.
Based on Eq. 2.15, the *integral flux* Ψ of Cosmic Rays can be calculated,

$$\Psi(E) = \int_E^\infty F(E')\,dE' \propto E^{\gamma+1} \tag{2.17}$$

and provides the entire number of Cosmic Ray particles per area dA, time dt, and solid angle $d\Omega$ with primary energies larger than the energy E. Therefore, the expected number of Cosmic Ray particles above the knee decreases by the factor $\sim 10^{-2}$ per energy decade.

2.3.2 Elemental composition

Figure 2.5 shows the mean logarithmic atomic mass of the Cosmic Ray primary as a function of its energy. At fixed energies, various experiments make different predictions for the mean elemental composition, even without overlapping error bars. This denotes systematic uncertainties persisting between the different experiments and the need for further composition studies. With increasing energies up to roughly 10^8 GeV, the mean logarithmic mass increases, which denotes an increase of the heavier component of Cosmic Rays.

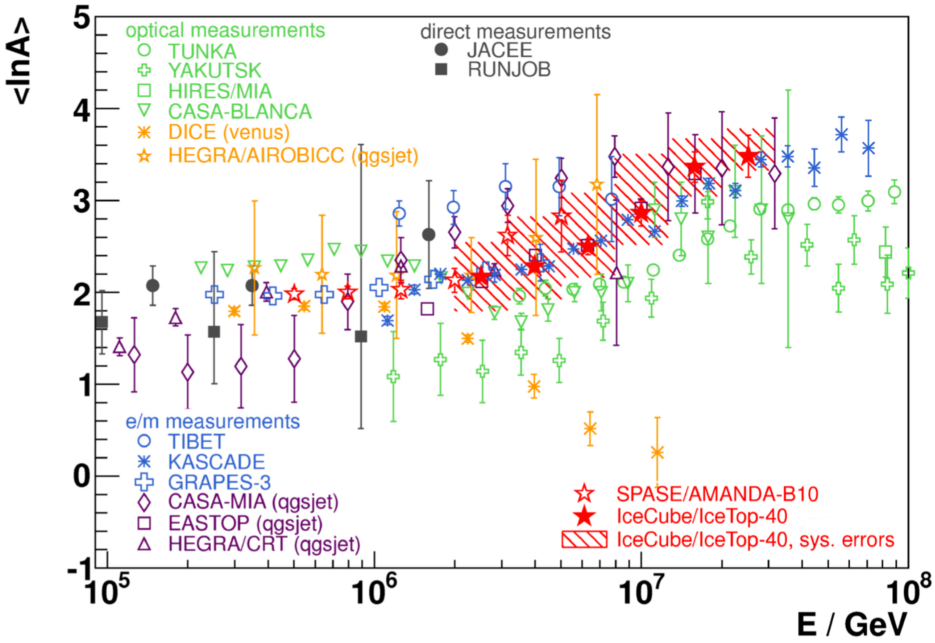

Figure 2.5: Mean logarithmic mass as a function of the primary energy for various experiments. Also included is the result for the former IceCube and IceTop configuration with 40 strings and 40 stations [9]. A mass composition of pure protons ($A = 1$) refers to $\ln A = 0$, while a mass composition of pure iron nuclei ($A = 56$) refers to $\ln A \approx 4$.

Chapter 3

The IceTop detector

In this chapter the focus is on the properties of the IceTop detector. After an explanation of the detector layout, the measurement and processing of experimental data is described. Finally, the dependance of experimental data on environmental conditions is accounted for.

Further and detailed information about the functionality of IceTop and the topics mentioned here can be found in Ref. [40].

3.1 Particles passing through matter

The passage of particles through matter is precisely described in Ref. [33]. In this work, the detection method bases on the measurement of so-called *Cherenkov light* which is explained in the following.

Electrically charged particles, crossing a medium such as ice, polarize nearby atoms. As a consequence, the atoms emit electromagnetic waves. If the velocity of the particle is greater than the phase velocity of light c_p in a medium [36] with refraction index n_p,

$$c_p = \frac{c}{n_p} \tag{3.1}$$

these waves coincide and one talks about Cherenkov light. The coincident waves, also called the *wavefront*, form a cone with opening angle θ_c which depends on the particle velocity v. As sketched in Fig. 3.1, the opening angle is given by

$$\cos\theta_c = \frac{1}{\beta n_p} \tag{3.2}$$

where $\beta = v/c$.

3.2 InIce and IceTop

The IceCube detector is located near the Amundsen-Scott-Station close to the geographical South Pole. It can be subdivided into two independent parts which are

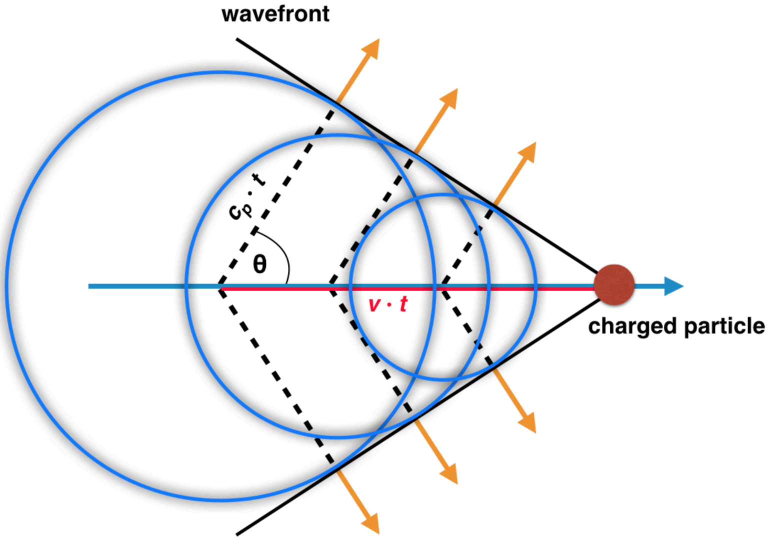

Figure 3.1: Schematic of Cherenkov light production.

called *InIce* and *IceTop*. The final configuration of IceCube, after deployment was finished in 2010, is depicted in Fig. 3.2. The sensors of IceCube are the *DOMs* (Digital Optimal Modules), presented in Fig. 3.3. They detect Cherenkov light emitted by charged particles traversing the ice.

The InIce detector comprises 86 strings deployed into the deep ice of the South Pole, each carrying 60 DOMs resulting in a total amount of 5160 DOMs. All DOMs are positioned in depths from 1.45 km to 2.45 km, together forming a 3-dimensional detector array. The ice at these depths has been chosen as detection medium since it is clear, and free of bubbles. Thus, this particular ice has excellent optical properties and serves as an appropriate medium for the detection of Cherenkov light. Additionally, the ice above the DOMs serves as a natural shield for most of the particles produced in extensive air showers which interact with the ice close to the surface and therefore do not reach InIce.
The goal of this work is the calculation of a muon number density using experimental data measured by IceTop. Therefore, experimental data from InIce is not used.

In Fig. 3.4 the layout of the IceTop detector is shown. It is located at an atmospheric depth of 680 g/cm^2 (2835 m a.s.l.) [40]. IceTop consists of 81 stations, where each station is composed of two tanks separated by 10 m, called the *Tank A* and the *Tank B*. As shown in Fig. 3.4, the IceTop geometry is not perfectly flat, there is a difference in altitude of $\sim$ 6 m between the highest and the lowest tank.Each IceTop tank includes two DOMs, resulting in a total amount of 324 DOMs.
The center of IceTop is the *IceCube Laboratory* (ICL) where experimental data from InIce strings and IceTop stations is collected. The deployment of stations having similar distances to each other, enables to measure primary energies between 100 TeV and 1 EeV [40] covering the knee of Cosmic Rays (see Sec. 2.3.1).

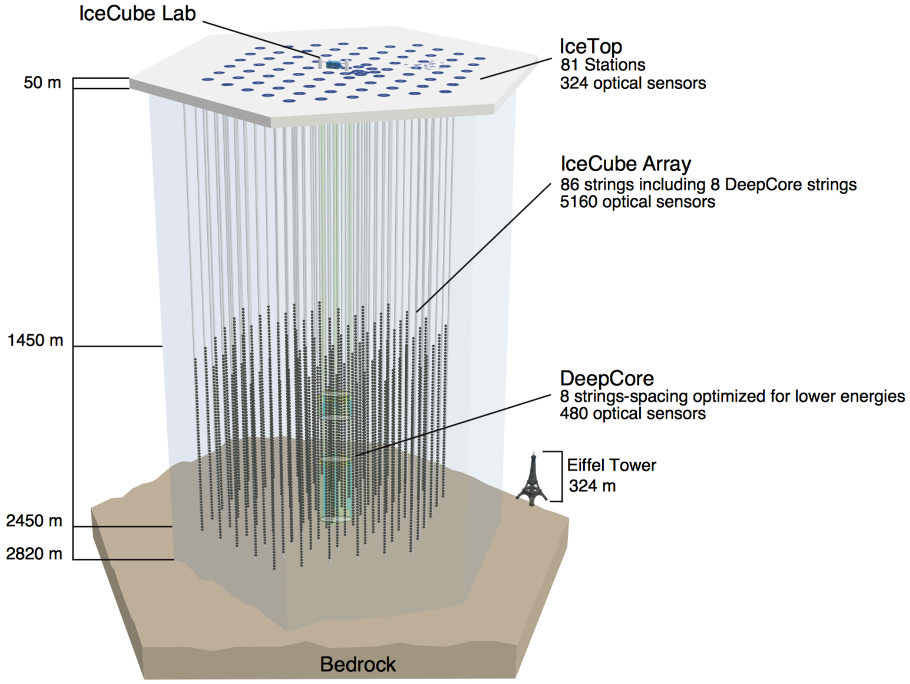
IceCube Lab
50 m
1450 m
2450 m
2820 m
IceTop
81 Stations
324 optical sensors
IceCube Array
86 strings including 8 DeepCore strings
5160 optical sensors
DeepCore
8 strings-spacing optimized for lower energies
480 optical sensors
Eiffel Tower
324 m
Bedrock

Figure 3.2: Schematic view of the IceCube detector [41]. InIce is the successor of the AMANDA detector, which took experimental data up to May 2009 [42, 43]. Also shown is DeepCore, a denser instrumented region lowering the minimum energy threshold for air shower detection. At the surface, IceCube is covered by the IceTop array which can be operated independently from InIce.

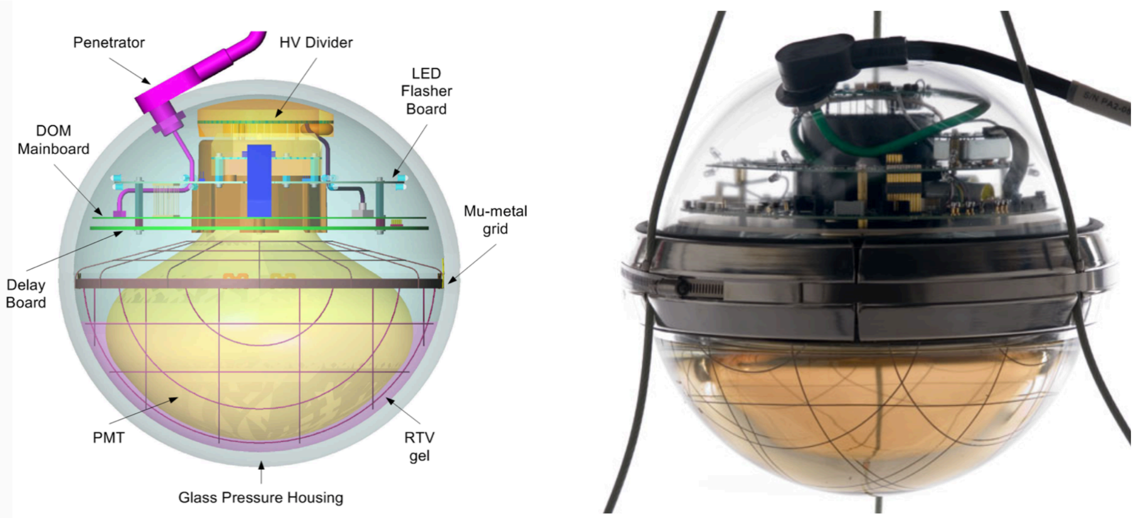
Penetrator
HV Divider
LED
Flasher
Board
DOM
Mainboard
Mu-metal
grid
Delay
Board
PMT
RTV
gel
Glass Pressure Housing

Figure 3.3: Digital optical modules (DOMs). Integrated into a glass pressure housing, a DOM consists of a *PMT* (Photomultiplier tube) connected to electronic components providing voltage supply for the PMT and readout elements. The materials of the DOM and the PMT are chosen such that Cherenkov light is favored to be detected (see Fig. 4.4 for the acceptance of a DOM).
Left: In this schematic of a DOM the nomenclature of its main components is shown [44].
Right: A photography of a DOM which is deployed in IceCube [45].

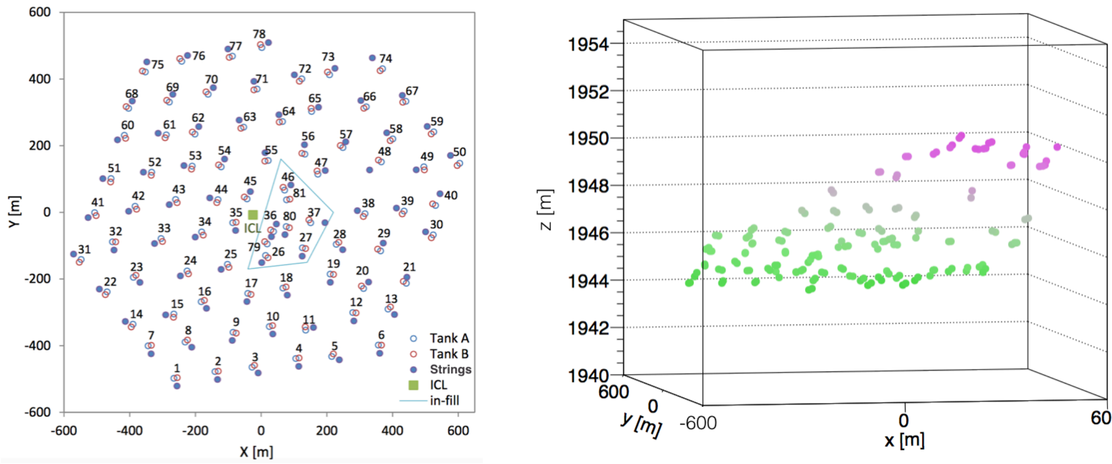

Figure 3.4: Different views of IceTop.
Left: In this 2-dimensional sketch the pattern of IceTop is shown [40]. Almost every pair of tanks is positioned near an InIce string. The trapeze represents the borders of IceTop in-fill.
Right: In this 3-dimensional sketch of IceTop it is shown that the z-coordinates of various tanks differ up to more than 5 m.

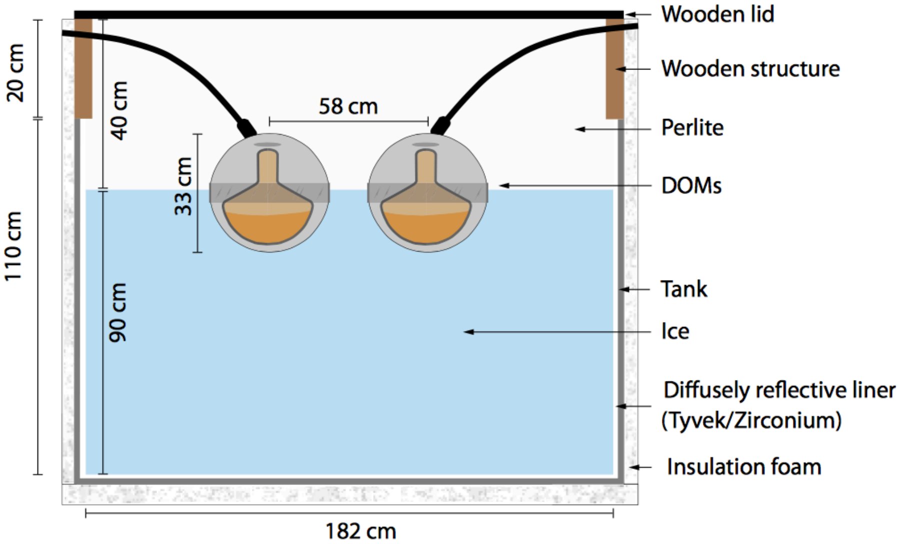
20 cm
110 cm
40 cm
90 cm
33 cm
58 cm
182 cm
Wooden lid
Wooden structure
Perlite
DOMs
Tank
Ice
Diffusely reflective liner
(Tyvek/Zirconium)
Insulation foam

Figure 3.5: Layout of an IceTop tank [40].

A denser instrumented region, called *IceTop in-fill*, allows the reconstruction of air showers with lower primary energies since stations are closer to each other. The stations 79, 80, and 81 are called the *in-fill-stations*, since they were especially deployed for in-fill, and all other stations are called the *standard-stations*.
In Fig. 3.5 the constituents of an IceTop tank are depicted. Two DOMs are frozen in ice, each containing a PMT able to measure Cherenkov light emitted by charged particles crossing the ice. The irradiation of the PMTs causes an ejection of *photoelectrons* at the PMT's photocathode. These photoelectrons are accelerated to dynodes of higher potential and knock off further electrons which are accelerated to the next dynode. After the number of photoelectrons increased exponentially, they are collected at the anode.
Light signals measured by IceTop vary significantly because the kinetic particle energy and density in an air shower varies depending on the distance to the shower axis. Because of this, the two PMTs in a tank are operated with different *gain*. The gain of a PMT is the ratio of anode to photoelectron current, and therefore depends on the voltage applied between the dynodes. The DOMs in a tank are called the *High Gain DOM* (HG) and the *Low Gain DOM* (LG) since they amplify the photoelectron current by gains of $5 \cdot 10^6$ and $1 \cdot 10^5$, respectively. Effects inherent to PMTs, such as pre-pulses, late pulses and afterpulses, are discussed in Ref. [46] and are not relevant here.
The output at the anode of a PMT is kept if the peak pulse voltage exceeds a particular discriminator threshold [40]. This value of this threshold is $\sim 23\,\mathrm{PE}$ for HG DOMs and $\sim 270\,\mathrm{PE}$ for LG DOMs, where a *PE* is the average charge value generated in a PMT by a *single photoelectron* (*SPE*) [40]. If the discriminator threshold is passed, one talks about a *DOMLaunch*. The recorded information following a DOMLaunch is called a *hit*. A DOMLaunch initiates the Data Acquisition System (DAQ) which digitizes the signal using two alternating ATWD (Analog Transient Waveform Digitizer) circuits, each with channels of different gains, and a fADC (fast Analog-to-Digital Converter). The ATWD channels sample the signal in 128 bins with a width of 3.33 ns. As a result, a *waveform* is recorded which is the superposition of all pulses holding the measured voltage as a function of time (see Fig. 3.6). The fADC samples the signal in 256 bins of each 25 ns size. In this work, only data from the ATWD stream is used. Data from the fADC stream has not yet been used in IceTop analyses [40]. As shown in Fig. 3.5, a tank is equipped with a liner consisting of material with a high diffuse reflectivity of Cherenkov light. This ensures only small light absorption. Most of the tank liners are made of zirconia, a few of Tyvek. A fraction of tanks with Tyvek liners, which have been deployed in 2005, have different diffuse reflectivities compared to the tanks deployed afterwards [40].

The deployment of two tanks per station has various important reasons [40]:

- As visible in the Cosmic Ray energy spectrum (Fig. 2.4), the flux of air showers of small primary energy is large. In IceTop, it can be suppressed by requiring that both tanks in a station see a signal within a particular time window (see Sec. 3.4). This favors the detection of air showers with higher primary energies.

- The reconstruction of air showers, explained in Ch. 5, is performed based on measured charge values in IceTop tanks. IceTop can be split into two subarrays, each using just one of both tanks. Since both subarrays are similar, fluctuations of the reconstructed quantities can be studied.

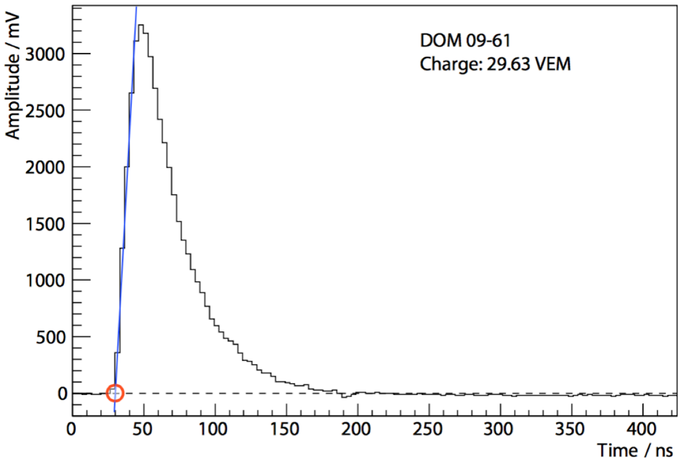

Figure 3.6: A waveform recorded by a particular DOM in IceTop [40]. The waveform is characterized by distinct features such as a steep rise, a peak, and an exponential drop-off. The blue line represents a linear description of the leading edge which is extrapolated up to the baseline (dashed line). The point of intersection (red circle) defines the time of the hit.

3.3 Experimental data

Experimental data in IceCube is organized in so-called *Runs*. This is a certain time period of usually eight hours, in which IceCube is taking data. During monitoring of experimental data, each Run is examined with focus on its quality for later analysis. The basic expectations of Runs recorded by IceCube are:

- The minimum livetime of a Run must be larger than ten minutes.

- The rates of monitored observables must be reasonable and have to match the rates belonging to Runs adjacent in time. This way, unexpected behavior due to a change in rate can be determined. DOMs which measure deviant rates or fail completely, are excluded from the Run.

- A failure of the electronics can cause peaks in measured rates or the failure of several DOMs. In this case, the entire Run may be excluded from data analysis.

- Runs with calibration purpose are excluded from usual data analysis. These are, for example, used to calibrate measured charge values in IceTop tanks (see Sec. 3.6), or ”flasher Runs”, where InIce is lit by an artificial light source.

After monitoring of each Run, it is decided which Runs are suitable for analysis and will be kept. For each Run, detailed information about the detector performance, such as the detector geometry including failed DOMs, snow heights on IceTop tanks, and calibration parameters of the DOMs, is stored. The default values needed for the charge calibration (see Sec. 3.6) are wrong, which propagates to wrong air shower reconstructions. The reconstructions were repeated using the correct values. For the simulation of air showers, a general configuration is used (see Sec. 4.3).
In this work IceTop experimental data is used which was measured in the time period from May 15th, 2012 to May 2nd, 2013, corresponding to the detector configuration *IT81-2*. This expression is commonly used for denoting that this particular data was recorded in the second year in which IceTop consists of 81 stations.

Experimental Data from IceTop and InIce can be used independently from each other or together. IceTop can, for example, be used as a veto. A measurement in InIce with a coincident measurement in IceTop, is likely caused by air shower particles. Such coincident analyses require the zenith angle of the incident Cosmic Rays to be smaller than 30° in order to have the shower axis contained in both detectors [10]. Since in this work only experimental data from IceTop is used, there is no such restriction on the zenith angle.

3.4 Local coincidences

As depicted in Fig. 3.7, each station in IceTop is equipped with wires connecting DOMs of different tanks. This allows the classification of signals in IceTop in two disjunct samples.
A launch in one of the HG DOMs initiates an electronic search for a launch in one of the DOMs in the other tank. If this exists within $1\,\mu$s, the participating DOMs are part of a *HLC* hit. In this case, the recorded waveforms are saved. If there is no DOM in the other tank having a coincident signal, one talks about a *SLC* hit at the tank with the initial signal. In this case, only the value of the integrated ATWD charge, and a timestamp are extracted by the DOMs firmware based on intrinsic information about PMT gain and baseline.
Exemplary distributions of HLC and SLC hits are shown in Fig. 3.8. Close to the shower axis, the particle density in the shower front is high. Thus, HLC hits mostly occur in this region. At particular distances which are dependent on the primary energy, the number of SLC hits becomes abundant. They occur when the density of particles in air showers is low, i.e. far away from the shower axis. Since SLC hits do not fulfill a local coincidence condition, they are isolated signals. Therefore, they can be used to identify single particles such as muons (see Ch. 6). This is important in this work since the calculation of a muon density requires signals in IceTop tanks with preferably low non-muonic contributions.

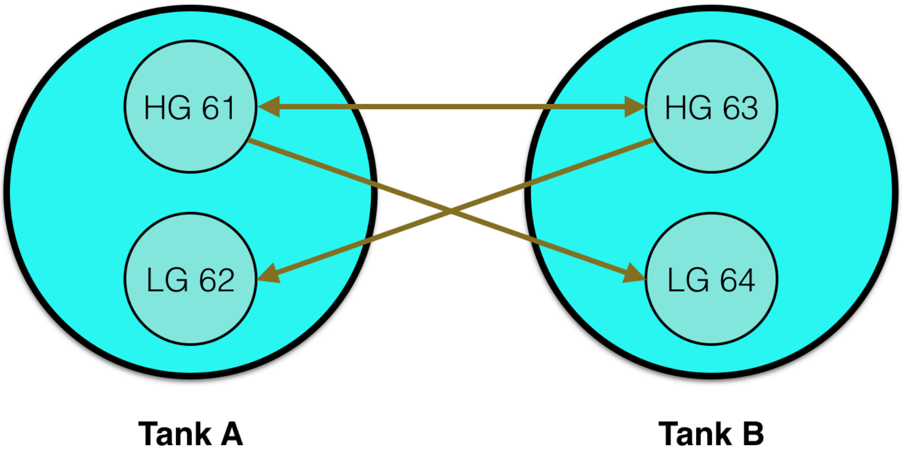

Figure 3.7: Cabling between the DOMs within an IceTop station. DOMs in IceTop are labeled with the numbers 61-64. If one of the HG DOMs passes a particular discriminator threshold, a search for a coincident signal in the other tank is initiated (see text).

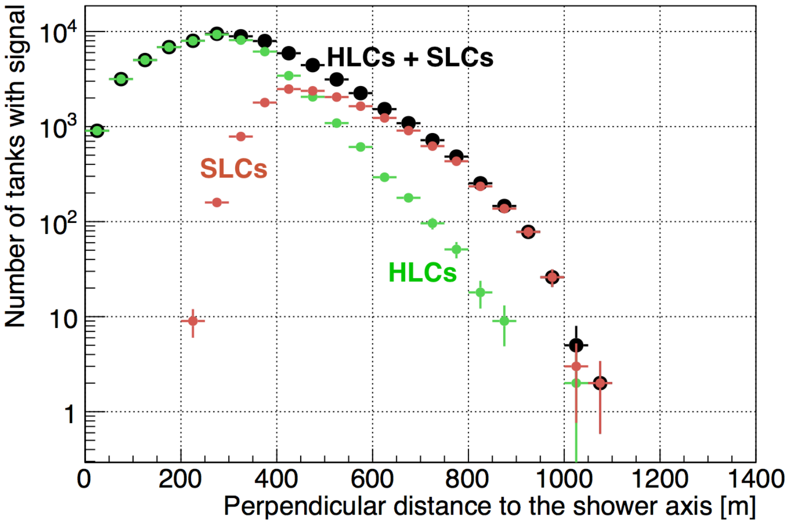

Figure 3.8: Exemplary distributions of HLC and SLC hits. They are produced using many simulated air showers with vertical iron primaries which have a primary energy of $E = 20\,\text{PeV}$. The shape of the distributions, including a peak, occurs due to counting in ring-like bins around the shower axis. This is explained in detail in Ch. 6.2.1.

3.5 IceTop Trigger and Filter

There are three different triggers running on IceTop experimental raw data [40] for the data taking during the IT81-2 season:

- The *Simple Majority Trigger* is a basic trigger for air shower physics and requires at least 6 HLC hits within 6 μs. Since HLC hits are used for the reconstruction of air shower parameters (see Ch. 5), the trigger also ensures a sufficient number of HLC hits for the reconstruction algorithm.
- The *Minimum Bias Trigger* collects all experimental data independently of any other trigger. In IceTop, it triggers on every 10^4 th event with at least one HLC hit, i.e. the corresponding event containing this particular HLC hit is kept. This condition ensures that every air showers is triggered with equal probability, i.e. independent from its properties such as primary energy. The Minimum Bias Trigger is particularly interesting for analyses studying background or searching for signatures which are not covered by other triggers.
- The *Calibration Trigger* triggers, for example, on *VEMCal Launches*. These are particular DOM launches at HG DOMs, besides usual data taking, where no local coincidence is found. This procedure suppresses electromagnetic background and emphasizes signals of single muons required for the calibration of measured charge values in units of VEM (explained in Sec. 3.6). In contrast to information provided by SLC hits, the full waveform is read out during a VEMCal launch.

Experimental data which is recorded by IceCube, and triggered by one of the above triggers, is transferred via satellite to the North for further processing. For reconstruction quality reasons (see Ch. 5), in this work only air showers are used where the total amount of HLC hits passed a filter called *IceTopSTA5*. This filter requires to pass the Simple Majority Trigger, and at least five standard-stations to record an HLC hit.

3.6 Charge calibration

PMT and DOM electronics are calibrated such that all DOMs have the same time reference, and the measured charge value is given in units of PE.

Since the various IceTop tanks have different optical properties, and different DOM efficiencies, the usage of PE as unit is not appropriate to compare charge values measured at different tanks. A global unit is introduced, which is used to calibrate each signal given in units of PE. This unit is the *VEM (Vertical Equivalent Muon)*. For calibration, the natural atmospheric flux of muons is used. The average energy of these muons at detector level is about 2 GeV - 2.5 GeV [40]. According to Ref. [33], muons of these energies are minimum ionizing, i.e. they have a mean energy loss in the tank which is close to the minimum of energy loss described by the Bethe equation [33]. Thus, muons with a particular track in an IceTop tank, depending on their zenith angle, always lose roughly the same amount of energy. For each HG DOM, the response of a vertical muon, given in units of PE, is defined as 1 VEM. During VEMCal launches the calibration is performed based on the measured charge distribution of a DOM, as for example shown in Fig. 3.9. In simulations, it has been shown that the bump in the exponential drop-off is due to muons [47].

The charge distribution is fit by a superposition of single muonic and electromagnetic contributions [40]:

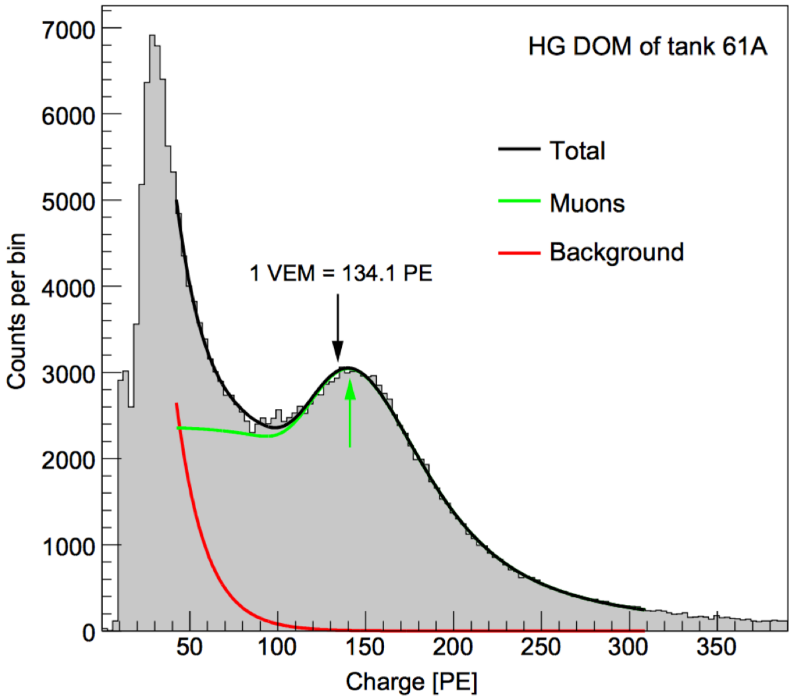

Figure 3.9: Charge distribution measured by a particular HG DOM, serving as example for the procedure of VEM calibration [40]. The term *Background* represents the electromagnetic contribution to the total measured charge. The fit functions (red and green lines) are described by Eq. 3.3.

$$f(x) = \underbrace{p_0 \left(L(x;p_1,p_2) + \frac{1.85}{p_1} \cdot \frac{1}{e^{\frac{x-p_1}{p_2}}+1} \right)}_{f_\mu(x)} + \underbrace{p_3 \cdot e^{p_4 \cdot x}}_{f_{\mathrm{em}}(x)} \tag{3.3}$$

where f has five free parameters $p_0 ... p_4$, f_μ represents the muonic part of the distribution, and f_{em} the elctromagnetic one. The function f_μ is composed of a Landau distribution L accounting for the signal of muons which enter IceTop tanks through the lid and leave through the bottom, while the second part of f_μ describes the signal of *edge-clipping* muons, i.e. muons which enter the tank through the edge and can have short tracks in the ice. The function f_{em} has an exponential form.

The definition of 1 VEM is given by 95% of the maximum of the peak in the fit

function f_μ (see Fig. 3.9). This value has been taken from measurements performed in the 2005/2006 season using a muon telescope which was placed directly on an IceTop tank. Only muons which had a nearly vertical track were able to pass this arrangement. The peak of the resulting charge spectrum appears to be at 95% of the peak measured without a muon telescope [40].
As mentioned before, the method described above is only used for HG DOMs. LG DOMs are cross-calibrated using relative differences in the measured charge values compared to the HG DOMs [40].

VEM calibration performed during detector simulation (see Sec. 4.3) slightly differs from the procedure performed in experimental data described above. In simulation, the number of Cherenkov photons per VEM, generated by simulated muons, is chosen such that the resulting charge spectrum at a tank (such as in Fig. 3.9) after detector simulation is equal to the one measured in experimental data (see [40] for a detailed explanation).
In the used software, 1 VEM is assigned to 32880 Cherenkov photons. A consequence of this method is that analyses in IceTop are not affected by DOM efficiencies.
The differences in VEM calibration between experimental data and simulation are handled as a systematic uncertainty in Sec. 7.6.

3.7 Snow on IceTop tanks

The height of snow on IceTop tanks yearly increases by $\sim 20\,\mathrm{cm}$ due to drifting and is impacted by slopes and buildings located on the IceTop array [40].
In IceTop, snow heights are determined using two methods depending on the season. During the antarctic summer (December - February), snow heights on IceTop tanks can be measured directly with a stick (in-situ).
In between direct measurements, snow heights can be estimated based on the VEM calibration spectra for each tank (see Sec. 3.6) [40]. This can be done because the electromagnetic shower component is attenuated within the snow while the muonic component is nearly unaffected [68]. According to Ref. [40], the functions f_μ and f_{em}, coming from the process of VEM calibration of charge values at a particular tank (Eq. 3.3), can be related to the snow height above this tank,

$$\frac{S_\mu}{B_{\mathrm{em}}} = \frac{\displaystyle\int_{0.3\,\mathrm{VEM}}^{2.0\,\mathrm{VEM}} f_\mu \, dS}{\displaystyle\int_{0.3\,\mathrm{VEM}}^{2.0\,\mathrm{VEM}} f_{\mathrm{em}} \, dS} \tag{3.4}$$

where S_μ/B_{em} is the ratio of the muonic component S_μ in the charge spectrum, treated as *signal*, to the electromagnetic component B_{em}, treated as *background*.
In Fig. 3.10 the dependance of this ratio on snow heights, measured during two antarctic summers, is shown.

The linear dependance of $\ln\left(S_\mu/B_{\text{em}}\right)$ on the observed snow heights h_{snow} is described by an exponential function which in turn provides a function for the snow height,

$$\begin{aligned} \frac{S_\mu}{B_{\text{em}}}\left(h_{\text{snow}}\right) &= e^{\frac{h_{\text{snow}}}{A}+B} \\ \Longleftrightarrow h_{\text{snow}}\left(\frac{S_\mu}{B_{\text{em}}}\right) &= A \cdot \left(\ln\left(\frac{S_\mu}{B_{\text{em}}}\right) - B\right) \end{aligned} \tag{3.5}$$

where A, B are free parameters.
In order to estimate snow heights above IceTop tanks during a particular antarctic winter, the parameter A is fixed to the average of all snow heights measured during the enclosing antarctic summers. The parameter B is calculated for each tank such that the calculated snow height, according to Eq. 3.5, is close to the snow heights above the corresponding tank, measured during the antarctic summer (see Ref. [40] for more information).

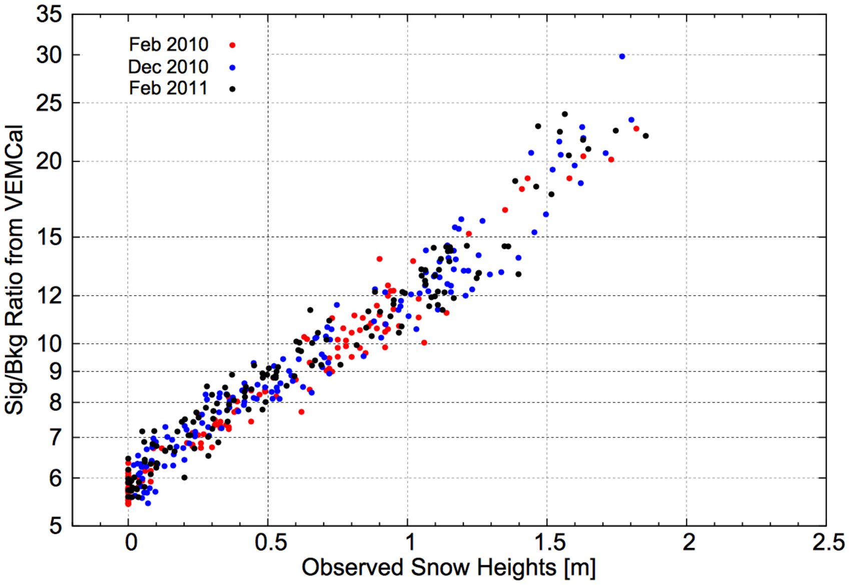

Figure 3.10: S_μ/B_{em} as a function of the snow heights above IceTop tanks directly measured during three different months [40]. The y-axis is presented in log-scale.

In Fig. 3.11 a sketch of the snow heights on IceTop tanks, measured in July 2012, and the interspaces is shown. These particular snow heights are used for detector simulation in this work (see Sec. 4.3). Snow also has an effect on the reconstruction of air showers, as explained in Ch. 5.

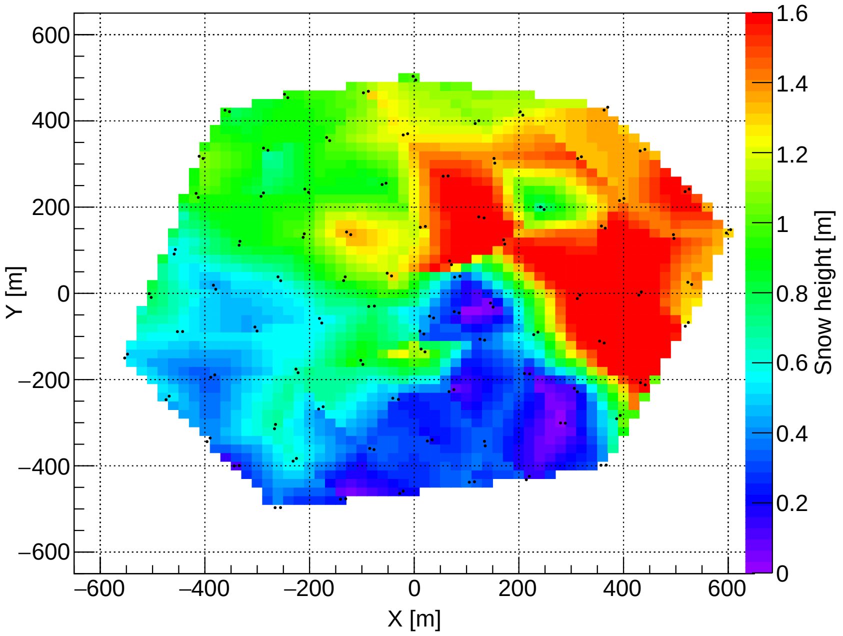

Figure 3.11: Snow height distribution on IceTop during July 2012. Each individual tank is denoted by a black dot. Snow heights between the tanks are due to interpolation.

3.8 Atmospheric conditions

While snow mainly affects the electromagnetic shower component [68], atmospheric conditions, such as density and ground pressure, influence the development of the entire air shower.

As shown in the upper plot of Fig. 3.12, the atmospheric temperature profile follows an annual cycle. Since temperature and density of the atmosphere are anti-correlated, the density follows an opposite profile. Furthermore, the atmosphere is denser during the antarctic winter than during the antarctic summer. Changes in atmospheric density lead to a change in the rate of lepton production [40]. Nevertheless, effects due to varying atmospheric densities cancel out in IceTop analyses, which use a full year of experimental data [40].

As can be seen in the lower plot of Fig. 3.12, the observed IceTop DOM launch rate is strongly anti-correlated with the pressure at the ground. For example, an increasing ground pressure leads to a stronger attenuation of an air shower and thus to a decrease of the DOM launch rate. During air shower reconstruction, a smaller DOM launch rate leads to an underestimation of the parameter S_{125} because it is correlated with the primary energy (see Sec. 5.3). Since a full year of data is used in this work, it is assumed that systematic uncertainties due to ground pressure cancel out.

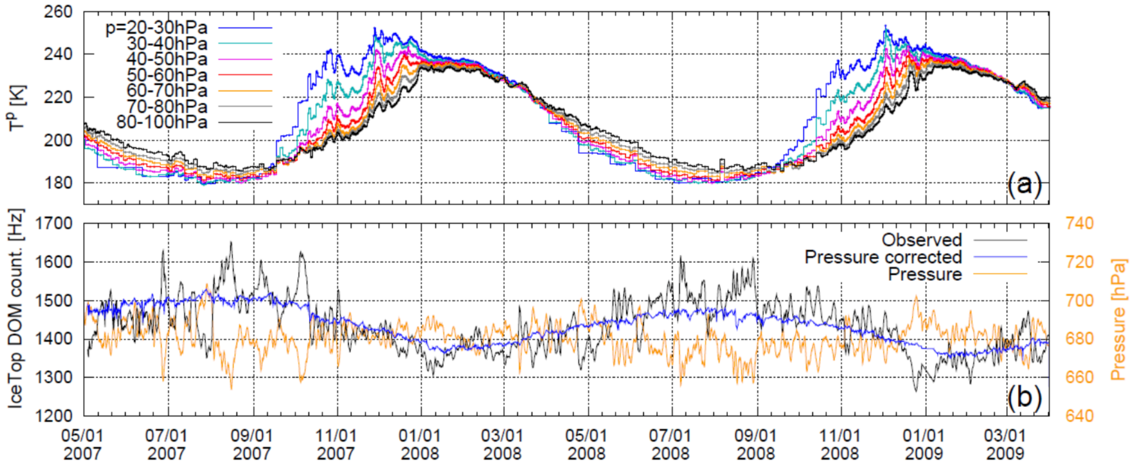

Figure 3.12: Temporal atmospheric conditions at the South Pole [40].
Top (a): The temperature profile for different layers of the stratosphere, given by their pressure, is depicted.
Bottom (b): The ground pressure is shown together with the observed IceTop DOM rate and a barometric correction of it (see Ref. [40]). The latter is not relevant here.

Chapter 4

Simulation

The goal of this work is the calculation of muon number densities based on hits recorded by IceTop. Such signals are usually produced by different particle types. Hits caused by muons can be separated based on selection cuts which are developed in Ch. 6. These cuts rely on the quantitative description of background which has to be estimated using air shower simulations.

This chapter deals with the properties of the simulation sets used in this work, and the treatment of particle signatures in the detector.

4.1 Air shower simulation

The properties of the simulation sets are presented in Tab. 4.1. They were provided by the IceCube collaboration and produced using version 73700 of the *CORSIKA* (COsmic Ray SImulations for KAscade) Monte Carlo code [48, 49] which is a common tool in astroparticle physics.

The adjustable parameters of the program are the environmental conditions of the development of the air shower. A wide range of properties can be simulated such as the primary particle type, primary energy range and spectrum, angular direction, magnetic field, atmosphere, and the height above the ground where particle tracking stops, called the *observation level*, i.e. where the air shower simulation finishes, and each particle information is saved. The used simulation sets comprise air showers initiated by proton and iron primaries, both representing extremes in primary mass. Developing an analysis on both simulation sets hypothetically guarantees that they bracket experimental data in important distributions.

The simulated energy and zenith angle spectra have particular properties:

- The simulated air showers follow a piecewise E^{-1} energy spectrum being flatter than the realistic one. It is not appropriate to simulate a realistic spectrum due to computing time reasons. Since the flux of Cosmic Rays is rapidly decreasing with primary energy, a large number of showers would have to be simulated to provide sufficient statistics at high energies. The simulated energy spectrum has 30 bins with a width of 0.1 in $\log_{10}(E\,[\mathrm{GeV}])$.

Interaction model	**SIBYLL2.1**	**QGSJetII04**	**EPOS LHC**
Provided nuclei	**p**, **Fe**		
Range in energy E	0.1 PeV - 100 PeV in 30 bins using a stepsize of $\log_{10}\left(\frac{E_{i+1}}{E_i}\right) = 0.1$		
Number of generated air showers	2000 per energy bin	200 per energy bin	
Energy spectrum	E^{-1}		
Range in zenith angle θ	$0°$ - $65°$ in a spectrum $\propto \sin\theta \cdot \cos\theta$		
Range in azimuth angle ϕ	$0°$ - $360°$ uniformly distributed		
Observation level	1953.08 m in IceCube coordinates ($\sim$2837 m a.s.l.)		
Magnetic field	$B_x = 16.4\,\mu\mathrm{T}$, $B_z = -53.4\,\mu\mathrm{T}$		
Atmosphere	Average April South Pole atmosphere profile with a pressure of $\sim$680 hPa at detector level		

Table 4.1: Properties of the air shower simulations. If not otherwise mentioned, the color code is used in this work to represent primary masses. **p** stands for the *hydrogen nucleus* (proton), and **Fe** for the *iron nucleus*. The IceCube coordinate system has its origin in the center of InIce ($\sim$ 884 m above sea level (a.s.l.)). The expressions B_x and B_z denote the horizontal and vertical components of Earth's magnetic field at the South Pole given in CORSIKA coordinates, as described in Ref. [49]. The South Pole atmosphere is described by a custom density profile which comprises five atmospheric layers and is derived from experimental data, as described in Ref. [59].

- The spectrum of the zenith angle θ for air showers distributed on a horizontal area, such as IceTop, is $\propto \sin\theta \cos\theta$. The term $\sin\theta$ originates from the solid angle differential ($d\Omega = \sin\theta \, d\theta \, d\phi$), and the $\cos\theta$ accounts for the orientation of the detector with respect to the incoming Cosmic Rays.

Essential for analyses relying on CORSIKA simulations is the hadronic interaction model the physics is based on. Low energy hadronic interactions have been well established by accelerator data and in IceCube they are simulated using FLUKA [50]. Since currently no man-made accelerator can reach up to the particle energies in high energy Cosmic Rays, the knowledge about quantities of high energy physical processes during the development of an air shower, such as interaction cross sections, ionization rates, or particle multiplicities, is sparse. Thus, the simulation of high energy interactions relies on theoretical models. For the description of high energy hadronic interactions the simulation production in IceCube usually relies on the SIBYLL2.1 model [51, 52, 53]. In this work, two additional simulation sets based on hadronic interactions calculated with the models QGSJetII04 [54] and EPOS LHC [55], provided by the IceCube collaboration, are used for studying systematic uncertainties. In the used hadronic interaction models charmed particles are not treated directly according to physical processes described in Sec. 2.2. Instead, the decay of charmed particles is implicitly handled by CORSIKA [49], for example by adjusting particle multiplicities [57].

QGSJetII-04 and EPOS LHC have been adapted to LHC data and significantly produce more muons than SIBYLL2.1. This has already been reported by the Pierre Auger collaboration, as shown in Fig. 4.1. Additionally, results for the conditions of the IceTop detector which have especially been produced for this work, are also presented in Fig. 4.1. Between 10 PeV and 100 PeV, the two bottom plots overlap however for IceTop the amount of muons is larger than for Auger. This is because the Auger experiment is closer to sea level ($\sim$ 1500 m a.s.l.) than IceTop ($\sim$ 2837 m a.s.l.), i.e. the observation level is different. In the latter case muons travel shorter distances through the atmosphere and thus less of them decay or interact before being detected.

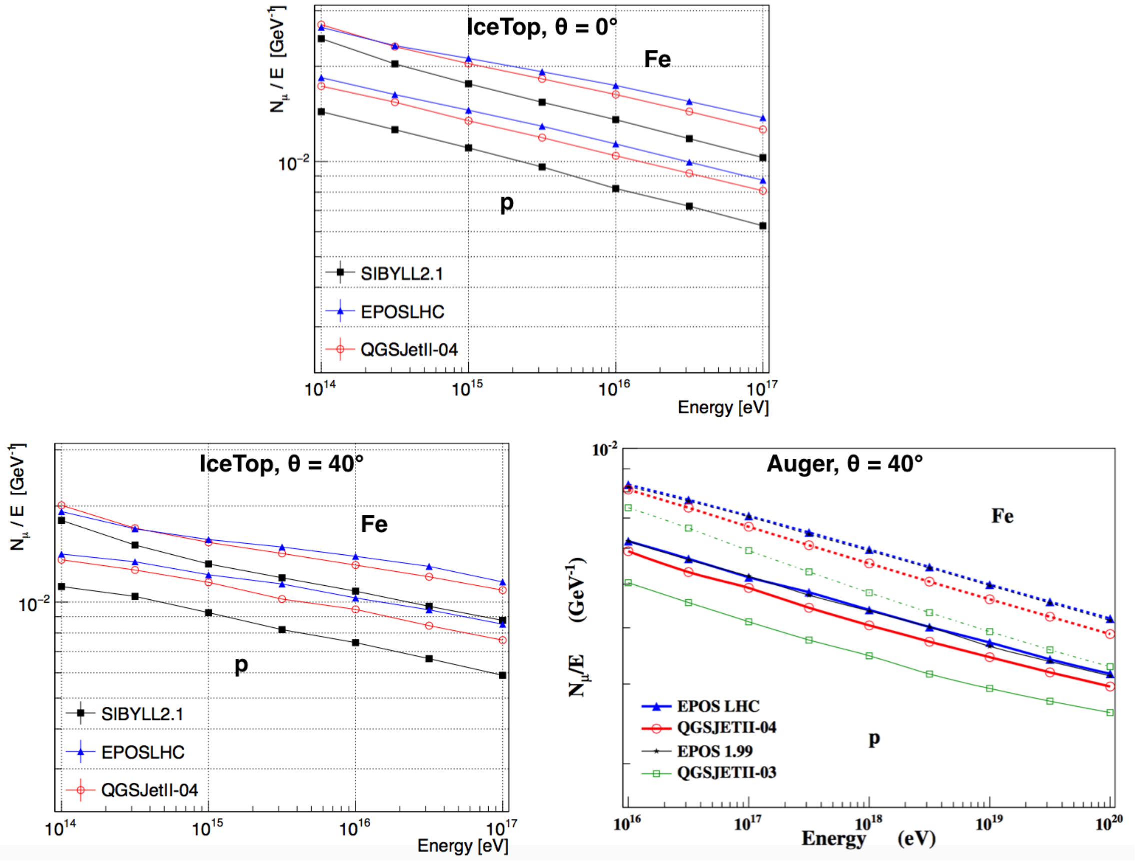

Figure 4.1: Number of muons per primary energy as a function of the primary energy for IceTop and Auger.
Top and bottom left: The produced simulation sets are similar to the sets used for analysis (see Tab. 4.1), but air showers are restricted to 0°, and 40°, and the electromagnetic component is suppressed.
Bottom right: This is the same as the plot on the bottom left, but for the Auger experiment and without SIBYLL2.1 [56].

4.2 Cosmic Ray flux model

The simulation of an E^{-1} spectrum makes the simulation sets different from experimental data where the energy spectrum is steeper. A *weighting* of all individual simulated air showers accounts for these differences. In differential form, the simulated and realistic spectra can be described as

$$dn_{\text{sim}} = \lambda \cdot E^{-1}\, dE \sin\theta \cos\theta \, d\theta \, d\phi \, dA \tag{4.1}$$

$$dn_{\text{real}} = k \cdot f(E)\, dE \sin\theta \cos\theta \, d\theta \, d\phi \, dA \tag{4.2}$$

where λ and k are normalization constants, $f(E)$ is the realistic energy spectrum as described in Sec. 2.3.1, and dn_{sim} (dn_{real}) is the differential number of simulated (realistic) air showers with slices of differential energy $E^{-1}\, dE$ ($f(E)\, dE$), zenith angle $\sin\theta \cos\theta \, d\theta$, azimuth angle $d\phi$ and detector area dA.
According to Eqs. 4.1 and 4.2 the required weight w is expressed by

$$w = \frac{dn_{\text{real}}}{dn_{\text{sim}}} = \frac{k}{\lambda} E f(E) \tag{4.3}$$

While k can be obtained from measurements of energy spectra, the calculation of λ requires multiple integration over the simulated quantities,

$$\int dn_{\text{sim}} = n_{\text{sim}} = \lambda \cdot \int_{E_{\text{min}}}^{E_{\text{max}}} E^{-1}\, dE \int_0^{\theta_{\text{max}}} \sin\theta \cos\theta \, d\theta \int_0^{2\pi} d\phi \int_0^{A_{\text{sim}}} dA \tag{4.4}$$

where n_{sim} is the number of simulated air showers in the energy bin $[E_{\text{min}}, E_{\text{max}}]$, zenith angle bin $[0, \theta_{\text{max}}]$, azimuth angle bin $[0, 2\pi]$, with shower cores distributed over the area A_{sim} which depends on primary energy (see Sec. 4.3.1). Integration leads to

$$\begin{aligned} n_{\text{sim}} &= \lambda \cdot \ln\left(\frac{E_{\text{max}}}{E_{\text{min}}}\right) \sin^2\theta_{\text{max}}\, \pi \, A_{\text{sim}} \\ \Leftrightarrow \frac{1}{\lambda} &= \ln\left(\frac{E_{\text{max}}}{E_{\text{min}}}\right) \sin^2\theta_{\text{max}}\, \pi \, \frac{A_{\text{sim}}}{n_{\text{sim}}} \end{aligned} \tag{4.5}$$

Inserting this into Eq. 4.3, the weight changes to

$$w = k\, E\, f(E) \ln\left(\frac{E_{\text{max}}}{E_{\text{min}}}\right) \sin^2\theta_{\text{max}}\, \pi \, \frac{A_{\text{sim}}}{n_{\text{sim}}} \tag{4.6}$$

This expression can be further modified. As described in Sec. 2.3.1, the energy spectrum of Cosmic Rays can be parametrized using the expression

$$f(E) = E_0^{\gamma_1} \left(1 + \left(\frac{E_0}{E_{\text{Knee}}}\right)^{\epsilon}\right)^{\frac{\gamma_2 - \gamma_1}{\epsilon}} \tag{4.7}$$

where E_{Knee} is the position of the knee in the Cosmic Ray energy spectrum (see Fig. 2.4). Due to the simulation of a piecewise energy spectrum with step sizes of 0.1 in $\log_{10}(E)$ for each of the 30 energy bins, it is

$$\ln\left(\frac{E_{\text{max}}}{E_{\text{min}}}\right) = \ln(10^{0.1}) = 0.1 \cdot \ln(10) = \frac{0.1}{\log(e)} \tag{4.8}$$

Using Eqs. 4.7 and 4.8, the weight changes to

$$w = \frac{0.1\,\pi}{\log(e)}\, k\, E_0^{\gamma_1+1} \left(1 + \left(\frac{E_0}{E_{\text{Knee}}}\right)^{\epsilon}\right)^{\frac{\gamma_2-\gamma_1}{\epsilon}} \frac{A_{\text{sim}} \sin^2\theta_{\text{max}}}{n_{\text{sim}}} \tag{4.9}$$

In this work, the weighting is performed with the energy E_0 given in units of PeV. Thus, the normalization constant k is the flux of Cosmic Rays at 1 PeV, extracted from real measurements shown in Fig. 1 of Ref. [74],

$$\begin{aligned} E^{2.5}\,\frac{dN}{dE}(1\,\text{PeV}) &= 2950\,\text{m}^{-2}\,\text{sr}^{-1}\,\text{s}^{-1}\,\text{GeV}^{1.5} \\ \Rightarrow k = \frac{dN}{dE}(1\,\text{PeV}) &= 2.95 \cdot 10^{-6}\,\text{m}^{-2}\,\text{sr}^{-1}\,\text{s}^{-1}\,\text{PeV}^{-1} \end{aligned} \tag{4.10}$$

In order to describe the energy spectrum, $\gamma_1 = -2.7$, $\gamma_2 = -3.1$, $E_{\text{Knee}} = 3.0\,\text{PeV}$ have been chosen (see Sec. 2.3.1). The curvature of Eq. 4.7 is described by $\epsilon = 2$. The small value of ϵ guarantees a smooth transition from one spectral index to the other [73].

4.3 Detector simulation

Simulations and reconstructions[†] of air showers are performed by *IceTray* which is an analysis framework based on a modular structure especially developed for IceCube. The modules are written in python and C++.

In this work, the official information about geometry, calibration, and detector status for the 2012 simulation production, is used for detector simulation. Default snow heights for each tank are wrong and thus overwritten by the snow heights of July 2012 (see Sec. 3.7). Furthermore, the DOM number 61 at station number 39 is not taking data. Thus, this particular DOM is not part of the simulation.

In the following sections, an overview of the most important properties used during detector simulation in this work is given. The simulation of the detector hardware, i.e. PMT and DOM electronics, is described in Refs. [40, 46].

[†] In this work version V04-01-07 of the simulation software *IceSim* and version V04-08-00 of the reconstruction software *IceRec* are used. After the reconstruction of air showers, C++ and the analysis framework ROOT [58] are used for further analysis.

4.3.1 Shower core placement

After an air shower simulation is finished at a particular observation level, the coordinates of the particles and the shower core, are arbitrary as long as the air shower is not referred to a particular detector. The placement of the shower core at particular coordinates determines the entire detector simulation.
The detector simulation itself is initiated by a module called the *I3TopSimulator*. The *I3CorsikaInjector* service module which is part of the I3TopSimulator module, places simulated air showers on the IceTop array by defining the coordinate vector of the shower core, $\vec{x}_\mathrm{C} = (x_\mathrm{C},\ y_\mathrm{C},\ z_\mathrm{ol})$, where z_ol is the observation level. In experimental data, showers are distributed isotropically over the detector area and beyond. Since simulations always try to get as close as possible to experimental data, this behavior is mimicked by randomly distributing shower core locations over a circular area with radius r_sim around the center of IceTop. Since air showers with increasing primary energy have increasing lateral spread, r_sim depends on the primary energy. It has to be chosen such that an air shower of a given primary energy with $|(x_\mathrm{C},\ y_\mathrm{C})| = r_\mathrm{sim}$ barely misses triggering the detector. This ensures correct simulation of air showers which trigger the detector even if the shower core is far away from the center of IceTop. The chosen values for r_sim are depicted in Fig. 4.2, and are commonly used in IceTop analyses.

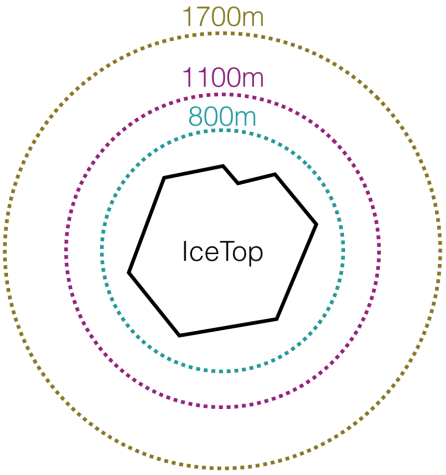

Figure 4.2: Radii in which shower cores are distributed on the IceTop array. The radii are different for each decade in primary energy (800 m for 0.1 PeV - 1.0 PeV, 1100 m for 1.0 PeV - 10 PeV, and 1700 m for 10 PeV - 100 PeV).

In order to gain statistics and to save computing time, each simulated air shower is placed within the corresponding circular area several times. This procedure keeps the entire information about the air shower, such as the zenith angle θ, the azimuth angle ϕ, and particle momenta and energies, but changes the x- and y-coordinates of the shower core location randomly within the corresponding area. The number of *Resamples* has been chosen to be 100 ensuring that on average each air shower triggers the detector once [40]. With this, and $A_{\mathrm{sim}} = \pi\, r_{\mathrm{sim}}^2$, the weight in Eq. 4.9 can be modified to

$$w = \frac{0.1\,\pi^2}{\log(e)}\, k\, E_0^{\gamma_1+1} \left(1 + \left(\frac{E_0}{E_{\mathrm{Knee}}}\right)^{\epsilon}\right)^{\frac{\gamma_2-\gamma_1}{\epsilon}} \frac{r_{\mathrm{sim}}^2 \sin^2\theta_{\mathrm{max}}}{100 \cdot n_{\mathrm{sim}}} \tag{4.11}$$

In the case of the SIBYLL2.1 simulation sets, the observation level of 1950.08 m is within the snow above several IceTop tanks, or even below the tank lids in a few cases, as shown in Fig. 4.3. This behavior leads to an incorrect detector response simulation (see Sec. 4.3.2).

In order to obtain correct results during the detector simulation, the observation level is raised artificially, i.e. the z-coordinate of each particle is raised in order to be above each snow height, instead of running air shower simulations with CORSIKA again. In particular, the observation level for air shower simulations based on SIBYLL2.1 is raised to 1953.08 m, which is equal to the observation level in the QGSJetII04 and EPOS LHC simulation sets.

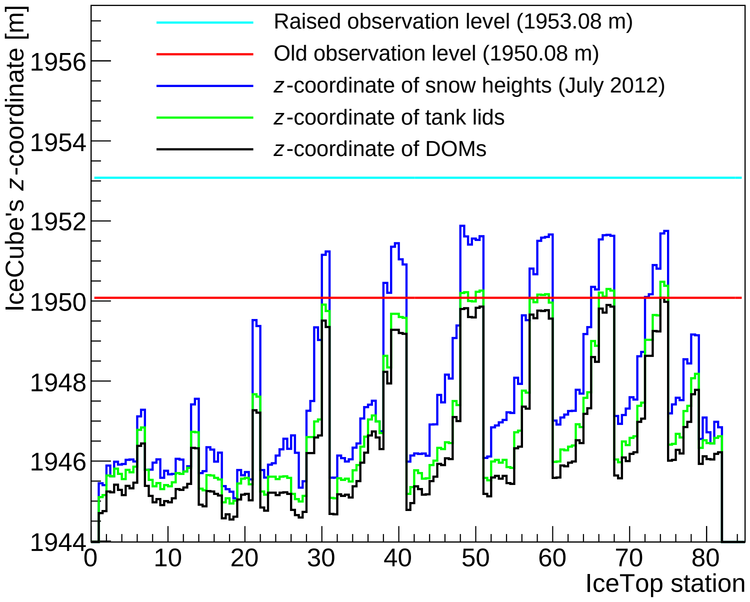

Figure 4.3: The observation level of the SIBYLL2.1 simulation sets compared to tank specific heights. Particles were propagated to an observation level (red line) which is too low compared to snow heights (dark blue line), and tank lids (green line). For a correct calculation of the detector response, the observation level is raised above all snow heights (light blue line). The structure of the distributions is due to height differences of IceTop tanks, as shown in the right plot in Fig. 3.4.

4.3.2 Particle tracks

The Geant4 simulation toolkit [60, 61, 62] is commonly used to calculate the passage and interaction of particles travelling through matter. In IceTop, Geant4 computes the particle tracks and energy losses during their propagation from observation level through the lower atmosphere, snow, ice, and the tank and DOM materials (see Fig. 3.3). A particle is excluded from detector simulation if its track is at least 30 cm from a tank volume.

4.3.3 Cherenkov photons

Within the ice of the tanks, Geant4 simulates the production of Cherenkov photons, and uses the conversion of a fixed number of photons to 1 VEM, as explained in Sec. 3.6. Due to computing time reasons, Cherenkov photons are not tracked. Instead, their arrival times at the photocathode are distributed exponentially in order to have similar waveform decay times in simulation and experimental data [40].
In Fig. 4.4, the acceptance of an IceTop DOM is shown. This is the fraction of recorded to incoming Cherenkov photons in consideration of the quantum efficiency of the PMT, and the materials the DOM consists of, such as glass and gel. Due to a high acceptance of wavelengths between 300 nm and 650 nm, only Cherenkov light matching these wavelengths is included in the detector simulation [40].

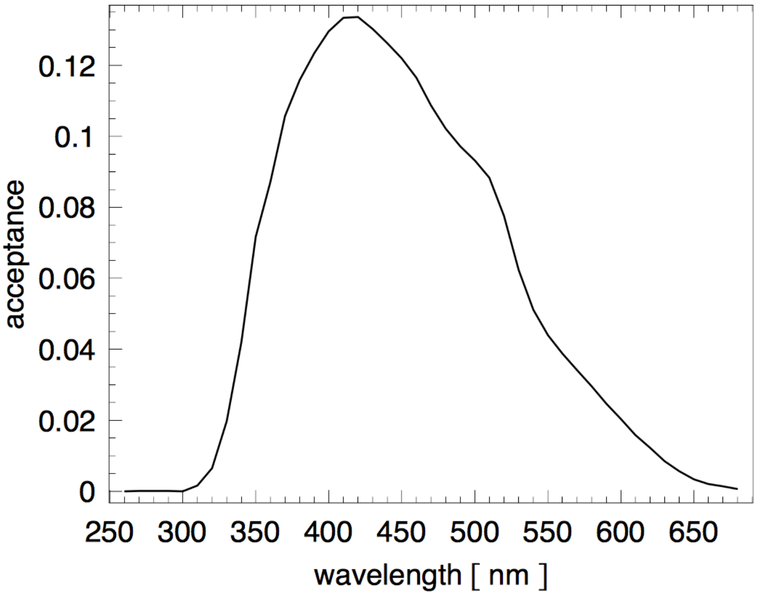

Figure 4.4: Acceptance of an IceTop DOM as a function of the incident wavelength [63].

4.3.4 Noise

Noise originating from different sources is present in experimental data and has to be modelled throughout the simulation. In IceCube, a project called *vuvuzela* is used to simulate

- the dark noise rate due to thermal fluctuations which lead to the emission of electrons in the dynodes.
- the so-called *correlated noise* which is caused by the radioactive decay of impurities. Within the glass housing of a DOM, the energy of the decay products is transferred into photons due to scintillation or luminescence effects.

The output of the I3TopSimulator module is the number of photoelectrons which were released from the photocathode, for each PMT. Noise due to the processes mentioned above is added to the output.

Chapter 5

Reconstruction and event selection

Air shower reconstruction in IceTop is performed by fitting theoretical models of the lateral air shower development and the shower front, to simulated or measured charge values and times of tanks with HLC hits. In this chapter, these models are described. From reconstruction, parameters characterizing the air shower are extracted. In the last part of this chapter, quality cuts and further improvements are applied on these parameters in order to ensure the quality of the used air showers.

5.1 Preparations of the used signals

Before the reconstruction of a particular air shower can be performed, the set of measured HLC hits has to be prepared. This is done by the so-called *cleaning* [40]:

- The charge value and time of a HLC hit are retrieved from the signal of the HG DOM. If the measured charge value exceeds a particular threshold, i.e. the DOM is saturated, the charge value from the LG DOM is used.

- Within a station the two tanks, A and B, must fulfill the following condition:

  $$|t_\mathrm{A} - t_\mathrm{B}| < \frac{|\vec{x}_\mathrm{A} - \vec{x}_\mathrm{B}|}{c} + 200\,\mathrm{ns} \tag{5.1}$$

  where t_i is the time the HLC hit is recorded at tank $i \in [\mathrm{A}, \mathrm{B}]$, and $\vec{x}_i$ is the coordinate vector of the corresponding tank. The first part of the right side of Rel. 5.1 is the time a horizontal air shower requires to travel from one tank to the other. For a distance of 10 m between both tanks of a station, at least roughly 30 ns are required. A time window of 200 ns is added because of two reasons. Firstly, it accounts for the time smearing of the propagation of Cherenkov light in the tank. Secondly, the interval is chosen to be small in order to remove unrelated signals. The corresponding station is discarded if Rel. 5.1 is not fulfilled.

- Different stations i and j with $i, j \in [1, 81]$ have to fulfil a condition similar to the previous one:
$$|t_i - t_j| < \frac{|\vec{x}_i - \vec{x}_j|}{c} + 200\,\mathrm{ns} \tag{5.2}$$
where $t_{i/j}$ is the signal time averaged for both tanks in a station, and $\vec{x}_{i/j}$ is the coordinate vector of the midpoint of both tanks. Stations which together fulfill the condition of Rel. 5.2 are merged to a so-called *cluster*. The appearance of more than one cluster represents coincident air showers in IceTop. Each cluster is kept and reconstructed separately.

An important variable in this work is the *perpendicular distance to the shower axis* $R_{T,p}$. As shown in Fig 5.1, this variable denotes the shortest distance of an IceTop tank T or a particle p to the reconstructed shower axis $\vec{A}$. The latter can be parametrized as a line intersecting the IceTop array at the position of the reconstructed shower core $\vec{x}_\mathrm{C}$,

$$\vec{A} = \vec{x}_\mathrm{C} + \alpha\vec{n} \tag{5.3}$$

where α is a scalar and the unit vector $\vec{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ is the reconstructed angular direction of the shower axis. The vector $\vec{x}_{T,p}$, connecting the center of an IceTop tank T or a particle p at observation level with the shower axis, is given by

$$\vec{A} - \vec{x}_{T,p} = \vec{x}_\mathrm{C} + \alpha\vec{n} - \vec{x}_{T,p} \tag{5.4}$$

and should be perpendicular to the direction of the shower axis:

$$\begin{aligned} 0 &= (\vec{x}_\mathrm{C} + \alpha\vec{n} - \vec{x}_{T,p}) \cdot \vec{n} \\ \Leftrightarrow \alpha &= (\vec{x}_{T,p} - \vec{x}_\mathrm{C}) \cdot \vec{n} \end{aligned} \tag{5.5}$$

Inserting α into Eq. 5.3 yields the point on the shower axis with shortest distance to $\vec{x}_{T,p}$, given by

$$R_{T,p} = |\vec{x}_\mathrm{C} + [(\vec{x}_{T,p} - \vec{x}_\mathrm{C}) \cdot \vec{n}]\,\vec{n} - \vec{x}_{T,p}| \tag{5.6}$$

The calculation of $\vec{x}_\mathrm{C}$ and $\vec{n}$ is described in Sec. 5.2.

5.2 Reconstruction procedure

The reconstruction of an air shower starts with first guess calculations of shower core and direction. Both are used as starting values for a likelihood minimization. The electromagnetic shower component is attenuated in non-uniform snow heights above IceTop tanks (see Fig. 3.11) which affect the height of measured charge values. This effect is accounted for by the so-called *snow correction*.

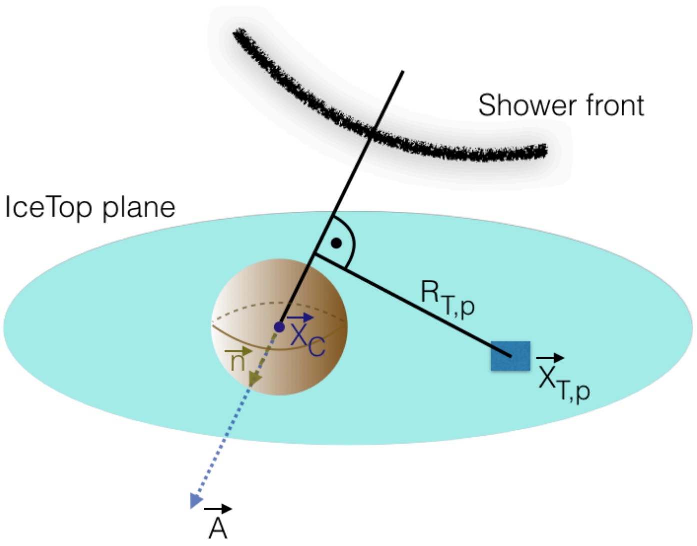

Figure 5.1: Perpendicular distance to the shower axis $R_{T,p}$.

5.2.1 First guess reconstructions

In order to estimate the position of the shower core $\vec{x}_\mathrm{C}$, the *center of gravity* $\vec{x}_\mathrm{COG}$ of the N tanks with the highest signals, is calculated. The tank coordinates $\vec{x}_i$ are weighted by the square root of the measured signal S_i, given in units of VEM,

$$\vec{x}_\mathrm{COG} = \frac{\sum_{i=1}^{N} \sqrt{S_i}\vec{\mathrm{x}}_i}{\sum_{i=1}^{N} \sqrt{S_i}} \tag{5.7}$$

The default number of tanks used for this calculation is N = 7 [40].

The approximate direction of the air shower axis, $\vec{n}$, is estimated by using the recorded hit times of the signal. The shower front is approximated as a plane propagating with the speed of light [40],

$$t\left(\vec{x}\right) = t_0 + \frac{1}{c}(\vec{x} - \vec{x}_\mathrm{COG}) \cdot \vec{n} \tag{5.8}$$

where t_0 marks the time when the estimated shower core $\vec{x}_\mathrm{COG}$ reaches the surface of IceTop. $\vec{n}$ is approximated by minimizing the value χ^2 using Eq. 5.8,

$$\chi^2 = \sum_{i}^{n_\mathrm{T}} \frac{(t_i^\mathrm{meas} - t_i(\vec{x}_i))^2}{\sigma_i^2} \tag{5.9}$$

where n_T is the number of hit and unsaturated tanks belonging to the largest cluster as described in Sec. 5.1, t_i^meas is the time the tank i measured a signal, and $\sigma = 5\,\mathrm{ns}$ is a time uncertainty assumed to be constant. The estimation of the shower direction $\vec{n}$ is performed in two steps. In the first step, the χ^2 in Eq. 5.9 is minimized assuming that all tanks have an equal z-coordinate. The resulting $\vec{n}$ is used as seed for a second minimization step which accounts for the different z-coordinates (the latter can be seen in Fig. 4.3).

5.2.2 Functions used for reconstruction

Reconstruction of air shower parameters in IceTop is performed by a so-called *likelihood minimization*[‡] of a log-likelihood function $\mathcal{L}$. This means that the reconstructed air shower parameters are calculated such that the probability for measuring a particular set of charge values and times at particular IceTop tanks, is at maximum. This means the reconstruction algorithm minimizes the difference between measured and expected signals.

In this section, only the most important properties of the reconstruction are mentioned. A detailed description is provided in Ref. [40]. Nevertheless, particular aspects of implementation can only be found in the source code.

The entire likelihood function is given by

$$\mathcal{L} = \mathcal{L}_\mathrm{q} + \mathcal{L}_\mathrm{t} + \mathcal{L}_0 + \mathcal{L}_\mathrm{s} \tag{5.10}$$

where the individual contributions are described in the following.

Charge likelihood function $\mathcal{L}_\mathrm{q}$

Measured charge values in IceTop tanks are described by a lateral distribution function [64, 65],

$$S(R) = S_\mathrm{ref} \left(\frac{R}{R_\mathrm{ref}} \right)^{-\beta - \kappa \log_{10}(R/R_\mathrm{ref})} \tag{5.11}$$

where $S(R)$ is the charge expectation value, in units of VEM, at perpendicular distance R to the shower axis, S_ref is the expected signal at reference distance R_ref given in units of VEM, β is a measure of the slope of the lateral distribution at R_ref, and $\kappa = 0.303$ is a measure of the curvature of the lateral distribution [40].

Equation 5.11 is also known as the *double logarithmic parabola*,

$$\log_{10} S(R) = \log_{10} S_\mathrm{ref} - \beta \log_{10} \left(\frac{R}{R_\mathrm{ref}} \right) - \kappa \log_{10}^2 \left(\frac{R}{R_\mathrm{ref}} \right) \tag{5.12}$$

since the term $\log_{10} \left(\frac{R}{R_\mathrm{ref}} \right)$ appears quadratically. In Ref. [66] it has been shown that during likelihood minimization, a reference distance of $R_\mathrm{ref} = 125\,\mathrm{m}$ minimizes the correlation between S_ref and β. Thus, $S_\mathrm{ref} = S_{125}$ is used in this work. S_{125} is also called the *shower size* since it is a measure of the energy of the primary particle (see Sec. 5.3 and Ref. [40]).

[‡] Usually, log-likelihood functions are maximized. The minimization of the same function, but with negative sign, is equivalent, and, for example, has technical advantages.

The likelihood function $\mathcal{L}_\mathrm{q}$ is given by [40]

$$\mathcal{L}_\mathrm{q} = -\sum_i^{n_T} \frac{(\log_{10} S_i^\mathrm{meas} - \log_{10} S_i^\mathrm{fit})^2}{2\sigma^2_{\log_{10} S}(S_i^\mathrm{fit})} - \sum_i^{n_T} \ln(\sigma_{\log_{10} S}(S_i^\mathrm{fit})) \tag{5.13}$$

where n_T is the number of hit and unsaturated tanks, S_i^meas is the charge value measured at tank i, S_i^fit is the expectation value according to Eq. 5.11 at tank i, and $\sigma_{\log_{10} S}(S_i^\mathrm{fit})$ is the fluctuation of the signal with a functional description provided in Ref. [40].

Timing likelihood function $\mathcal{L}_\mathrm{t}$

During likelihood minimization, measured times are described by a function modeling a curved shower front [40] which is more realistic than Eq. 5.8,

$$t(\vec{x}) = t_0 + \frac{1}{c}(\vec{x} - \vec{x}_\mathrm{C}) \cdot \vec{n} + \Delta t(R) \tag{5.14}$$

where t is the time expectation value at $\vec{x}$, t_0 is the time when the shower core $\vec{x}_\mathrm{C}$ reaches the surface of IceTop, the unit vector $\vec{n}$ is the direction of the shower, R is the perpendicular distance to the shower axis, and $\Delta t(R)$ is the model of the shower front which can be described by a parabola and a gaussian,

$$\Delta t(R) = aR^2 + b \cdot \left(1 - \exp\left(-\frac{R^2}{2\sigma^2}\right)\right) \tag{5.15}$$

where $a = 4.823 \cdot 10^{-4}\,\mathrm{ns/m}^2$, $b = 19.41\,\mathrm{ns}$, and $\sigma = 83.5\,\mathrm{m}$ are constants [40]. The timing likelihood function is given by [40]

$$\mathcal{L}_\mathrm{t} = -\sum_i^{n_T} \frac{(t_i - t(\vec{\mathrm{x}}_i))^2}{2\sigma^2_\mathrm{t}(R_i)} - \sum_i^{n_T} \ln(\sigma_\mathrm{t}(R_i)) \tag{5.16}$$

where n_T is the number of hit and unsaturated tanks, t_i is the time the tank i measured a signal, $t(\vec{x}_i)$ is the expectation value according to Eq. 5.14, and $\sigma_t(R_i) = 2.92\,\mathrm{ns} + 3.77 \cdot 10^{-4}\,\mathrm{ns} \cdot (R_i/\mathrm{m})^2$ denotes the fluctuation of the arrival times [40].

No-hit likelihood function $\mathcal{L}_0$

$\mathcal{L}_0$ accounts for IceTop stations which do not trigger, i.e. do not measure a signal above discriminator threshold (see Ref. [40] for more information).

Saturation likelihood function $\mathcal{L}_\mathrm{s}$

$\mathcal{L}_\mathrm{s}$ is a likelihood function which is used in addition to the three functions previously shown. It was developed in Ref. [67] and accounts for all tanks which are saturated. This has to be done because the measured charge value in a saturated tank is an underestimation of the true charge value.

5.2.3 Shower size correction due to snow

The electromagnetic shower component is attenuated due to snow above the IceTop tanks [68]. Thus, measured charge values are smaller compared to the case of absent snow. Since snow is not distributed uniformly over the IceTop array, there is a stronger attenuation of charge values in IceTop regions with more snow. Consequently, air showers having their core in regions with more snow are reconstructed with a smaller shower size S_{125} than in regions with less snow. This shift in S_{125} leads to a lower rate of measured air showers in regions with more snow. This is accounted for by applying a correction to each of the fit values S_i^{fit} in Eq. 5.13:

$$S_i^{\text{fit, corr}} = S_i^{\text{fit}} \exp\left(-\frac{d_i^{\text{snow}}}{\lambda_{\text{snow}} \cos\theta}\right) \tag{5.17}$$

where $S_i^{\text{fit, corr}}$ is the corrected fit expectation value at tank i, d_i^{snow} is the corresponding snow height measured directly or estimated during VEM calibration (see Sec. 3.7), θ is the reconstructed zenith angle of the air shower, and λ_{snow} is the attenuation length of electrons in snow. Thus, the correction performed by Eq. 5.17 does not change measured charge values, but rather the reconstructed shower size S_{125}. Furthermore, the correction of the fit expectation values does not address air showers which did not trigger the detector in regions with more snow which mainly happens at lower energies, where air showers only have few HLC hits, i.e. if $S_{125} \lesssim 0$. The attenuation length is set to $\lambda_{\text{snow}} = 2.1\,\text{m}$ as shown in Ref. [68]. In the case of air shower simulations, snow heights from July 2012 are used for the correction (see Sec. 3.7), whereas in experimental data the correction is performed using snow heights of the particular month the data was recorded.

5.2.4 Overview of the reconstruction

The first guess calculations, performed as described in Sec. 5.2.1, are used as a seed for the log-likelihood minimization of Eq. 5.10. The minimization provides

- The shower size S_{125}
- The measure of the slope of the lateral distribution function β
- The coordinates of the shower core $\vec{x}_{\text{C}}$ and the time t_0 it reaches the surface of IceTop
- The direction $\vec{n}$ of the air shower given by the zenith and azimuth angles θ and ϕ

The minimization is performed in an iterative process. In each step, tanks are excluded if their distance to the hypothetical core is less than 11 m. These tanks can have large signals or even saturate which has negative effects on the resolution of the shower core [66]. Furthermore, the lateral distribution function (Eq. 5.11) sharply rises at small distances which probably results in an imprecise charge expectation.

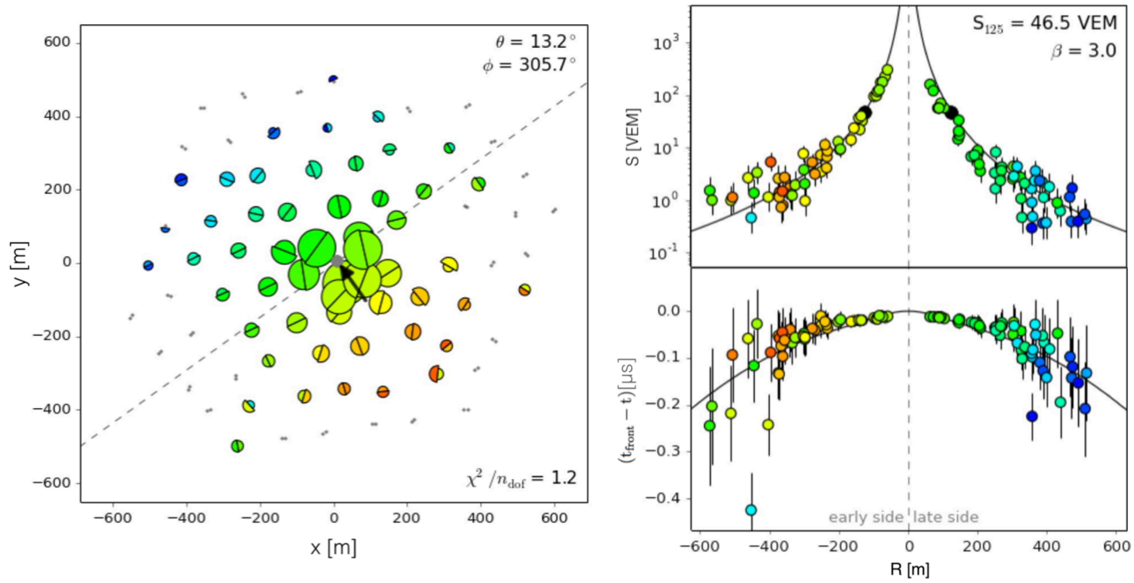

Figure 5.2: Visualization of air shower reconstruction in IceTop. The primary particle has a simulated energy of $E = 35\,\mathrm{PeV}$ and a simulated core position of $(0,\,0,\,z_{\mathrm{ol}})$. In the middle and bottom plots, measured quantities are shown together with the fit functions.
Left: Air showers in IceTop are reconstructed using timing information and charge values measured in tanks which are part of an HLC hit. Each tank is represented by a semicircle where two of them constitute a station. The radius of the semicircles is proportional to the measured charge value in the corresponding tank. Timing information is represented by the color code, where red tanks are hit early, and blue are hit late. This information is used to reconstruct the shower core position (grey dot) and the shower direction (black arrow). The gray dashed line visualizes the direction of the shower front projected on the IceTop plane, where the tanks on the side of the black arrow are hit before the shower core reaches the surface of IceTop. Thus, the expressions *early side* and *late side* refer to the arrival times relative to the shower core.
Top right: The measured charge values as a function of the perpendicular distance to the shower axis R, are parametrized by the lateral distribution function (Eq. 5.11). The values for R are defined as negative for the early side.
Bottom right: The model of the shower front (Eq. 5.14) is fit to measured *time residuals.* These are obtained by subtracting the time the shower core reaches the surface of IceTop, t_{front}, by measured times t at the tanks.

An example of air shower reconstruction, using a simulated air shower with a primary energy of $E = 35\,\mathrm{PeV}$, is shown in Fig. 5.2. Fit errors are not treated by the used reconstruction software.

5.3 Event selection

In this section, various standard IceTop cuts developed by the IceCube collaboration are explained. These cuts improve the quality of simulated and measured air showers which are used for analysis, and are applied on observables, or quantities provided by the air shower reconstruction. Since the cuts improve the quality of the air shower sample, they are called the *quality cuts.* The sample of air showers remaining after all quality cuts is called the *final event sample.* The following cuts are part of the standard event selection in IceTop.

IceTopSTA5 filter

The IceTopSTA5 filter keeps events which have at least 6 HLC hits within 6 μs (Simple Majority Trigger) and at least five standard-stations which record an HLC hit (also see Sec. 3.5). The minimum number of HLC hits is needed to fit a curved shower front (see Sec. 5.2.2).

Fit status

Air shower reconstructions fail in $\sim 0.2\,\%$ of all events because during log-likelihood minimization a minimum cannot be determined explicitly. This can for example happen if the stations participating in the reconstruction constitute particular geometrical patterns, such as a straight line, or a pattern of stations where some are clustered and some are separated from them with at least one layer of stations in between. In all these cases the reconstruction fails because the geometrical requirement for fitting the likelihood functions, namely a uniform distribution of stations around the shower core, is not given (see Sec. 5.2.2). Only events are kept where the reconstruction succeeded.

Shower size S_{125}

In Fig. 3 of Ref. [69], the differential number of reconstructed air showers is shown as function of S_{125}. The contribution to the spectrum becomes small for $\log_{10}(S_{125}) < 0$, i.e. the detector efficiency for detecting events with five or more stations is small. In order to avoid large uncertainties on the detection efficiency, in this work it is required that

$$\log_{10}(S_{125}) \geq 0. \tag{5.18}$$

In Fig. 5.3 the dependance of the true primary energy on the shower size S_{125} is shown. The linear correlation is explained in detail and used for the conversion of S_{125} to primary energy in Ch. 7.

Zenith angle θ

As presented in Tab. 4.1, air showers were simulated with zenith angles between 0° and 65°. In Fig. 5.4, the resolutions of the shower core and direction are depicted as a function of the primary energy, and for two bins in zenith angle.

The resolutions get worse for increasing zenith angles. In order to ensure the high quality of the final event sample, the following restriction on reconstructed zenith angles is used:

$$0.8 \leq \cos\theta \leq 1.0 \Rightarrow 0 \leq \theta \leq 36.87° \tag{5.19}$$

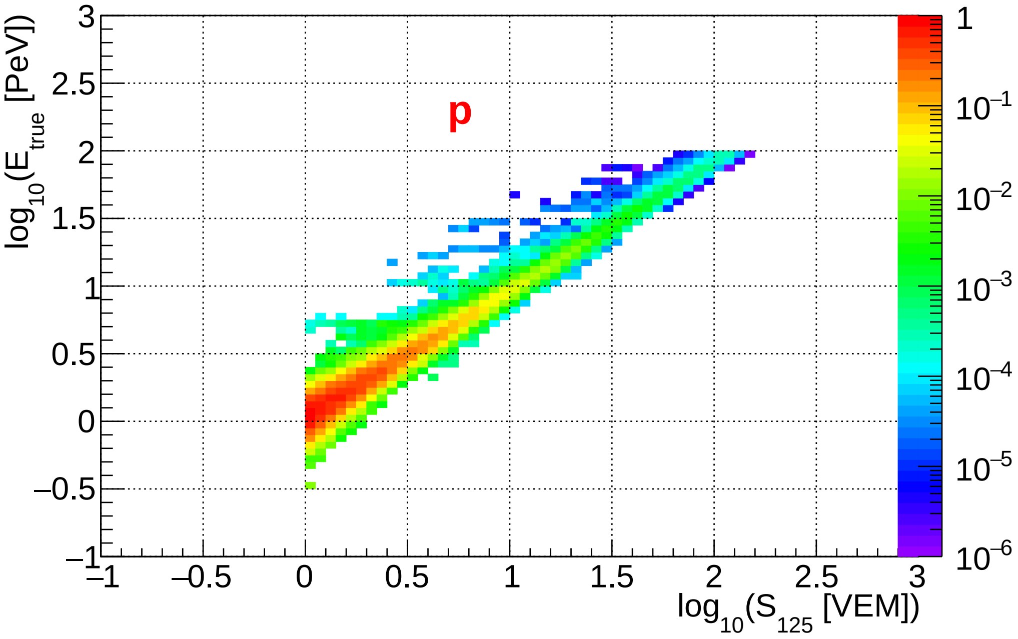

Figure 5.3: Relation between true primary energy and the shower size S_{125} for vertical showers ($\cos\theta \geq 0.95$) initiated by proton primaries. The axis labels are contrary to the convention that the x-axis is labelled with E_true since the latter is going to be calculated as function of the shower size S_{125} in Ch. 7.
The simulation sets used in this work, and presented in Sec. 4.1, comprise primary energies which fulfill $-1 \leq \log_{10}(E_\text{true}\,[\text{PeV}]) \leq 2$. The quantity $\log_{10}(E_\text{true})$ linearly depends on $\log_{10}(S_{125})$ such that the interval $0 \leq \log_{10}(E_\text{true}\,[\text{PeV}]) \leq 2$ roughly corresponds to $0 \leq \log_{10}(S_{125}\,[\text{VEM}]) \leq 2$. Air showers with smaller values for $\log_{10}(S_{125})$, are rejected by requiring $\log_{10}(S_{125}) \geq 0$. Each entry of the histogram is weighted according to the used Cosmic Ray flux model (see Sec. 4.2). The contribution to the mentioned energy range, coming from air showers with primary energies larger than 100 PeV, which are reconstructed with smaller shower size S_{125}, can be neglected since the integral flux of Cosmic Rays above 100 PeV is very small (see Sec 2.3.1).

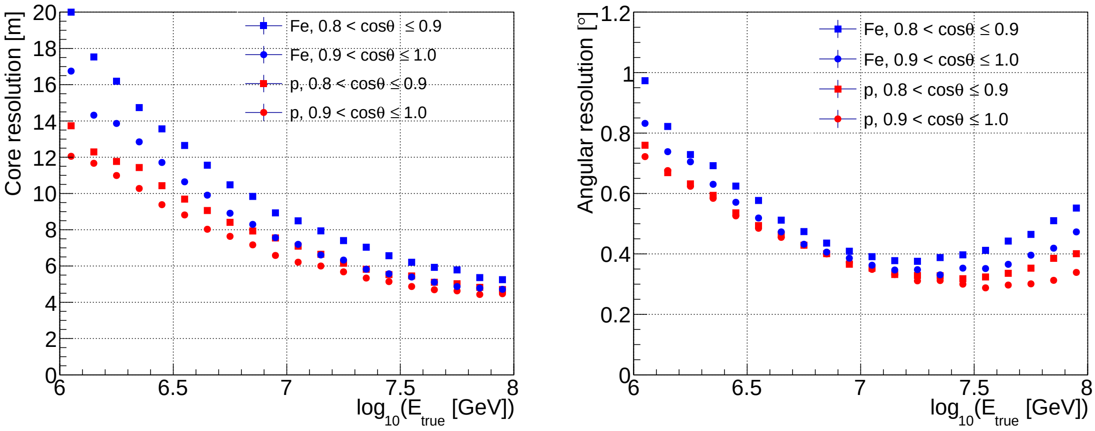

Figure 5.4: Quality of the reconstructed shower core and direction as function of the primary energy, shown for proton and iron initiated air showers, and for two zenith angle bins. For a given bin in primary energy, the resolution is defined as 68% of the cumulative distribution of ΔX, where ΔX is the distance between true and reconstructed shower core, or the angle between true and reconstructed shower direction, respectively. This means that in each energy bin 68% of the participating events have resolutions smaller than the quoted value.
Left / Right: At primary energies of a few PeV, the resolutions worse compared to higher energies. This is because the number of HLC hits used for reconstruction is small, and thus the reconstructed parameters are less accurate. For higher primary energies within the studied energy range, the core resolution gets much better than 10 m, and the angular resolution improves up to $\sim 0.3°$. For higher energies, the resolutions get worse again since stations begin to saturate which complicates the reconstruction. Furthermore, the resolutions in the case of iron showers are worse than for proton showers.

Containment

Air showers with cores landing outside of the IceTop array or at its edge, i.e. next to the outer stations, are difficult to reconstruct since large parts of the air shower are not visible by IceTop. The reconstruction of these air showers provides inaccurate values for the shower core location and the shower size S_{125}. In order to account for this problem, only air showers with reconstructed cores within a so-called *geometric containment* are selected. In this work, containment is described by the parameter i_{scale}. This is a scaling factor which decreases the area of the IceTop array by keeping its shape. For example, air showers with $i_{\text{scale}} = 0$ have reconstructed cores at the center of IceTop, and air showers with $i_{\text{scale}} = 1$ have reconstructed cores at the outer row of tanks. In this work, geometric containment is required by

$$i_{\text{scale}} \leq 0.9 \tag{5.20}$$

The effect of the containment cut is visualized in Fig. 5.5.

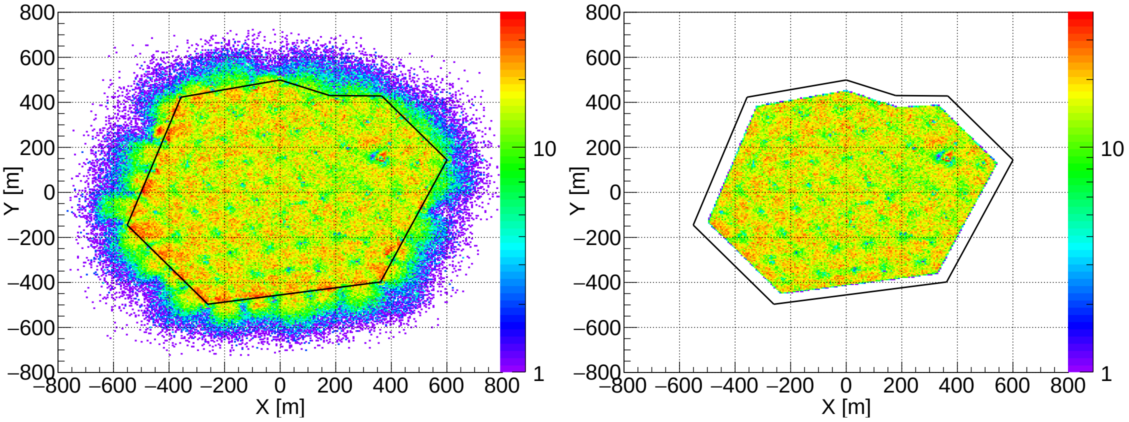

Figure 5.5: Visualization of the containment cut. Reconstructed shower core locations are shown for a small subset of experimental data which passed the cuts on the shower size S_{125} and zenith angle θ. The dips visible throughout the detector area originate from the reconstruction procedure in which tanks are excluded if they are closer to the shower axis than 11 m (see Sec. 5.2.4). The shape of the entire IceTop detector, as given by Fig. 3.4 is represented by the black line.
Left: The locations of reconstructed shower cores are distributed also beyond IceTop.
Right: This is the same as the plot on the left, but with the containment cut applied.

Migration of shower cores

Air showers passing the containment cut can still have true shower cores located outside of the IceTop array, but reconstructed inside. In this case the shower size S_{125} is underestimated since only the part of the air shower is measured where the particle density is smaller than in the shower core. In order to account for these air showers, different cuts are applied:

- The *loudest station* is the station where the tank which measures the highest HLC charge, compared to all hit tanks belonging to the same air shower, is

contained. In this work, the highest charge value is called Q_L. It is expected that in an air shower with a core landing outside of IceTop, Q_L is measured in a station at the edge of IceTop. If this is the case, the corresponding event is rejected.
However, at higher primary energies, air showers remain which produce comparably small signals and where Q_L is not located at the edge. In this case the shower core is far beyond the IceTop array. It is accounted for this by requiring that Q_L exceeds a minimum value. It is required that [72]

$$Q_\mathrm{L} \geq 6\,\mathrm{VEM} \tag{5.21}$$

- For the parameter β, which is a measure of the slope of the logarithmic LDF (Eq. 5.12) at reference distance R_ref, it is required that

  $$2.0 \leq \beta \leq 4.5 \tag{5.22}$$

  On the one hand, air showers with values of β outside the above interval, are badly reconstructed. On the other hand, the lower border of the interval prevents the migration of shower cores. In Fig. 5.2 it can be seen that the lateral distribution of an air shower has a flat curvature at large distances from the shower axis. Thus, air showers with cores landing outside the IceTop array, and reconstructed inside, may have small values of β. These air showers are excluded by Rel. 5.22.

The effect of the cuts mentioned in this subsection is shown in Fig. 5.6.

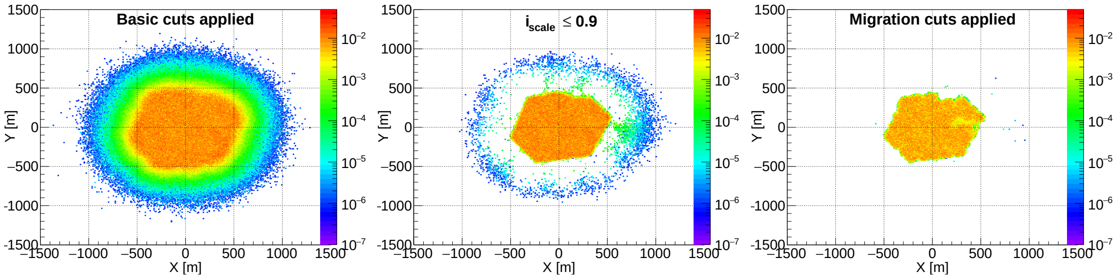

Figure 5.6: Reduction of shower core migration. All plots show the distribution of true shower cores for the case of iron primaries.
Left: Only basic cuts have been applied, i.e. cuts on the shower size S_{125} and zenith angle θ.
Middle: Events which pass the containment cut, have reconstructed shower cores within the containment area. As can be seen here, a fraction of these air showers are falsely reconstructed since they have simulated shower cores outside the containment area.
Right: The cuts related to shower core migration almost eliminate every falsely reconstructed air shower. Explicit rates are provided in the next subsection.

5.4 Comparison between simulation and data

In general, air shower and detector simulations of any kind try to reproduce physical processes and detector performances in oder to model and better understand experimental data. Thus, simulations are required to agree with experimental data.

5.4.1 Passing rates

Passing rates

The air shower passing rates for individual cuts are presented in Tab. 5.1 for simulation and experimental data. The passing rates are normalized to the rate passing the first set of cuts. All cumulative passing rates for simulation and experimental data differ in less than 1 % denoting a reasonable event selection.

	Passing Rate [%]		**Cumulative** [%]	
Condition	**Sim**	**Data**	**Sim**	**Data**
IceTopSTA5 Fit status OK $\log_{10}(S_{125}) \geq 0$ $\theta \leq 36.87°$	100	100	100	100
$i_{\text{scale}} \leq 0.9$	65.92	65.24	65.92	65.24
Q_{L} not on edge	94.10	94.06	62.03	61.37
$Q_{\text{L}} > 6\,\text{VEM}$	99.15	99.04	61.50	60.78
$2.0 \leq \beta \leq 4.5$	99.49	99.18	61.19	60.28

Table 5.1: Relative and cumulative passing rates of individual cuts calculated for simulation and experimental data.

5.4.2 Observables

Since in this work experimental data is composed of nuclei of different mass, and the used simulations are composed of hydrogen and iron nuclei, it is expected that the corresponding distributions of reconstructed observables enclose those for experimental data. In Fig. 5.7, the distributions of the reconstructed parameters are compared for simulation and experimental data. They show a good agreement.

5.5 Enhancement of events

There are two issues regarding IceTop hits which also have to be accounted for. Firstly, the measured charge value of SLC hits can be inaccurate when there is a change in the PMT baseline. Secondly, measured HLC and SLC hits have to be selected according to their time compatibility with expected signals given by the air shower reconstruction.

5.5.1 Charge recalibration

As explained in Sec. 3.4, waveforms are stored only in the case of HLC hits. For SLC hits, only the integrated ATWD charge value and a time stamp are available. The

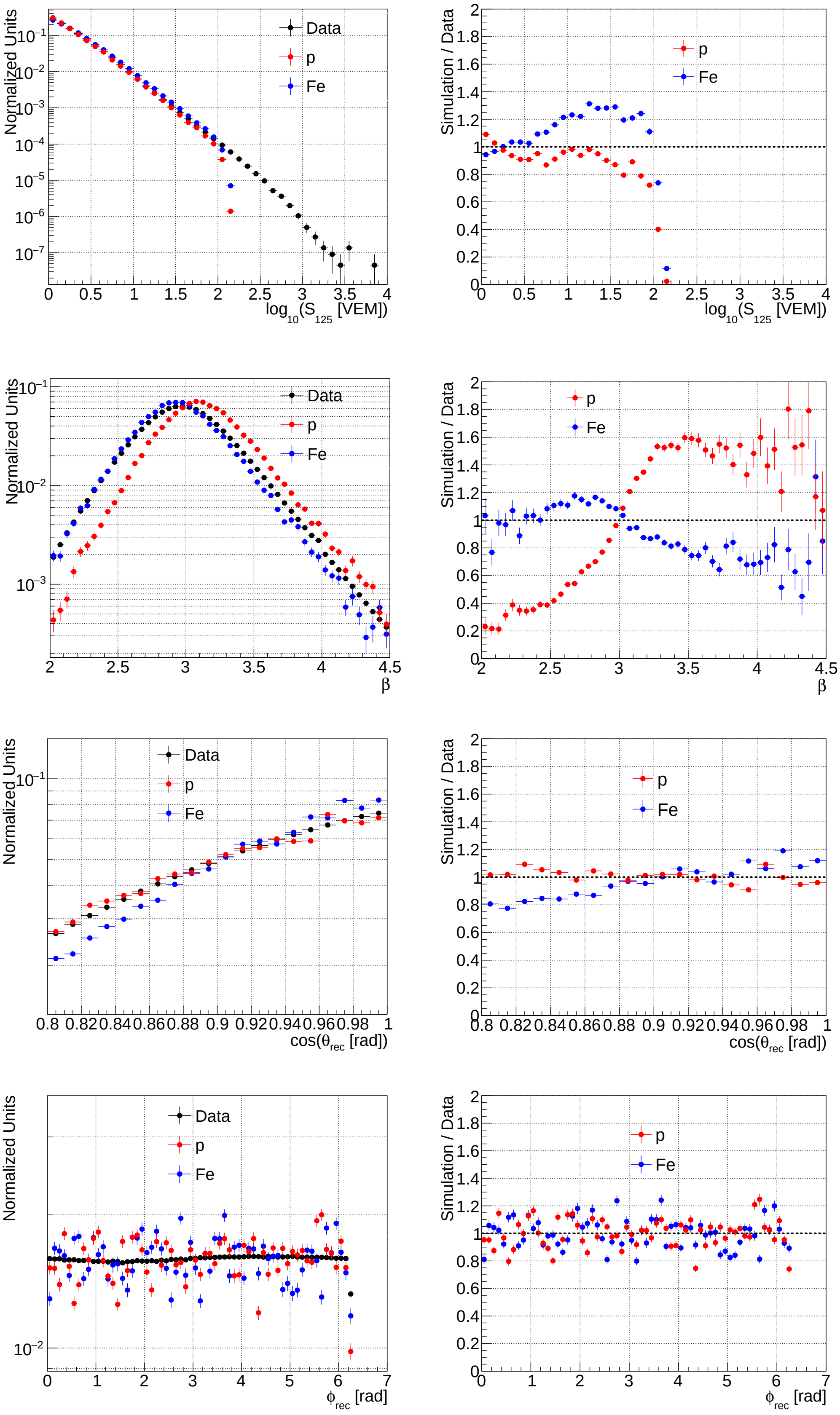

Figure 5.7: Comparison of reconstructed observables between simulation and experimental data. The variables S_{125}, β, zenith angle θ and azimuth angle ϕ are shown line-by-line.

calculation of the SLC hit charge relies on known values for the gain and baseline of the PMT. If the baseline changes, e.g. due to PMT droop [40], the extracted tank charge value is not precise and has to be corrected. For this purpose, the IceCube collaboration developed a particular project which performs a recalibration of SLC hit charge values [71]. The recalibration is done in two steps:

1. For each ATWD in each DOM, a charge value from a measured HLC hit, is extracted. Additionally, this HLC hit is mimicked to be an SLC hit. This allows the extraction of the integrated ATWD charge. The comparison of both charge values yields a linear correlation given by

   $$S_{\mathrm{HLC}} = c_0 + c_1 S_{\mathrm{SLC,\,mim}} \tag{5.23}$$

   where S_{HLC} is the calibrated charge value of the HLC hit, $S_{\mathrm{SLC,\,mim}}$ is the uncalibrated charge value of the HLC hit mimicked to be an SLC hit, and c_0, c_1 are calibration constants which are obtained by a fit and differ for each ATWD in each DOM.

2. The values for c_0 and c_1 for each ATWD in each DOM are used for the recalibration of SLC hit charge values

   $$S_{\mathrm{SLC,\,cal}} = c_0 + c_1 S_{\mathrm{SLC,\,uncal}} \tag{5.24}$$

   where $S_{\mathrm{SLC,\,cal}}$ and $S_{\mathrm{SLC,\,uncal}}$ refer to the calibrated and uncalibrated SLC hit charge values, respectively.

In this work, an existing table of values for c_0 and c_1 is used for recalibration.

5.5.2 Time residuals

For each hit in IceTop, the time of the launch, t_{launch}, is compared with the expectation of the time according to the reconstructed shower front, t_{reco} (see Sec. 5.2.2). In Fig. 5.8, the distribution of time residuals is shown for experimental data.
In this work, hits with time residuals within the *signal window*, given by the interval [-200 ns, 1800 ns], are kept for further analysis.

Hits with smaller time residuals constitute background since they chronologically occur before the shower front arrives. These hits are due to so-called *uncorrelated particles* which are not part of the particular air shower but nevertheless produce a hit in the detector. The used simulation software does not cover background given by uncorrelated particles, thus this type of background can only be observed in experimental data. Background has to be subtracted from experimental data measured in the signal window in order to compare it with simulations.
In this work, background is estimated using the plateau which is earlier than the signal window, comprising residuals within the so-called *background window* [-8000 ns, -2000 ns]. For a reliable background estimation, the background window is three

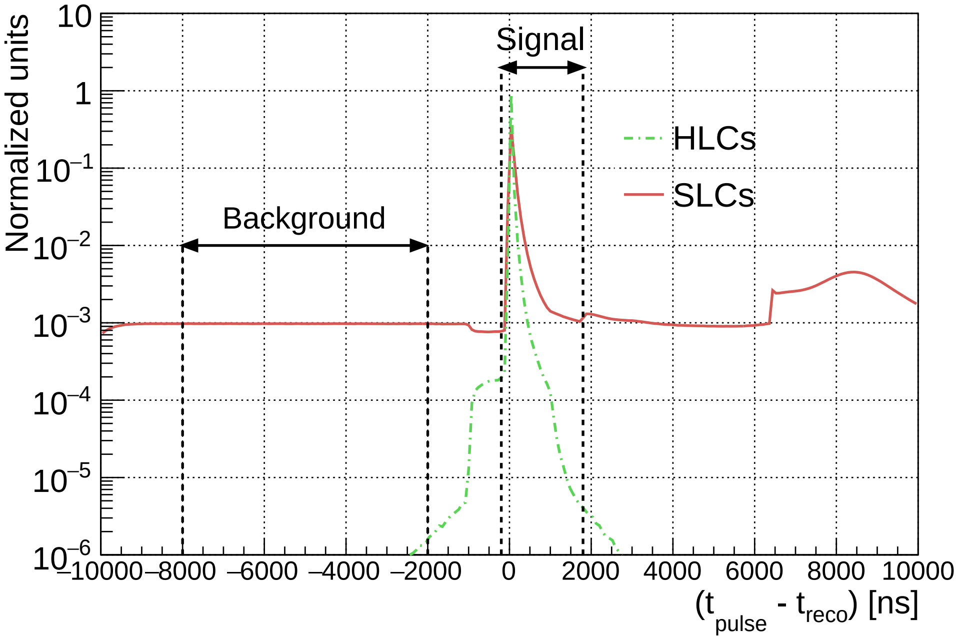

Figure 5.8: Time Residuals in experimental data, shown for HLC and SLC hits. Background due to uncorrelated particles is visible only in the case of SLC hits since single particles do not trigger HLC hits. The structure to the right of the signal window is due to effects inherent to PMTs [46].

times larger than the signal window, but finally rescaled. For each upcoming distribution showing experimental data, background due to uncorrelated particles is subtracted.

Experimental data and simulations are compared for time residuals within the signal window in Fig. 5.9. Within the interval [0 ns, 400 ns], the distributions agree within a few percent for both SLC and HLC hits. For SLC hits, this corresponds to roughly 96.4 % of the hits in experimental data, and 94.2 % of the hits in simulation, and 98.8 % and 99.1 % in the case of HLC hits, respectively. For large time residuals, e.g. $t_{\text{pulse}} - t_{\text{reco}} \gtrsim 400\,\text{ns}$ especially in the case of SLC hits, there are huge differences between simulation and experimental data. Nevertheless, if there is an effect due to this inconsistency, it is expected to be marginal due to the above numbers.

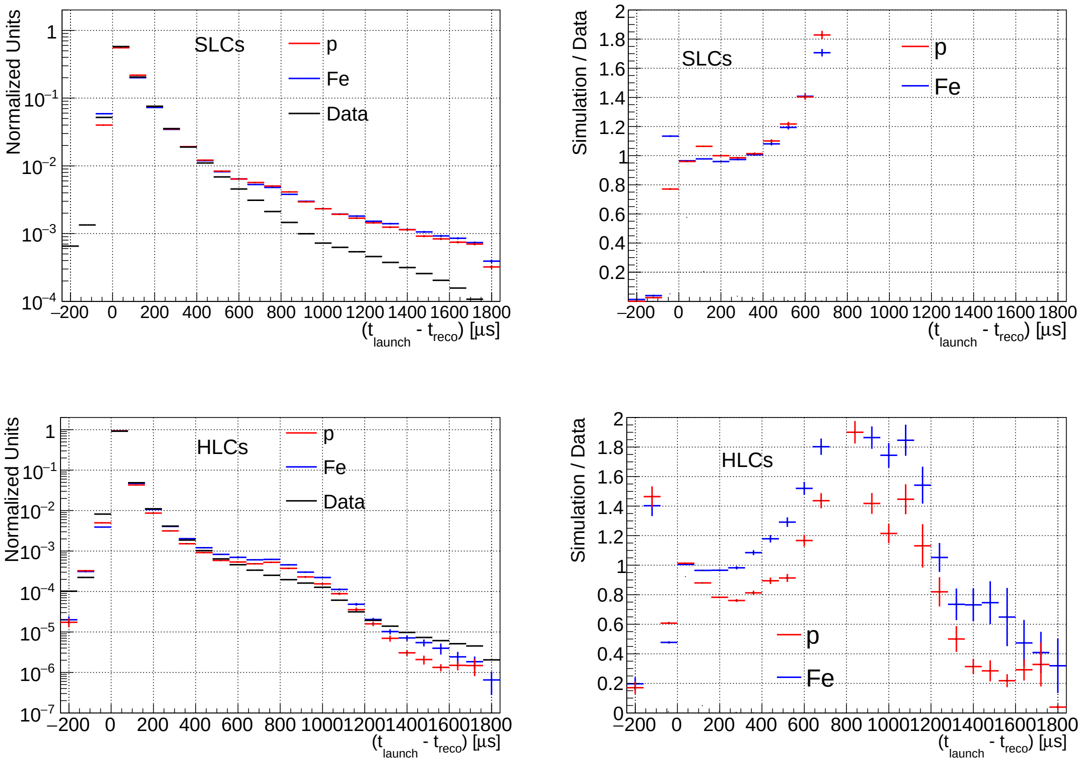

Figure 5.9: Comparison of time residuals for all hits within experimental data and simulation.

Chapter 6

Muon number estimation

Hits in IceTop, introduced in Sec. 3.4, are caused by individual particles such as muons, electrons, and charged hadrons, or by a mixture of them. As was shown in Fig. 3.8, HLC hits are predominantly close to the shower axis, while the contribution of SLC hits to the absolute amount of hits becomes significant in the outer region of an air shower. In this chapter it is shown that the distance at which the amount of HLC hits decreases in favor of SLC hits, is dependent on the primary energy.

Single SLC hits are caused by few particles only, thus it seems plausible to use information from SLC hits in order to calculate muon number densities. However, it is shown that HLC hits provide a significant contribution and therefore can not be neglected. Instead, it is important that the hits exhibit a muonic signal which is not superimposed by too much background. In this chapter, HLC and SLC hits are grouped according to the particle types crossing the corresponding tanks. Based on this knowledge, cuts on tank variables, which are the value of the measured charge, and distance to the shower axis, are developed. It is shown, that these cuts separate hits of muonic origin. Furthermore, these cuts are the basis for the construction of two muon number estimators which are studied in detail.

6.1 Motivation

The distributions of tank charge values produced by simulated, individual particle types vertically traversing an IceTop tank, are investigated in Ref. [75]. For electrons and muons, the distributions are depicted in Fig. 6.1.

The different shapes of the tank charge distributions arise from different energy loss mechanisms and penetration depths of particles in matter. In general, the average energy loss of relativistic heavy, charged particles with energies in the order of GeV, such as muons traversing IceTop tanks (see Fig. 6.2), is dominated by atomic excitation and ionization, and described by the Bethe-Bloch equation [33]. For electrons the Bethe-Bloch equation has to be modified since the incident electrons can not be distinguished from the atomic electrons when leaving the material. Furthermore, the development of electromagnetic cascades, a sequence of bremsstrahlung and pair production, has to be accounted for.

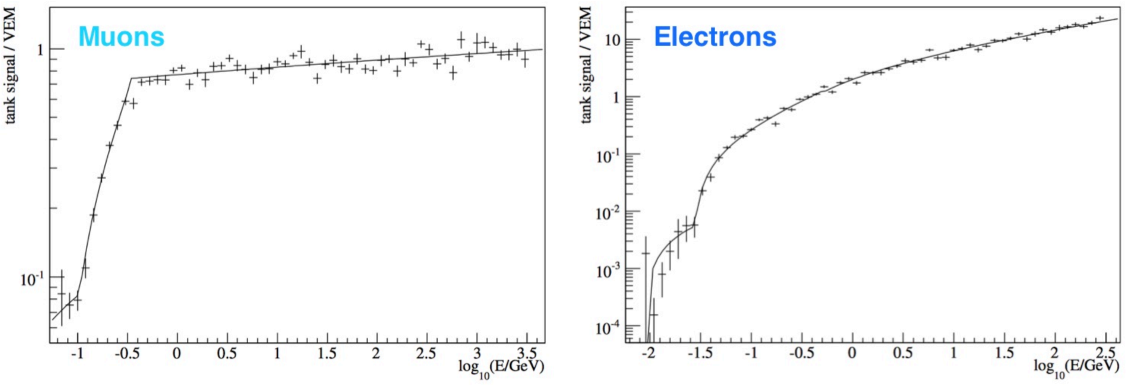

Figure 6.1: Average light yield of electrons and muons vertically traversing an IceTop tank as function of their kinetic energy [75].
Left: The light yield of muons with energies above $\sim 0.3\,\mathrm{GeV}$ is roughly constant at $\leq 1\,\mathrm{VEM}$.
Right: In the case of electrons, the light yield increases with energy.

Studies show, that the amount of energy loss due to bremsstrahlung increases with the particle energy [33], and that electromagnetic cascades provide a large fraction of the detected Cherenkov photons [76].

In the IceTop detector simulation performed in this work, relevant information is kept for each particle which produced any signal in the tanks. For these electrons and muons, their average kinetic energy at observation level as function of the distance to the shower axis is presented in Fig. 6.2. In the case of muons, the outer shower region is dominated by muons of comparatively low energy, which is around $1\,\mathrm{GeV} - 3\,\mathrm{GeV}$. According to the left plot of Fig. 6.1, this corresponds to a measured charge value which is close to $1\,\mathrm{VEM}$. The average energy of the electromagnetic shower component (electrons, positrons, and photons) is between $0.01\,\mathrm{GeV}$ and $0.1\,\mathrm{GeV}$ which corresponds to a measured charge value of significantly less than $1\,\mathrm{VEM}$, where in turn the measurement of $1\,\mathrm{VEM}$ requires a kinetic energy of approximately $10^{-0.4}\,\mathrm{GeV} = 0.4\,\mathrm{GeV}$, according to the right plot of Fig. 6.1.
The different values of the tank charges for both the muonic and electromagnetic shower components can for example be visualized using a 2-dim contour plot in which each of the measured tank charge values with its corresponding distance to the shower axis is included. Such a histogram is presented in Fig. 6.3. At large distances to the shower axis, it can be assumed that the charge distribution is a superposition of at least two distributions, which can be classified as muonic at approximately $1\,\mathrm{VEM}$ and electromagnetic at roughly $0.2\,\mathrm{VEM}$ according to the descriptions above.

In conclusion, the brightness of IceTop hits originating from different particle types offers a possibility to separate them.

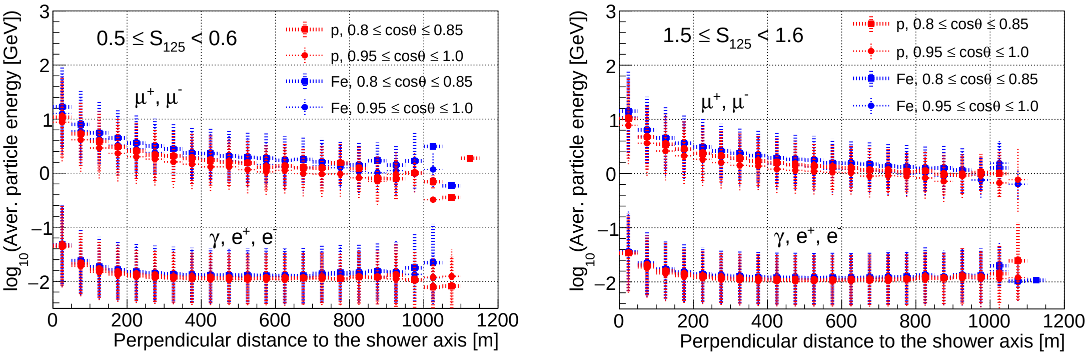

Figure 6.2: Average kinetic particle energy for the muonic and electromagnetic shower components at observation level, as function of the perpendicular distance to the shower axis, shown for both primaries, and two bins in zenith angle and shower size S_{125}. The average kinetic particle energy of the shower components differ in more than two orders of magnitude and show little dependance on primary mass and zenith angle.

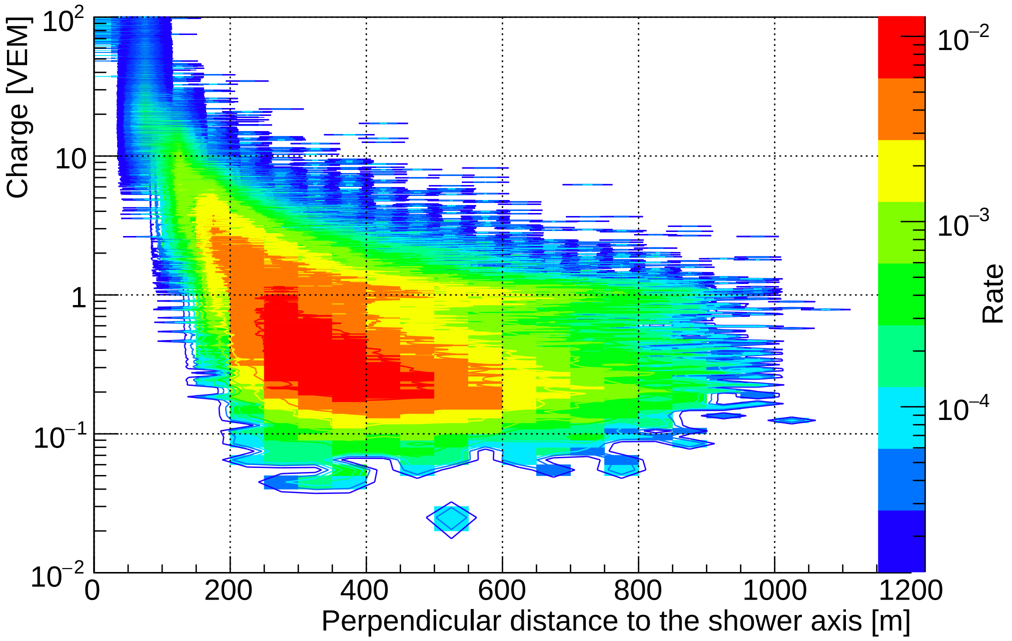

Figure 6.3: Visualization of electromagnetic and muonic distributions in a 2-dim contour plot, for vertical, iron induced air showers with $1.0 \leq S_{125} \leq 1.1$.

6.2 Separation of signal

The consecutive use of air shower and detector simulations described in Ch. 4 technically allows to extract all particles which are responsible for triggering IceTop hits. Thus, each hit can be classified as muonic, electromagnetic, hadronic or a mixture of them. The contributions of hits caused by various particle combinations, to all hits, are presented in Fig. 6.4 for a subsample of simulations.

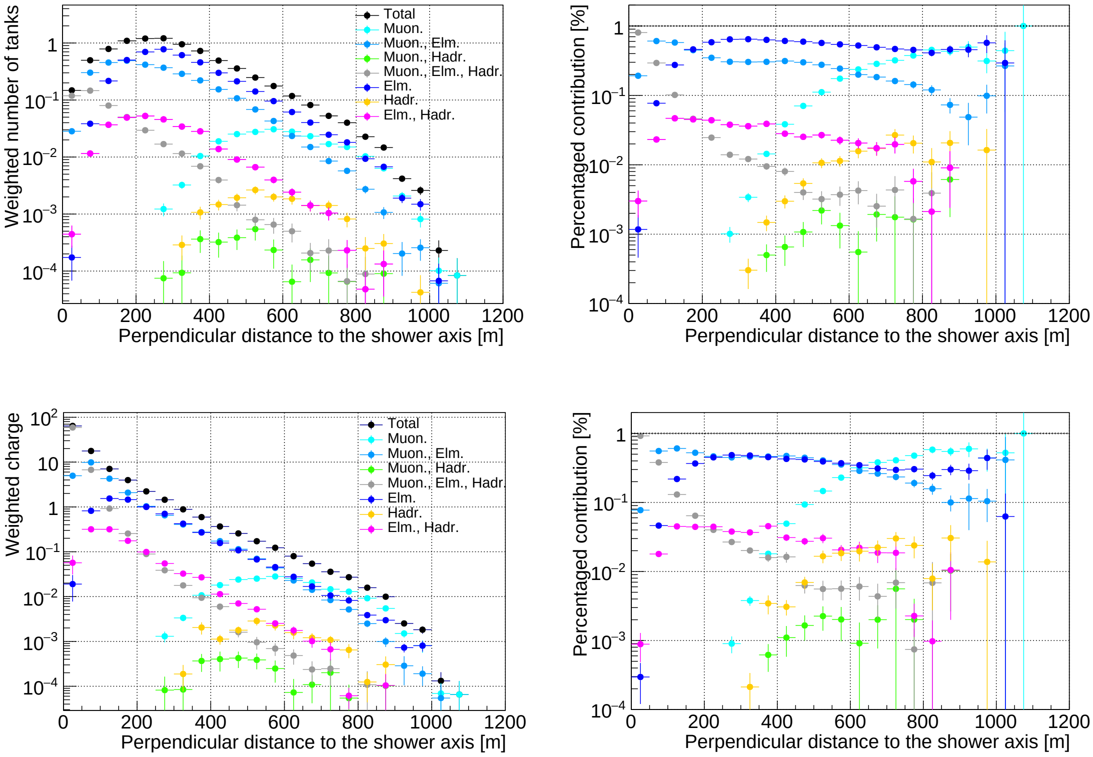

Figure 6.4: Particle combinations which cause hits in IceTop, shown for vertical, iron induced air showers with $1.0 \leq \log_{10}(S_{125}) \leq 1.1$.
Top and bottom row: Hits are caused by individual particle types, i.e. muonic (μ^+, μ^-), electromagnetic (e^+, e^-, γ), and hadronic particles (p, n, $\pi^{+/-/0}, ...$), or superpositions. At large distances, hits of electromagnetic and muonic origin dominate.

From now on, hits which are caused by at least one muon are treated as *signal*, otherwise as *background*. This classification is used to construct two cuts which select hits with charges above a particular value, and above a particular distance, where these values are dependent on primary energy and zenith angle.

6.2.1 Properties of signal and background

In order to study signal and background distributions, and to get a guess for adequate cuts to emphasize hits with muonic contribution, a parameter called the *average charge* is defined. This is the total charge value measured in IceTop hits within a ring R_{ij} perpendicular to the shower axis divided by all hits within this ring. Since the average charge parameter is used as preparation for a cut on the value of the hit charge, and this cut only affects tanks which are hit, tanks without a hit are not considered in the calculation. The idea is depicted in Fig. 6.5.

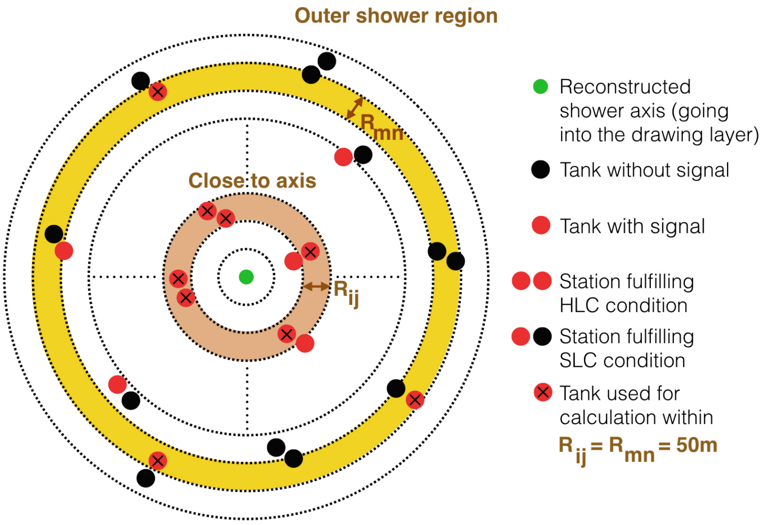
Outer shower region
R_mn
Close to axis
R_ij
Reconstructed shower axis (going into the drawing layer)
Tank without signal
Tank with signal
Station fulfilling HLC condition
Station fulfilling SLC condition
Tank used for calculation within
R_ij = R_mn = 50m

Figure 6.5: Sketch of two exemplary rings in which the average charge is calculated. Each tank is projected on a plane perpendicular to the shower axis. The plane is sliced into rings of 50 m thickness around the axis. A tank with distance R from its center to the shower axis is within a ring R_{ij} if $R_i \leq R < R_j$. The dashed lines and circles symbolize that the entire detector can be subdivided into rings. Tanks, which are hit, are presented as red dots, and all other tanks as black dots. The average charge is calculated by dividing the weighted charge of all hit tanks in a ring by the weighted number of hit tanks in this ring. The calculation is independent from zenith angle θ.

The average charge $\langle Q(R_{ij})\rangle_{s/b}$ measured in the ring R_{ij} produced by signal s or background b can be calculated by

$$\langle Q(R_{ij})\rangle_{s/b} = \frac{\sum_{t=1}^{T_{s/b}(R_{ij})} w_t \cdot Q_{\mathrm{t}}}{\sum_{t=1}^{T_{s/b}(R_{ij})} w_t} \tag{6.1}$$

where $T_{s/b}(R_{ij})$ is the number of hit tanks produced by signal or background with tank centers located within the ring R_{ij}, w_t is the weight of each tank according to the Cosmic Ray flux model (see Sec. 4.2), and Q_T is the charge value measured at tank T. The particular values obtained in the numerator and denominator of Eq. 6.1, for different R_{ij} as a function of R_{ij}, represent the lateral distribution of total weighted measured charge, and the weighted number of hit tanks within the ring R_{ij}. From now on, it is defined that $L_{\mathrm{ch}}^{\mathrm{true}}(R_{ij}) := \sum_{t=1}^{T_{s/b}(R_{ij})} w_t \cdot Q_{\mathrm{t}}$, and $L_{\mathrm{h}}^{\mathrm{true}}(R_{ij}) := \sum_{t=1}^{T_{s/b}(R_{ij})} w_t$, where L represents lateral distributions derived from simulations.

Exemplary distributions are shown in Fig. 6.6. for both signal and background, both HLC and SLC hits, and together with the average charge $\langle Q(R_{ij}) \rangle_{s/b}$.
Close to the shower axis, the number of high energetic particles is huge, thus the lateral distributions of total weighted hit charge, and weighted number of hits, are dominated by HLC hits containing high charge values. With increasing distances, the kinetic energy and density of the particles decrease, leading to a drop of the HLC hit distributions in favor of the SLC hit distributions. The shape of the distributions, especially the maxima in the case of SLC hits, can be explained by the ring-like binning of the tanks (see Fig. 6.5). Close to the shower axis, the rings contain few tanks, thus the distributions start with low values. With higher distance to the shower axis, i.e. with increasing indices i and j of R_{ij}, the number of tanks in a ring gets larger. The maxima in the distributions occur since the number of hit tanks decreases for large values of i and j. As a consequence, the ratio of HLC to SLC hits always reflects the particle density of air showers within a particular energy range.

In Fig. 6.6, the two columns represent two bins in shower size S_{125} which roughly differ by one order of magnitude in primary energy. With higher primary energies, thus higher values of S_{125}, the particle density and the spread of an air shower increases. In the right column, corresponding to air showers of higher primary energy than in the left column, this effect can be observed. The tails of the HLC distributions extend to larger distances, and the SLC distributions start to be recorded at larger distances than in the left column, i.e. the head of the distribution is shifted to larger distances. Thus, for the calculation of a muon number density at particular distances from the shower axis, HLC hits give a significant contribution and can not be neglected.

The difference of signal and background hits becomes visible when the average charge distributions, $\langle Q(R_{ij}) \rangle_{s/b}$, are drawn. As can be seen, above particular distances the average charge of signal is constant at roughly 1 VEM, while the average charge of background is significantly smaller. The average charge becomes nearly constant when less particles are participating in the signal production, i.e. when the particle density is low. Therefore, the minimum distance above which the average charge roughly remains constant, also depends on primary energy. This can be seen when the average charge distributions of the left and the right column are compared.

These observations are the basis for the introduction of two cuts which increase hits of muonic origin within all hits. They are defined as

- a cut on the hit charge $Q \geq Q_{\mathrm{min}}$, and
- a cut on the perpendicular distance of a hit to the shower axis, $R \geq R_{\mathrm{min}}$.

For proton and iron showers, different sets of cuts are produced. Since in experimental data the primary particle of an air shower is not known, both sets of cuts are applied on experimental data. This introduces a systematic uncertainty on the final result (see Sec. 6.3.3 for further information and Sec. 7.8 for the final result).

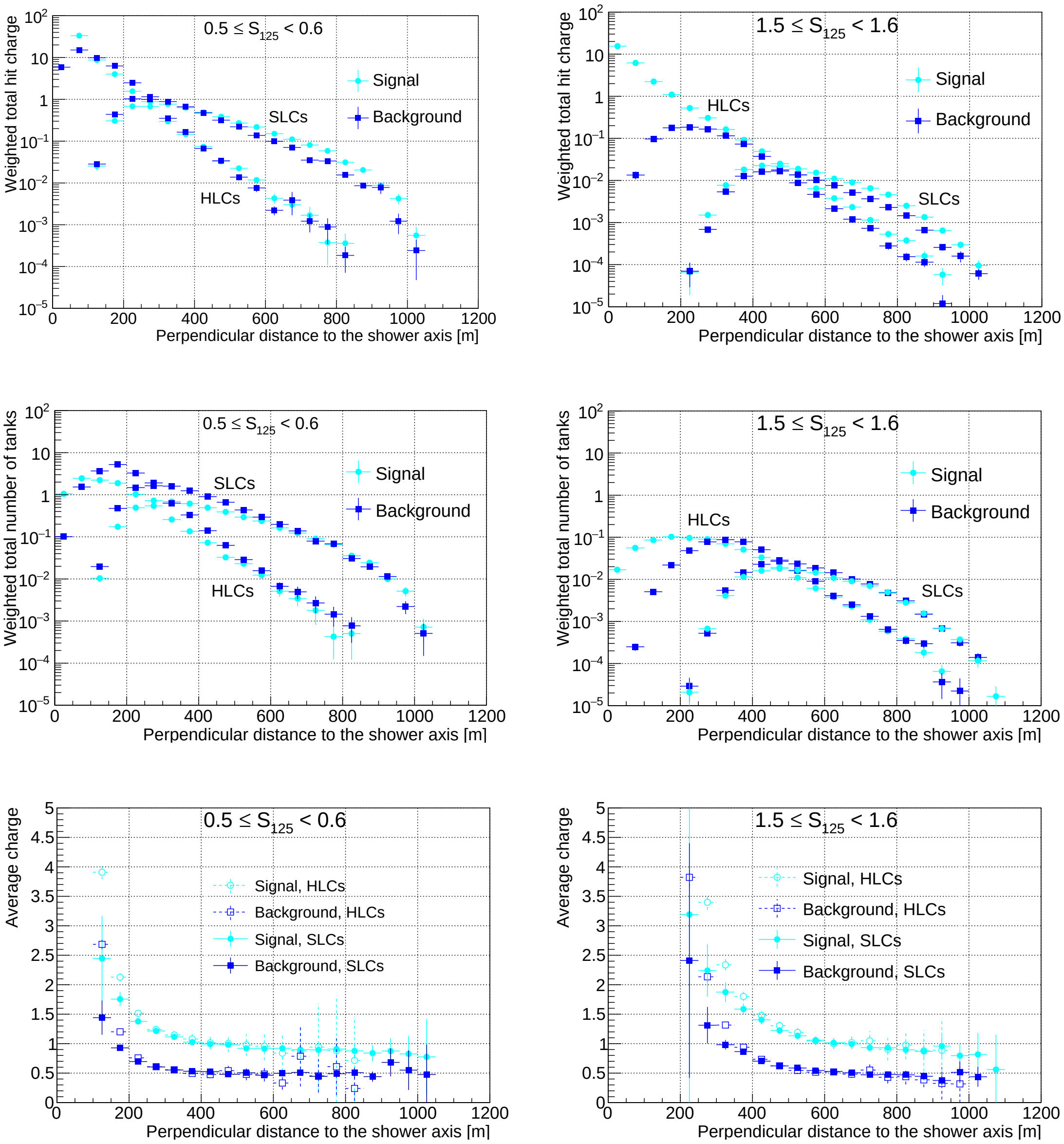

Figure 6.6: Total weighted hit charge, total weighted number of hit tanks, and the average charge as a function of the perpendicular distance to the shower axis for both signal and background, for both HLC and SLC hits, detected in vertical air showers initiated by iron primaries, shown for two bins in shower size S_{125} (see columns).
Top row: In each distance bin, the weighted measured hit charge is summed up.
Middle row: In each distance bin, the weighted number of hit tanks is accumulated.
Bottom row: The average charge is calculated according to Eq. 6.1. In the case of HLC hits the error bars of the average charge are large. Thus, the decreasing number of HLC hits at larger distances can be observed.

6.2.2 Cut on the hit charge

The value for the cut on the measured hit charge, Q_{min}, is obtained by using both the charge distributions of signal and background. Exemplary distributions are shown in Fig. 6.7. The charge distribution of signal deviates from a sharp peak at a value of 1 VEM. This is due to muons traversing the ice in a tank with an inclined zenith angle, e.g. edge-clipping muons, and therefore with a smaller or longer track lengths compared to vertically arriving muons. Additionally, for both charge distributions of signal and background, there is a difference in the peak position which is expected according to the discussions in the previous sections.
A cut on hit charge value is chosen such that most of the background is discarded while keeping most of the signal. This is achieved by choosing the particular charge value where the differences of the cumulative charge distributions is at maximum (see Fig. 6.7). This procedure results in different values of Q_{min} for each bin in perpendicular distance (see Fig. 6.7). A simple exponential function given by $f(x) = e^{p_0 + p_1 \cdot x} + p_2$, with fit parameters p_0, p_1, and p_2, is used to model Q_{min} as a function of perpendicular distance. The fit is performed in an interval between 0 m and 800 m. Higher distances are excluded from the fit due to low statistics. Such an exponential fit function is determined for both proton and iron primaries, and for each bin in zenith angle, and shower size S_{125}.

6.2.3 Cut on the hit distance

A cut on the perpendicular distance R of a hit suppresses HLC hits with large signals close to the shower axis, and therefore hits where the muonic signal is superimposed by background. Furthermore, it neglects regions of the detector where the fraction of hit tanks to all tanks in a particular ring around the shower axis is 1 (see Fig. 6.12 for an example). This issue becomes important in Ch. 7.
In this work, the value R_{min} is chosen such that after the application of the cut on hit charge, as described in Sec. 6.2.2, the relative contribution of HLC hits to the total number of hits is less than 50 %, i.e. for $R \geq R_{\mathrm{min}}$ SLC hits dominate. The region of an air shower where $R \geq R_{\mathrm{min}}$ is called the *shower periphery*. According to this definition, the shower periphery depends on primary energy, and thus on the shower size S_{125} (see Sec. 6.2.1). The existence of R_{min} is ensured since the particle density decreases with increasing perpendicular distance.
For both proton and iron primaries, and for each bin in zenith angle, the distribution $R_{\mathrm{min}}\,(S_{125})$ is created, as exemplary shown in Fig. 6.8. For analysis, the used value for R_{min} at given S_{125} and zenith angle bin, is the average of the R_{min} values for the case of proton and iron primaries.

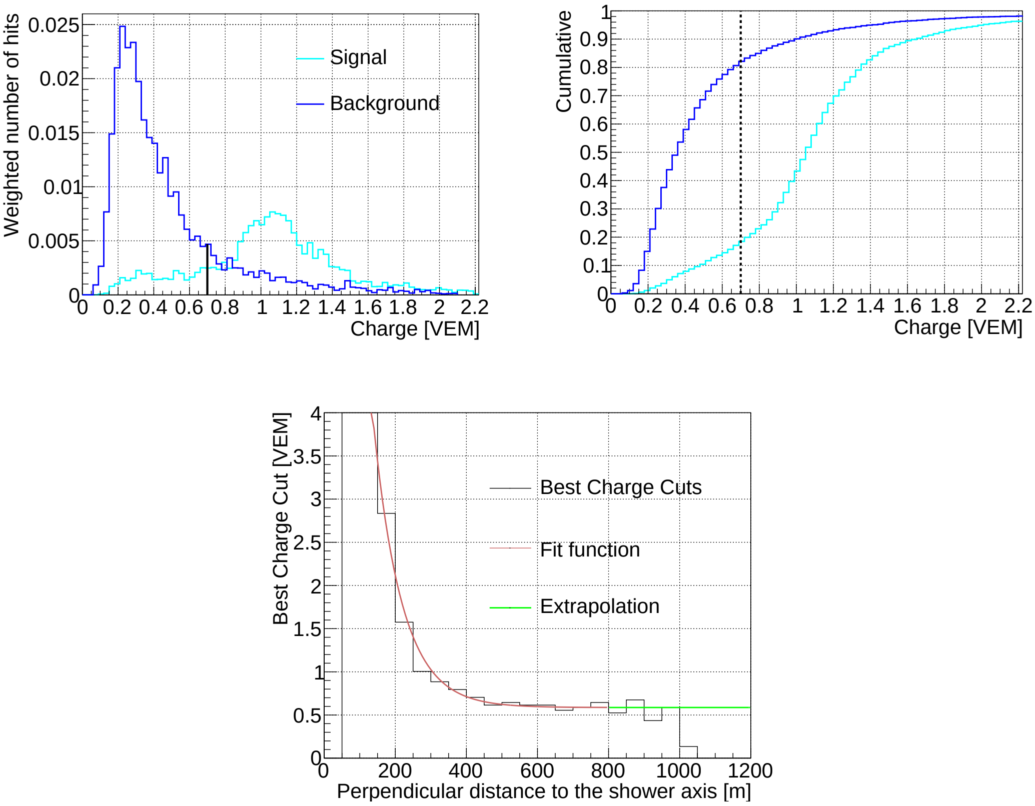

Figure 6.7: Determination of Q_{min} for vertical iron showers, and $1.0 \leq \log_{10}(S_{125}) \leq 1.1$.
Upper row: The charge distributions of signal and background, and their cumulative distributions are shown for distances of 400 m−450 m. The shape of the distributions can be compared to Fig. 3.9. In this example, the best value for Q_{min} is at 0.7 VEM, as denoted by the black lines in the upper plots.
Bottom plot: An exponential function is fit to the values of Q_{min} at all perpendicular distances. The bin at a perpendicular distance of 400 m is represented by the upper plots.

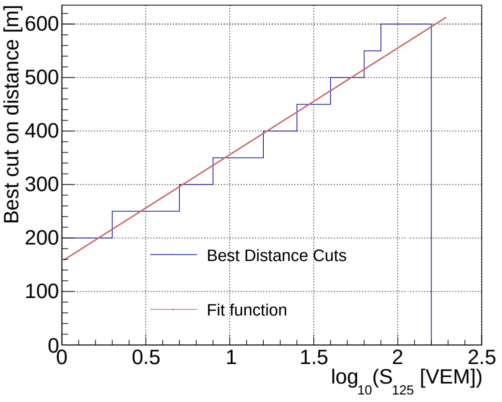

Figure 6.8: Example of $R_{\mathrm{min}}\,(S_{125})$, and an appropriate linear fit for vertical iron showers.

6.2.4 Discussion

The application of both cuts on hit charge, and perpendicular distance to the shower axis, as defined in the last two sections, lead to different effects when the relative contributions of signal and background, either to the total weighted measured charge in IceTop hits, or to their weighted number, are considered. The relative contributions of signal before and after the application of both cuts are depicted in Fig. 6.9.

If no cut is applied, at small distances the contribution of signal decreases from almost 100 % to its lowest value. This is consistent with the shape of the HLC distributions shown in the two upper right plots of Fig. 6.6, where the distributions of signal clearly are above background only for small distances. Furthermore, the high contribution of signal is due to the existence of various particle types close to the shower axis. This can for example be observed by the gray and azur markers in Fig. 6.4.

The effect of the cut on the hit distance, R_{min}, can be seen in Fig. 6.9, when the contributions of signal after the application of both cuts, are compared. The two columns of plots denote different bins in the shower size S_{125}, and thus in primary energy. As explained in section 6.2.3, the cut on distance depends on S_{125}, thus in Fig. 6.9 the distributions start at different distance values.

The effect of the cut on hit charge differs depending on which variable is considered, e.g. total weighted hit charge or weighted number of hits. In the first case, the contribution of signal increases by roughly $15\,\% - 20\,\%$.

In the second case, the contribution increases by about $30\,\% - 40\,\%$. This difference can be explained by the charge distributions of signal and background (see the upper left plot of Fig. 6.7). Above Q_{min}, the total number of remaining hits is small, but they carry higher charge values than Q_{min}. Thus, the summed charge for hits with $\mathrm{Q} > \mathrm{Q}_{\mathrm{min}}$ provides a high contribution to the total charge.

Using Fig. 6.9, one can also derive a few aspects regarding the impact of the cuts on primary mass. Firstly, the contribution of signal in the case of iron initiated air showers is roughly $5\,\% - 10\,\%$ higher than for proton initiated air showers. This can be explained by the fact that iron initiated air showers, compared to proton or gamma initiated air showers, on the one hand have less particles at ground level, but on the other hand can have a higher fraction of muonic and electromagnetic particles at kinetic energies of around 1 GeV (see Ref.[75]). Since in this work the term *signal* is defined as a hit which contains at least one muon, the higher fraction of signal in the case of iron initiated air showers, as shown in Fig. 6.9, is reasonable. Secondly, by Fig. 6.9 it is demonstrated that the mentioned difference of $5\,\% - 10\,\%$ is not affected by the cut on hit charge. Thus, this cut has only little dependance on primary mass (also see appendix).

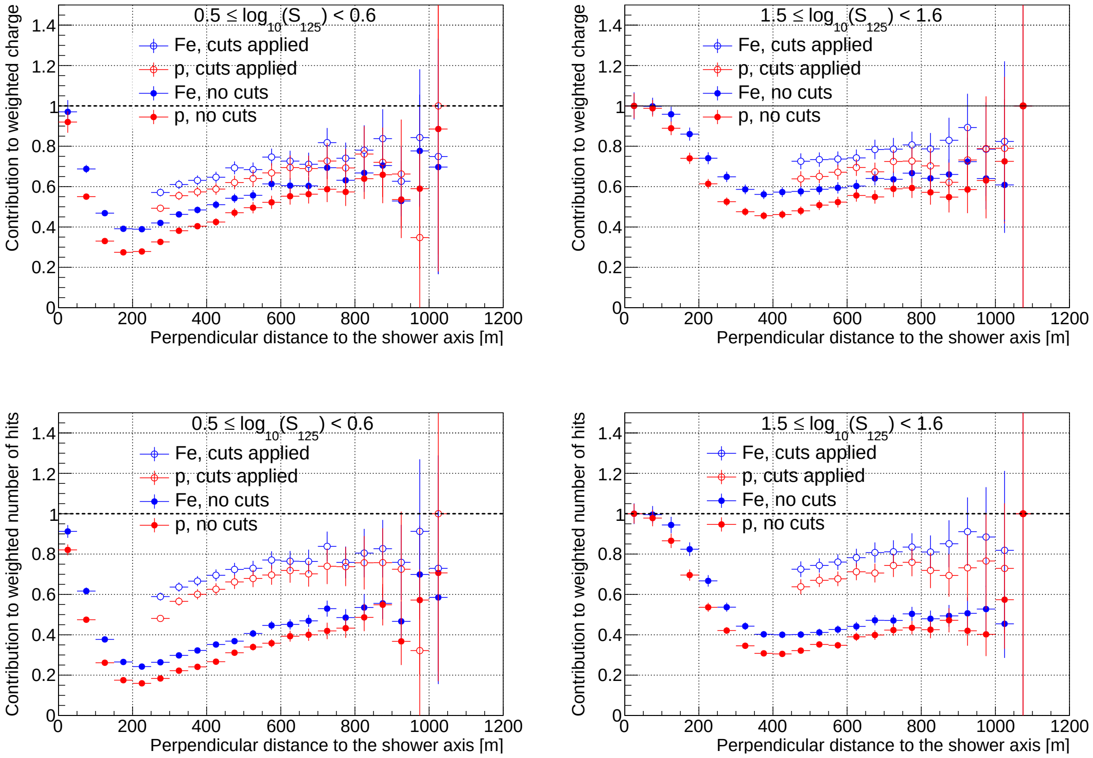

Figure 6.9: Percentaged contribution ($1 \equiv 100\%$) of signal to all hits, shown for vertical air showers induced by proton and iron primaries, and two bins in shower size S_{125}, before and after applying cuts on hit charge and distance.

6.2.5 Consistency check

As explained in Ch. 5, a particular air shower is reconstructed by an algorithm which provides information about the shower axis. This information include specific values for the shower core, the shower size S_{125}, the slope β, the zenith angle θ, and the azimuth angle ϕ. With these parameters, the air shower is clearly defined. Furthermore, the distance of each tank to the shower axis is fixed. For a single air shower with reconstructed values of S_{125}, and β, the lateral distribution function of HLC hits in IceTop is given by (Eq. 5.11)

$$S_{\mathrm{exp}}(R) = S_{125}\left(\frac{R}{125\,\mathrm{m}}\right)^{-\beta-\kappa\log_{10}(R/125\,\mathrm{m})} \tag{6.2}$$

with the parameter R, denoting the perpendicular distance to the shower axis. This function is a fit through the measured amount of charge of HLC hits (see Fig. 5.2 for an example). Thus, by using Eq. 6.2, the expected charge value at particular distances R, $S_{\mathrm{exp}}\,(R)$, can be obtained.

In turn, if small values for the signal expectation are selected, only hits in the outer region of an air shower can fulfill Eq. 6.2. In the shower periphery, SLC hits dominate, and since SLC hits are not accounted for in the air shower reconstruction, the difference between the expected and measured charge values can be large. In Fig. 6.10, several charge distributions of the measured hits are illustrated for four different threshold values according to Eq. 6.2, and for simulation and experimental data. For all charge distributions, proton and iron induced air showers agree with experimental data within the statistical errors.

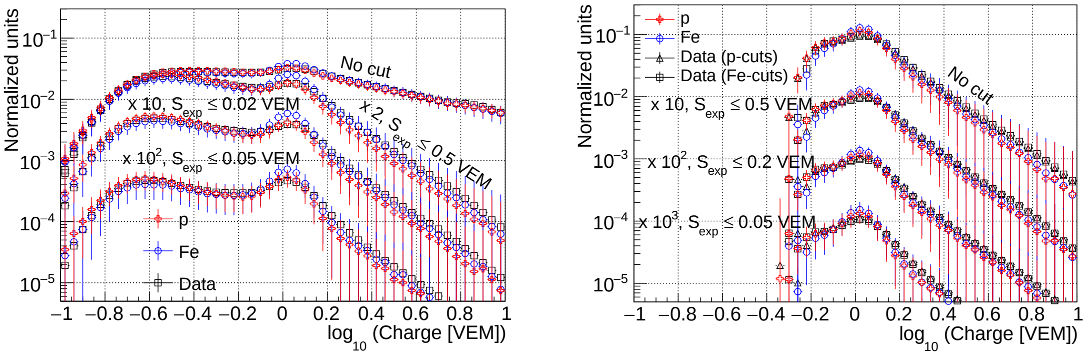

Figure 6.10: Charge distributions for four different thresholds for the charge expectation S_{exp}. For better visualization, the distributions are scaled by appropriate factors. The term “No cuts” means that every pulse, independent of any threshold, is considered.
Left: The charge distributions are shown without applying cuts on hit charge and distance.
Right: Cuts on hit charge and distance are applied.

6.3 Muon number estimators

According to Fig. 6.9, the application of cuts can be visualized,

$$L^{\mathrm{true}}_{\mathrm{ch,\,h}}\left(R\right) = S^{\mathrm{true}}_{\mathrm{ch,\,h}}\left(R\right) + B^{\mathrm{true}}_{\mathrm{ch,\,h}}\left(R\right) \xrightarrow{\mathrm{cuts}} L'^{\,\mathrm{true}}_{\mathrm{ch,\,h}}\left(R\right) = S'^{\,\mathrm{true}}_{\mathrm{ch,\,h}}\left(R\right) + B'^{\,\mathrm{true}}_{\mathrm{ch,\,h}}\left(R\right) \tag{6.3}$$

where $L^{\mathrm{true}}_{\mathrm{ch,\,h}}\left(R\right)$ represents the simulated lateral distribution of weighted total charge $L^{\mathrm{true}}_{\mathrm{ch}}\left(R\right)$ at distance R from the shower axis, or weighted number of hits $L^{\mathrm{true}}_{\mathrm{h}}\left(R\right)$, for all HLC and SLC hits, given by the sum of signal $S^{\mathrm{true}}_{\mathrm{ch,\,h}}$ and background $B^{\mathrm{true}}_{\mathrm{ch,\,h}}$, before the application of cuts, and $L'^{\,\mathrm{true}}_{\mathrm{ch,\,h}}$, $S'^{\,\mathrm{true}}_{\mathrm{ch,\,h}}$, and $B'^{\,\mathrm{true}}_{\mathrm{ch,\,h}}$ are the corresponding distributions after the application of cuts.
In the following, the variables $\mu^{\mathrm{est}}_{\mathrm{ch}}\left(R\right)$ and $\mu^{\mathrm{est}}_{\mathrm{h}}\left(R\right)$ are defined as estimators for the simulated muon number

$$\mu^{\mathrm{est}}_{\mathrm{ch,\,true}}\left(R\right) \coloneqq L'^{\,\mathrm{true}}_{\mathrm{ch}}\left(R\right) \tag{6.4}$$

and

$$\mu^{\mathrm{est}}_{\mathrm{h,\,true}}\left(R\right) \coloneqq L'^{\,\mathrm{true}}_{\mathrm{h}}\left(R\right) \tag{6.5}$$

The muon number estimators, $\mu^{\mathrm{est}}_{\mathrm{ch,\,true}}\left(R\right)$, and $\mu^{\mathrm{est}}_{\mathrm{h,\,true}}\left(R\right)$, correspond to the signal-like lateral distributions of charges and hits, after applying cuts on hit charge and distance, i.e. every hit was caused by at least one muon.
In this section, the muon number estimators $\mu^{\mathrm{est}}_{\mathrm{ch,\,true}}\left(R\right)$, and $\mu^{\mathrm{est}}_{\mathrm{h,\,true}}\left(R\right)$, are studied in detail, with focus on

- their difference to the true muon number provided by the simulation,
- and their comparison with muon number estimators in data.

6.3.1 Discussion

Both $\mu^{\mathrm{est}}_{\mathrm{ch,\,true}}\left(R\right)$ and $\mu^{\mathrm{est}}_{\mathrm{h,\,true}}\left(R\right)$ can be interpreted in an oversimplified picture:
In the shower periphery, the SLC hits dominate which are usually caused by few particles only. If at least one muon is participating in a hit, charge contributions of roughly 1 VEM due to a muon can be assumed according to Fig. 6.1. Thus, under the assumption that the charge contribution of other particle types is comparatively small, in the muon number estimator $\mu^{\mathrm{est}}_{\mathrm{ch,\,true}}\left(R\right)$ a charge value given in units of VEM can roughly be interpreted as the number of muons participating in this hit. Furthermore, due to low particle density in the shower periphery, the number of hits given in the muon number estimator $\mu^{\mathrm{est}}_{\mathrm{h,\,true}}\left(R\right)$ can be interpreted as the number of muons.

6.3.2 The true muon number $\mu^{\text{true}}\,(R_{ij})$

In each ring R_{ij} around the shower axis, the simulated muon number $\mu^{\text{true}}\,(R_{ij})$ is given by

$$\mu^{\text{true}}\,(R_{ij}) = \sum_{m=1}^{N_\mu(R_{ij})} w_m \tag{6.6}$$

where $N_\mu(R_{ij})$ is the number of muons producing hits in tanks located within R_{ij}, and w_m the weight according to the Cosmic Ray flux model (see Sec. 4.2).
In Fig. 6.11, the muon number estimators $\mu^{\text{est}}_{\text{ch, true}}\,(R)$, and $\mu^{\text{est}}_{\text{h, true}}\,(R)$, are compared to the simulated muon number, $\mu^{\text{true}}\,(R)$, derived for the SIBYLL2.1 simulation sets. The differences of the estimators to the true muon number at low distances to the shower axis occur since the muon number estimators still are composed of both signal and background hits, which in total are more abundant than the number of muons.

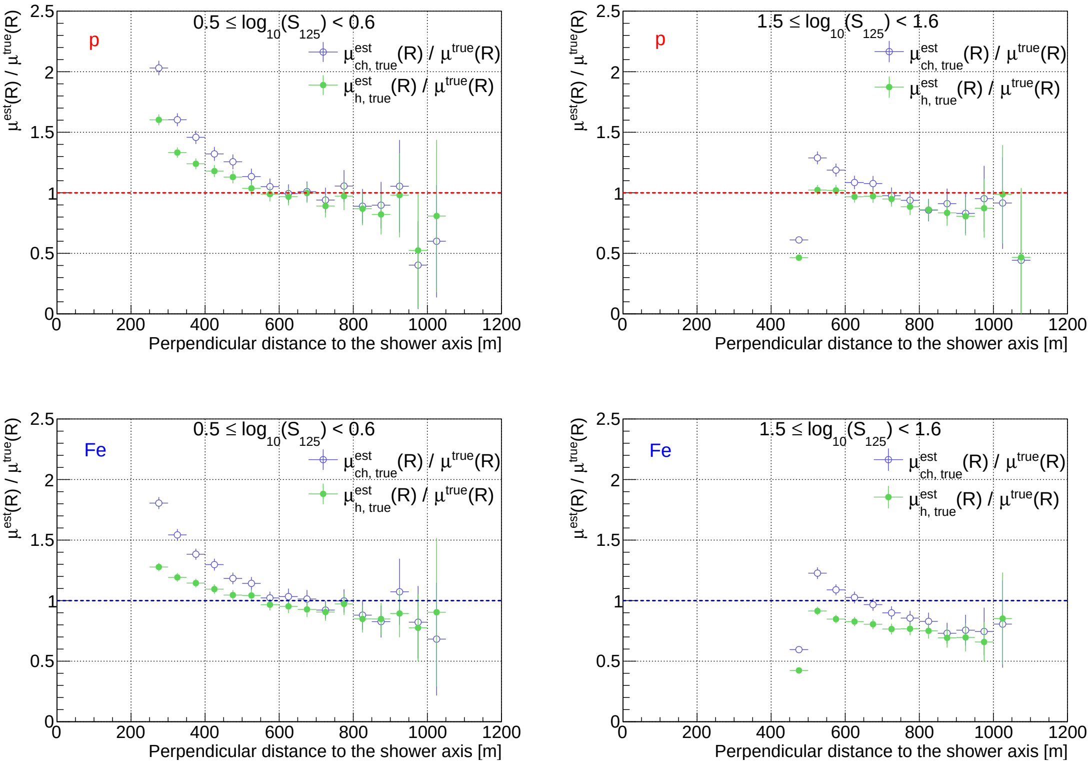

Figure 6.11: Muon number estimators $\mu^{\text{est}}_{\text{ch, true}}\,(R)$, and $\mu^{\text{est}}_{\text{h, true}}\,(R)$, compared to the simulated muon number $\mu^{\text{true}}\,(R)$, shown for vertical proton an iron showers, and two bins in shower size $\log_{10}\,(S_{125})$.

Furthermore, the muon number estimators do not perfectly agree with the true muon number in simulation, $\mu^{\mathrm{true}}(R)$, because muons hit the tank under various zenith angles, thus the hit charge can be significantly different from 1 VEM, and because other particles contribute. However, in Ch. 7 it is shown that the muon number estimators can be used to calculate a muon number density.

6.3.3 Consistency check

The muon number estimators, $\mu^{\mathrm{est}}_{\mathrm{ch,\,true}}(R)$, and $\mu^{\mathrm{est}}_{\mathrm{h,\,true}}(R)$, are intended to be used to calculate a muon number density. Their quality is checked by ensuring that corresponding distributions for proton and iron induced air showers enclose those for experimental data. For this purpose, muon number estimators for experimental data, $\mu^{\mathrm{est}}_{\mathrm{ch,\,data}}(R)$, and $\mu^{\mathrm{est}}_{\mathrm{h,\,Data}}(R)$, are defined as

$$\mu^{\mathrm{est}}_{\mathrm{ch,\,data}}(R) := L'^{\,\mathrm{data}}_{\mathrm{ch}}(R) \tag{6.7}$$

and

$$\mu^{\mathrm{est}}_{\mathrm{h,\,Data}}(R) := L'^{\,\mathrm{Data}}_{\mathrm{h}}(R) \tag{6.8}$$

where $L'^{\,\mathrm{Data}}_{\mathrm{ch}}(R)$, and $L'^{\,\mathrm{Data}}_{\mathrm{h}}(R)$, are the measured lateral distributions of total charge and number of hits after the application of the cuts on hit charge and distance. As shown in Sec. 6.2.2, the cuts on hit charge slightly depend on the primary mass. Thus, for each primary mass, there is a muon number estimator in data. This fact shows that the cut on hit charge is not only used to emphasize hits of muonic origin, but also to reduce the differences of the in background B originating from proton and iron induced air showers (also see appendix).

In order to compare the muon number estimators derived from simulation and data, $\mu^{\mathrm{est}}_{\mathrm{ch,\,true}}(R)$, $\mu^{\mathrm{est}}_{\mathrm{h,\,true}}(R)$, $\mu^{\mathrm{est}}_{\mathrm{ch,\,data}}(R)$, and $\mu^{\mathrm{est}}_{\mathrm{h,\,data}}(R)$, in each bin of distance to the shower axis, each estimator is divided by the total number of existing tanks in this bin. This number is counted for each event, and thus is higher in case of data.

In Fig. 6.12, it is illustrated that there is good agreement between simulation and data.

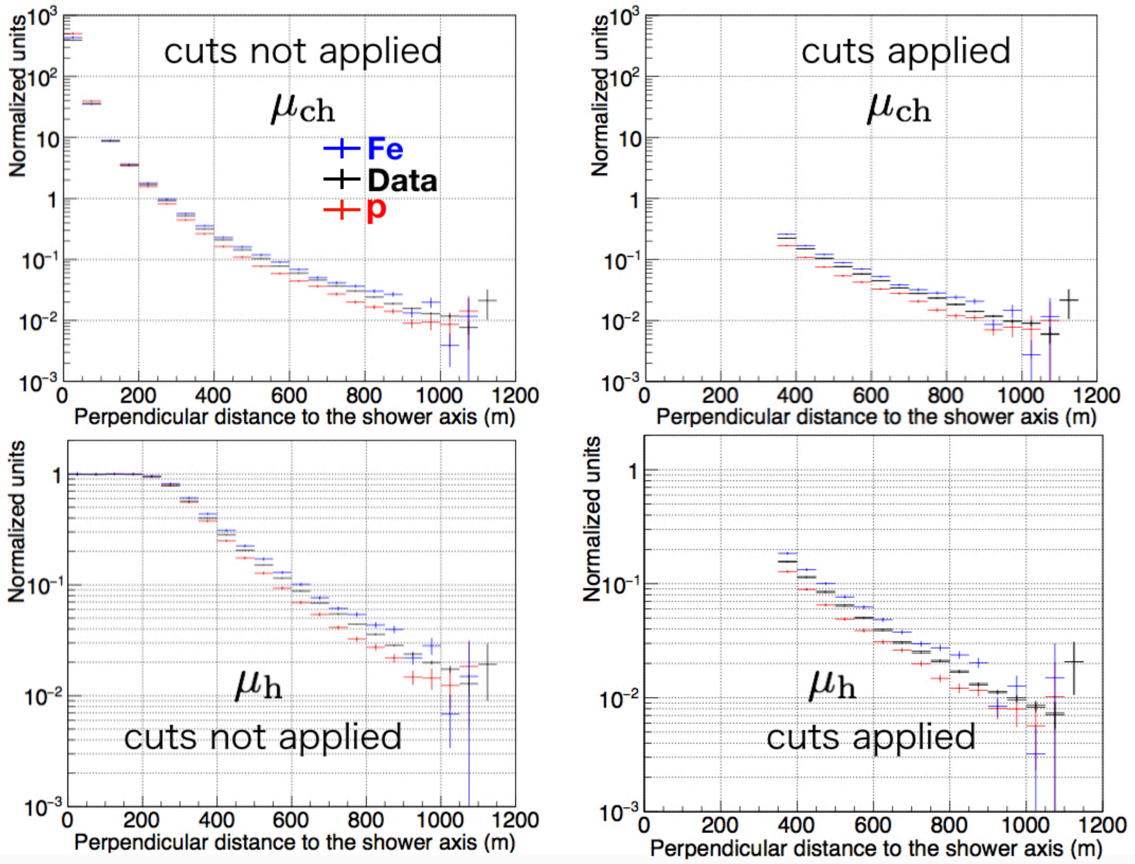

Figure 6.12: Comparisons between the muon number estimators $\mu^{\rm est}_{\rm ch,\,h,\,true}\,(R)$ and $\mu^{\rm est}_{\rm ch,\,h,\,Data}\,(R)$, divided by all tanks (see text), for two bins in shower size S_{125}.

6.4 Summary

In this chapter, it is shown that, using the IceTop simulation and reconstruction software, measured hits can be grouped according to specific particle types. If a hit contains at least one muon, the hit is defined as signal-like, otherwise as background-like.

The calculation of muon number densities requires a set of hits which contains as little background as possible in order to get a pure result. For this purpose, it is shown that the charge distribution of IceTop hits is a superposition of at least two distinct distributions, which have their peak positions at different charge values. The distribution of signal has its peak value at roughly 1 VEM which is the average charge a low energy muon creates by passing an IceTop tank. The average charge of background is far below 1 VEM.

This issue is used to construct a cut on the hit charge, which is performed for each bin in shower size S_{125} and zenith angle θ. Furthermore, a cut on the perpendicular distance of a hit to the shower axis ensures that the fraction of hit tanks to all tanks is smaller than 1. It is also illustrated, that the contribution of signal increases due to the application of these cuts.

The lateral distribution of hit charges and hit distances are used to create two muon number estimators. For this purpose, the cuts on hit charge and distance are applied. The resulting distributions, called $\mu^{\mathrm{est}}_{\mathrm{ch,\,true}}\,(R)$, and $\mu^{\mathrm{est}}_{\mathrm{h,\,true}}\,(R)$, compare very well with muon number estimators derived from experimental data.

Chapter 7

Calculation of muon number densities

In this chapter, a method is presented which allows the derivation of the muon number density $\rho_{\,\mathrm{data}}\,(R)$ based on the muon number estimators $\mu^{\mathrm{est}}_{\mathrm{ch,\,true}}\,(R)$, $\mu^{\mathrm{est}}_{\mathrm{h,\,true}}\,(R)$, $\mu^{\mathrm{est}}_{\mathrm{ch,\,data}}\,(R)$, and $\mu^{\mathrm{est}}_{\mathrm{h,\,data}}\,(R)$. Firstly, $\mu^{\mathrm{est}}_{\mathrm{ch,\,true}}\,(R)$, and $\mu^{\mathrm{est}}_{\mathrm{h,\,true}}\,(R)$ are transformed into muon number density estimators, called $\rho^{\mathrm{est}}_{\mathrm{ch,\,true}}\,(R)$, and $\rho^{\mathrm{est}}_{\mathrm{h,\,true}}\,(R)$. Secondly, using an appropriate conversion, these muon number densities are related to the true muon number density known from air shower simulations, called $\rho_{\,\mathrm{true}}\,(R)$. The robustness of the conversion is furthermore studied since the statistical and systematic uncertainties of the final result depend on those of the conversion. Afterwards, the conversion is applied on muon number estimators derived from experimental data, $\mu^{\mathrm{est}}_{\mathrm{ch,\,data}}\,(R)$, and $\mu^{\mathrm{est}}_{\mathrm{h,\,data}}\,(R)$, and yields the muon number density $\rho_{\,\mathrm{data}}\,(R)$.
The final result is the muon number density presented as a function of the perpendicular distance to the shower axis, and at a fixed distance as function of the reconstructed energy E_{rec}.

7.1 The true muon number density $\rho_{\,\mathrm{true}}\,(R)$

The true muon number density $\rho_{\,\mathrm{True}}\,(R_{ij})$, in a ring R_{ij} around the shower axis, can be calculated by

$$\rho_{\,\mathrm{true}}\,(R_{ij}) = \frac{1}{N_{\mathrm{at,\,true}}(R_{ij})} \cdot \sum_{p=1}^{N_\mu(R_{ij})} \frac{w_p}{A_{\mathrm{top}} \cdot \cos\theta_{\mu,\,p} + A_{\mathrm{side}} \cdot \sin\theta_{\mu,\,p}} \tag{7.1}$$

where $N_{\mathrm{at,\,true}}(R_{ij})$ is the weighted simulated number of all tanks in the ring R_{ij}, regardless if a tank is hit or not, $N_\mu(R_{ij})$ is the number of muons producing hits in these tanks, w_p the particle weight according to the Cosmic Ray flux model (see Sec. 4.2), and $A_{\mathrm{top}} \cdot \cos\theta_{\mu,\,p} + A_{\mathrm{side}} \cdot \sin\theta_{\mu,\,p}$ is the area of the ice cylinder seen in the arrival direction of the particle p, with the top of the ice block $A_{\mathrm{top}} = \pi r^2$ and its cross-sectional area $A_{\mathrm{side}} = 2rh$. (radius $r = 0.91\,\mathrm{m}$, height $h = 0.9\,\mathrm{m}$, see Fig. 7.1.)

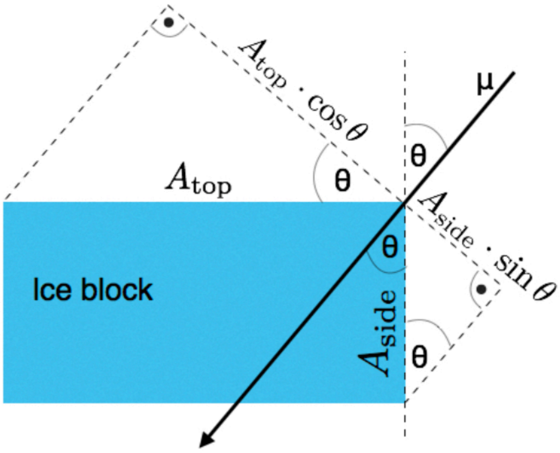

Figure 7.1: Visualization of an ice block and an incoming muon.

7.2 Discussion

Equation 7.1 does not consider particles whose track may cross two tanks, i.e. particles which produce a charge in two different tanks. Since the centers of the two tanks of a station are separated by 10 m, an individual particle could only be counted twice, if it would have an zenith angle of $\theta \gtrsim 83,72°$, i.e. $\cos\theta \lesssim 0.11$. The fraction of particles with these zenith angles is negligible because of the used zenith angle range ($0.8 \leq \cos\theta \leq 1.0$).

Exemplary distributions of the true muon number density $\rho_{\text{true}}(R_{ij})$, as described by Eq. 7.1, are illustrated in Fig. 7.2, for three bins in true energy E_{true}. For the calculation, true instead of reconstructed parameters, such as primary energy, core, direction, are used. It can be seen that there is a significant difference between the muon content of air showers initiated by different primaries, and that the muon content strongly depends on the primary energy.

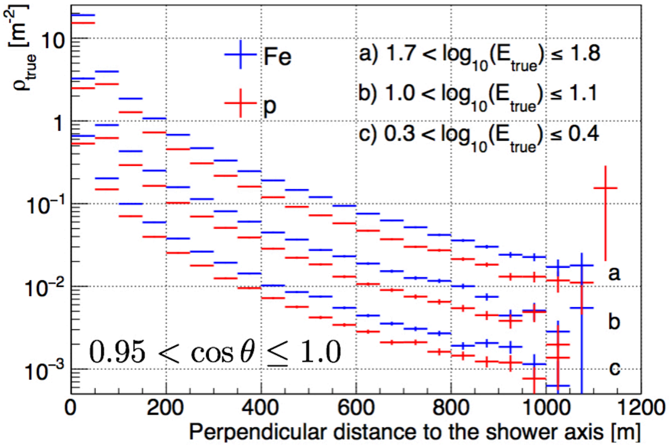

Figure 7.2: The muon number density $\rho_{\text{true}}(R)$ calculated for vertical proton and iron showers, and three bins in true energy E_{true}.

7.3 Energy conversion

The conversion from shower size S_{125} to the primary energy E_{rec}, is done using so-called *conversion functions*. These functions can be derived by evaluating the dependance of S_{125} on E_{true}, as for example shown in Fig. 5.3. For each bin in S_{125}, the distribution of true energy E_{true} is fit with a gaussian function, and the mean of the gaussian is used as energy estimation of the current bin. The values of the mean of all bins in S_{125} are fit by linear functions which are the conversion functions,

$$\log_{10}\left(E_{\mathrm{rec}}\right) = P_1 \cdot \log_{10}\left(S_{125}\right) + P_0 \tag{7.2}$$

where the parameters P_0 and P_1 are parameters describing the linear fit, and depend on zenith angle θ. More information about the method is described in Ref. [69]. The parameters for the conversion functions used in this work are taken from Ref. [79], and are shown in Tab. 7.1.

Zenith angle bin	P_0	P_1
$0.95 < \cos\theta \leq 1.0$	6.010569	0.933316
$0.90 < \cos\theta \leq 0.95$	6.054677	0.923860
$0.85 < \cos\theta \leq 0.9$	6.109777	0.914971
$0.80 < \cos\theta \leq 0.85$	6.177271	0.907456

Table 7.1: Parameters of the functions converting S_{125} to E_{rec}.

The numbers were derived using the H4a model [81]. Dependent on the assumption of the Cosmic Ray composition, the numbers differ, as was shown in Ref. [69]. Thus, the conversion from shower size S_{125} to reconstructed energy E_{rec} suffers from a systematic uncertainty. The numbers shown in Tab. 7.1 do not match with those shown in Ref. [69] because the latter are not correct due to a wrong observation level (see Sec. 4.3.1). Currently, correct numbers are available internally for the H4a model only. The uncertainty due to the composition assumption is estimated by 7 %, as was also an estimation in Ref. [69].

7.4 Development on simulation

Both the true muon number density, $\rho_{\mathrm{true}}\left(R_{ij}\right)$, and the muon number estimators, $\mu^{\mathrm{est}}_{\mathrm{ch,\,h,\,true}}\left(R\right)$ can be used to define the so-called *conversion distributions* $C_{\mathrm{ch,\,h}}\left(R\right)$. With these histograms the muon number density $\rho_{\mathrm{true}}\left(R\right)$ can be calculated, if only the muon number estimators $\mu^{\mathrm{est}}_{\mathrm{ch,\,h,\,true}}\left(R\right)$ are known. The conversion distributions $C_{\mathrm{ch,\,h}}\left(R\right)$ are derived step-by-step using the following procedure:

1. The true muon number density $\rho_{\mathrm{true}}\left(R_{ij}\right)$, as for example presented in Fig. 7.2, is calculated using true instead of reconstructed parameters (primary energy, core, direction). The muon numbers estimators are derived using reconstructed

quantities (S_{125}, reconstructed core, direction). In a first step, S_{125} is translated into primary energy using the procedure described in Sec. 7.3, such that the muon number estimators are given in bins of E_{rec}.

2. The muon number estimators $\mu^{\mathrm{est}}_{\mathrm{ch,\,h,\,true}}\left(R_{ij}\right)$ are transformed into *muon number density estimators* $\rho^{\mathrm{est}}_{\mathrm{ch,\,h,\,true}}\left(R_{ij}\right)$ by

$$\rho^{\mathrm{est}}_{\mathrm{ch,\,h,\,true}}\left(R_{ij}\right) = \frac{1}{N_{\mathrm{at,\,true}}(R_{ij})} \cdot \frac{\mu^{\mathrm{est}}_{\mathrm{ch,\,h,\,true}}\left(R_{ij}\right)}{A_{\mathrm{top}} \cdot \cos\langle\theta\rangle + A_{\mathrm{side}} \cdot \sin\langle\theta\rangle} \tag{7.3}$$

   where $N_{\mathrm{at,\,true}}(R_{ij})$ is the simulated number of all tanks in the ring R_{ij} (based on E_{rec}), and $\langle\theta\rangle$ is the average zenith angle of the zenith angle range in which $\mu^{\mathrm{est}}_{\mathrm{ch,\,h,\,true}}$ are considered, e.g. $0.95 < \cos\theta \leq 1.0$ in the case of vertical air showers.

3. The ratio between the true muon density, $\rho_{\mathrm{true}}\left(R\right)$, and the muon number density estimators, $\rho^{\mathrm{est}}_{\mathrm{ch,\,h,\,true}}\left(R_{ij}\right)$ define the conversion distributions $C_{\mathrm{ch,\,h}}\left(R\right)$,

$$C_{\mathrm{ch,\,h}}\left(R_{ij}\right) = \rho_{\mathrm{true}}\left(R_{ij}\right)/\rho^{\mathrm{est}}_{\mathrm{ch,\,h,\,true}}\left(R_{ij}\right) \tag{7.4}$$

The conversion distributions $C_{\mathrm{ch,\,h}}\left(R\right)$ consist of factors for each bin in perpendicular distance to the shower axis. These factors are multiplied with corresponding values of the muon number density estimators $\rho^{\mathrm{est}}_{\mathrm{ch,\,h,\,true}}\left(R\right)$ and yield the true muon density in that bin. Exemplary distributions for $C_{\mathrm{ch,\,h}}\left(R\right)$ are illustrated in Sec. 7.6.

7.5 Application on experimental data

The conversion distributions, $C_{\mathrm{ch,\,h}}\left(R\right)$, as presented in the last section, are applied on the muon number estimators derived from data, $\mu^{\mathrm{est}}_{\mathrm{ch,\,h,\,data}}\left(R\right)$:

1. Similar to the procedure applied on simulation, convert S_{125} to E_{rec}.

2. Based on $\mu^{\mathrm{est}}_{\mathrm{ch,\,h,\,data}}\left(R\right)$, a muon number density estimator for experimental data is calculated via

$$\rho^{\mathrm{est}}_{\mathrm{ch,\,h,\,data}}\left(R_{ij}\right) = \frac{1}{N_{\mathrm{at,\,data}}(R_{ij})} \cdot \frac{\mu^{\mathrm{est}}_{\mathrm{ch,\,h,\,data}}\left(R\right)}{A_{\mathrm{top}} \cdot \cos\langle\theta\rangle + A_{\mathrm{side}} \cdot \sin\langle\theta\rangle} \tag{7.5}$$

   The calculation is similar to Eq. 7.3, but with $N_{\mathrm{at,data}}(R_{ij})$ based on experimental data being different from $N_{\mathrm{at,true}}(R_{ij})$ because the used data sample has much more events than the used simulation sets.

3. The conversion factors $C_{\mathrm{ch,\,h}}\left(R\right)$ are applied on the estimated muon number densities $\rho^{\mathrm{est}}_{\mathrm{ch,\,h,\,data}}\left(R_{ij}\right)$ in order to derive the muon number densities in experimental data,

$$\rho_{\mathrm{ch,\,h,\,data}}\left(R_{ij}\right) = C_{\mathrm{ch,\,h}}\left(R_{ij}\right) \cdot \rho^{\mathrm{est}}_{\mathrm{ch,\,h,\,data}}\left(R_{ij}\right) = \rho_{\mathrm{true}}\left(R_{ij}\right) \cdot \frac{N_{\mathrm{at,\,true}}(R_{ij})}{N_{\mathrm{at,\,data}}(R_{ij})} \cdot \frac{\mu^{\mathrm{est}}_{\mathrm{ch,\,h,\,data}}\left(R_{ij}\right)}{\mu^{\mathrm{est}}_{\mathrm{ch,\,h,\,true}}\left(R_{ij}\right)} \tag{7.6}$$

7.6 Systematic uncertainties

In simulations, systematic uncertainties arise from

- **the hadronic interaction models**
  As presented in Fig. 4.1, air showers simulated using the high energy interaction models QGSJetII-04, and EPOS LHC, produce a different amount of muons than SIBYLL2.1. This affects the amount of signal hits. In this work, the final result which is the muon number density as function of the reconstructed energy, is presented separately for each interaction model.

- **the snow heights on IceTop tanks**
  The reconstruction of air shower parameters, such as the shower size S_{125}, takes the snow heights on IceTop tanks into account (see Sec. 5.2.3). Uncertainties are provided by the attenuation length λ_{snow} of electrons in snow (see Eq. 5.17) which is varied within the given uncertainties $\lambda_{\mathrm{snow}} = 2.1\,\mathrm{m} \pm 0.2\,\mathrm{m}$ [69]. The shower size S_{125}, and thus the energy scale, is affected by this uncertainty.

- **the VEM calibration of the hit charge**
  As shown in Sec. 3.6, the VEM calibration is performed differently for experimental data and simulation. Thus, the definition of 1 VEM in a particular tank may differ in both cases. In order to cover potential differences, simulated charges in IceTop tanks are varied by $\pm 3\,\%$ before air shower reconstruction [78]. The uncertainty on the VEM calibration propagates to an uncertainty on the shower size S_{125}, thus on the energy scale.

- **the composition of Cosmic Rays**
  The conversion from the shower size S_{125} to the reconstructed primary energy E_{rec} depends on the assumed mass composition. The uncertainty is estimated by 7% and is added to the final results shown in Sec. 7.8.

Parameters which are obtained from simulations, and which are chosen to represent those derived from experimental data, are required to be *robust*. A parameter is robust, if it is stable against changes in the simulation, i.e. if the sum of its statistical and systematic uncertainties is small. In this section, the robustness of the conversion distributions $C_{\mathrm{ch,\,h}}\,(R)$ is investigated since these distributions are applied on measurements performed with experimental data and thus define the systematic uncertainty of the final result. In Fig. 7.3, the statistical, and systematic uncertainties due to the VEM calibration and snow correction, on the conversion distribution $C_{\mathrm{h}}\,(R)$, for a particular bin in reconstructed primary energy E_{rec}, and for the case of air showers induced by iron primaries, is presented. The conversion distribution $C_{\mathrm{h}}(R)$ was calculated according to Eq. 7.4. The cut values on hit charge and distance were always used from the non-systematic dataset. All statistical and systematic errors are summed quadratically.

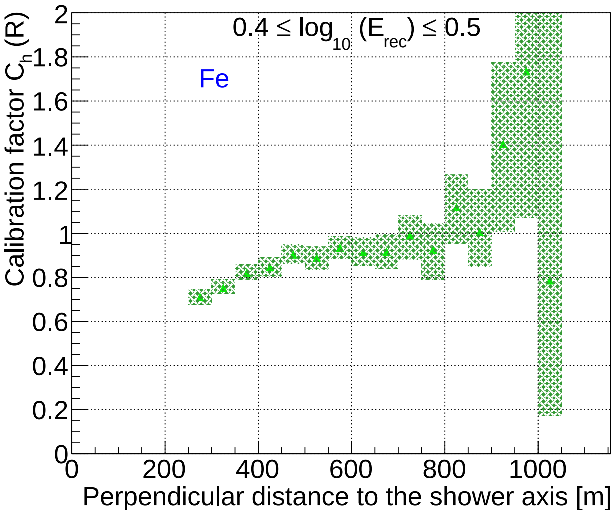

Figure 7.3: Uncertainties on the conversion distributions $C_{\mathrm{h}}(R)$, for vertical air showers induced by iron primaries. The green points denote the values of the non-systematic conversion distribution $C_{\mathrm{h}}(R)$.

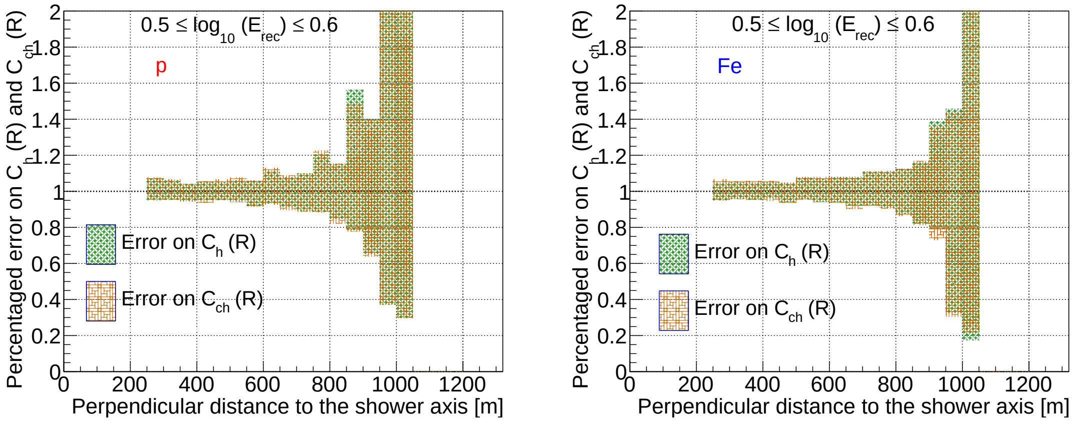

Figure 7.4: Total uncertainties on the conversion distributions $C_{\mathrm{ch,\,h}}(R)$ for proton, and iron induced air showers, and two bins in reconstructed energy E_{rec}.

In Fig. 7.4, the relative total error of the conversion distributions $C_{ch,\,h}(R)$ is shown. The increasing error values are due to the lack of statistcs at high distances from the shower axis. The total uncertainties, in the case of the conversion distributions based on measured charge values, $C_{\mathrm{ch}}(R)$, exceed the corresponding errors for $C_{\mathrm{h}}(R)$ in all bins of reconstructed energy. Thus, to achieve the final results, $C_{\mathrm{ch}}(R)$ is not used.

7.7 Consistency check

The method used to derive a muon number density, as presented in Secs. 7.4 and 7.5, is also checked using simulations only. A simulation set is separated into two disjunct samples. The first one is treated as “simulation-like”, and is used to construct the conversion distributions $C_\mathrm{h}\left(R\right)$ according to Sec. 7.4. The second sample is treated as “data-like”, on which the conversion distributions are applied as explained in Sec. 7.5. As can be seen in Fig. 7.5, the muon density from the “simulation-like” sample is well reproduced by this procedure which was expected when only statistical uncertainties vary the datasets.

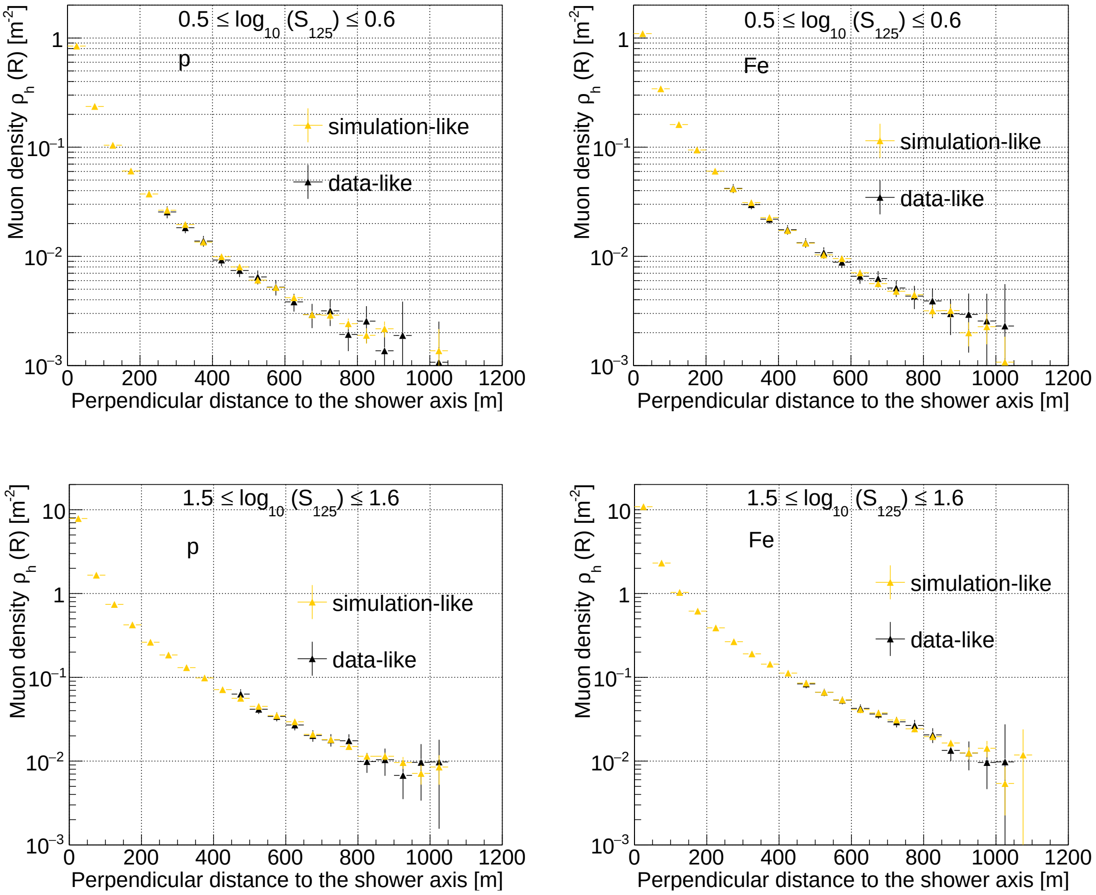

Figure 7.5: Simulation-like and data-like consistency check.

7.8 Results

In this section, the final results of this work are presented. These are the muon number density as function of perpendicular distance to the shower axis, and as function of reconstructed energy E_rec for distances of 400 m, and 600 m.

7.8.1 Muon density as function of distance R

The reconstructed muon density $\rho_{\text{data}}\,(R)$ is presented in Fig. 7.6 for various reconstructed energies E_{rec}. Due to the energy dependent cut on hit distance, the distributions start at different values. It can be seen that the difference in the muon density, based on mass dependent cuts on the hit charge, is small. Thus, the results could be used to develop and check analytic functions for the lateral distribution of the muon density as function of the primary energy. At distances from the shower axis larger than $\sim 800\,\text{m}$ fluctuations become larger due to low statistics.

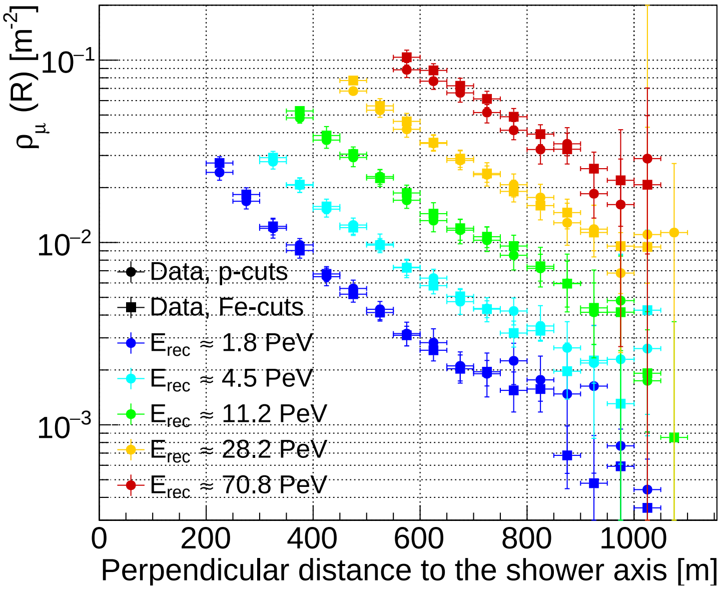

Figure 7.6: Muon number density $\rho_{\text{data}}\,(R)$ for air showers with zenith angles θ within $0.95 < \cos\theta \leq 1.0$. The derivation is based on the SIBYLL2.1 interaction model. The circles represent the density of muons calculated using the cuts on hit charge and distance developed using air showers induced by protons, and the squares refer to the cuts based on iron induced air showers.

7.8.2 Muon density as function of E_{rec}

In Fig. 7.7, the muon number density is shown as function of the the reconstructed primary energy E_{rec}, for two fixed distances of roughly 400 m, and 600 m to the shower axis. The distribution for 400 m cuts off at $\log_{10}(\text{E}\,[\text{PeV}]) \simeq 1.4$ due to the energy dependance of the cut on hit distance (see Sec. 6.2.3). The distributions for 400 m, and 600 m, are separated as expected due to the decreasing muon density with distance to the shower axis.

The muon number density calculated using the non-systematic simulation set SIBYLL2.1, is presented in Fig. 7.8, together with muon number densities extracted from air shower simulations, and results for the systematic simulation sets.

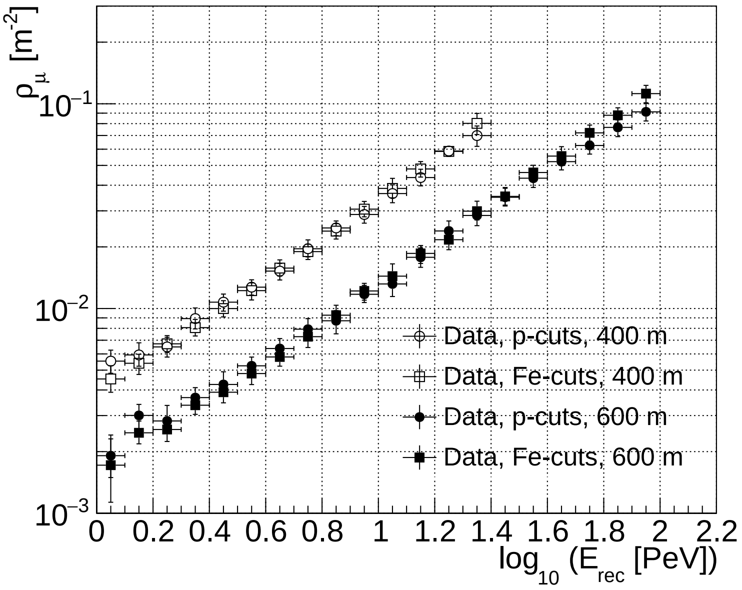

Figure 7.7: Muon number densities at 400 m, and 600 m perpendicular distance to the shower axis, for vertical air showers as a function of E_{rec}. The derivation is based on the SIBYLL2.1 interaction model.

In the case of QGSJetII04, and EPOS LHC, the entire analysis chain was run again. Regarding the interaction model SIBYLL2.1, the true muon density brackets the data points in all cases, and indicates a heavy composition at high energies. For each bin in reconstructed energy, the data points differ due to the cuts on hit charge, and serve as lower and upper bound for the muon density. In the case of QGSJetII04, and EPOS LHC, the number of muons is rather overestimated. In all of the plots shown in Fig. 7.8, there is no visible muon excess as recorded by Auger, it is maybe indicated for SIBYLL2.1 at high primary energies. The results presented in Fig. 7.8 are consistent with results of a similar analysis performed by two members of IceCube [80]. Nevertheless, the systematic simulation sets only have 15% statistics compared to the non-systematic one, because during this work resources for additional air shower simulations were not available. Further simulations can lower the errors in the bottom plots of Fig. 7.8 so that further conclusions can be drawn.

7.9 Summary

In this chapter, a method is described which allows to calculate the muon number density in experimental data. Based on muon number density estimators derived from air shower simulations, a conversion is defined which relates the muon number density estimator to the true muon number density known from simulations. This conversion is applied to muon number estimators derived from data.

The results are the muon number densities as function of the perpendicular distance to the shower axis, and reconstructed primary energy E_{rec}.

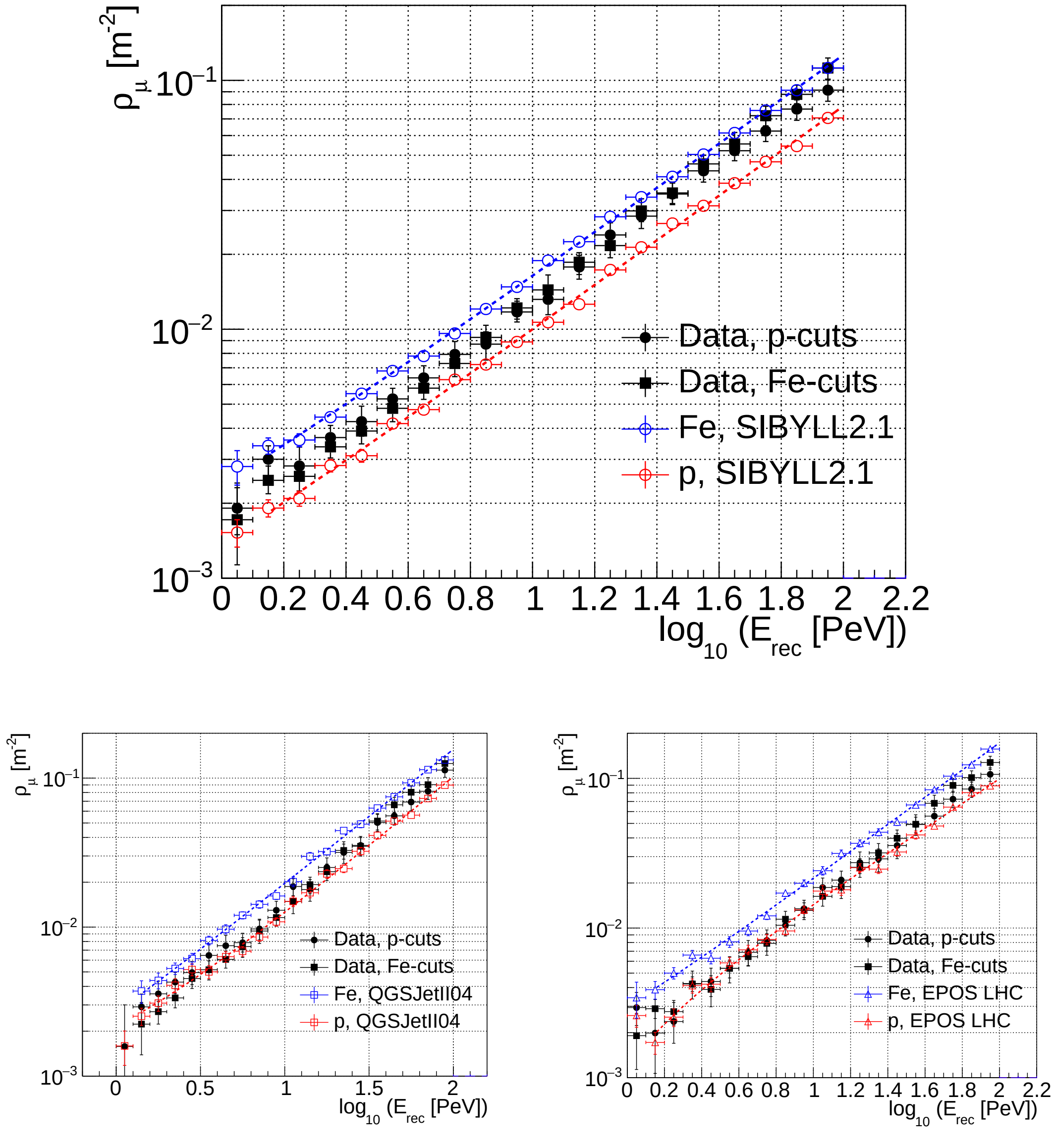

Figure 7.8: Muon number densities for vertical air showers as a function of E_{rec}, and for three interaction models. The horizontal errors denote the bin size.
Upper plot: Results based on the non-systematic simulation set SIBYLL2.1. Numbers are provided in Tabs. 7.2 and 7.3.
Lower plots: Results regarding the systematic simulation sets QGSJetII04, and EPOS LHC.

$\log_{10}(\mathrm{E}_{\mathrm{rec}}[\mathrm{PeV}])$	$\mathrm{E}_{\mathrm{rec}}[\mathrm{PeV}])$	value ± uncertainty
0.05	1.12	$0.00191^{+26.16\%}_{-21.80\%}$
0.15	1.41	$0.00300^{+13.17\%}_{-14.06\%}$
0.25	1.78	$0.00282^{+19.09\%}_{-12.76\%}$
0.35	2.24	$0.00367^{+11.86\%}_{-12.23\%}$
0.45	2.82	$0.00426^{+15.34\%}_{-12.58\%}$
0.55	3.55	$0.00525^{+10.32\%}_{-10.81\%}$
0.65	4.47	$0.00638^{+11.83\%}_{-10.98\%}$
0.75	5.62	$0.00792^{+12.64\%}_{-9.70\%}$
0.85	7.08	$0.00871^{+12.05\%}_{-13.66\%}$
0.95	8.91	$0.01173^{+10.50\%}_{-8.86\%}$
1.05	11.22	$0.01318^{+11.98\%}_{-13.18\%}$
1.15	14.13	$0.01779^{+11.05\%}_{-10.64\%}$
1.25	17.78	$0.02392^{+11.88\%}_{-9.25\%}$
1.35	22.39	$0.02847^{+8.77\%}_{-10.81\%}$
1.45	28.18	$0.03494^{+11.60\%}_{-9.36\%}$
1.55	35.48	$0.04329^{+10.69\%}_{-9.89\%}$
1.65	44.67	$0.05216^{+9.40\%}_{-8.85\%}$
1.75	56.23	$0.06256^{+10.07\%}_{-9.22\%}$
1.85	70.79	$0.07667^{+8.84\%}_{-9.87\%}$
1.95	89.13	$0.09133^{+9.93\%}_{-9.76\%}$

Table 7.2: Uncertainties on the data points labeled with “p-cuts” in Fig. 7.8 in the case of SIBYLL2.1.

$\log_{10}(\mathrm{E}_{\mathrm{rec}}[\mathrm{PeV}])$	$\mathrm{E}_{\mathrm{rec}}[\mathrm{PeV}])$	value ± uncertainty
0.05	1.12	$0.00172^{+34.46\%}_{-34.10\%}$
0.15	1.41	$0.00247^{+14.12\%}_{-11.67\%}$
0.25	1.78	$0.00256^{+12.11\%}_{-12.60\%}$
0.35	2.24	$0.00337^{+11.38\%}_{-9.76\%}$
0.45	2.82	$0.00390^{+11.06\%}_{-11.23\%}$
0.55	3.55	$0.00481^{+11.69\%}_{-11.64\%}$
0.65	4.47	$0.00580^{+10.67\%}_{-9.85\%}$
0.75	5.62	$0.00729^{+10.46\%}_{-11.46\%}$
0.85	7.08	$0.00928^{+11.83\%}_{-9.07\%}$
0.95	8.91	$0.01218^{+8.92\%}_{-9.87\%}$
1.05	11.22	$0.01439^{+14.81\%}_{-10.14\%}$
1.15	14.13	$0.01858^{+9.29\%}_{-10.70\%}$
1.25	17.78	$0.02168^{+10.23\%}_{-10.64\%}$
1.35	22.39	$0.02989^{+11.85\%}_{-9.44\%}$
1.45	28.18	$0.03529^{+9.57\%}_{-9.34\%}$
1.55	35.48	$0.04616^{+8.47\%}_{-9.35\%}$
1.65	44.67	$0.05553^{+11.19\%}_{-9.65\%}$
1.75	56.23	$0.07211^{+9.13\%}_{-8.95\%}$
1.85	70.79	$0.08785^{+8.82\%}_{-8.30\%}$
1.95	89.13	$0.11210^{+9.71\%}_{-9.79\%}$

Table 7.3: Uncertainties on the data points labeled with "Fe-cuts" in Fig. 7.8 in the case of SIBYLL2.1.

Chapter 8

Discussion and outlook

The goal of this work was the derivation of a method to measure the muon number density in experimental data. In this work, air shower simulations are used to construct estimators for the muon number density. As was shown in Sec. 7.8, this procedure yields a reasonable result for the muon number density. Especially important is the fact, that compared to the used interaction models SIBYLL2.1, QGSJetII04, and EPOS LHC, there is no muon excess visible in the considered energy range.

In subsequent analyses, larger statistics for the systematic datasets, QGSJet and EPOS LHC, could be produced. With these, a more precise statement about the validity of high energy interaction models can be achieved so that the Cosmic Ray composition can be studied with IceTop.

There is a number of improvements which could be done in order to raise the quality of the results:

- The high energy interaction model SIBYLL2.3 could be used as standard simulation, or as additional systematic. It is the successor of SIBYLL2.1 and takes new measurements into account.

- The cuts on hit charge and distance were calculated separately for proton and iron induced air showers. Using simulation sets representing primary masses different from proton and iron, or an assumption of composition mixture, such as in the H4a model [81], can yield cut values which represent all primary masses.

- Lately, the software versions used for detector simulation and air shower reconstruction, have been improved and extended. This includes for example the possibility to extract errors on the fit parameters derived during air shower reconstruction, i.e. on the shower size S_{125}, β, zenith angle θ, azimuth angle ϕ, and the position of the shower core $\vec{x}_\mathrm{C}$. Furthermore, software has been developed which allows the simulation of uncorrelated background which was not available for this work. Additionally, a more advanced cleaning is available which selects HLC hits appropriate for air shower reconstruction. This could improve the core and angular resolution.

- There are a few issues regarding the simulation of SLCs in IceTop. For example, in the case of HLCs, the time of the hit is calculated using the waveform and the leading edge (see Fig. 3.6). In the case of SLCs, the extracted timestamp does not provide a precise value. This issue is currently studied [82]. Once this is solved, information of SLCs can for example be used for air shower reconstruction.

- As presented in Sec. 5.2.3, the reconstruction of air showers includes a correction of the shower size parameter S_{125} due to snow, and the attenuation length of electrons has been varied according to Sec. 7.6. However, this correction does not account for the change of DOM launch rates due to snow, i.e. for large snow heights, HLC hits could occur as SLC hits, whereas SLC hits could be suppressed. Thus, the change in rate could be studied by performing additional detector simulations in which the snow heights are artificially lowered or raised by an specific amount. A desirable improvement of the reconstruction software could be the implementation of the probability of getting a discriminator trigger. Another possibility is to separate the IceTop detector into two parts, where the first one contains the tanks deployed in the early stage of the detector. These cover more snow than those deployed later [40] which are comprised by the second part. This way, the effect can be studied using experimental data.

In conclusion, the results of this work show potential to be the basis for a number of subsequent analyses. These can for example be studies of Cosmic Ray composition using the IceTop detector without InIce.

Appendix

- Exemplary cuts on hit charge and distance, are presented in Fig. 6.7, and Fig. 6.8. In Fig. 1, all cuts on hit distance are shown for vertical and inclined air showers. In Fig. 2, the best cuts on the hit charge are presented for vertical and inclined air showers, and for two bins in distance.

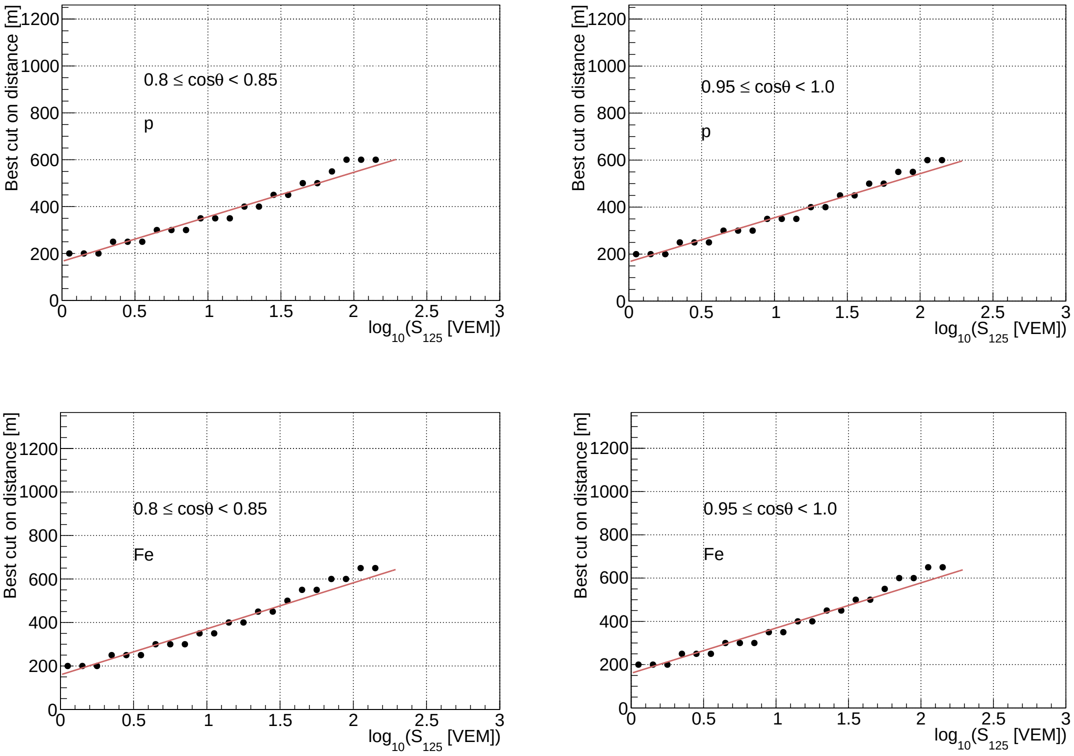

Figure 1: Best cuts on hit distance to the shower axis.

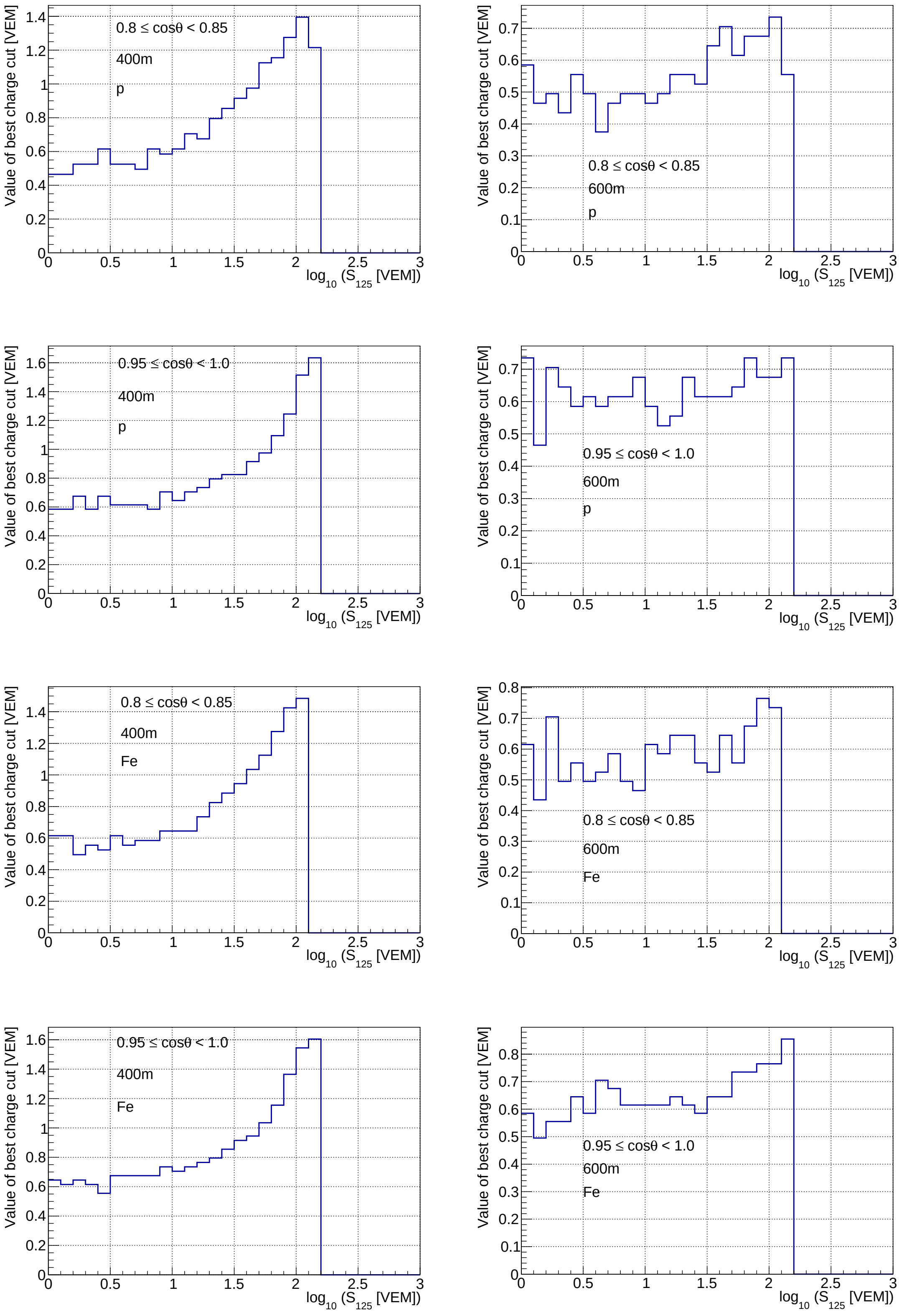

Figure 2: Best cuts on hit charge

- The muon densities, as shown in Secs. 7.8.1 and 7.8.2, are calculated for the case of vertical air showers ($0.95 < \cos\theta \leq 1.0$). In Fig. 3, Fig. 4, and Fig. 5, they are also shown for inclined air showers with $0.80 < \cos\theta \leq 0.85$.

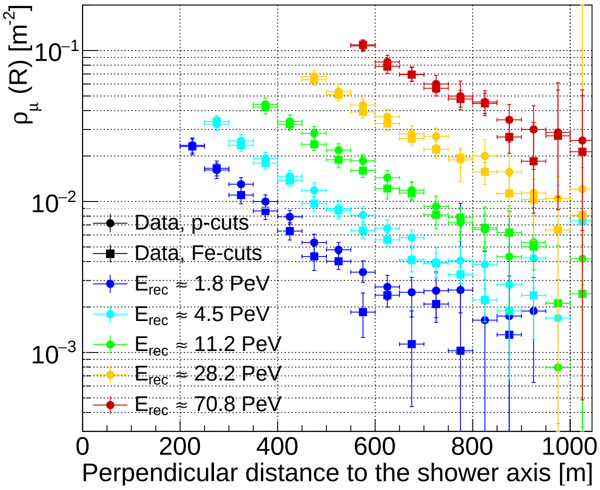

Figure 3: Muon number density $\rho_{\mathrm{data}}(R)$ for air showers with zenith angles θ within $0.80 < \cos\theta \leq 0.85$. The derivation is based on the SIBYLL2.1 interaction model.

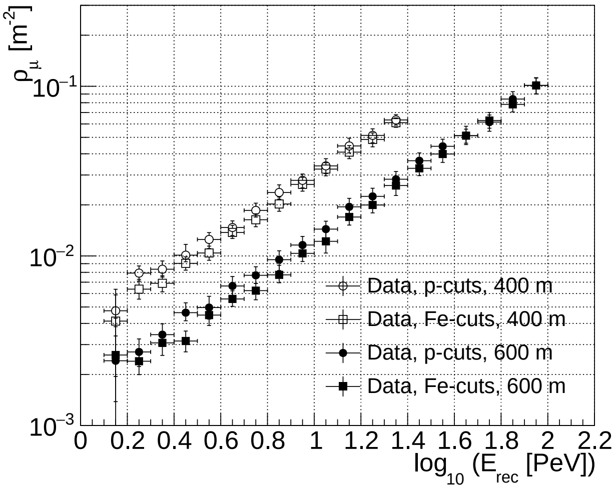

Figure 4: Muon number density $\rho_\mu(E_{\mathrm{rec}})$ for perpendicular distances of 400 m and 600 m to the shower axis. The derivation is based on the SIBYLL2.1 interaction model.

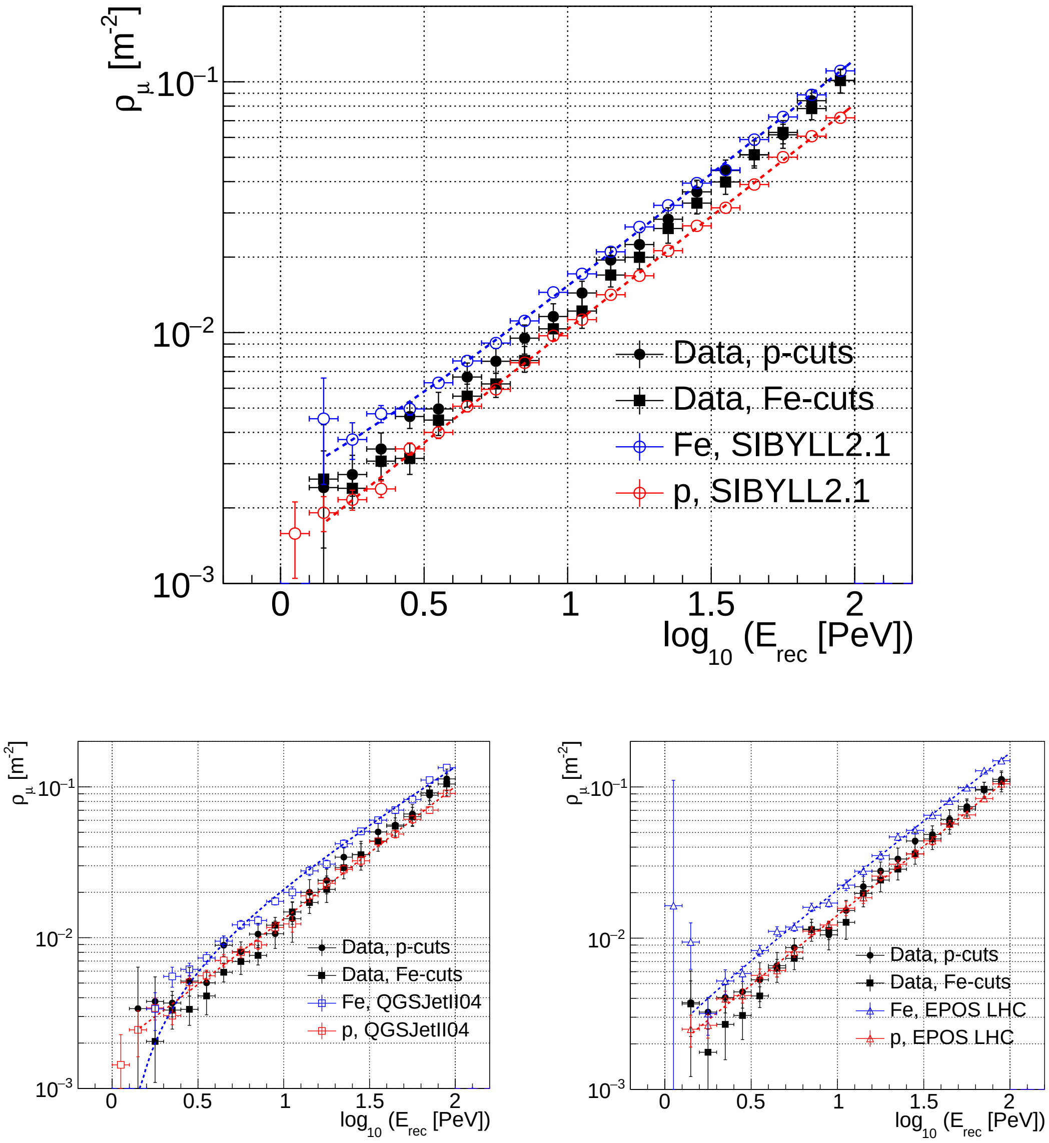

Figure 5: Muon number densities for inclined air showers as a function of E_{rec}, and for three interaction models.

- The importance of both the cuts on hit charge and distance can be visualized by deriving the results without the application of these cuts. A comparison is shown in Fig. 6.

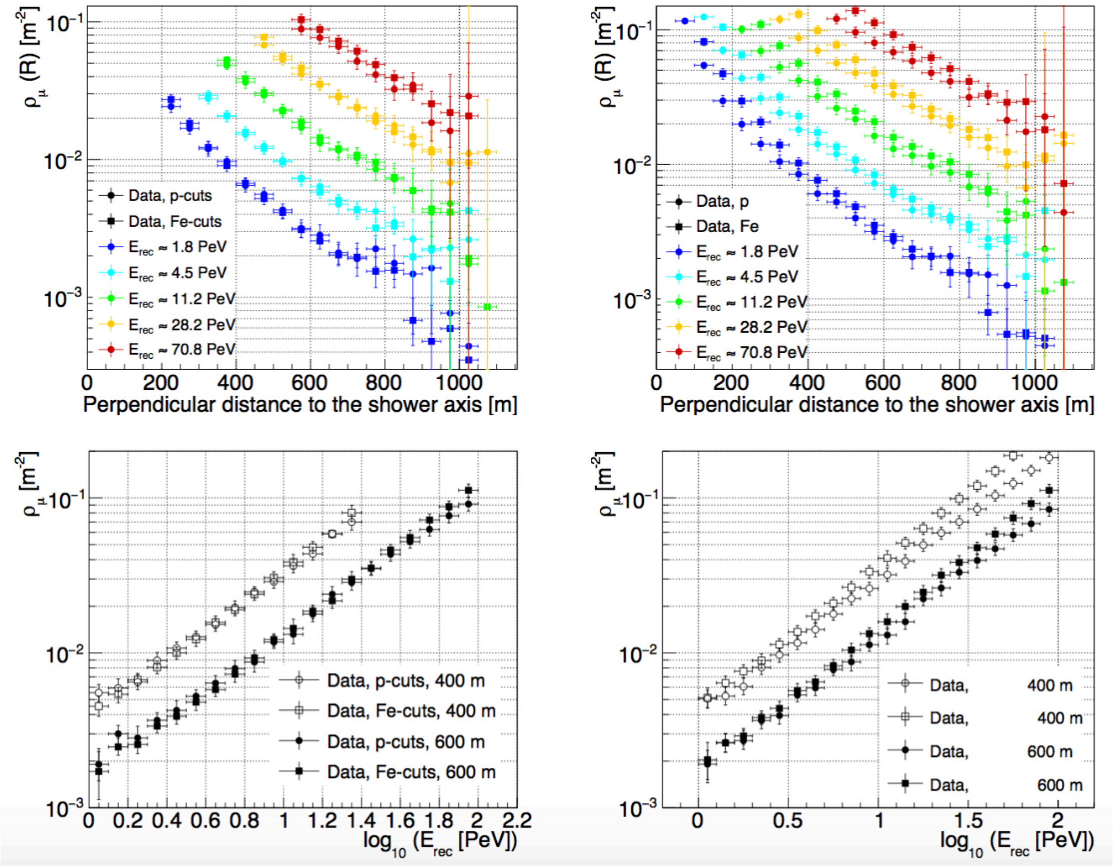

Figure 6: Results produced with the application of the cuts on hit charge and distance (left column), and without (right column).

Bibliography

[1] V. F. Hess, „Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten“, *Phys. Z.* 13 (1912) 1084

[2] Per Carlson, “A century of Cosmic Rays”, *Phys. Today* 65(2) (2012) 30

[3] P. Carlson, A. De Angelis, “Nationalism and internationalism in science: The case of the discovery of Cosmic Rays”, *Eur. Phys. J.* H 35 (2010) 309, arXiv:physics.hist-ph/1012.5068

[4] W. Kolhörster *et al.*, „Gekoppelte Höhenstrahlen“, *Naturwissenschaften* 26(35) (1938) 576

[5] P. Auger *et al.*, “Extensive Cosmic-Ray Showers”, *Rev. Mod. Phys.* 11 (1939) 288

[6] K.-H. Kampert, “Ultra-High Energy Cosmic Rays: Results and Prospects”, arXiv:astro-ph.HE/1305.2363

[7] K.-H. Kampert, M. Unger, “Measurements of the Cosmic Ray Composition with Air Shower Experiments”, arXiv:astro-ph.HE/1201.0018

[8] A. Aab *et al.*, Pierre Auger Collaboration, “Muons in air showers at the Pierre Auger Observatory: Mean number in highly inclined events”, *Phys. Rev.* D 91 (2015) 032003, arXiv:astro-ph.HE/1408.1421

[9] R. Abbasi *et al.*, IceCube Collaboration, “Cosmic Ray Composition and Energy Spectrum from 1-30 PeV Using the 40-String Configuration of IceTop and IceCube”, arXiv:astro-ph.HE/1207.3455

[10] J. Gonzalez, IceCube Collaboration, “Inclined Cosmic Ray Air Showers in IceCube”, *Proceedings of the 33th International Cosmic Ray Conference* (2013), 0973

[11] T. Antoni *et al.*, KASCADE Collaboration, “The cosmic-ray experiment KASCADE”, *Nucl. Instrum. Methods* A 513 (2003) 490

[12] W.D. Apel *et al.*, KASCADE-Grande Collaboration, “The KASCADE-Grande experiment”, *Nucl. Instrum. Methods* A 620 (2010) 202

[13] W.D. Apel *et al.*, KASCADE-Grande Collaboration, "The KASCADE-Grande energy spectrum of Cosmic Rays and the role of hadronic interaction models", *Advances in Space Research* 53 (2014) 1456

[14] W.D. Apel *et al.*, KASCADE-Grande Collaboration, "KASCADE-Grande measurements of energy spectra for elemental groups of Cosmic Rays", *Astropart. Phys.* 47 (2013) 54

[15] J. Abraham *et al.*, Pierre Auger Collaboration, "Properties and performance of the prototype instrument for the Pierre Auger Observatory", *Nucl. Instrum. Methods* A 523 (2004) 50

[16] I. Allekotte *et al.*, Pierre Auger Collaboration, "The Surface Detector System of the Pierre Auger Observatory", *Nucl. Instrum. Meth.* A 586 (2008) 409, arXiv:astro-ph/0712.2832

[17] A. Aab *et al.*, Pierre Auger Collaboration, "The Pierre Auger Observatory: Contributions to the 33rd International Cosmic Ray Conference (ICRC 2013)", arXiv:astro-ph.HE/1307.5059v1

[18] R. M. Baltrusaitis *et al.*, Fly's Eye Collaboration, "The Utah Fly's Eye Detector", *Nucl. Instrum. Methods* A 240 (1985) 410

[19] D. J. Bird *et al.*, Fly's Eye Collaboration, "Detection of a Cosmic Ray with measured energy well beyond the expected signal cutoff due to Cosmic Microwave Radiation", *Astrophys. J.* 441 (1995) 144, arXiv:astro-ph/9410067

[20] P. Biermann and G. Sigl, "Introduction to Cosmic Rays", *Lect. Notes Phys.* 576 (2001) 1, arXiv:astro-ph/0202425

[21] M. Kachelrieß, "Lecture Notes on High Energy Cosmic Rays", arXiv:astro-ph/0801.4376

[22] M. Bustamante *et al.*, "High-energy cosmic-ray acceleration", https://cdsweb.cern.ch/record/1249755/files/p533.pdf, accessed on 2016/04/07

[23] M. G. Baring, "Diffusive Shock Acceleration: the Fermi Mechanism", arXiv:astro-ph/9711177

[24] A. M. Hillas, "The Origin of Ultra-High-Energy Cosmic Rays", *Ann. Rev. Astron. Astrophys.* 22 (1984) 425

[25] "Ultra-High Energy Cosmic Rays", http://hires.physics.utah.edu/reading/uhecr.html, accessed on 2016/04/08

[26] P. K. F. Grieder, "Extensive Air Showers", Springer-Verlag Berlin Heidelberg 2010, ISBN 978-3-540-76940-8

[27] T. Antoni *et al.*, KASCADE Collaboration, “Electron, muon, and hadron lateral distributions measured in air showers by the KASCADE experiment”, *Astropart. Phys.* 14 (2001) 245, arXiv:astro-ph/0004233

[28] T. Stanev, “High Energy Cosmic Rays”, Springer Verlag Berlin Heidelberg New York 2004, ISBN 3-540-40653-0

[29] T. K. Gaisser, “Cosmic Rays and Particle Physics”, Cambridge University Press (1990), ISBN 0-521-33931-6

[30] J. Blümer, R. Engel, and J. R. Hörandel, “Cosmic Rays from the Knee to the Highest Energies”, *Progress in Particle and Nuclear Physics* 63 (2009) 293, arXiv:astro-ph.HE/0904.0725

[31] J. Matthews, “A Heitler model of extensive air showers”, *Astropart. Phys.* 22 (2005) 387

[32] D. F. Torres and L. A. Anchordoqui, “Astrophysical origins of ultrahigh energy cosmic rays”, *Rept. Prog. Phys.* 67 (2004) 1663, arXiv:astro-ph/0402371

[33] K. A. Olive *et al.*, Particle Data Group, “Review of Particle Physics”, *Chin. Phys.* C 38 (2014) 090001

[34] R.C. Weast, “Handbook of Chemistry and Physics”, *The Chemical Rubber Company, Cleveland (USA)*, F141 (1986)

[35] D. Kümpel, “Multivariate Search for a Directional Excess of EeV Photons with the Pierre Auger Observatory”, PhD thesis (2011), Bergische Universität Wuppertal

[36] A. A. Watson, “The discovery of Cherenkov radiation and its use in the detection of extensive air showers ”, arXiv:physics.hist-ph/1101.4535

[37] J.Ridky, D.Nosek, P.Travnicek, P.Necesal, “Prompt muons in extended air showers”, arXiv:astro-ph/0706.2145

[38] L.V. Volkova and O. Saavedra, “Prompt muon production in cosmic rays”, *Astropart. Phys.* 32 (2009) 136

[39] A. Fedynitch, Ralph Engel, Thomas K. Gaisser, Felix Riehn, Todor Stanev, “Calculation of conventional and prompt lepton fluxes at very high energy”, arXiv:hep-ph/1503.00544

[40] R. Abbasi *et al.*, IceCube Collaboration, ‘ÌceTop: The surface component of IceCube”, *Nucl. Instrum. Methods* A 700 (2012) 188, arXiv:astro-ph.IM/1207.6326

[41] https://gallery.icecube.wisc.edu, internal web page, accessed on 2016/01/15

[42] E. Andres *et al.*, “The AMANDA Neutrino Telescope: Principle of Operation and First Results”, *Astropart. Phys.* 13 (2000) 1, arXiv:astro-ph/9906203

[43] R. Abbasi *et al.*, IceCube Collaboration, "The Energy Spectrum of Atmospheric Neutrinos between 2 and 200 TeV with the AMANDA-II Detector", *Astropart. Phys.* 34 (2010) 48, arXiv:astro-ph.HE/1004.2357

[44] R. Abbasi *et al.*, IceCube Collaboration, "The IceCube data aquisition system: Signal capture, digitization and timestamping", *Nucl. Instrum. Meth.* A 601 (2009) 294, arXiv:physics.ins-det/0810.4930

[45] http://www.psl.wisc.edu/projects/large/icecube/more-icecube/dom, accessed on 2016/01/15

[46] R. Abbasi *et al.*, IceCube Collaboration, "Calibration and Characterization of the IceCube Photomultiplier tube", *Nucl. Instrum. Methods* A 618 (2010) 139, arXiv:astro-ph.IM/1002.2442

[47] L. Demirörs *et al.*, IceCube Collaboration, "IceTop tank response to muons", *Proceedings of the 30th International Cosmic Ray Conference* 5 (2007), 1059

[48] D. Heck *et al.*, "CORSIKA : A Monte Carlo Code to Simulate Extensive Air Showers", *Report FZKA* 6019 (1998), Forschungszentrum Karlsruhe;

[49] D. Heck, T. Pierog, "Extensive Air Shower Simulation with CORSIKA: A User's Guide", https://web.ikp.kit.edu/corsika/usersguide/usersguide.pdf, accessed on 2016/01/15

[50] A. Fassò *et al.*, "The FLUKA code: present applications and future developments", arXiv:physics.comp-ph/0306162

[51] R. S. Fletcher, T. K. Gaisser, P. Lipari and T. Stanev, "SIBYLL: An event generator for simulation of high energy cosmic ray cascades", *Phys. Rev.* D 50 (1994) 5710

[52] E.-J. Ahn, R. Engel, T. K. Gaisser, P. Lipari and T. Stanev, "Cosmic ray interaction event generator SIBYLL2.1", *Phys. Rev.* D 80 (2009) 094003

[53] E.-J. Ahn, R. Engel, T. K. Gaisser, P. Lipari, F. Riehn and T. Stanev, "LHC Update of the Hadronic Interaction Model SIBYLL2.1", *Proceedings of the 33rd International Cosmic Ray Conference* (2013) 0803, http://www.cbpf.br/~icrc2013/papers/icrc2013-0803.pdf, accessed on 2016/01/15

[54] S. Ostapchenko, "Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: I. QGSJet-II model", *Phys. Rev.* D 83 (2011) 014018, arXiv:hep-ph/1010.1869

[55] T. Pierog, Iu. Karpenko, J. M. Katzy, E. Yatsenko and K. Werner, "EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider", *Phys. Rev.* C 92 (2015) 034906, arXiv:hep-ph/1306.0121

[56] T. Pierog, "Effect of LHC data on Extensive Air Shower Development", *Proceedings of the 48th Rencontres de Moriond on Very High Energy Phenomena in the Universe*, http://moriond.in2p3.fr/Proceedings/2013/Moriond_VHEPU_2013.pdf (2013) 39, accessed on 2016/01/18

[57] Personal communication with Dieter Heck, Institute for Nuclear Physics Karlsruhe Institute of Technology, 2016/04/20

[58] R. Brun, F. Rademakers, "An Object Oriented Data Analysis Framework", *Nucl. Instrum. Methods* A 389 (1997) 81, https://root.cern.ch/publication-list

[59] https://wiki.icecube.wisc.edu/index.php/Real_atmosphere_for_CORSIKA, internal web page, accessed on 2016/03/08

[60] Geant4 Collaboration, User documentation and User's Guide, http://geant4.cern.ch/support/userdocuments.shtml (2014), accessed on 2016/01/15

[61] S. Agostinelli *et al.*, Geant4 Collaboration, "GEANT4 - a simulation toolkit", *Nucl. Instrum. Methods* A 506 (2003) 250

[62] K. Amako *et al.*, Geant4 Collaboration, "Geant4 developments and applications", *IEEE Trans. on Nucl. Sc.* 53 No.1 (2006) 270

[63] M. G. Aartsen *et al.*, IceCube Collaboration, "Measurement of South Pole ice transparency with the IceCube LED calibration system", *Nucl. Instrum. Methods* A 711 (2013) 73, arXiv:astro-ph.IM/1301.5361

[64] S. Klepser *et al.*, IceCube Collaboration, "Lateral Distribution of Air Shower Signals and Initial Energy Spectrum above 1 PeV from IceTop", *Proceedings of the 30th International Cosmic Ray Conference* 4 (2007), 0858

[65] S. Klepser, "Reconstruction of Extensive Air Showers and Measurement of the Cosmic Ray Energy Spectrum in the Range of 1 – 80 PeV at the South Pole", PhD thesis (2008), Humboldt-Universität zu Berlin, http://edoc.hu-berlin.de/dissertationen/klepser-stefan-2008-06-24/PDF/klepser.pdf, accessed on 2016/01/22

[66] R. Abbasi *et al.*, IceCube Collaboration, "All-particle cosmic ray energy spectrum measured with 26 IceTop stations", arXiv:astro-ph.HE/1202.3039

[67] T. Feusels, "Measurement of cosmic ray composition and energy spectrum between 1 PeV and 1 EeV with IceTop and IceCube", PhD thesis (2013), Gent University, as cited in [70].

[68] Katherine Rawlins, IceCube Collaboration, "The Effect of Snow Accumulation on Signals in IceTop", *Proceedings of the 33th International Cosmic Ray Conference* 9 (2013) 1106, arXiv:astro-ph.HE/1309.7006

[69] M. G. Aartsen *et al.*, IceCube Collaboration, "Measurement of the cosmic ray energy spectrum with IceTop-73", arXiv:astro-ph.HE/1307.3795

[70] A. A. K. Haj Ismail, "Measurement of the Cosmic Ray Energy Spectrum between 500 TeV and 100 PeV with IceTop", PhD thesis (2015), Gent University, https://biblio.ugent.be/publication/5852240, accessed on 2016/01/25

[71] https://wiki.icecube.wisc.edu/index.php/IceTop_SLC_Charge_Calibration, accessed on 2016/09/18

[72] Internal communication with T. Feusels (2014)

[73] J.R. Hörandel, "On the knee in the energy spectrum of cosmic rays", *Astropart. Phys.* 19 (2003) 193; arXiv:astro-ph/0210453

[74] Thomas K. Gaisser, "Spectrum of cosmic-ray nucleons, kaon production and the atmospheric muon charge ratio", *Astropart. Phys.* 35 (2012) 801, arXiv:astro-ph.HE/1111.6675v2

[75] S. Klepser, F. Kislat, A. van Overloop, "A Lateral Distribution Function and Fluctuation Parametrisation for IceTop", IceCube Internal Report icecube/200702001 (2007), as cited in [77]

[76] L. Rädel, C. Wiebusch, "Calculation of the Cherenkov light yield from electromagnetic cascades in ice with Geant4", *Astropart. Phys.* 44 (2013) 102, arXiv:astro-ph.IM/1210.5140

[77] A. Lucke, 'Analysis of Signals in the IceTop Detector to identify Muons in Extensive Air Showers", Diploma thesis (2008), Desy, https://astro.desy.de/neutrino_astronomy/icecube/publications/diploma_and_master_theses/e597/infoboxContent752/adams-thesis.pdf

[78] A. van Overloop, IceCube Collaboration, "Simulation of IceTop VEM calibration and the dependency on the snow layer", *Proceedings of the 32nd International Cosmic Ray Conference* (2011), arXiv:astro-ph.HE/1111.2735

[79] https://wiki.icecube.wisc.edu/index.php/Combined_3_year_spectrum:_IceTop-alone_analysis_details, internal web page, accessed on 2017/11/16

[80] https://wiki.icecube.wisc.edu/index.php/Muon_LDF, internal web page, accessed on 2019/01/24

[81] T. K. Gaisser, "Spectrum of cosmic-ray nucleons, kaon production, and the atmospheric muon charge ratio", *Astropart. Phys.* 35 (2012) 801, arXiv:astro-ph.HE/1111.6675

[82] https://wiki.icecube.wisc.edu/index.php/IceTop_SLC_Times, accessed on 2018/07/24

Glossary

acceptance, 36
AMANDA, 17
Amundsen-Scott-Station, 15
angular resolution, 48
ankle, 11
atmosphere, 28
atmospheric
  density, 28
  ground pressure, 28
  muons, 10
  temperature profile, 28
attenuation length of electrons, 44
ATWD (Analog Transient Waveform Digitizer), 19
average charge $\langle Q(R_{ij}) \rangle_{s/b}$, 61

background, 60
background window, 53
bremsstrahlung, 10, 57

conversion distributions $C_{\mathrm{ch,\,h}}\,(R)$, 77
Calibration Trigger, 23
center of gravity $\vec{x}_{\mathrm{COG}}$, 41
charge
  calibration, 23
  likelihood $\mathcal{L}_{\mathrm{q}}$, 42
charmed particles, 10
Cherenkov light, 15
cleaning, 39
cluster, 40
conventional muons, 10
core resolution, 48
correlated noise, 37
CORSIKA, 29
Cosmic Rays, 1, 5
cut on hit charge Q_{min}, 62
cut on hit distance R_{min}, 62

DAQ (Data Acquisition System), 19
dark noise rate, 37
decay processes, 8
DeepCore, 17
detector simulation, 33
differential flux $F(E)$, 11
direct measurements, 11
DOM (Digital Optical Module), 16
DOMLaunch, 19
double logarithmic parabola, 42

early side, 45
EAS (extensive air shower), 1, 8
edge-clipping muons, 24
electromagnetic air shower component , 10
elemental composition, 13
energy spectrum, 11
EPOS LHC, 30
event selection, 45
expected charge value at particular distances R ($S_{\mathrm{exp}}\,(R)$), 68

fADC (fast Analog-to-Digital Converter) , 19
Fermi acceleration, 5
final event sample, 45
first order, 6
flux model, 32

gain, 19
Geant4 simulation toolkit, 36
geometric containment, 49

- hadronic
    - air shower component, 8
    - interaction models, 30
- Heitler model, 10
- HG (High Gain), 19
- Hillas criterion, 6
- HiRes, 11
- hit, 19
- HLC hit (Hard Local Coincidence), 21

- I3CorsikaInjector, 34
- I3TopSimulator, 34
- IceCube detector, 15
- IceRec, 33
- IceSim, 33
- IceTop, 16
- IceTopSTA5 filter, 23
- IceTray, 33
- ICL (IceCube Laboratory), 16
- in-fill, 19
- in-fill-stations, 19
- indirect measurements, 11
- InIce, 16
- integral flux $\Psi(E)$, 13
- interaction length λ, 8
- interstellar clouds, 5
- ionization, 10
- IT81-2, 21

- knee, 11

- Larmor radius, 12
- late side, 45
- lateral
    - background distribution of weighted total charge after cuts $B_{\mathrm{ch}}^{\prime\,\mathrm{true}}$, 69
    - background distribution of weighted total charge before cuts $B_{\mathrm{ch}}^{\mathrm{true}}$, 69
    - background distribution of weighted total number of hits after cuts $B_{\mathrm{h}}^{\prime\,\mathrm{true}}$, 69
    - background distribution of weighted total number of hits before cuts $B_{\mathrm{h}}^{\mathrm{true}}$, 69
    - distribution of total charge after cuts $L_{\mathrm{ch}}^{\prime\,\mathrm{Data}}(R)$, 71
    - distribution of total number of hits after cuts $L_{\mathrm{h}}^{\prime\,\mathrm{Data}}(R)$, 71
    - distribution of weighted total charge after cuts $L_{\mathrm{ch}}^{\prime\,\mathrm{true}}$, 69
    - distribution of weighted total charge before cuts $L_{\mathrm{ch}}^{\mathrm{true}}(R)$, 69
    - distribution of weighted total number of hits after cuts $L_{\mathrm{h}}^{\prime\,\mathrm{true}}$, 69
    - distribution of weighted total number of hits before cuts $L_{\mathrm{h}}^{\mathrm{true}}(R)$, 69
    - signal distribution of weighted total charge after cuts $S_{\mathrm{ch}}^{\prime\,\mathrm{true}}$, 69
    - signal distribution of weighted total charge before cuts $S_{\mathrm{ch}}^{\mathrm{true}}$, 69
    - signal distribution of weighted total number of hits after cuts $S_{\mathrm{h}}^{\prime\,\mathrm{true}}$, 69
    - signal distribution of weighted total number of hits before cuts $S_{\mathrm{h}}^{\mathrm{true}}$, 69
- LG (Low Gain), 19
- likelihood minimization, 40, 42
- local coincidences, 21
- loudest station, 49

- Minimum Bias Trigger, 23
- muon number density
    - estimator in exp. data $\rho_{\mathrm{Data}}^{\mathrm{est}}(R_{ij})$, 78
    - estimator in simulation $\rho_{\mathrm{true}}^{\mathrm{est}}(R_{ij})$, 78
    - in exp. data $\rho_{\mathrm{Data}}(R_{ij})$, 75, 82
- muon number estimator
    - $\mu_{\mathrm{ch,\,data}}^{\mathrm{est}}(R)$ using measured charges, 71
    - $\mu_{\mathrm{ch,\,true}}^{\mathrm{est}}(R)$ using simulated charges, 69
    - $\mu_{\mathrm{h,\,data}}^{\mathrm{est}}(R)$ using measured hits, 71
    - $\mu_{\mathrm{h,\,true}}^{\mathrm{est}}(R)$ using simulated hits, 69
- muon problem, 3
- muonic air shower component, 10

- no-hit likelihood $\mathcal{L}_0$, 43
- noise, 37

- observation level, 29

- pair production, 10, 57

passing rates, 51
PE, 19
perpendicular distance to the shower axis , 40
phase velocity of light c_p, 15
photoelectrons, 19
Pierre Auger collaboration 31
PMT (Photomultiplier tube), 17
poly-gonato model, 12
primary particle, 8
prompt muons, 10

QGSJetII04, 30
quality cuts, 45

reconstruction, 39, 40
refraction index n_p, 15
resamples, 35
resampling radius $\mathrm{r}_{\mathrm{sim}}$, 34
robustness, 79
Run, 20

saturation likelihood $\mathcal{L}_{\mathrm{s}}$, 43
second
  knee, 11
  order, 5
secondary particles, 8
shower
  axis, 8
  core, 8
  front, 8
  periphery, 64
SIBYLL2.1, 30
signal, 60
signal lateral distribution function, 42
signal window, 53
Simple Majority Trigger, 23
simulated muon number $\mu^{\mathrm{true}}(R_{ij})$, 70
simulation, 29
SLC hit (Soft Local Coincidence), 21
SLC hit charge recalibration, 53
snow
  correction, 44
  height estimation, 25
  height measurement, 25
  on IceTop tanks, 25
sources, 5
SPE (Single Photoelectron), 19
spectral index γ, 11
standard-stations, 19
stations, 16
subshowers, 10

tanks, 16
time according to the reconstructed shower front t_{reco}, 53
time of the launch t_{launch}, 53
time residuals, 45, 53
timestamp, 21
timing likelihood $\mathcal{L}_{\mathrm{t}}$, 43

Ultra High Energy Cosmic Rays, 6
uncorrelated particles, 53

VEM (Vertical Equivalent Muon), 23
VEM calibration, 23
VEMCal Launch, 23
vuvuzela, 37

waveform, 19
wavefront, 15
weighting, 32