DESIGN, CONSTRUCTION AND COMMISSIONING
OF THE EMMA EXPERIMENT

JUHO SARKAMO

REPORT SERIES IN PHYSICAL SCIENCES
Report No. 92 (2014)
DESIGN, CONSTRUCTION AND COMMISSIONING OF

THE EMMA EXPERIMENT

JUHO SARKAMO

Department of Physics
University of Oulu
Finland

Academic dissertation to be presented, with the permission of the Faculty of Science of the University of Oulu, for public discussion in the Auditorium TS101, Linnanmaa, on November 7th, 2014, at 12 o’clock noon.
**Opponent**
Dr. Bruno Alessandro  
National Institute for Nuclear Physics, Italy

**Reviewers**
Dr. Andreas Haungs  
Karlsruhe Institute of Technology, Germany  
Dr. Jacek Szabelski  
National Centre for Nuclear Research, Poland

**Custos**
Prof. Kalevi Mursula  
University of Oulu, Finland

ISSN 1239-4327

Juvenes Print – Suomen Yliopistopaino Oy  
Oulu 2014
Sarkamo, Juho.
Design, construction and commissioning of the EMMA experiment

Department of Physics and Oulu Southern Institute, University of Oulu, Finland. Report No. 92 (2014)

Abstract

The work describes the design, construction and commissioning of the underground cosmic-ray experiment Experiment with MultiMuon Array (EMMA). The experiment is built into the Pyhäsalmi mine, in the town of Pyhäjärvi, Finland. The aim of EMMA is to determine the elemental composition of cosmic rays at an energy region around 4 PeV, the energy region called the 'knee' region. This is achieved by measuring the lateral density distribution of high-energy muons originating from Extensive Air Showers (EAS).

The design calculations for the EMMA experiment, which are based on the use of the parametrization of the lateral density distribution of muons, the method of shower reconstruction, and the energy and composition indicators, are presented. A strategy for reconstructing the composition of the cosmic rays is presented and it demonstrates the potential of applying unfolding techniques to the EMMA data. The effect of an array extension on the performance of EMMA is studied.

The hardware used in the EMMA experiment is presented starting with an overview of the array and its detector stations. The EMMA array employs three different particle detectors, for which the main technical properties are given, and their use in the EMMA array is presented. A description of the infrastructure of the experiment is given and the rock overburden at the EMMA site at the depth of 80 metres is documented.

The work contains the latest analysis of EAS data recorded by the tracking detectors of the experiment, which demonstrates that the experiment is taking data as planned and that the data are according to EAS physics expectations. Methods for event selection and tracking efficiency correction are presented, after which the analysis results of measured track multiplicity spectra are given. The shape of the recorded multiplicity spectrum indicates that the simplest model of a knee-like spectrum with a pure proton composition can not explain the data and that further analysis of the spectrum is required.

Keywords: EMMA, Experiment with MultiMuon Array, underground physics, Pyhäsalmi mine, cosmic rays, muons, high-energy muons, knee, composition
I started my studies in underground physics during the year 2000 and I have been working for the EMMA experiment from the year 2005 onwards. The work has been carried out at the Oulu Southern Institute and at the Department of Physics of the University of Oulu. Through the years, I have been privileged to work with so many happy and hard working people that it is difficult to remember you all. My warmest gratitude
go to all former and current members in CUPP and EMMA, especially Dr. Timo Enqvist for providing the possibility to work for the experiment and Dr. Wladyslaw Trzaska for all the guidance he has given through the years and for providing the possibility to work in the ALICE experiment in CERN. I also want to express my gratitude to Dr. Juha Peltoniemi, Dr. Pasi Kuusiniemi, Dr. Petteri Keränen, Dr. Tomi Räihä, Mr. Changquan Shen, Mr. Jari Joutsenvaara, Mr. Maciej Slupecki, Mr. Kai Loo, Mr. Joonas Karjalainen, Mr. Tuomo Kalliokoski, Mr. Johannes Hisa, Mr. Antto Virkajärvi, Mr. Janne Narkilahti, Dr. Marko Aittola and Mrs. Johanna Kutuniva. My best wishes go also to the former CUPP members: Mr. Mikko Mutanen, Mr. Mika Lehtola, Mr. Teppo Jämsén, Mr. Sami Nurmenniemi, Mr. Matti Vaittinen, Ms. Annika Mattila and Ms. Hanna Remes.

The finalisation of this thesis would not have been possible without the many useful comments and corrections I received. Thanks to Dr. Wladyslaw Trzaska, Prof. Kalevi Mursula, Dr. Timo Enqvist, Dr. Pasi Kuusiniemi, Dr. Tomi Räihä and Mr. Maciej Slupecki for reading the manuscript. Thanks also to Mr. Kai Loo for the help in the commissioning analysis. Thanks to the pre-examiners Dr. Jacek Szabelski and Dr. Andreas Haungs whose comments finalised the work.

The biggest thanks go to my supervisors Dr. Wladyslaw Trzaska and Prof. Kalevi Mursula. Thanks to the Oulu Southern Institute, Dr. Eelis Kokko and Dr. Timo Enqvist for the continuing support and thanks to the Physics Department of Oulu University for offering me the workplace for the final years.

I would also like to thank the colleagues from the Institute of Nuclear Research, especially L. Bezrukov, V. Petkov, L. Inzhechik and B. Lubsandorzhiev. I also thank all the members of the Scientific Advisory Boards of CUPP and EMMA, especially K.H. Kampert and F. Dydak for useful suggestions and help during the EMMA design phase. Thanks to A. Haungs and F. Cossavella for their hospi-
tality when I visited the KASCADE experiment. Greetings to T. Karavicheva, A. Mayevskaya and F. Krizek in ALICE-T0, as well as to the rest of the ALICE collaboration.

The financial support received from Jenny and Antti Wihuri Foundation, Kerttu Saalasti Foundation, Oulu University Scholarship Foundation, Faculty of Science of the University of Oulu, Vilho, Yrjö and Kalle Väisälä Foundation, and Graduate School of Particle and Nuclear Physics is acknowledged.

Finally, the warmest thanks go to my family and friends.

Oulu, October 8th, 2014

Juho Sarkamo
Contents

Abstract
Acknowledgements
Contents
1. Introduction ................................................. 1
   1.1 Experiment with MultiMuon Array (EMMA) .......... 1
   1.2 Author’s contribution to EMMA ....................... 2
   1.3 Dissertation structure ................................. 3
2. Scientific background for EMMA ............................. 5
   2.1 Energy spectrum and composition of cosmic rays .......... 5
      2.1.1 Energy spectrum and composition below $10^{15}$ eV . 5
      2.1.2 Energy spectrum and composition above $10^{15}$ eV . 7
      2.1.3 Astrophysical interpretations ...................... 9
   2.2 Extensive Air Showers ................................ 12
   2.3 Experimental methods and results of the knee region .... 15
      2.3.1 Compilation of composition measurements ........ 15
      2.3.2 Selected experiments, methods and results ....... 15
   2.4 Conclusions for EMMA ................................... 20
3. EMMA design ............................................. 21
   3.1 Background to the design work ......................... 21
   3.2 Array size and location ................................ 22
   3.3 Muon yields and composition sensitivity .............. 24
   3.4 Layout design ........................................... 27
      3.4.1 Shower reconstruction ............................. 27
      3.4.2 Array geometries and shower axis position reconstruction . 30
      3.4.3 Optimization of station placements ............. 32
      3.4.4 Layout at 80 metres depth ....................... 34
   3.5 Composition reconstruction strategy .................. 36
   3.6 Array extension ......................................... 44
   3.7 Summary and conclusion .................................. 56
4. EMMA experiment ......................................... 57
   4.1 Overview of EMMA ....................................... 57
   4.2 Drift chambers ......................................... 59
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.1</td>
<td>Technical description</td>
<td>59</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Calibration of drift chambers</td>
<td>61</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Tracking detector</td>
<td>63</td>
</tr>
<tr>
<td>4.3</td>
<td>SC16 scintillators</td>
<td>65</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Technical description</td>
<td>65</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Performance estimates for the scintillator setup</td>
<td>68</td>
</tr>
<tr>
<td>4.4</td>
<td>Limited Streamer Tubes</td>
<td>72</td>
</tr>
<tr>
<td>4.5</td>
<td>Underground infrastructure</td>
<td>73</td>
</tr>
<tr>
<td>4.6</td>
<td>Rock overburden</td>
<td>74</td>
</tr>
<tr>
<td>4.7</td>
<td>Summary</td>
<td>77</td>
</tr>
<tr>
<td>5.</td>
<td>EMMA commissioning study</td>
<td>79</td>
</tr>
<tr>
<td>5.1</td>
<td>Measurement configuration and data taking</td>
<td>79</td>
</tr>
<tr>
<td>5.2</td>
<td>Event reconstruction</td>
<td>80</td>
</tr>
<tr>
<td>5.3</td>
<td>Data set</td>
<td>80</td>
</tr>
<tr>
<td>5.4</td>
<td>Simulations and efficiency correction</td>
<td>84</td>
</tr>
<tr>
<td>5.5</td>
<td>Shower direction reconstruction performance</td>
<td>92</td>
</tr>
<tr>
<td>5.6</td>
<td>Multiplicity spectra</td>
<td>94</td>
</tr>
<tr>
<td>5.7</td>
<td>Hypothesis tests on the shape of the spectrum</td>
<td>101</td>
</tr>
<tr>
<td>5.8</td>
<td>Conclusions</td>
<td>104</td>
</tr>
<tr>
<td>6.</td>
<td>Discussion and outlook</td>
<td>105</td>
</tr>
<tr>
<td>7.</td>
<td>Summary of thesis</td>
<td>109</td>
</tr>
</tbody>
</table>
1. Introduction

This work describes the design, construction and commissioning of the underground experiment Experiment with MultiMuon Array (EMMA) situating in the tunnels of the Pyhäsalmi mine in the town of Pyhäsalmi, Finland. The experiment is designed and built to measure the cosmic-ray composition in the so-called 'knee' region of the energy spectrum of high-energy cosmic rays.

1.1 Experiment with MultiMuon Array (EMMA)

Primary cosmic rays are fully ionized nuclei that propagate in interstellar space and arrive to Earth’s vicinity or the atmosphere, where they are measured. At low energies, the energy spectrum of cosmic rays follows a power law of about $dN/dE \propto E^{-2.7}$ and the elemental composition has been found to be close to the elemental composition of the Solar System. At a primary energy above $10^{15}$ eV, a distinct feature in the all-particle primary cosmic-ray energy spectrum is the knee, where the spectrum steepens following a power law of about $dN/dE \propto E^{-3.1}$ and where the elemental composition is observed to change. The reason behind the feature has long been studied and most common explanations for the knee are related to the acceleration of particles in the galactic supernovae or leakage of cosmic rays outside the galaxy while propagating from their source but the explanation of the knee remains open and further measurements of cosmic rays in the knee region using different approaches are needed (see Chapter 2).

A pioneering experiment situated in the underground laboratory in Pyhäsalmi is the Experiment with MultiMuon Array (EMMA), which aims to study the composition of primary cosmic rays at the knee region. The experiment measures the high-energy muon component of Extensive Air Showers (EAS), i.e. particle cascades which are created as the primary cosmic rays interact with the nuclei of Earth’s atmosphere. The EMMA experiment is the main activity of scientific research carried out in Centre for Underground Physics in Pyhäsalmi (CUPP). First concepts for a multi-muon experiment in Pyhäsalmi were born in the late 1990s, the design for the EMMA experiment was created in 2005–2007 and a fully devoted project for the construction of the experiment started in 2009. As of 2014, the construction of the experiment is nearly complete and the experiment is
already taking data in a partial configuration for the analysis of EAS and primary cosmic-ray spectrum.

1.2 Author’s contribution to EMMA

The author of this thesis has contributed to different aspects of the experiment working within the EMMA collaboration from its start in 2005. The author has executed design calculations for the array’s geometry and created methodology whereby the EMMA experiment can fulfill its aim: the measurement of cosmic-ray composition (see Chapter 3). The author has contributed to the successful realization of the experiment, that is to the planning, to the construction and to the operation of the underground detector stations, which measure the muon yields of Extensive Air Showers (Chapter 4). The author has committed analyses for the commissioning of EMMA and has found the first physics result (Chapter 5).

During 2005–2006 the collaboration worked for a proposal for a multi-muon experiment in Pyhäsalmi mine, which, after feedback from the Scientific Advisory Board of CUPP, was modified and the plans for EMMA experiment in its fulfilled form were set. Author’s main input to the planning of EMMA consists of work carried out mainly by numerical simulations introducing array’s geometry and the principles and methods by which the challenging aim of the experiment can be achieved. The studies conducted for the proposal stage have been later updated for an array extension design. Besides the numerical analyses for the EMMA array, the author has gained knowledge in the field of shower reconstruction from participation in the development of numerical simulation and analysis methods for the CARPET array in Baksan, Russia.

The construction of the first array stations of EMMA begun in 2006. The author participated in the investigations of the possible underground sites for the array, in the manual labour of constructing the underground detector stations and associated infrastructure and in the day-to-day operations in running the underground and surface laboratories in Pyhäjärvi. The author participated in the first tests of the delay line response of drift chambers and in devising calibration and efficiency determination schemes for them. The author has executed first demonstrative tracking analysis results with data taken in surface laboratory and underground test installations. The author has created software for delay matching routines and logging of environmental data.

A high-granularity scintillator setup was designed in order to improve EMMA’s capability in measuring high muon densities. The author participated in defining the functional goals and the technical requirements for the scintillator detectors, in the measurements with first scintillator prototypes and in the determination of the time resolution of the detectors. The author has investigated the expected performance of the setup by numerical estimates of shower arrival direction determination and guiding a M.Sc. thesis related to the muon number measurement capability of the setup.

Since 2010, the author’s main contribution has been in data analysis. The author has created principles and software for data quality monitoring. The main results in the analysis of data have been the invention of the method for selecting the EAS events and a method of tracking efficiency correction. These have been
invented by analysing the first data of underground tracking detectors and by executing the EAS simulations necessary for a successful interpretation of the recorded data. The analysis of data recorded in 2013 shows that the recorded muon multiplicity spectrum is consistent with cosmic-ray composition changing towards heavy elements in the knee region and that further data taking is required to investigate the spectrum further.

The author has presented the idea, the status and the analysis results of the experiment in international and national schools and conferences and during research visit to Karlsruhe Institute of Technology, and also to the Scientific Advisory Board of EMMA having regular meetings close-by the Pyhäsalmi mine. These efforts have helped to strengthen the collaboration and bring EMMA to the interest of the international community.

The author has contributed in the commissioning, operation and analysis of the T0 detector of the ALICE experiment during 2008–2013. The work committed in CERN has been synergistic to the work within EMMA: it gave the author an opportunity to get familiar with the best and largest scientific institute in the world, learn the latest software solutions and benefit from the unique scientific environment.

1.3 Dissertation structure

A scientific background to the EMMA experiment is given in Chapter 2. The chapter introduces the cosmic rays, summarizes some of the results about the measurements of cosmic-ray spectrum and composition and presents the astrophysical motivation for the studies. The phenomenology of Extensive Air Showers, as they are the physical process by which the investigation of cosmic-ray spectrum is possible, is introduced and the principle of cosmic-ray spectral measurements via the EAS measurements is explained. The chapter contains a list of experimental results about spectrum and composition in the knee region and the scientific motivation for EMMA is given in the chapter conclusions.

Chapter 3 presents the design of EMMA. The chapter summarizes the design calculations executed during 2005–2007, the period when the construction of the experiment started and explains why the EMMA experiment emerged in its current form. The geometry of the array has been planned by carrying out numerical simulations of shower reconstruction and it is concluded that the array geometry of EMMA is optimal for the composition studies. The chapter also contains an update of shower reconstruction calculations, which accommodate an array extension of EMMA.

Chapter 4 introduces the EMMA experiment. It presents the array and its detector stations. The three different particle detector types used in EMMA are introduced and the main technical parameters and properties are given. The chapter contains a description of the experimental infrastructure and documents the rock overburden at the EMMA site.

Chapter 5 describes the first physics analysis of the experiment. The analysis is about data recorded by array’s three central muon tracking detectors in 2013. The chapter presents the methods of event selection and tracking efficiency correction after which the measured muon multiplicity spectra are compared with models of
the cosmic-ray spectrum.

The methods and results presented in the work are discussed in Chapter 6 and the thesis is summarized in Chapter 7.
2. Scientific background for EMMA

The field of high-energy cosmic-ray physics aims to reveal details about astrophysical origin and propagation of cosmic rays. In this chapter, a brief overview of the methodology and results of high-energy cosmic-ray flux measurements is given together with interpretations of the results. The focus is on the knee region, i.e. the energy region, which the EMMA experiment is designed to investigate.

Section 2.1 introduces the field and describes the main features of the cosmic-ray spectrum and the physics motivation behind the science. In Section 2.2, the phenomenon of Extensive Air Showers (EAS) is presented and measurement principles of cosmic-ray spectrum and composition via measurements of EAS are explained. In Section 2.3, experiments, methods and results of the knee region are given. The chapter is concluded in Section 2.4.

2.1 Energy spectrum and composition of cosmic rays

Primary cosmic rays consist of fully ionized nuclei and originate from galactic and extragalactic sources [1–7]. Their energies extend from the $10^6$ eV range up to about $10^{20}$ eV (see Figure 2.1). The flux of cosmic rays decreases rapidly as a function of energy, yielding 1000 particles per second per square metre at $10^9$ eV, but only about 1 particle per year per square metre at about $10^{15}$ eV. Up to the energy of $10^{15}$ eV, cosmic rays are measured directly by satellite- or balloon-borne experiments. Above $10^{15}$ eV, the flux of cosmic rays is too low for the direct measurements and cosmic rays can only be measured indirectly by measuring the properties of EAS, i.e. particle cascades originating from interactions between the cosmic rays and Earth’s atmosphere.

2.1.1 Energy spectrum and composition below $10^{15}$ eV

The cosmic rays can be measured directly up to the energy of about $10^{15}$ eV. Direct experiments are typically balloon- or space-borne mass spectrometers or calorimeters able to measure the nucleorum energy and mass. The energy spectrum, as it is measured in the upper atmosphere or in the space in the vicinity of Earth, and as is evident in Figure 2.1, is characterized by a power law $dN/dE \propto E^{-\gamma}$ with a value of the spectral index $\gamma \approx -2.7$. Only the low-energy region ($E \lesssim 10$ GeV) is modulated by solar activity and deviates from this general law.
Fig. 2.1: Upper panel: The all-particle energy spectrum of cosmic rays. At low energies, the primary proton spectrum is shown. The arrowheaded lines indicate the energy region studied by the direct and the EAS experiments. Lower panel: The all-particle energy spectrum multiplied by $E^{2.5}$. Symbols correspond to different experiments. The collision energies reachable by different particle accelerators are indicated by the arrows. Both figures are taken from [3].
The elemental composition of the cosmic rays, as it is measured with the direct experiments, resembles the overall elemental composition of the solar system and each nucleus follows a power law with spectral indices $\gamma \approx -2.7$ above energies modulated by the Sun, see Figure 2.2. Protons are the most abundant nuclei (fraction of about 87%), helium is also abundant (12%) and other nuclei with charges $Z = 3 - 26$ contribute most of the rest (1%). Nuclei heavier than iron are also present but contribute in minimal amounts. The direct experiments have also revealed that abundances of certain nuclei (such as Li, B, Be) are higher in galactic cosmic rays than that in the solar system.

### 2.1.2 Energy spectrum and composition above $10^{15}$ eV

The cosmic rays above the energy of $10^{15}$ eV can only be measured by indirect means, which encompass the measurement of Extended Air Showers. EAS experiments measure the properties of these particle cascades, which originate from interactions between the cosmic rays and Earth’s atmosphere. The spectrum can be deduced from these indirect measurements, which have revealed changes in the spectral index of cosmic-ray spectrum above $10^{15}$ eV (see Figure 2.1). These features are referred to as ‘the knee’, ‘the second knee’ and ‘the ankle’.

The cosmic-ray energy spectrum steepens at the knee energy around $4 \times 10^{15}$ eV. The spectral index $\gamma$ changes from a value $\gamma \approx 2.7$ below the knee energy to a value $\gamma \approx 3.1$ above the knee energy. Besides the measurement of a steepening flux, several experiments have measured that the composition of the cosmic rays changes from a light composition (proton-like) towards a heavy composition (iron-like) in the energy interval from $10^{15}$ to $10^{17}$ eV. Results of the measurements of the knee region are discussed in more detail in Section 2.3.

The second knee at $10^{17.5}$ eV where the spectrum steepens from $\gamma \approx -3.1$ to $\gamma \approx -3.3$ has been observed by several experiments, the results are summarized in [9]. Composition measurements in the energies above $10^{17.5}$ eV, as presented in [3], indicate a decrease in the average mass of the primary cosmic rays (see Figure 2.6 in Section 2.3).

The ankle at $10^{18.5}$ eV, where the spectrum flattens from $\gamma \approx -3.3$ to $\gamma \approx -2.8$, has been recently confirmed by the experiments PAO [10, 11], HiRes [12] and Telescope Array [13]. The results of HiRes and preliminary results of Telescope Array are consistent with a proton dominant composition at the energy region of $10^{18} - 5 \times 10^{19}$ eV [12, 13] while the measurements of PAO support an interpretation of changing composition from light towards heavy nuclei [14]. A flux suppression after $10^{19.5}$ eV is observed by all the three experiments [12, 13, 15].
Fig. 2.2: Upper panel: The relative abundances of nuclei in cosmic rays as a function of nuclear charge number $Z$ around the energies of 1 GeV per nucleon. The symbols correspond to different direct experiments and are compared with the elemental composition of the solar system. Figure is taken from [3]. Lower panel: The spectra of individual nuclei as measured with different direct experiments in the energy interval $10^9 - 10^{15}$ eV. The spectra are scaled for clarity as indicated in the labels. The symbols correspond to different experiments discussed in [8] from where the figure is taken.
2.1.3 Astrophysical interpretations

Direct measurements of the cosmic-ray spectrum indicate at least three main points about their origin and propagation [3]. First, the power-law spectrum of the cosmic rays, as was presented above, indicates that the sources are not thermal. A power-law spectrum can be explained by a mechanism called diffusive first order shock acceleration where particles gain energy in cycles while passing through a shock and back. Second, the similarity of cosmic-ray composition with that of the solar system, as was also presented, indicates that the cosmic rays originate from samples of well-mixed galactic matter. Third, the overabundance of certain nuclei, namely Li, Be, B and Sc, Ti, V, Cr, Mn, in the spectrum in comparison to the abundance in the Solar System, can be explained if the cosmic-ray paths are confined within the galactic disk and its halo by galactic magnetic fields. In such a case, the nuclei can be explained as being the spallation products of interactions between the propagating cosmic rays and the interstellar medium.

The most common scenario entailing the points above is that the high-energy cosmic rays are originated from galactic supernova remnants (SNR), in which the interstellar matter can be accelerated in supernova shock fronts resulting in the injection of high-energy cosmic rays. The accelerated particles propagate in the galaxy, confined by its magnetic field, and arrive to Earth, where they are observed. Since supernova shock acceleration can only accelerate protons up to $10^{14} \sim 5 \times 10^{15}$ eV and only particles with energies $E \lesssim 10^{18}$ eV can be confined within the galaxy, the highest-energy cosmic rays are commonly explained to be of extragalactic origin, the source being the Active Galactic Nuclei for example.

The observations of spectral steepening and increasing mass composition at the energy of the knee, at $4 \times 10^{15}$ eV, can be explained by a break in the spectrum of proton. In short, the knee energy is the maximum energy attainable for protons when accelerated in SNRs. The acceleration of particles depends on their rigidity $R = p/Z$, where $p$ is the momentum and $Z$ is the charge of the nucleus, and the break-off energy scales as $E \propto Z$. The average mass of cosmic-ray primaries increases as a function of energy as the light elements break off at lower energies than the heavy elements. Rigidity-dependent acceleration models can therefore explain the steepening of the spectrum and the composition getting heavier in the knee region.

Another set of models presented are the propagation models in which the cosmic rays are confined within the galaxy due to the galactic magnetic fields, but leak outside the galaxy if the characteristic gyroradius of the particle is larger than the characteristic galactic size. Such models predict rigidity-dependent break-off energies for different elements as well. Additionally, models that take into account both the input spectrum from SNR and propagation effects are presented. Also other acceleration models and models encompassing a single source or interactions of cosmic rays with background particles have been presented as the explanations of the knee. The consistency of the models with experimental results is discussed in [3] and references therein.

On the basis of the rigidity-dependent model, one can try to explain the second knee at $10^{17.5}$ eV as a break-off of the last element of the spectrum. Assuming a proton knee at $E_{\text{knee}} \approx 4 \times 10^{15}$ eV, an iron knee would be expected to be at
26 \times E_{\text{knee}} \approx 1 \times 10^{17} \text{ eV}, which is however below the energy of the observed second knee. It is possible that significant amounts of ultra-heavy elements, i.e. elements up to uranium, could be present in the spectrum [16]. A uranium knee would be at an energy $92 \times E_{\text{knee}} \approx 4 \times 10^{17} \text{ eV}$, the energy which is equal to that of the observed second knee. Another model [17] suggests that the spectrum has two galactic components: component A with a rigidity-dependent spectrum containing elements up to iron and another component B, which would be caused by an accelerator able to accelerate the particles even above the energy of the second knee. The model as an example explanation of the cosmic-ray spectrum is described in Figure 2.3.

Fig. 2.3: An example of an astrophysical interpretation of the cosmic-ray spectrum: the cosmic-ray all-particle spectrum as measured with EAS experiments and an interpretation by Hillas [17]. The symbols denote the experimental results of the all-particle flux, the line which fit through the symbols is the sum spectrum of the model components. Component A is a galactic source with a break-off energy for protons (p) at $3 \times 10^{15} \text{ eV}$ and rigidity-dependent breaks for heavier elements (He, Fe). Component B is another galactic source with a break-off energy of about an order of magnitude larger than in A. The total extragalactic component (H+He) is denoted by line labeled EGT, the proton spectrum is denoted by EGp. The line “EG no losses” is the extragalactic spectrum without GZK and other energy losses. The dashed line Q is a model expectation without the component B. Figure is taken from [17].

The galactic component is expected to end around the energy $10^{18} \text{ eV}$ and the flattening of the spectrum at the ankle at $10^{18.5} \text{ eV}$ is explained being due to the onset of extragalactic cosmic rays. The onset of extragalactic protons or helium
may also explain the observed decrease of the average mass of the cosmic rays above the energy $10^{17.5}$ eV. The flux suppression at around $10^{20}$ eV can be explained as the maximum attainable energy in extragalactic sources, but can also be explained by the so-called Greisen-Zatsepin-Kuzmin limit [18]. At energies larger than $10^{19.7}$ eV, high-energy protons may interact with cosmic microwave background photons via resonance production, $\gamma_{\text{CMB}} + p \rightarrow p + \pi^0$ and $\gamma_{\text{CMB}} + p \rightarrow n + \pi^+$. After propagating a distance of approximately 50 Mpc, the protons would have lost their excess energy and no protons with energies above the limit would be seen on Earth. Similarly, also nuclei of the highest energy would be broken due to photodisintegration.
2.2 Extensive Air Showers

Extensive Air Showers are particle cascades that develop in Earth’s atmosphere [3, 7, 19]. EAS are initiated by high-energy primary cosmic rays, which enter the atmosphere from the space and interact with the nuclei of the air. A schematic diagram of EAS is shown in Figure 2.4. The first interaction occurs at heights of 10 – 50 kilometres in the atmosphere and particles are created. The particles are hadronic (p, n, π, K, etc.), electromagnetic (e, γ) and muonic (µ), and they propagate at relativistic speeds in the direction of the shower axis, i.e. the direction of the primary cosmic ray. These particles either decay or interact further with air nuclei and, subsequently, the particle numbers in the shower increase until a shower maximum is reached. After the shower maximum, the electromagnetic component attenuates partly reaching the ground level and is absorbed while the high-energy muons propagate to the ground level and penetrate into the ground even for several kilometres. EAS are a rapid process (about 100µs) and the lateral size of the shower at the ground level can be up to several kilometres while total particle number can be up to billions.

The energy of the first interaction between the incoming primary and the air nucleus may correspond to energies reachable in particle accelerators and also above it. In these high-energy interactions mesons, dominantly pions, are produced. The

Fig. 2.4: A schematic view of an Extensive Air Shower together with measurement techniques.
pions may either decay or interact with air nuclei. Neutral pions are very short lived ($\tau = 8.4 \times 10^{-17}$ s) and they dominantly decay ($\pi \to \gamma + \gamma$ with consequent interactions $\gamma \to e^+ + e^-$, $e^- \to \gamma + e^-$ and $e^+ + e^- \to \gamma + \gamma$). This produces the electromagnetic component of EAS. The number of gammas and electrons increase due to the radiative interactions while their energies decrease until radiative energy losses become less significant and ionization losses dominate the transport, the electromagnetic component attenuates, the electrons are absorbed, and particle numbers decrease. Charged pions have a life-time of $\tau = 2.6 \times 10^{-8}$ s and since their energies are relativistic they may interact with air nuclei before decaying. The decays are dominantly $\pi^+ \to \mu^+ + \nu_\mu$ and $\pi^- \to \mu^- + \bar{\nu}_\mu$, from which the muon component of EAS arises. Hadronic interactions between pions and air nuclei cause more mesons to be produced, which again, after decays and interactions, 'feed' the electromagnetic and muonic components. A competition between the decays and interactions is sensitive to the energies of the mesons and the density of the atmosphere and this competition affects the production of the muons, such as their numbers and energy spectrum. The muons have a relatively long lifetime ($\tau = 2.2 \times 10^{-6}$ s) and due to their relativistic energies they can reach the ground level before decaying, the survival probability being close to 1 for muon energies above 10 GeV.

A simple model of a hadron-initiated EAS is presented and discussed in [20]. According to the model, the depth of the shower maximum $X_{\text{max}}$, which is expressed in the units of g/cm$^2$ and calculated from the top of the atmosphere, scales according to $X_{\text{max}} \propto \ln(E/A)$, where $A$ and $E$ are the mass number and energy of the primary cosmic ray. The electron number at the shower maximum scales as $N_{\text{max}} \propto E^{1.03}$ and the total number of muons as $N_\mu \propto E^{0.85}A^{0.15}$. The model expresses the average properties of EAS development: i) the number of particles at the shower maximum depends on the primary energy but is insensitive to the primary mass, ii) the number of muons is larger in heavy-nucleus-initiated showers than in proton showers and iii) the depth of the shower maximum increases as a function of the logarithm of the primary energy being less for heavy nuclei than proton. These also lead to a difference between proton- and heavy-nuclei-initiated showers: as the shower maximum is higher for heavy nuclei than proton but electron numbers at the maximum equal, at an observation level after the shower maximum nuclei-initiated showers will be attenuated more than proton-initiated and electron numbers at the observation level will be smaller for heavy nuclei than proton.

After the shower reaches the ground level, the electromagnetic and hadronic components are absorbed by a few metres of ground but the muon component of the shower advances. While traversing through rock, the propagation is influenced by ionization and radiative energy losses and the muon energies decrease. The muon energies at the ground level may be from GeV up to TeVs corresponding to ranges in rock from metres to several kilometres. The deeper the muons are measured, the higher their energies at ground level were and the earlier in shower development the muons are originated from. The muons are scattered in rock, which leads to spatial displacement with respect to an unscattered track and to a deviation in the muon direction. Muon interactions in the medium may also lead
to the production of electromagnetic and hadronic cascades. (See [21] for more details and for simulations of these effects for EMMA.)

The production of Cherenkov light in EAS follows from the velocities of the charged shower particles that may surpass the velocity of light in air. Most of the Cherenkov radiation is caused by electrons and positrons and is focused in a narrow cone (approximately $1.3^\circ$) in the air. Fluorescent light originates from excitation of atmospheric nitrogen caused by the charged shower particles, is dominantly in the near UV region and is emitted isotropically.

In more details, the phenomena of EAS are modelled numerically. A modelling takes into account propagation, interaction and decay of particles in the atmosphere, as does the widely used CORSIKA [22] program. In the left panel of Figure 2.5, an example evolution of the simulated particle numbers as a function of the shower depth is presented. The determination of cosmic-ray spectrum and composition commonly follow from the use of computer simulations of EAS together with selected energy and composition indicators for the measured EAS. An example of such indicators, the electron number $N_e$ and the muon number $N_\mu$ at sea level, is shown in the right panel of Figure 2.5. It is evident from the figure, that a simultaneous measurement of $N_e$ and $N_\mu$ of an air shower makes it statistically possible to determine the primary energy and mass, the details of which are interaction-model dependent, while the overall scheme remains the same for different models applied.

![Fig. 2.5: Left panel: The longitudinal development of EAS. The number of electrons and muons as a function of shower depth for an example case of proton- and iron-initiated air showers of $E = 10^{19}$ eV. Figure is taken from [3]. Right panel: The electron and muon numbers of EAS at sea level as obtained with three different hadronic interaction models (named in the legend). The contour lines denote the FWHM of distributions expected from primaries of fixed energy and mass. Figure is taken and adapted from [3].](image-url)
2.3 Experimental methods and results of the knee region

The mass and energy of the primary cosmic ray can be interpreted from the measurements of the properties of EAS. Typical cosmic-ray experiments are large area surface arrays, which measure the electromagnetic, muonic and hadronic components of EAS. The sizes of these arrays vary from 0.001 km$^2$ to 1000 km$^2$ depending on the energy region they are designed to measure. Several detector stations arranged into a detector array (see Figure 2.4) measure the local particle densities of the shower. From the measured particle densities, tracks, and arrival times the shower arrival angle and shower axis position are determined together with the lateral density distribution of the particles. From these distributions an energy and a composition estimator can be evaluated. Besides the particle density measurements, the measurement of the depth of the shower maximum from the detection of Cherenkov or fluorescent light is used. Multicomponent methods, which apply a simultaneous measurement of different shower components, are preferred.

2.3.1 Compilation of composition measurements

EAS experiments have studied the mass composition in the knee region. Figure 2.6 shows a compilation of results of the average cosmic-ray mass composition in the knee region. The average mass composition is given as a mean logarithmic mass: $\langle \ln A \rangle = \Sigma_i r_i \ln A_i$, where $r_i$ is a fraction of nuclei $i$ in the spectrum and $A_i$ is the mass number of nucleus $i$. Results based on the measurements of $X_{\text{max}}$ and particle numbers on ground and underground are given. Several EAS experiments, which measure the particle numbers, have reported a clear increase in the average primary mass as a function of increasing energy in the energy interval from $10^{15}$ eV to $10^{17}$ eV. The composition changes from a light (proton-like) towards a heavy (iron-like) spectrum. Experiments have also measured the shower fluorescent or Cherenkov light yields and derived the composition from the measurement of the shower maximum. These measurements indicate a rise in the average mass of the primaries in the energy interval of $4 \times 10^{15}$ to $4 \times 10^{16}$ eV, but the magnitude of the rise is less in comparison to the results from particle density measurements and the absolute values of the average masses differ. The distinct decrease of average mass below the energy $4 \times 10^{15}$ eV is not seen in particle number measurements. The differences between the results of the $X_{\text{max}}$ and particle density measurements can not be explained in detail but tuning of the hadronic interaction models used in the analysis can make the discrepancy less, as is stated in [3].

2.3.2 Selected experiments, methods and results

There are many approaches for the measurement of cosmic-ray spectrum and composition. A list of experiments, analyses and results with the emphasis of relevance to this work are given below.

Surface array measurements
The KASCADE multicomponent air shower array [23], in Karlsruhe, Germany, measured the cosmic rays of the knee region. The surface array contained 252 detector stations arranged into a rectangular grid of 13 metre grid spacing on an
Fig. 2.6: Left panel: Mean logarithmic mass of cosmic rays as derived from the average depth of shower maximum with several experiments (QGSJET01c). Right panel: Mean logarithmic mass as derived from the measurements of electrons, muons and hadrons with several experiments (QGSJET01c). The compilations are taken from [3] in which the details are presented.

Area of 200 \times 200 \text{ m}^2. A measurement of the electron and muon densities (threshold \( E_\mu > 230 \text{ MeV} \)) in a shower made it possible to reconstruct the position of the shower axis and local particle densities, after which the total number of electrons and muons in the shower were reconstructed.

Two-dimensional shower size spectrum is used in the analyses in [24, 25]. Electromagnetic shower size \( \lg N_e \) and truncated muon size \( \lg N_{\mu tr} \) are reconstructed for each EAS. The results for elemental spectra follow from an analysis of the distribution \( (\lg N_e, \lg N_{\mu tr}) \). The number of events in a bin, \( N_j \), of the distribution can be expressed as a sum of the contributions of individual nuclei to that bin, i.e.

\[
N_j = C \sum_{A=1}^{N_A} \int_{-\infty}^{\infty} dJ_A/d\lg E \ p_A \ d\lg E
\]

where \( p_A = p_A(\lg N_{ej}, \lg N_{\mu tr} | \lg E) \) is the probability for a cosmic-ray of energy \( E \) and mass \( A \) to produce an EAS of shower sizes in a given bin \( j \), \( dJ_A/d\lg E \) is the spectrum of nuclei \( A \), \( C \) is a normalization and the summation runs over a sample of selected nuclei. The probability \( p_A \) is estimated from an EAS simulation, which includes the modelling of EAS and the array. The method allows a reconstruction of the cosmic-ray spectrum for samples of selected nuclei separately. The resulting spectra in the knee region were presented for five nuclei (p, He, C, S, Fe) in the energy interval from \( 10^{15} \text{ eV} \) to \( 10^{17} \text{ eV} \) and were given for the high-energy interaction models QGSJET [26] and SIBYLL [27] and low-energy interaction models FLUKA [28] and GHEISHA [29]. The results indicate that the proton and helium spectra exhibit knee-like features and that the proportion of heavy nuclei in the cosmic-ray spectrum increases above the knee energy but that the exact proportions of elements are model dependent. The result for the knee energy of all-particle spectrum is \((4.0 \pm 0.8) \times 10^{15} \text{ eV} \) (QGSJET01) and the results for the
spectral indices below and above the knee energy are $2.70 \pm 0.01$ and $3.10 \pm 0.07$, respectively. For the SIBYLL hadronic interaction model, the corresponding values are $(5.7 \pm 1.6) \times 10^{15}$ eV, $2.70 \pm 0.06$ and $3.14 \pm 0.06$.

The successor of KASCADE was KASCADE-Grande [30]. The experiment was expanded by 37 scintillator detector stations (each 10 m$^2$) over an area of $700 \times 700$ m$^2$. This extension provided the experiment a sufficient resolution for spectral studies in the energy range from $10^{16}$ eV to $10^{18}$ eV. The main results are an observation of a steepening in the cosmic-ray energy spectrum of heavy primary particles at about $8 \times 10^{16}$ eV [31], a hardening of spectrum of light primaries at about $10^{17}$ eV [32], and an all-particle energy spectrum [33] involving a hardening at $2 \times 10^{16}$ eV and a steepening at $8 \times 10^{16}$ eV (QGSJET-II [34]). The intensity of the all-particle spectrum, the positions of the hardening and steepening of the spectrum, and the relative abundance of the elements depend on the hadronic interaction model used in the interpretation of the data [35].

Other approaches in surface measurements exist also, for example the all-particle lateral density distribution on ground is sensitive to the primary mass and the features of the spectrum. This is used to interpret the particle density measurements of the CARPET array in Baksan, Russia. The CARPET array contains a central scintillator detector of the area of 200 m$^2$ and six satellite stations of the areas 18 m$^2$ at 20 – 40 metres distance from the array centre. The progress in the analysis of the lateral density distribution is presented in [36].

**Deep underground detector (TeV muons)**

The deep underground detector MACRO at Gran Sasso, Italy, recorded very high-energy muons ($E_{\mu} \gtrsim 1.3$ TeV) over a detector volume of $76.6 \times 12 \times 4.8$ m$^3$ in coincidence with the EAS-TOP surface array [37]. The data were analysed for electromagnetic shower size $N_e$ and high-energy muon number $N_{\mu}^{TeV}$ correlations and models of cosmic-ray spectrum. The results of the analysis are an increase in the average mass from $\langle \ln A \rangle \approx 2.1$ to $\langle \ln A \rangle \approx 3.2$ and a steepening of the spectrum of light elements in the primary energy interval from $1.5 \times 10^{15}$ eV to $1.5 \times 10^{16}$ eV (QGSJET). The results are also in overall agreement with results from EAS-TOP surface measurements using the electromagnetic and muonic shower sizes $N_e - N_{\mu}^{GeV}$ [38].

**Medium depth (100 – 1000 GeV muons)**

The SPASE array measured EAS in coincidence with the AMANDA under-ice neutrino detector at the South Pole [39]. SPASE consisted of 30 scintillator stations organized into a triangular grid with 30 m grid spacing. It measured the shower arrival direction, shower axis position and the shower size at the ground level. AMANDA consisted of a group of photomultipliers attached to strings, which were deployed under-ice into a depth of about 1.5 kilometres corresponding to a muon cut-off energy of $E = 300$ GeV, and measured Cherenkov light caused by the high-energy muons penetrating into the deep ice. The analysis of cosmic-ray composition relied on the use of two parameters: $S(30)$, particle density at 30 metres from the shower axis, and $K(50)$, a parameter related to the total energy loss of all the muons penetrating into the AMANDA volume. The analysis result
is that $\langle \ln A \rangle$ increases from approximately 2.0 to 2.8 in the energy interval from $1 \times 10^{15}$ to $6 \times 10^{15}$ eV (QGSJET98).

The successors of AMANDA/SPASE are the ICECUBE under-ice detector and the ICETOP surface array. In the analysis [40] of 30 days of data, parameters $S_{125}$, surface lateral density distribution at 125 m, and $K_{70}$, underground lateral density at 70 m, were used. The results are cosmic-ray spectral index values of $\gamma = 2.61 \pm 0.07$ below and $\gamma = 3.23 \pm 0.09$ above a knee at $(4.75 \pm 0.59) \times 10^{15}$ eV, and the average mass $\langle \ln A \rangle$ increasing from approximately 2.2 to 3.5 in the energy interval $1 \times 10^{15} - 30 \times 10^{15}$ eV (SIBYLL).

The Baksan Underground Scintillating Telescope (BUST) in Baksan, Russia, measures the number of high-energy muons ($E > 230$ GeV) in coincidence with the Andyrchy surface array. An analysis of shower size spectrum and muon data [41] results in a primary composition getting heavier in the knee region, but also point to a discrepancy between the EAS size spectrum and muon data (QGSJET-II).

**Shallow depths (10 – 100 GeV muons)**

The LEP experiments DELPHI and ALEPH at CERN measured also the high-energy muon bundles of cosmic origin. The multiplicity spectrum of the muons depends on the cosmic-ray energy spectrum and mass composition and can be compared with models of cosmic-ray spectrum. The ALEPH detector, located 140 metres underground, measured the multiplicity spectrum of high-energy ($E > 70$ GeV) muon bundles over an area of 16 m$^2$. In the analysis of about 20 days of data [42], the measured multiplicity spectrum was compared with expectations from cosmic-ray spectral models and it was concluded that the multiplicity distribution favours a chemical composition that changes from light to heavy composition in the knee region (QGSJET). The detector also recorded five high-multiplicity muon events, which occurred at a frequency almost an order of magnitude higher than expected. The DELPHI detector with an area of 75 m$^2$, at a depth of 100 metres and a cutoff energy of $E = 50$ GeV, found in an analysis of about 19 days of data [43] that the multiplicity distribution cannot be explained in a satisfactory way and that the QGSJET model fails to describe the abundance of high-multiplicity events, seven of which saturated the detector. The multiplicity distributions of both DELPHI and ALEPH are shown in Figure 2.7.

The ALICE experiment, at a depth of 30 metres and a vertical cutoff of 15 GeV, of the Large Hadron Collider (LHC) is recording cosmic-ray muon data with an area of approximately 15 m$^2$. After measurements during the years 2010–2012, the ALICE experiment has reported the observation of high-multiplicity events as well [44], as also shown in Figure 2.7.
Fig. 2.7: The muon multiplicity spectra compared with expectations from pure proton and iron cosmic-ray spectrum by the experiments DELPHI [43] (top panel, integrated multiplicity), ALEPH [42] (middle panel) and ALICE [44] (bottom panel).
2.4 Conclusions for EMMA

The cosmic-ray spectrum follows a power law with spectral index $-2.7$ up to the knee energy. Below this energy the spectra of individual nuclei are measured directly. At and above the knee energy the knowledge of the cosmic-ray composition relies on the measurements of Extensive Air Showers, the modelling of which is of crucial importance in the analysis of the spectrum. The knee and the second knee are explained as being breaks due to the galactic cosmic-ray sources, propagation or interactions. In order to find out the most accurate astrophysical explanation for the spectral features observed, the measurements of the spectrum and mass composition should be continued towards more accurate and less interaction-model dependent results.

The basic picture of the knee region is well established. At the energy of about $4 \times 10^{15}$ eV, the spectrum steepens and the average mass composition rises. Several experiments employing different methods have confirmed this change. However, as it is evident from the figures and quoted values in Section 2.3, the exact value of the knee energy, spectral index and the exact elemental spectra in the knee region remain to be studied further. The results of the knee region still suffer from the methodological limitations and the results differ for different experiments or different hadronic interaction models applied. At this point no model or any experiment alone can explain the EAS of the knee energy in all the details. Although the current main efforts in the measurements are on the transition region from $10^{17} - 10^{18}$ eV and around the extragalactic region of $10^{19}$ eV, the study of the knee energy region from $10^{15}$ to $10^{17}$ eV continues to be of essential importance since this energy region is the lowest energy where the drastic changes in the spectral features emerge.

The experiments, which involve the measurement of the high-energy muon component of EAS, are in good agreement with other experiments. However, the measurements of the muon yields at shallow depths are scarce and have, as shown in Section 2.3, resulted in interesting outcome. The experiments in CERN have reported anomalously high muon multiplicities. The DELPHI and ALEPH experiments took data only for a period of a few weeks and ALICE cosmic data taking is also limited by its main purpose in LHC. Therefore a high-statistics experiment specially devoted for the study of the muon events at shallow depths is required to further investigate the issue.

The EMMA experiment, which is the subject of this thesis, is designed to study the knee region of cosmic rays with a novel approach. The experiment aims to measure the composition of cosmic rays in the knee region by measuring the lateral density distribution of high-energy muons at shallow depths underground. EMMA brings further light into resolving the exact features of the spectrum of cosmic rays and is also expected to bring input for the detailed modelling of EAS.
3. EMMA design

The main aim of the EMMA experiment is to measure the cosmic-ray spectrum and composition in the knee region with a new method. Cosmic-ray composition measurements in the energies of the knee rely on the use of various models of Extensive Air Showers (EAS), the exact results on the composition depending on the interaction model applied in an analysis. Several experiments at shallow depths underground have measured muon multiplicities higher than expected and therefore it is important to approach the composition study by further measuring the muon yields of EAS. The novel approach of EMMA is to determine the composition from the measurements of the lateral density distribution of high-energy muons at shallow depths underground.

The EMMA underground array is designed in order to achieve the aim. For a composition study, the experiment should measure EAS in sufficient numbers, determine the arrival directions of the showers and record local muon densities. The shower axis position and muon density lateral distribution are to be determined from the measured muon densities of EAS, which then allows to deduce the energy and mass of the primary cosmic ray.

The design study of EMMA has been conducted parallel to the construction of the experiment and its background is briefly presented in Section 3.1. Sections 3.2 – 3.5 summarize the main simulation work executed during the planning of the experiment indicating that the designed array geometry of EMMA at the depth of 80 metres is the optimal and presents a composition analysis strategy for the experiment. Section 3.6 describes the performance of the array extension of EMMA. In Section 3.7, a summary of the array design studies is given.

3.1 Background to the design work

Experimental ideas related to EMMA date back to the late 1990’s. An approach to an underground experiment measuring the cosmic-ray composition was presented in 1999 [45]. This proposal described an underground detector array of hexagonal geometry consisting of 7 underground muon tracking stations of the size of 100 m² each. The experiment was planned at an underground depth of 50 metres. It was proposed that a difference between the high-energy muon lateral density
distributions of proton- and iron-initiated air showers could be used for cosmic-ray composition analysis. Another proposal with a similar aim was the CORAL proposal [46] at CERN. The proposal consisted of an underground muon tracking detector of the size of $21 \times 23 \, \text{m}^2$ accompanied with a surface array. The aim of the experiment was to determine the composition of cosmic rays at the knee region, study cosmic-ray anomalies and extend the study of high-multiplicity muon events.

MUon Barrel chambers (MUBs) from the disassembled DELPHI [47] experiment in CERN offered a relatively low-cost detector option to be used in a multi-muon experiment in the Pyhäsalmi mine. First set of detectors were retrieved to Oulu in 2002 for testing and development of DAQ, and the majority of the detectors were transported from CERN to Pyhäjärvi in late 2005. They have an overall sensitive detector area of approximately $200 \, \text{m}^2$. These detectors form the backbone of the EMMA experiment and the basic array design was made based on the number of these detectors.

Optimal underground locations for EMMA were searched in the Pyhäsalmi mine. Since there was no funding to make new excavations for the experiment, underground detector stations for the high-energy muon detection were to be built within the existing tunnels, caverns or maintenance spaces in a part of the mine where no active mining operations were ongoing or anticipated. A surface array which would measure the electron size of a shower in coincidence with the muon measurements was seen also as an option but active mining operations and funding limitations discarded any reasonable surface array designs. Therefore it was found that the experiment should be designed based on underground detector stations only and, without a surface detector, shower axis location and cosmic-ray composition indicators should be determined solely from the measurements of the underground array.

The design work aimed to find an optimal configuration for the underground array. A construction site for the array needed to be selected and the geometry of the array needed to be designed. The sensitive detector area of EMMA is rather moderate in comparison to many other experiments studying the knee region and a method for the cosmic-ray composition determination needed to be quantified. The studies of the array geometry design and composition analysis strategy are summarized in Sections 3.2 – 3.5. The main findings of these sections have been previously reported in [48] and internally in [49].

After the construction of the EMMA array had already started, two new detector options arose. High-granularity scintillators [50], which were designed to improve the accuracy of the muon number measurements within EMMA detector stations, were delivered to Pyhäjärvi in 2010. A number of Limited Streamer Tubes (LSTs) [51] were acquired from the disassembled KASCADE experiment in 2012. With the LSTs, a possibility to extend the array with new detector stations emerged. A study on the array extension and a performance update of shower reconstruction is given in Section 3.6.

### 3.2 Array size and location

The optimal size and the location of the array in the mine were to be determined.

Figure 3.1 shows the number of EAS to be recorded by an array per year.
In order to study the cosmic rays in the interval from $10^{15}$ eV to $10^{17}$ eV the acceptance of the array should be at least 100 m$^2$sr to guarantee sufficient number of recorded EAS. Expecting that the array stations could determine the shower arrival directions in a zenith angle interval $0 - 35$ degrees, solid angle corresponding to approximately 1 sr, the fiducial area of the array should be equal or larger than 100 m$^2$.

![Fig. 3.1: The integrated flux of EAS $(I(E > E_0))$ per year per steradian as a function of $E_0$. Lines denote the number of EAS to be recorded by an array of a fiducial area as indicated in the legend. The knee is assumed to be at $4 \times 10^{15}$ eV.](image)

There were several possible underground sites for the experiment to be built in. In the beginning, several locations in the mine between the depths of 80 and 400 metres were considered. Two of the locations, a tunnel crossing at a depth of 80 metres ($E_{\text{cut}} \approx 50$ GeV), and the tunnels and halls around a maintenance level at a depth of 210 metres ($E_{\text{cut}} \approx 150$ GeV) appeared as the prime candidate locations for the experiment. Figure 3.2 shows the simulated muon densities of vertical EAS of the primary energy of $10^{15}$ eV and $10^{16}$ eV for these cutoffs and they are compared with muon densities at ground level. As the muon energy cutoff increases, the muon densities in the tail of the showers are decreasing more rapidly than in the shower core and the separation between proton- and iron-initiated showers is enhancing. The muon densities to be recorded are, however, smaller for higher cutoff energies, namely approximately 0.8, 0.5, 0.3 m$^{-2}$ at the shower core and about 0.3, 0.1, 0.02 m$^{-2}$ at a shower tail at the distance of 20 metres for 1 PeV showers for the cutoffs 0, 50 and 150 GeV, respectively. For the energy of 10 PeV, the ratios of the values are similar, but the muon densities are larger by about a factor of 10. It was concluded that the site at the depth of 80 metres is preferred, since there the muon yields over a distance interval 0 – 50 metres are larger by factors of 2 to 10 in comparison to the site at the depth of 210 m.
The EMMA experiment was to be built at a depth of 80 metres. At that depth one preferred location for the array existed, namely a crossing of three tunnels about 40 metres from mine’s inclined access road. An underground array of the optimal dimensions could be built at the location. The site is relatively easy to access and maintain, since it is close to the surface and close to the main inclined driveway of the mine. Detector stations of the area of \( A \approx 15 \text{ m}^2 \), which would fit into the tunnels, were planned. In order to guarantee measurable muon yields from EAS of the knee energy with that size of detector stations, distances of the array stations from the array centre were to be less than 50 metres. The sensitive detector area to be used in the experiment was approximately 200 m\(^2\), but since at least a part of the detector stations should be equipped with several detector layers in order to track the shower muons and determine shower arrival angles, it was feasible to design the array based on nine detector stations. This corresponds to an overall muon detection area of 135 m\(^2\).

### 3.3 Muon yields and composition sensitivity

An array station of the size 15 m\(^2\) at a depth of 80 metres, which corresponds to the cutoff of 50 GeV, will record muon numbers as presented in Figure 3.3. For showers of energies below the knee, namely 1 PeV, the recorded muon numbers will be scarce. At the shower axis, on the average approximately 6 muons will be recorded and around a distance of 20 metres the average numbers are approximately 1 and 2 for proton and iron-induced showers. The number of recorded muons increases with increasing energy. For the energy of 10 PeV, approximately 60 muons at the shower core and numbers around 10 and 1 at the shower tail distances of 20 and 50 metres will be recorded. The muon yields of EAS of fixed energy fluctuate. The number distribution widths are also shown in Figure 3.3. For example, the
fluctuations at the shower core are approximately 30% and 15% for 10 PeV proton and iron showers, respectively.

![Fig. 3.3: The mean and the standard deviation of the muon numbers to be recorded by a detector station of the area of 15 m$^2$ at the depth of 80 metres as a function of station distance from the shower axis. Shown for proton (red) and iron-induced (blue) showers of the fixed energies of 1 PeV (lower symbols) and 10 PeV (upper symbols). 2000 simulated vertical EAS per primary type (CORSIKA-QGSJET01C). The symbols for iron are slightly shifted at the x-axis for clarity. The mean and standard deviation are shown even though the distributions are asymmetric when mean muon numbers are close to zero.](image)

A composition sensitivity as a function of distance from the shower axis is shown for three primary energies in Figure 3.4. The composition sensitivity is calculated as $\sigma_{p-Fe} = (\mu_p - \mu_{Fe})/((\sigma_p^2 + \sigma_{Fe}^2)/2)^{1/2}$, where $\mu_i$ are the mean number of muons detected in a detector area of 45 m$^2$ and $\sigma_i$ are the standard deviations of the number distributions of fixed energy. The muon numbers of iron showers are normalized to yield equal density with proton showers of the same energy at 0 metres (a factor between 0.8 – 0.9 depending on the energy). The area of 45 m$^2$ corresponds to an area of three detector stations, which, given a shower axis position at the centre of a three-arm layout, are at the same distance from the axis. (A three-arm layout is presented in the next section.) The composition sensitivity rises with increasing distances up to the distance of 10 – 15 metres, after which the sensitivity stays nearly constant. If the shower axis position, the core muon density and the energy of the primary are known, it is optimal to measure the shower muons at distances of 10 – 40 metres from the shower axis, the sensitivity
being rather independent of the exact distance.

Fig. 3.4: The composition sensitivity between proton- and iron-induced showers as a function of the detector distance from the shower axis. $\sigma_{p-Fe}$ is a measure of the statistical separation between expected number of detected muons from proton and iron showers for an area of 45 m$^2$. The sensitivity is drawn for three different energies. Curves are drawn for $\mu_p \geq 3$ (thin) and $\mu_p \geq 9$ (thick), and correspond to primary energies as given in the legend.
3.4 Layout design

Layout design was performed by using CORSIKA-QGSJET01 simulations. The design studies for the array geometry were based on the numerical simulations of EAS, the measurement and reconstruction of them, on the assumption of 135 m\(^2\) of detector area and the use of the existing tunnels at the depth of 80 metres.

Aspects of the optimal design were studied with different approaches. The section starts with an overall description of shower reconstruction and parametrization of the muon lateral density distribution, which yields the energy and composition indicators for the experiment. Different array geometry options are presented and their effect on the shower reconstruction is discussed, after which an optimization study of station positioning is presented. The section ends with a description of the EMMA layout at the depth of 80 metres.

3.4.1 Shower reconstruction

The average lateral density distributions of proton- and iron-initiated showers of different energies are shown in Figure 3.5. The muon numbers in EAS increase as a function of primary energy. It is noticeable that the muon density at the shower core is nearly independent on the primary mass (i.e. an energy estimator) and the shape of the lateral density distribution differs for proton- and iron-initiated showers, proton showers having a sharper distribution than iron. Interpretation of the energy and mass of the primary cosmic ray is possible if the shower axis position of EAS is reconstructed and the muon number density is measured as a function of the distance from the shower axis.

Fig. 3.5: The simulated average lateral density distributions for different energies and masses for 50 GeV cutoff energy (QGSJET01c) and for vertical EAS. The number of simulated EAS is 2000 for each fixed energy and mass.
Fig. 3.6: A set of simulated air showers in the \((\rho(1), R_0)\)-plane (QGSJET01c). The number of showers is 200 for each fixed energy and mass.

The lateral density distribution is parametrized using the following lateral density function:

\[
\rho(r) = \frac{N}{(2\pi \times 0.11099 \times R_0^2)} \times \left(\frac{r}{R_0}\right)^{-0.4} \times \left(1 + \frac{r}{R_0}\right)^{-5}, \quad (3.1)
\]

where \(r\) is the distance from the shower axis, \(N\) is the size of the shower and \(R_0\) is a parameter related to the gradient of the lateral density distribution. A parameter related to the core density is \(\rho(1)\), which is the reconstructed muon density at 1 metre from the axis.

Figure 3.6 illustrates the fluctuations of the showers in the \((\rho(1), R_0)\)-plane for different primaries of fixed energy and mass. The scattered points represent fluctuations from shower to shower for each primary. It is evident that primaries with different energies and masses populate different regions in the plane. As a first approximation, \(\rho(1)\) correlates with energy and \(R_0\) correlates with mass, thus being good indicators for energy and mass, respectively. For the analysis of shower reconstruction, a two-component description of the cosmic-ray composition was used. The lateral density distributions for primaries with mass numbers between proton and iron have lateral shapes between the two extremes.

An underground array must be able to measure the shower axis position and the muon densities at different distances from the axis in order to estimate \(\rho(1)\) and \(R_0\) and determine the energy and composition of the primary cosmic ray. The following simplifications are used in the array simulations: i) the showers are assumed vertical, ii) the rock overburden is assumed uniform and iii) detector properties are assumed ideal, i.e. each muon traversing a detector area is recorded.
A number of showers with given input shower axis positions \((r_{\text{input}})\) are simulated and the coordinates of the muons, which have an energy \(E_\mu > 50\) GeV and pass through a detector area are stored. These detected muons are referred to as 'hits'. A fit of the lateral density function (Eq. 3.1) is applied to the muon hits shower by shower. The fit maximizes a log-likelihood of the numbers of muon hits in binned data (approximative bin sizes of \(1 \times 1\) m\(^2\)). The MINUIT package [52] is used for the fit. In the fit, the reconstructed shower axis position \(r_{\text{rec}}\), the shape \(R_0\) and size \(N\) are free variables and the initial guess for \(r_{\text{rec}}\) is given by the mass centre of the muon hits. The best fit values are recorded and the procedure is repeated a number of times with different showers of the same energy and mass resulting in a distribution which includes the shower fluctuations and the reconstruction uncertainties (See Figure 3.6 for EAS reconstructions given an infinite detector area).

In the analyses of the next subsections, nine detector stations of the size of approximately 15 m\(^2\) are assumed and various layout options are compared in different analyses. The approach is to study the relations between the reconstruction variables, i.e. shower axis positions \(r_{\text{rec}} = (x_{\text{rec}}, y_{\text{rec}})\), core densities \(\rho(1)\) and shapes \(R_0\), and the input shower axis positions, energies and primary masses. An example detector layout and a simulated shower together with its reconstruction are shown in Figure 3.7.

![Fig. 3.7: Left panel: A top view of an example simulated vertical 4 PeV proton shower and one possible EMMA layout in the mine tunnels. The shower muons having energy greater than 50 GeV are shown as gray dots, the squares indicate the detector stations and the 'detected' muons (hits) are shown as large dots. The black cross indicates the input shower axis position and the red cross is the reconstructed axis position. The lines denote the tunnel walls at the EMMA site at the depth of 80 metres. Right panel: The shower density profile (blue dotted line) compared with the reconstructed one (red solid line).](image-url)
3.4.2 Array geometries and shower axis position reconstruction

The shower axis reconstruction performance for several array geometries were studied in order to find out the preferred array geometry for EMMA. The tunnels at the EMMA site at the depth of 80 metres are depicted together with examples of theoretically possible array layouts in Figure 3.8. The figure also shows example distributions of input axis positions of selected showers. The selection is based on the reconstructed shower axis position and a selection area equal to the area of the centremost station is used. The examples are shown for 6.3 PeV vertical proton showers. The corresponding shower axis uncertainties, the uncertainty defined as $|r_{\text{input}} - r_{\text{rec}}|$, are shown in Figure 3.9.

By comparing the options (a), (c) and (d), it is evident that if the muon densities are measured only at the smallest possible distances around the shower axis as in option (a), the resulting axis reconstruction uncertainties are the smallest but the spectrum also exhibit a tail contribution. This tail arises from local density fluctuations in the EAS. A fraction of showers are 'misinterpreted', typically, as showers of lower $\rho(1)$ and higher $R_0$ than the 'true' shower. If the grid spacing between the stations is increased as in options (c) and (d), the average shower axis reconstruction uncertainties increase but the tail contribution disappears. The grid spacing will also affect the fiducial area of the experiment and thus the measured number of events: the larger the grid spacing, the larger the fiducial area. A symmetric grid with an optimal grid spacing is theoretically optimal for the shower axis reconstruction but no such option was possible for EMMA.

Three options (b, e and f) depict those that were possible to construct in the existing tunnels. The tail contribution arises in all of the layout geometries. Some showers having their input shower axis positions in the region where there are no detector stations can be reconstructed as showers with axis at the central part of the array and are 'misinterpreted'. In a line-type geometry (b) the shower axis position in the perpendicular direction to the alignment of the array is poorly reconstructed and this leads to larger axis reconstruction errors than in options (e) and (f). Therefore array geometries like option (b) were rejected. Two of the options, (e) and (f), rely on the use of the three tunnels extending from the crossing, where the centre of the array is. An option with concentrated detector area at the center of the array (f) yields a higher proportion of EAS with small axis uncertainties compared to option (e), but has also a tail contribution similar to (e) and deploys less sensitive detection area at the shower tails, where the muon yields are more sensitive to the mass of the primary.

For other shower energies and masses the behavior of the reconstruction is similar and the detector stations of EMMA were decided to be placed in the three arms, like in the options (e) and (f) which were further studied.
Fig. 3.8: Various layout options together with an example shower axis distribution for 6.3 PeV proton showers as viewed from above. The input shower axis positions (red dots) of the showers, which were reconstructed as showers with shower axis position within an area of the centremost unit, and the tunnel walls of the site at the depth of 80 metres (lines) and detector stations (black crosses) are shown. The layout options are: a) one large detector unit, b) line-line layout, c) grid with a separation of 10 m between the detector units, d) as “c” but with a separation of 20 m, e) three-arm layout, and f) as “e” but with a central detector.
Fig. 3.9: Examples of shower axis uncertainty spectra for the layouts given in Figure 3.8. The data are shown for 6.3 PeV proton-induced showers which are reconstructed as showers with their axis position in the area of the centremost station. The spectra are normalized to yield an integral of 1.

### 3.4.3 Optimization of station placements

The layout of EMMA was decided to be a three-arm layout, since such a layout will result in the smallest shower axis uncertainties. In order to optimize for the composition sensitivity, the muons should be measured at the distances between 10 and 40 metres from the shower axis, but the exact positioning of the stations in the array affects also the accuracy of shower axis reconstruction. It was studied what should be the optimal placement of the stations and whether a central detector at the centre of the layout could be better than a grid of equally spaced stations.

The following describes an analysis on the composition sensitivity, which includes the shower axis reconstruction uncertainties. Samples of 2000 showers of the energy of 2.5 PeV, both proton and iron, were simulated. The shower axes were distributed uniformly inside a triangular area of $1/2 \times 20 \times 12 = 120 \text{ m}^2$ at the centre of the layout. This area is the minimal requirement for the fiducial area of the array. Each shower was reconstructed keeping all the parameters of the fit function free.

The analysis was carried out for the layouts presented in Figure 3.10. The options were: (A) the layouts with a central detector with distances to the outer stations between 10 and 45 metres, and (B) the layouts with a triangular central geometry of 10 metres distance between the stations and with distances 10 – 45 metres to the outer stations.

The showers are reconstructed given that the total number of muon hits is larger
Fig. 3.10: Left panel: The layouts A with central detector and distances $r_1$ and $r_2$ to the outer stations. Right panel: The layouts B with no central detector and distances $r_1$ and $r_2$ to the outer stations. The crosses denote the positions and dimensions of the stations and the area where shower axis positions are distributed in is denoted by the dotted triangular region.

than 5. The results of the composition sensitivity and reconstructed core density ($\rho(1)$) width are shown in Table 3.1. A number in each cell of Table 3.1 describes a separation goodness $I$. It is calculated based on the sum

$$S = \Sigma_{R_0} |N(R_0)^P - N(R_0)^{Fe}|,$$

where $N(R_0)^i$ refers to the number of events of particular $R_0$ in binned data and the summation is over all the bins. $I$ is defined as $I = (1 - S/N)$, where $N$ is the total number of simulated showers. Therefore $I$ describes the proportion of the overlapping area between the $R_0$-distributions of proton- and iron-induced showers. The smaller the value of $I$, the less is the overlap between proton- and iron-induced showers, and therefore the better is the composition sensitivity of the layout. The values $W_p$ in the Table 3.1 denote a reconstructed core density ($\rho(1)$) width for proton. The distribution width is defined as $W_p = \log_{10} \rho(1)^{75\%} - \log_{10} \rho(1)^{25\%}$, where $\rho(1)^{i\%}$ are distribution’s $i\%$ percentiles. The smaller the value of $W_p$, the better reconstruction of the primary cosmic-ray energy is foreseen. The distances $r_1$ and $r_2$ refer to the distances to the outer array stations as defined in Figure 3.10.

By comparing the values of Table 3.1, the following is concluded. The composition sensitivity is optimized by type B layouts with distances to outer stations from 20 to 45 metres. Layouts with smaller distances are less sensitive to composition but gain in the $\rho(1)$-reconstruction, which worsens with increasing station distances. Layouts around B20–30 are close to optimal for both the composition and $\rho(1)$-reconstruction. The type A layouts yield worse composition sensitivity than type B layouts, only the option A10–35 being close to that of B. All the layouts are suitable for cosmic-ray composition measurements, but type A layouts with distances $r_1 > 10$ metres and type B layouts with distances $r_2 < 25$ metres
Table 3.1: The composition sensitivity values and the core density distribution widths for proton for various layouts defined in the main text. The values given in each cell are $I$, $W_p$ and are listed for the geometry options A and B separately. The distances $r_1$ (rows) and $r_2$ (columns) refer to the distances defined in Figure 3.10. The sampling error of $I$ is estimated to be approximately 0.010. The error of $W_p$ is not estimated.

<table>
<thead>
<tr>
<th>$r_1/r_2$</th>
<th>15m</th>
<th>20m</th>
<th>25m</th>
<th>30m</th>
<th>35m</th>
<th>40m</th>
<th>45m</th>
</tr>
</thead>
<tbody>
<tr>
<td>A10m</td>
<td>0.775,0.30</td>
<td>0.768,0.32</td>
<td>0.744,0.32</td>
<td>0.735,0.31</td>
<td>0.731,0.31</td>
<td>0.737,0.31</td>
<td>0.749,0.32</td>
</tr>
<tr>
<td>A15m</td>
<td>0.748,0.41</td>
<td>0.775,0.39</td>
<td>0.755,0.36</td>
<td>0.744,0.40</td>
<td>0.761,0.38</td>
<td>0.734,0.42</td>
<td></td>
</tr>
<tr>
<td>A20m</td>
<td>0.781,0.50</td>
<td>0.775,0.52</td>
<td>0.766,0.51</td>
<td>0.767,0.52</td>
<td>0.778,0.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A25m</td>
<td>0.767,0.62</td>
<td>0.771,0.64</td>
<td>0.755,0.62</td>
<td>0.780,0.63</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A30m</td>
<td>0.774,0.83</td>
<td>0.792,0.83</td>
<td>0.759,0.76</td>
<td>0.790,0.96</td>
<td>0.776,1.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A35m</td>
<td>0.774,0.83</td>
<td>0.792,0.83</td>
<td>0.759,0.76</td>
<td>0.790,1.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A40m</td>
<td>0.774,0.83</td>
<td>0.792,0.83</td>
<td>0.759,0.76</td>
<td>0.790,1.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B10m</td>
<td>0.767,0.33</td>
<td>0.776,0.32</td>
<td>0.746,0.31</td>
<td>0.729,0.31</td>
<td>0.747,0.31</td>
<td>0.735,0.32</td>
<td>0.733,0.34</td>
</tr>
<tr>
<td>B15m</td>
<td>0.749,0.34</td>
<td>0.726,0.34</td>
<td>0.735,0.33</td>
<td>0.729,0.32</td>
<td>0.725,0.33</td>
<td>0.717,0.32</td>
<td></td>
</tr>
<tr>
<td>B20m</td>
<td>0.716,0.39</td>
<td>0.707,0.35</td>
<td>0.713,0.35</td>
<td>0.709,0.37</td>
<td>0.719,0.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B25m</td>
<td>0.710,0.40</td>
<td>0.702,0.39</td>
<td>0.735,0.39</td>
<td>0.708,0.40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B30m</td>
<td>0.713,0.43</td>
<td>0.740,0.43</td>
<td>0.708,0.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B35m</td>
<td>0.713,0.46</td>
<td>0.714,0.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B40m</td>
<td>0.712,0.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

can be discarded due to their worsening composition sensitivity, and options with $r_1 > 25$ metres can be discarded due to their increasing $\rho(1)$-distribution widths. All in all, type B layouts are preferred and at least the outermost stations should be placed at distances larger than 20 metres.

### 3.4.4 Layout at 80 metres depth

Studies on the optimum design for EMMA were conducted and yielded the following results. The stations of the array should be arranged in a grid, which extends over an area. The stations should measure the muon yield at the shower core and at distances from 10 to 40 metres from the shower axis in order to optimize the cosmic-ray composition measurement sensitivity. For an acceptance larger than 100 m$^2$sr, an equally separated grid of detector stations is favoured in comparison to an option with a large area central detector. The composition sensitivity is insensitive to small changes in the station locations and, for the decision of the final layout, practical arguments could be used with strong emphasis.

The final design at the depth of 80 metres is depicted in Figure 3.11. The relative distance between the detector stations was decided to be approximately 10 metres. This grid spacing between stations was optimal not only for physics but also for practical reasons, i.e. for the construction of the stations. Several metres of space was decided to be left clear for access and maintenance purposes at the tunnel entrance, which limited the station locations of one arm. The array
was decided not to be extended over 25 metres on the two other arms due to more hazardous environment in the tunnels and since, for practical analysis purposes, it would be good to preserve as much symmetry as possible. The layout of EMMA is close to the option B15–25 in Section 3.4.3.

Fig. 3.11: A top view of the EMMA layout at 80 metres below ground. The designed layout consists of nine detector stations indicated by blue rectangles, each having a separation of approximately 10 metres from the neighbouring stations. The tunnel walls are indicated by black lines. The shaded tunnels describe the parts of the tunnel, which are not enforced nor scaled for loose rocks. The shaded area next to the inclined access driveway is reserved for access to the site.
3.5 Composition reconstruction strategy

A strategy to reconstruct the primary energy spectrum and composition is described in this section. The strategy is described through an example simulation study, which encompasses the simulations of air showers, shower reconstruction, event selection and unfolding.

The chosen array layout in this exercise is the one nearest to the optimal in Table 3.1, namely the option B20–30. (The analysis was committed before the finalization of the EMMA layout.) The input shower axes are uniformly distributed over an area of $160 \times 160$ m$^2$ around the array and two cuts are applied to select a part of the simulated data for the analysis. The simulation area of $160 \times 160$ m$^2$ is chosen in a way that no shower, whose input shower axis is outside the area, passes the first cut. The first cut, i.e. a multiplicity cut, is a condition that a shower produces more than 10 muon hits in at least one detector station. The showers surviving the cut are reconstructed after which a second cut is applied. The second cut, i.e. a shower axis cut, is a condition that the reconstructed shower axis position is within a shower selection area chosen to be a triangular-shaped area of $A \approx 150$ m$^2$ at the central part of the array.

A reconstruction bias factor is defined as the ratio between the number of showers with reconstructed axis position within the triangular shower selection area and the number of showers with input axis position within the selection area. This bias factor of shower reconstruction is drawn in Figure 3.12 as a function of primary energy and is given for proton and iron showers separately. In an ideal case the value of the bias factor is equal to 1 over the studied energy region for both primary masses. The bias factor deviates from the ideal due to the multiplicity cut and 'inflow' and 'outflow' of shower axes to and from the selection area. From the figure, it is evident that the values are different between proton- and iron-initiated showers with primary energies less than 10 PeV. This is due to the fact that muon numbers in iron showers are larger than the numbers in proton showers of the same energy. The values of the factors below 1 in the energy interval $1 - 5$ PeV are mainly due to the multiplicity cut: showers which yield muon numbers less than the one defined in the multiplicity cut are not selected to the analysis. Above 5 PeV the bias factors approach an approximative value of 1.25. Factors larger than 1 follow from an asymmetry in the fluctuations of reconstructed shower axis positions: the reconstruction algorithm tends to reconstruct more shower axis positions towards the central part of the array than out of the central region. These biases of shower reconstruction can be taken into account and corrected for by applying the following strategy in the analysis.

In a numerical example realistic data is simulated from 0.8 PeV to 20 PeV with spectral index $\gamma = -2.7$. The distribution $N_{\text{data}}(\rho(1), R_0)$ denotes the number of showers with given $\rho(1)$ and $R_0$. Examples of simulated data are shown in Figure 3.13 in which the two selection cuts have been applied.

Number distributions $f_{i,j}(\rho(1), R_0)$ of EAS of fixed energy and mass are simulated for the unfolding of composition. The energies are $i = 1.0, 1.6, 2.5, 4.0, 6.3, 10.0, 15.8$ PeV and masses $j = p, \text{Fe}$. Figure 3.14 shows the number distributions $f_{i,j}(\rho(1), R_0)$ for three different example energies. By comparing the figures, it is evident that the maxima of the distributions move towards increasing $\rho(1)$ with
increasing energy, iron-initiated showers having larger $\rho(1)$-values than those of proton. The proton- and iron-induced showers populate also different regions in the direction of $R_0$, the maxima of iron showers having larger values of $R_0$ than those of proton. The distributions have noticeable tails following from the inaccuracies of the reconstruction of showers. The normalized number distributions, i.e. single-energy probability distributions $f'_{i,j}(\rho(1), R_0)$, denote the probability of a shower which is of certain energy and mass and which passed the cuts to be reconstructed as a shower in a given bin in the $(\rho(1), R_0)$-plane.

The aim of the reconstruction is to express the data as a linear combination of the single-energy probability distributions. The reconstruction is carried out by finding suitable $a_{i,j}$ such that

$$N_{\text{data}}(\rho(1), R_0) \approx \sum_{i,j} a_{i,j} \times f'_{i,j}(\rho(1), R_0),$$

(3.3)

where the factors $a_{i,j}$ denote the number of showers with energy $i$ and mass $j$ in the data. Numerical examples of composition reconstructions are shown in Figures 3.15–3.17. The figures demonstrates that the fractions of proton- and iron-initiated EAS in the simulated data can be reconstructed with statistical accuracy of $10 - 30\%$ given a data taking time of 1.5 years. The accuracy depends on the input composition, since the fluctuations of $R_0$ and $\rho(1)$ depend on the primary particle. It is also demonstrated that changes in the composition as a function of energy could be measured, one specific numerical example is given in the lower panel of Figure 3.17. The reconstructed numbers of showers can be corrected for the biases by dividing the numbers with the reconstruction bias factors. After the bias correction one arrives to the reconstructed cosmic-ray spectrum as a function of energy. The simulation was carried out with a spectral index $-2.7$. For a knee-like spectrum, the number of events in the data would be slightly less: for the highest simulated energy, where the effect of a more sharper spectrum is the highest, the statistics is about two thirds of the values shown in the figures.

The composition reconstruction method presented above relies on the simulation of the single-energy probability distributions, which should take into account all aspects of the measurement: the air showers, the detection efficiencies of muons and the biases of the reconstruction algorithm. For the analysis of real data, it will be therefore necessary to describe the operation of the array in a simulation environment. These include the simulation of muons traversing the rock, knowledge of the detector performance, namely efficiencies and muon number saturation, and the properties of the reconstruction algorithm itself. This is especially important for the analysis of showers of energies below the knee because the recorded muon numbers are scarce and notable biases between proton- and iron-initiated showers exist.
Fig. 3.12: The reconstruction bias factors as a function of primary cosmic-ray energy shown for proton- and iron-induced showers for the case of the two specific cuts described in the main text.
Fig. 3.13: The $(\rho(1), R_0)$-distributions of simulated data (continuous spectrum 0.8 – 20 PeV). Composition scenarios: (upper left) 100% proton, (upper right) 100% iron, (bottom left) 50% proton and 50% iron and (bottom right) 80% proton and 20% iron below $E < 3$ PeV and 20% proton and 80% iron above. Each colour refer to the number of counts (equal or larger than the numbers indicated in the legend) in a bin for 1.5 years of data. The tails of the distributions extend to values of $R_0 > 300$ m.
Fig. 3.14: The distribution of events in the \((\rho(1), R_0)\)-plane for 1.6, 4.0 and 10.0 PeV protons (left, lowest energy at bottom) and irons (right). Each colour refers to the number of counts (equal or larger than the number indicated in the legend) in a bin. The number of events is determined by the simulated sample. The tail of a distribution may extend to values of \(R_0 > 300\) m.
Fig. 3.15: Examples of composition reconstruction. The histograms in each figure denote the input proportions of proton and iron showers (solid red and dashed blue) as a function of energy in the simulated data. The symbols denote the reconstructed number of proton and iron events as a function of energy. Input compositions are: 100% proton (upper panel), and 100% iron (lower panel). Error bars denote 1-sigma uncertainties. Uncertainties for reconstructed numbers equal to zero are not plotted (singular errors returned by fit algorithm).
Fig. 3.16: Examples of composition reconstruction. The histograms in each figure denote the input proportions of proton and iron showers (solid red and dashed blue) as a function of energy in the simulated data. The symbols denote the reconstructed number of proton and iron events as a function of energy. Input compositions are: 80% − 20% proton-iron (upper panel), and 50% − 50% proton-iron (lower panel). Error bars denote 1-sigma uncertainties.
Fig. 3.17: Examples of composition reconstruction. The histograms in each figure denote the input proportions of proton and iron showers (solid red and dashed blue) as a function of energy in the simulated data. The symbols denote the reconstructed number of proton and iron events as a function of energy. Input compositions are: 20% −80% proton-iron (upper panel) and 80% −20% proton-iron below 3 PeV and 20% −80% proton-iron above 3 PeV (lower panel). Error bars denote 1-sigma uncertainties.
3.6 Array extension

Limited Streamer Tube detectors (LSTs) used in the central detector of the now dissambled KASCADE array were seen as an intriguing possibility for improving the performance of EMMA. This initiated an investigation on possibilities to expand the array by additional detector stations and study the resulting improvements in its performance. The study relied on an updated code of EAS simulation and shower reconstruction. (This section is a revision of simulations executed for [53].)

Previous investigations have proved the possibility to reconstruct the composition with the array layout presented. However, one of the disadvantages of the array geometry constrained by the mine tunnels is that the shower axis position reconstruction for showers with 'true' axis positions in the regions where there are no detector stations is biased. A significant number of EAS with their shower axis in these regions can be reconstructed as showers with their reconstructed axis at the centre part of the array. To overcome the effects of these misreconstructions, a possibility to build an array extension at a separate mine tunnel at a depth of 40 metres was studied (see the array layout in Figure 3.18).

Two stations at the depth of 40 metres were planned. Both of the stations will have a detection area of approximately 11 m$^2$ each. The place for the stations is at a horizontal distance of about 35 metres from the array centre. The average muon yields in the stations when the shower axis is at the centre of the array are approximately 0.9 and 1.6 for 1 PeV and 7.5 and 17 for 10 PeV vertical proton and iron showers, respectively, and thus the area is sufficient for the detection of muons in coincidence with the array stations at the depth of 80 metres. Because of muon-induced cascades which may extend over an area of a station and local muon density fluctuations in the showers, the stations were decided to be separated by 10 metres.

The following procedure is used to investigate the layout performance. It consists of two parts: the datamaker and the reconstructor. The datamaker approximates the real data taking. A database of discrete energy CORSIKA air showers with zenith angles between 0 and 35 degrees and energies 4 PeV and 10 PeV is simulated. The shower arrival angles are distributed isotropically and according to a flat detector geometry. The shower axis positions are distributed uniformly inside a circular region, which has the centre at the centre of the array and radius of 80 metres. The radius of the distribution area is large enough to accommodate all possible EAS, which survive a multiplicity cut in reconstruction. The positions of the detector stations (locations including the depth), orientations and dimensions are as in the experiment, and the muon energy cutoff is approximated and scaled according to the zenith angle $\theta$ of the shower: $E_{\text{cut}} = 45 \text{ GeV} / \cos \theta$. For the stations at the depth of 40 metres the cutoff used is $E_{\text{cut}} = 30 \text{ GeV} / \cos(\theta)$. Other effects related to the muon propagation, namely the muon scattering in rock and muon-induced particle showers (cascades), are not included, and efficiency of muon detection is assumed to be 100%.

A muon hit is defined if a CORSIKA muon output coordinate is inside a detector station and the muon momentum is larger than the energy cutoff. Each station is divided into $4 \times 4$ bins of the size of approximately 1 m$^2$. The numbers of
hits in each bin are counted. The detector plane is by definition at the EMMA depth of 80 metres against which all recorded muon hits are projected according to the shower arrival direction. The shower arrival direction is assumed to be known and the coordinates of the detector stations and the bin centers and bin areas are transformed from the detector plane to shower plane, i.e. the plane perpendicular to the shower axis. The mutual distances of the stations on the shower plane change according to the shower arrival direction. This is due to the depth differences and is most pronounced for the stations of the array extension. Since the stations of the extension are 40 metres higher in the mine than the other stations and the muon cutoff energy is lower for these stations, a weighting factor is used to scale down the number of muons recorded, as if the stations would be at the same depth as the others. The ratio of muon yields between the cutoffs 45 GeV and 30 GeV varies as a function of primary energy, mass and distance from the shower axis. This ratio is $w \approx 0.6 - 0.9$ for energy and distance regions of relevance. In the reconstructor, a constant weighting factor ($w = 0.8$) is used for all the showers.

![Efficiency of shower selection](image)

Fig. 3.18: The efficiency of the shower selection as a function of true shower axis position for 4 PeV proton (left) and iron (right) showers. Red color indicates higher than 90% efficiency, orange higher than 50% and yellow higher than 10%. The rectangles indicate the detector stations as viewed directly from above. The two stations composing the array extension at a horizontal distance of about 35 metres from the array centre are shown as shaded rectangles. $X$ and $Y$ refer to the coordinates in the detector plane.

The first cut applied to the simulated data is a multiplicity cut. In the reconstructor only events with more than four hits in any of the three centermost
stations are selected for shower reconstruction. The resulting selection efficiencies are shown in Figure 3.18. For example for 4 PeV EAS, the efficiency is approximately 96% for proton and larger than 99% for iron in a circular selection area of radius 10 metres centered on the array centre. The efficiencies are higher for iron showers than for proton showers since the muon yields in iron-initiated showers are larger than in proton showers. This efficiency is closer to 100% for energies higher than 4 PeV and for smaller selection radii than 10 metres.

The function in Equation 3.1 is fitted to the binned muon data in the shower plane. The algorithm calculates the maximum likelihood for the parameters $x_{\text{rec}}$, $y_{\text{rec}}$, $N$ and $R_0$ using a minimization routine from the MINUIT [52] library. In the reconstruction all the parameters of the fit-function are free. After the shower reconstruction the shower is described by four parameters: the coordinates of reconstructed shower axis position, the density at the shower core $\rho(1)$ which to a first approximation relates to the energy of the cosmic ray, and the shape $R_0$ which to a first approximation relates to the mass.

The position spectra of reconstructed shower axes of an example case of 10 PeV iron showers are shown in Figure 3.19 for the two layouts: 'EMMA' at 80 m and 'EMMA+EXT' the layout with additional detector stations at 40 m. The input shower axis have been uniformly distributed with a density of approximately $13 \, \text{m}^{-2}$. The bin size in the figure is $2 \times 2 \, \text{m}^2$, which corresponds to approximately 52 showers per bin in an ideal reconstruction. The shower axes tend to be reconstructed within or close-by the detector stations and the numbers of showers there are highly biased. By comparing the two cases, it is evident that the biases at the array centre reduce, if additional detector stations exist.

A second cut, i.e. a shower axis cut, is applied to the data based on the reconstructed shower axis positions. The reconstruction is best for the region at the array centre and therefore a selection area of circular region of radius 10 metres and centered on the centre of the array (between the three centremost stations) is used. The selection area is approximately $310 \, \text{m}^2$.

A third cut, an analysis threshold cut, is based on the reconstructed values of $\rho(1)$. Only EAS with reconstructed values of $\log_{10} \rho(1)/\text{m}^2 > -0.2$ are taken into analysis since the discarded showers would correspond to reconstructed energies below a threshold of composition reconstruction, i.e. the energy of approximately 1 PeV. The input shower axis positions of the selected showers for iron primaries of the energy 10 PeV are shown in Figure 3.20 for the layouts EMMA and EMMA+EXT. The 'bulk' of the reconstructed events have their input shower axis positions within the selection area, while the three tails of the distribution are a consequence of the three-arm geometry of the array. Showers with input shower axis position in the region of the tails are 'misreconstructed' as showers with their axis in the centre of the array. In the case of the reconstruction which includes the muon hits in the stations of the array extension one of the tails is significantly reduced. In Figure 3.21, projections of the axis position distributions along the orientation of the six aligned detector stations are shown. By building the array extension, the tail can be suppressed by a factor of 3.4 and the asymmetry in the shower axis reconstruction is reduced. The spectrum can be compared with an ideal case: the spectrum of EMMA+EXT consists of a bulk of 'proper
Fig. 3.19: The spectrum of reconstructed shower axis positions at the detector plane for all 10 PeV iron showers surviving the multiplicity cut. Upper panel: The layout EMMA. Lower panel: The layout EMMA+EXT. The number of showers in each bin is indicated by the colours. The array stations are overlayed on the spectrum. The red circle denotes the boundary of the shower selection area according to the reconstructed shower axis positions. $X$ and $Y$ refer to the coordinates in the detector plane.
Fig. 3.20: The spectrum of input shower axis positions for 10 PeV iron showers surviving the multiplicity cut, the shower axis cut and the analysis threshold cut. X and Y refer to the coordinates in the detector plane. Upper panel: The layout EMMA. Lower panel: The layout EMMA+EXT. The number of showers in each bin is indicated by the colours. The array stations are overlayed on the spectrum. The red circle denotes the boundary of the shower selection area according to the reconstructed shower axis positions.
reconstructions', but contains 36% of the reconstructions in the tail. At the central region of the selection area, the spectrum is lower than in an ideal case, due to ‘outflow’ of reconstructed shower axis positions. The outflow of reconstructed shower axis positions from the selection area with respect to the input positions is 16% (84% of EAS, which have their input shower axis in the selection area, are reconstructed as showers with shower axis in the area) and the ‘inflow’ (the ratio of the number of showers with input shower axis positions outside the area and with the reconstructed ones inside, to the number of input showers in the area) is 48%.

Fig. 3.21: The projections of input shower axis position spectra given in Figure 3.20 along the orientation of the two parallel arms (the tunnel of six stations). $Y'$ is the direction coordinate perpendicular to the orientation and centered in the centre of selection area. Spectra are shown for the layout EMMA (blue longdash) and for the layout EMMA+EXT (red solid), which are compared to that of an ideal case, i.e. a cut on the input shower axis positions (black dotdash).

The shower axis uncertainty spectra are given in Figure 3.22 for 4 PeV and 10 PeV proton and iron showers. The modes of the spectra are approximately 1 m for 10 PeV and 2 m for 4 PeV EAS. The array extension reduces the uncertainties in all cases, the reduction being evident in the tail of the uncertainty distribution. For 10 PeV iron, the mean value of the uncertainty (the average of the spectrum in Figure 3.22) is reduced from $13.3 \pm 0.2$ to $9.0 \pm 0.2$ metres. The corresponding values for other simulated cases are: $8.6 \pm 0.2 \rightarrow 5.7 \pm 0.2$ metres (10 PeV p),
8.9 ± 0.1 → 7.3 ± 0.1 metres (4 PeV Fe) and 6.1 ± 0.1 → 5.3 ± 0.1 metres (4 PeV p). The reduction of the tail is evident in all cases but is the strongest for 10 PeV EAS.

The reconstruction bias factor is defined as the ratio between the number of reconstructed showers and the number of input showers in a predefined selection area. For example, over the selection area of radius 10 metres centered on the centre of the array, the reconstruction bias factor is reduced from 1.83 to 1.32 given that the array extension is built. For the case of 10 PeV proton shower the reduction of the reconstruction bias factor is 1.49 → 1.25 and for the energy of 4 PeV the reductions are 1.45 → 1.31 and 1.08 → 1.04 for iron and proton, respectively. The estimated errors of the calculated ratios are 0.02 for 10 PeV and 0.01 for 4 PeV. In case the array extension exists, the reconstruction biases are less and possible systematical uncertainties related to the correction of the biases are reduced. The biases are listed and brok-down in Table 3.2.
Table 3.2: A break-down of the reconstruction biases following from the multiplicity cut, the shower axis cut and the analysis threshold cut. $E$ is the primary energy in simulation, $A$ is the primary mass, 'Layout' refers to the two layout options presented in the main text and $\epsilon_{sel}$ is the selection efficiency. $f_{in-in}$ is the ratio between the number of showers, which have their reconstructed and input shower axes within the selection area, and the number of showers, which have their input shower axis position in the selection area. $f_{out-in}$ is the ratio between the number of showers, which have their reconstructed axis within the selection area, but input shower axis outside, and the number of showers, which have their input shower axis position in the selection area. The effect of the analysis cut is included in these values. RBF is the reconstruction bias factor, $RBF = \epsilon_{sel} \times (f_{in-in} + f_{out-in})$. In an ideal case: $\epsilon_{sel} = 1.00$, $f_{in-in} = 1.00$, $f_{out-in} = 0.00$ and $RBF = 1.00$.

For an EAS array of symmetric and grid-like geometry, the values are optimally $f_{in-in} + f_{out-in} = 1.00$ and $\epsilon_{sel} = 1.00$ for an energy interval studied. The cuts used are: more than 4 muon hits in one the three centremost stations, reconstructed shower axis position within a selection area of a radius of 10 metres centered in the centre of the array and $\log_{10} \rho(1)/m^2 > -0.2$.

The array extension affects the energy and composition determination of the experiment. The single-energy number distributions in the $(\rho(1), R_0)$-plane are shown for 10 PeV iron showers for the cases of EMMA and EMMA+EXT in Figure 3.23. The projections of these distributions in $\rho(1)$ are shown in Figure 3.24 together with projections of 4 PeV showers and proton showers, respectively. For 10 PeV, two peaks are seen, the 'signal' peak around $\log_{10} \rho(1)/m^2 = 0.6$ and a peak of 'misreconstructed' showers around $\log_{10} \rho(1)/m^2 = -0.3$. The suppression of the misreconstructions is most pronounced for the energies of 10 PeV, for which the suppression of the tail is the highest, even after applying an analysis cut $\log_{10} \rho(1)/m^2 > -0.2$. Gaussian fits to the 'signal' peak of the distributions reveal that, for the layout EMMA+EXT, the spectrum is approximately 11% narrower than for the layout EMMA in the case of 10 PeV iron (9% in the case of proton showers). For the energy of 4 PeV, the 'signal' and 'misreconstruction' peaks are not separated so clear as in the case of 10 PeV. The 'signal' peaks are, however, more pronounced if the array extension exists.
The $R_0$-spectra of 4 PeV and 10 PeV proton and iron showers are shown in Figure 3.25. Fits of Landau distributions to the spectra are also shown. By visual comparison of the distributions, the separation of proton and iron-induced showers appears more enhanced in the direction of $R_0$ for the energy of 10 PeV if the array extension exists. For the energy of 4 PeV, no improvement is observed.

The array extension increases the symmetricity of the array and shower reconstruction, reduces the number of misreconstructed showers, i.e. the tail of the shower axis uncertainty spectrum, which reduces the biases in shower reconstruction and possible systematical uncertainties following from them. It leads to narrower $\rho(1)$-distributions indicating better energy and composition indicators for the experiment, the improvement being larger at the energy of 10 PeV compared to 4 PeV. Furthermore, the extension increases the acceptance of the experiment.
Fig. 3.23: The distribution of 10 PeV iron events in the \((\rho(1), R_0)\)-plane after the multiplicity cut and the shower axis cut. Upper panel: The spectrum for the layout EMMA. Lower panel: The spectrum for the layout EMMA+EXT. The number of showers in each bin is indicated by the colours.
Fig. 3.24: Upper panel: The $\rho(1)$-projections of 4 and 10 PeV iron events after the multiplicity cut and the shower axis cut. Lower panel: The $\rho(1)$-projections of 4 and 10 PeV proton events after the multiplicity cut and the shower axis cut. Blue lines: The projections for the layout EMMA. Red lines: The projections for the layout EMMA+EXT. Solid lines are for 10 PeV and dashed lines are for 4 PeV. The spectra are normalized to integral of 1.
Fig. 3.25: The projections in $R_0$ of 10 PeV (two upper panels) and 4 PeV (two lower panels) proton (red) and iron (blue) events after the multiplicity cut, the shower axis cut and the analysis threshold cut for the layouts EMMA and EMMA+EXT. First (upmost) panel: EMMA (10 PeV). Second panel: EMMA+EXT (10 PeV). Third panel: EMMA (4 PeV). Fourth (lowest) panel: EMMA+EXT (4 PeV). The solid lines indicate Landau fits in the interval $0 - 200$ m. The spectra are normalized to an integral of 1. The tails extend to values $R_0 > 500$ m.
3.7 Summary and conclusion

A short background to array’s geometry design was given in Section 3.1. The experiment has been planned in phases, before and during the construction, as detector options arose. In Section 3.2 it was described that different possibilities for the site of the experiment were looked for, and the site was fixed at the depth of 80 metres. In Section 3.3, the expected muon yields were presented and the optimal size of the array was determined. In Section 3.4, a parametrization of the lateral density distribution of muons, a method of shower reconstruction and the energy and composition indicators were introduced. Different array geometries and performance of the reconstruction of shower axis positions and composition and energy indicators were compared. It was concluded that the designed three-arm geometry, where each arm extends to 25 metres from the centre of the array and where each detector station is separated 10 metres from the neighbours, is the optimal. A strategy to reconstruct the composition of the cosmic rays from the measurements of a designed three-arm array layout was given in Section 3.5 and demonstrated the potential of applying unfolding techniques to EMMA data. In Section 3.6 updates on the shower reconstruction simulation algorithms were presented and the effect of an array extension was studied. It was concluded that an extension of the array in the tunnel at 40 metres of depth in the mine will further increase the acceptance of the array, decrease the biases of shower reconstruction and improve the shower axis position reconstruction, which improve the expected performance of composition measurement by the experiment.

The EMMA experiment measures the cosmic-ray composition at and above the knee. The reconstruction of cosmic-ray composition will be limited by a threshold energy of shower reconstruction around $1 - 4 \text{ PeV}$. The EMMA experiment has an acceptance of at least $300 \text{ m}^2 \text{ sr}$ and, with a data taking of three years, will reconstruct around 3000 EAS from above the energy of 4 PeV, 500 above 10 PeV, 30 above 40 PeV and 5 above 100 PeV.
4. EMMA experiment

The EMMA experiment is introduced in this chapter starting with an overview of the array and its detector stations in Section 4.1. A brief technical description of the drift chambers of EMMA is given in Section 4.2 together with descriptions of drift chamber calibration and the tracking detector of the central station. In Section 4.3, technical details of SC16 scintillators are given and the high-granularity scintillator setup is introduced. Summaries of two studies concerning the use of the high-granularity scintillator setup in EMMA are also presented. Limited Streamer Tube (LST) detectors are introduced in Section 4.4. A description of underground infrastructure is given in Section 4.5 and the rock overburden at the EMMA site at the depth of 80 metres is presented in Section 4.6. The chapter is summarized in 4.7.

4.1 Overview of EMMA

The geometry of the EMMA array and its detector stations are presented in Figure 4.1. The array consists of two types of detector stations: central stations and outer stations positioned in the mine tunnels as shown in the figure. The central stations are for measuring the shower core and the outer stations are for measuring the tails of EAS.

A tracking detector of a central station consists of drift chambers in three detector planes, which are used for tracking of the muons, allowing the determination of muon density at the shower core and shower arrival direction. The dimensions of the tracking detector are $3.65 \times 4.22 \times 2.25$ m$^3$ and the geometrical acceptance for three-plane single-muon tracking is 18 m$^2$ sr. Central stations will also contain the high-granularity scintillator setup, which is especially designed to improve the determination of muon densities at high muon densities. It will consist of scintillator planes of the area of 6 m$^2$ in three stations and has a granularity of 64 scintillators per m$^2$. The scintillator setup will also contribute to the determination of the shower arrival direction by the timing of the individual muons and will provide a start time for the drift chambers.

The outer stations are designed to consist of two planes of drift chambers or LSTs separated vertically by approximately 1.1 metres. Their purpose is to de-
Fig. 4.1: A top view of the EMMA array consisting of eleven stations (squares), nine at the depth of 80 metres and two at the depth of 40 metres. The stations are labeled with letters. The stations are separated by 10 metres from their neighbours except the array extension being at 35 metre distance from the centre. The array situates at the tunnels of the old mine (black lines). The array center consists of three central stations (C, F and G), which will contain three planes of drift chambers, a high-granularity scintillator setup and a layer of LSTs. The outer stations and the array extension will contain two planes of drift chambers or LSTs. The inclined access road of the mine allowing access to all stations is on the left.

termine muon densities at shower tails by sampling muon numbers given a reconstructed shower direction from the central stations. Their area will be $11 - 15 \text{ m}^2$ depending on the final configuration of detectors in each station.

As of 2014, the tracking detectors of the central stations are taking data (see Chapter 5) and underground tests and calibration runs with SC16 detectors are ongoing. All stations have been constructed, drift chambers are installed into the stations and LSTs are prepared to be installed.
4.2 Drift chambers

Muon Barrel Chambers (MUBs) of the former DELPHI experiment [47] are used for muon detection in the central and outer stations. In total, 84 units of MUBs have been transported to Pyhäjärvi. This section gives a brief technical description of MUBs and their calibration together with a description of the tracking detector of the central station.

4.2.1 Technical description

Schematic diagrams of MUBs used in EMMA are shown in Figure 4.2. Drift chambers are assembled into groups of seven, which form one detector unit hereby referred to as a ‘plank’. The weight of a plank is approximately 130 kg. The gas volume of a drift chamber has a length of 365 cm, a width of 20 cm and a height of 1.5 cm, and is shielded in an aluminum frame. The tungsten anode wire is 47 µm of diametre and has a characteristic impedance of 500 Ω. The wire runs through the central axis of a chamber and is supported by plastic holders placed 1.2 metres apart. Grading copper strips, each 0.4 cm wide, are glued on the chamber walls. One of the central gradings acts also as a delay line. It is a copper winding with characteristic impedance of 600 Ω.

The MUBs are operated with +6 kV anode and +4 kV grading voltages. The anode voltage is fed over a 100 MΩ resistor. The pre-amplificator for the anode signal consists of a (270Ω) resistor and a capacitor (1 nF) in series and a terminating resistor (560Ω) coupled to ground. The current is less than 1 µA. The grading voltage is fed over a 1 MΩ resistor and a voltage divider, which consists of a chain of seven 22 MΩ resistors. The voltage is divided into 26 grading strips, each having their specific voltage and closing to ground voltage near the ends of the chamber. This voltage division guarantees an approximately uniform field (400 V/cm) inside the chamber and results in constant drift velocity of approximately $v_{\text{drift}} \approx 4 \text{ cm/\mu s}$ towards the anode wire. An avalanche near the anode induces a signal into the delay line which propagates at a velocity $v_{\text{delay}} \approx 0.5 \text{ cm/ns}$ towards both ends of the chamber. The delay line pre-amplification electronics consists of a resistor 560 MΩ, a capacitor (1 nF) and a pre-amplificator 8607CN1916. The high-voltage is supplied by dual channel power supplies made by iseg [54], one supply serving typically 35 chambers via high-voltage distribution boxes.

Three signals are read out from each chamber: one from the anode wire (A) and two from the two ends of the delay line, i.e. the near signal (N) and the far signal (F). The signals are fed into Front Electronics Boxes (FEBs), which host amplificator and discriminator cards providing an ECL-level output to CAEN V767B TDC via twisted pair cable. V767B has 128 channels, a least significant bit resolution of 0.8 ns and a double hit resolution of 10 ns in single channel.

For the EMMA experiment, a self-trigger logic was designed in order to operate the drift chambers independently from any external trigger. In the self-trigger system, the anode signals are split in the discriminator cards providing a logical OR of all anode signals and is used as an input for the trigger electronics. In this system the trigger signal is delayed by a variable time in comparison to data flow. The Stop trigger matching mode of the V767B TDC is used, and the width and the offset of the data acquisition window, typically 8 µs and 3.5 µs, are set to...
In EMMA the gas mixture used with MUBs is Ar:CO\textsubscript{2} in percentages of 92\%:8\%. This gas mixture was selected since it has no flammable component (methane was left out from the original mixture) and the selected percentage was measured to give the best efficiency.

**Efficiency**

Typical detection efficiencies of the drift chambers are better than 90\% and anti-correlate with air pressure, see Figure 4.3. The dependence on air pressure is individual for each chamber. Further details are to be reported in [56].

**Position resolution**

If an external start time provided by a scintillator is applicable, the position coordinate perpendicular to the anode wire, $x$, is calculated from the difference of
Fig. 4.3: The efficiency of three example chambers (symbols) during a calibration run in surface laboratory. The efficiencies were determined with reference to tracks reconstructed from signals in five other chambers. Atmospheric pressure is indicated by the solid line.

The position coordinate is determined as $x \approx 5 \text{ cm} + v_{\text{drift}} \times (T_{AX} - T_{AY})$, where $T_{AX}$ and $T_{AY}$ are the recorded anode times from two overlapping chambers. The resolution is approximately 1 cm.

4.2.2 Calibration of drift chambers

The calibration relations $f_i$ were determined with calibration setup in the surface laboratory. A set of reference drift chambers were chosen, whose delay lines were calibrated by using a radioactive sodium source ($^{22}\text{Na}$). The source was placed in several positions above a chamber and singles data was recorded for 20 minutes for each position. The source was transported along the length of the chamber in steps of 10 centimetres to map whole of the delay line response.

A testing and calibration bench was assembled, see Figure 4.4. Four reference planks were set 48 centimetres apart in the vertical direction and a set of planks to be calibrated was positioned between the reference planks. Cosmic-ray muons were...
Fig. 4.4: A photograph of the surface calibration setup in which all drift chambers were tested and calibrated.

recorded and reference tracks were reconstructed from the positions determined in the reference chambers allowing the calibration of other chambers. The geometrical acceptance of the setup was 0.96 sr m² and the trigger rate was approximately 70 Hz. During one month of effective running time, each chamber was passed by 10⁵ muons per one centimetre of delay line direction. This statistics was very suitable for committing the calibration.

In the data taking from an underground station, the used electronics are different to those used in the calibration procedure and the recorded spectra may appear shifted in comparison to calibration reference spectra, i.e. the spectrum from which the calibration relations are derived from. Additional constants $C_i$ are applied to correct the calibration relation: $f_i(x) \rightarrow f_i(x + C_i)$, where $C_i$ are typically less than a few nanoseconds corresponding to a distance correction of the order of one centimetre. The procedure is called delay matching and allows the
correction of constant electronics delays.

4.2.3 Tracking detector

The configuration of drift chambers in the tracking detectors of the central stations is shown in Figure 4.5. The dimensions of the tracking detector are 4.22×3.65×2.25 m$^3$ and the geometric acceptance for three-plane single-muon tracking is approximately 18 m$^2$ sr.

![Fig. 4.5: The configuration of drift chambers in a tracking detector of a central station. The dimensions of the detector are 4220 × 3650 × 2250 mm$^3$. Each tracking detector consists of three detector planes (top, middle, bottom), separated vertically by 1125 mm, and each consisting of two layers of chambers (X- and Y-layers). Each plane contains 35 chambers, 15 in Y-layer and 20 in X-layer. The detector consists of 105 chambers in total. The chamber positions and sizes (width 200 mm and height 20 mm) are denoted by horizontal bars. The lengths of the chambers are 3650 mm.](image)

An example of track-fit residuals of tracks reconstructed in the tracking detector of a central station is given in Figure 4.6. Tracks are fitted through the positions recorded in the drift chambers and the residuals are calculated as $x_{hit} - x_{fit}$, where $x_{hit}$ is the position coordinate reconstructed in the chamber and $x_{fit}$ is the track position in a chamber. The analysis is for single tracks that pass through six drift chambers.

The track and shower arrival direction reconstruction in the tracking detector for multi-muon events has been designed and studied with Geant4 simulations [21]. It is estimated that in the final configuration utilizing a scintillator start time, the most probable value of shower direction reconstruction uncertainty will
Fig. 4.6: Example spectra of track-fit residuals in the delay line direction (Delay) and drift direction (Anode). Analysis is for single tracks that pass through six drift chambers. The spectra are sums of the chambers in the top detector plane of Station C and include all track directions.

be approximately $0.1^\circ$ and that the uncertainty will be less than $1^\circ$ for almost all showers with zenith angles less than $35^\circ$.

Analysis of data recorded by the tracking detectors together with preliminary results about muon number determination capability and shower arrival direction reconstruction are presented in Chapter 5.
4.3 SC16 scintillators

The high-granularity scintillator setup of EMMA serves three purposes in the array: muon number estimation, measurement of EAS arrival angle and start time generation. The SC16 scintillator detectors and measurement electronics were manufactured by Institute of Nuclear Research of the Russian Academy of Sciences, Russia after a joint design by RAS, Univ. Oulu and Univ. Jyväskylä. The scintillators have been documented in [50, 57–61]. In total 97 scintillators (24 m$^2$) were delivered to Pyhäjärvi in 2010. This section gives a brief technical description of the detectors and summarizes the findings of two studies about their use in the EMMA array.

4.3.1 Technical description

An SC1 scintillator is a doped polystyrene block of the size of $122 \times 122 \times 30$ mm$^3$. They are coated with reflector and employ a wavelength shifting fibre for light collection and an avalanche photodiode (APD) for readout, see Figure 4.7. The SC1s are grouped into units, called SC16, which consists of 16 SC1s and the electronics for signal processing and APD-voltage adjustment. The scintillating light is collected by a single Y11 (Kuraray) wavelength shifting (WLS) fibre, which is placed on a carved groove. The fibre is viewed from one end by the APD and the other end is covered by an aluminized mylar reflector. The fibre geometry guarantees light collection uniformly over the entire detector. Outer surface of the scintillator block is etched with a chemical agent acting as a diffusive reflector. The scintillator block and the WLS fibre are wrapped inside an additional Tyvek reflector.

Multi-pixel APDs of 556 pixels operating in a limited Geiger mode are used for the light readout. The typical operating voltage of the APD is from 23 to 35 V
providing a gain of $2 - 5 \times 10^5$. The maximum photon detection efficiency is 23\% at WLS fibre emission spectrum of 515 nm. The recovery time is approximately 1 $\mu$s.

One SC16 module contains 16 scintillator counters grouped into a $4 \times 4$ matrix and packed into a steel box of $0.50 \times 0.50 \times 0.12$ m$^3$. The SC16 also hosts boards for 16 pre-amplifiers, discriminators and an analogue summer together with servicing circuits. A novel aspect of the electronics is a feedback system for APD counting rate stabilization by bias voltage regulation. Each SC16 detector forms a 16 bit binary word with information about the state of each of the APDs. At state change of at least one SC1 this status word is transmitted to serial output socket. Each SC16 has a coaxial (BNC) socket for digital output (OR of any of the 16 SC1s), 8 lamellas socket for 4 twisted pairs of the serial differential output and one coaxial socket (analog sum) for testing purposes. In addition, one socket is for the power supplies.

The digital outputs from the SC16s are connected to Timing Board (TB) via 30 m coaxial cables. The digital signal of each SC16 is converted to LVDS and send further to CAEN V1190A TDC, which has a channel resolution of 100 ps. The Timing Board hosts a programmable FPGA chip for trigger output generation. The inputs to the trigger are the fast digital outputs of the SC16s arranged in to four sets (24 channels each). The trigger generated by TB is of NIM standard.

Hodoscope Boards (HB) further process the SC1 state information. HB receives the state info from the detectors on an SC1 state change via twisted pair cables and stores it to registers. A trigger can be generated internally in the HB or external trigger from the TB can be used. After receiving the trigger the HB communicates
with a CAEN V1495 module via a series of commands transmitting the SC1 status words to the V1495.

**Efficiency and time resolution**

The detection efficiency of SC16 is typically $98 \pm 1\%$ and the time resolution of one SC16 is typically approximately $1.7$ ns. An example time difference spectrum is shown in Figure 4.9. Further details are to be reported in [56].

![Gaussian fit](image)

Fig. 4.9: Example time difference spectrum of coincidences of two SC16s placed directly on top of each other (black). The trigger was OR of either of the channels and coincident events were selected by software. Gaussian fit (red) results in a width ($2.2$ ns) corresponding to $\sqrt{2}$ times the time resolution of one SC16.

**High-granularity scintillator setup**

Three planes of scintillators will be placed in the three central stations. The three planes will consist of 24 SC16s ($A = 6$ m$^2$) each. A schematic view of the high-granularity scintillator setup is shown in Figure 4.10.
4.3.2 Performance estimates for the scintillator setup

Two performance studies concerning the use of the high-granularity scintillator setup were done and resulted in estimates for the shower arrival direction reconstruction performance and muon number determination capability. The studies assumed 96 scintillators in the setup but the conclusions of the studies hold also for the current plan (24 SC16s are now reserved to be used in other experiments and the setup will consist of 72 SC16s).

Determination of shower arrival direction

The high-granularity scintillator setup can be used for the determination of shower arrival direction. The measurement of the shower arrival direction is based on the recorded particle passage times by individual SC16s as the advancing shower plane will pass through the detectors at different times depending on the angle of shower arrival. The uncertainties in shower angle determination follow from the geometry of the scintillator setup, the baseline distance of 10 metres and the time resolution of each SC16. Also uncertainties in the relative electronics delays between different scintillators may have an effect.

A simple numerical model of the scintillator setup was built to estimate the uncertainties in shower angle reconstruction. In the model, the locations of fired SC16 detectors are randomized according to uniform spatial distribution and the time recorded in each fired SC16 detector is generated given an input shower arrival direction, the time resolution and additional uncertainty of the relative...
delays between the individual detectors. The shower arrival angle is reconstructed by minimizing a sum $\sum_i (|Ax_d + By_d + D| - cT_i)^2$, where $x_i$ and $y_i$ are the detector coordinates, $T_i > 0$ is the time recorded by the detector $i$, $c$ is the speed of light, summation runs over the fired SC16s, and $A$, $B$ and $D$ are the fit parameters of the shower plane. The shower zenith ($\theta$) and azimuthal ($\phi$) angles are then calculated.

The average uncertainties of the shower arrival angle reconstruction as given by the numerical model are shown in Figure 4.11. (The setup used in the simulation is the same as in Figure 4.10, except that 48 SC16s were assumed in one of the scintillator planes.) The zenith angle dependence of the uncertainty is shown in the left panel. For fixed multiplicity of five fired scintillators in each scintillator plane, which is a typical multiplicity for events with shower axis position between the planes and with primary energy of about 4 PeV, the mean uncertainty is approximately 2.5 degrees for zenith angle range 0 – 35 degrees being insensitive to the angle in this range and worsening for larger zenith angles. The uncertainty decreases as a function of the number of fired scintillators as shown in the right panel and approaches to a mean uncertainty of 1° for 20 fired scintillators in each plane.

![Fig. 4.11: The expected mean uncertainty of shower arrival direction reconstruction with the high-granularity scintillator setup. Left panel: The mean uncertainty as a function of shower zenith angle for events with five fired scintillators in each plane. Right panel: The mean uncertainty as a function of the number of fired SC16s (in each plane) in an event for vertical showers. Estimates are based on a numerical model with SC16 time resolution of $\sigma = 1.7$ ns and delay uncertainties of $\pm 1$ ns between the individual SC16s.]

The scintillator setup brings useful redundancy in the shower direction reconstruction by the EMMA array allowing cross-checks with shower directions reconstructed in the tracking detectors.
**Scintillator response**

The response of a scintillator plane consisting of 48 SC16s was studied with Geant4 simulations [62]. The study aimed to estimate the relation between the number of muons passing through the scintillator plane and the number of fired SC1s for shower cores. The simulations were executed with CORSIKA air showers and included muon-induced secondary production, i.e. the electrons and gammas contributing to the total energy loss in the scintillators. Scintillators were fired if the total energy loss in the scintillator was larger than 3 MeV approximating the used threshold.

An example of a simulated response of the high-granularity scintillator plane is shown in Figure 4.12. The simulation is carried for a scenario were the scintillators are positioned below the lowest detector plane of the tracking detector. It is evident that the number of fired pixels correlates with the muon number and that the number of fired scintillators is larger than the number of muons. The difference is because of the muon-induced secondaries also firing the scintillators. The number of fired scintillators is typically approximately 25% larger than the number of muons, fluctuates for fixed muon multiplicity by approximately 20% and may contain events with several times more fired SC1s than muons. In the study [62], it is also concluded that the fluctuations in the number of fired scintillators can be reduced significantly if the scintillators are shielded with lead (shielding with 3 centimetres of lead will reduce the secondary contribution by approximately 70%) or if they are placed directly below the roof of the tunnel (almost complete removal of scintillators fired by secondaries).

A high-granularity scintillator plane consisting of 48 SC16s in total is expected to reach a full saturation limit, defined as $N_{\text{fired SC1s}} = 0.5 \times N_{\mu}$, only at the muon multiplicities of approximately 1000. Such muon densities are expected for shower cores from the primary energies above 100 PeV, which are expected to be recorded less than once during several years of operation time. The conclusion holds also for the setup with 24 SC16s.
Fig. 4.12: A simulated response of a high-granularity scintillator plane consisting of 48 SC16s. Upper panel: The relation between the number of muons and number of fired SC1s for 10 PeV proton shower cores. The number of events is indicated by the colours. The line presents the identity relation. Lower panel: Scaled projection of the distribution in the upper panel for muon numbers between 60 − 61.
4.4 Limited Streamer Tubes

The Limited Streamer Tubes (LST) were originally deployed in the Central Detector of the KASCADE array [51]. In total 66 LST modules were transported to Pyhäjärvi in the year 2012 and they are currently being tested for the development of DAQ. The LSTs are to be used as detector layers in the outer stations, in the array extension and as additional layers in the central stations.

An LST module has a length of 2900 mm, out of which 2750 mm is sensitive, and a width of 1000 mm (see photograph in Figure 4.13). The weight of one module is approximately 20 kg, so their handling is easy in underground conditions. An LST module consists of 6 streamer chambers. One streamer chamber consists of 16 tubes of cross section $9 \times 9 \text{ mm}^2$. An anode wire runs through the central axis of the tube. Tubes are enclosed in a cathode profile and are insulated. Streamer chamber has a length of 2800 mm, width of 166.7 mm and height of 13.4 mm. A voltage is applied to the cathode ($-4.8 \text{ kV}$) and the anode is at ground potential.

Adjacent anode outputs are grouped together in pairs leaving 8 channels for readout per one chamber. Position resolution perpendicular to wires is therefore approximately the size of two tubes, i.e. 2 cm. Influence pads are attached on a polystyrol plate above the streamer chambers. Their width is equal to one chamber (162 mm) and their length 82 mm determines the position resolution along wire direction. 32 pads cover the length of a tube. In total there are 48 readout channels for the wires and 192 channels for the pads per module.

Fig. 4.13: A photograph of an LST module with its aluminum cover removed undergoing testing in Pyhäjärvi surface laboratory in 2014. The acquisition boards of the module stand on top of a copper plate, below which the six streamer chambers are.
4.5 Underground infrastructure

The array is located just by the inclined access road of the mine at a crossing of old tunnels of the height of about 6 metres and width of about 5 metres. Several operations were done in the tunnels for the experiment. A wall separating one tunnel from the others was dismantled and the ceilings were scaled to remove loose rocks and old concrete. A wide crack extending across a tunnel was covered with steel wire and further enforced with sulfur proof concrete.

At the depth of 80 metres, the temperature remains fairly stable at 10° C and the air pressure correlates with surface changes. The tunnel air has a relative humidity of 100% and acid water ($pH = 3$) is dripping from the tunnel ceiling and walls. A station housing shields the detectors and associated hardware from the tunnel air and supports the load of the detectors. The frame of the housing is made of acid-proof steel, which stands on several legs on a concrete groundwork. The walls and the roof are covered with farming plywood and double plastic sheets. Radiators are used to increase the temperature inside the stations reducing the relative humidity and air purifiers are used to filter dust from the air inside the stations. Measurement conditions ($T \approx 20° C$, $RH \approx 50–60\%$) are achieved inside the stations. Humidity, temperature and pressure sensors are used to monitor and log the environmental conditions inside the stations.

Each detector station is equipped with two electric lines. One is for the radiators and lights and another is for the measurement electronics which are behind Uninterruptible Power Supplies (UPS). The measurement electronics and high-voltage supplies situate inside the stations together with DAQ and measurement computers. The underground computers are connected to surface operating room by optical fibre. Cable canals accommodate data, trigger and network cables between the stations.

The gas-mixture for the drift chambers is delivered from a surface gas-mixing and distribution centre hosting argon tanks and carbon-dioxide bottles. The gas-mixture flows to the underground site via a 100-metre-long main supply line, which runs through the rock from the surface to the EMMA site. At the underground site the main supply line connects to a closet where the main supply line is divided to three lines supplying the three stations of each of three arms. The gas flow is adjusted for each station individually. The gas is blown to the tunnels at the end of the lines.
4.6 Rock overburden

The rock around the EMMA site has been modeled in detail [63] and the geography of the EMMA site is visualized in Figure 4.15. The EMMA tunnels are in the border of felsic vulcanite ($\rho = 2.6 \text{ g/cm}^3$) and sericite schist ($\rho = 2.75 \text{ g/cm}^3$). A column of macif vulcanite rock ($\rho = 2.9 \text{ g/cm}^3$) stands directly on the array centre. The ground level above the site is flat besides an open pit, whose edge locates horizontally about 60 metres to East from the array centre position. The rock overburden $X$ is known to different directions and the cutoff energy is estimated by $E_{\text{cut}} = a(e^{X/b} - 1)$, where $a = 500 \text{ GeV}$ and $b = 2.5 \times 10^{-5} \text{ g/cm}^3$ [2, 64]. The cutoff energies vary between the detector stations by a few GeVs. The dependence of the cutoff energy on the azimuthal direction is shown in Figure 4.16 for the detector stations at 80 metres. Most pronounced effect in the azimuthal dependence of the cutoff energy is due to the open pit (relatively low cutoff in eastward direction as seen in Figure 4.16) but its effect is suppressed or non-existent for zenith angles less than 35 degrees.
Fig. 4.15: The rock geology around the EMMA site. The dashed lines denote the boundaries of different rock types at a coordinate depth 80 m. 1 - Felsic vulcanite ($\rho = 2.6 \text{ g/cm}^3$), 2 - Mafic vulcanite ($\rho = 2.9 \text{ g/cm}^3$), 3 - Sericite schist ($\rho = 2.75 \text{ g/cm}^3$), 4 - Cordierite-anthophylite gneiss ($\rho = 2.9 \text{ g/cm}^3$), 5 - Open pit fill ($\rho = 1.7 \text{ g/cm}^3$), 6 - Ore fill ($\rho = 1.9 \text{ g/cm}^3$). The dot-dashed line approximates the edge of the open pit on ground level. The tunnels are drawn with solid lines and the capital letters indicate the positions of the EMMA detector stations. The grid spacing is 50 metres. The coordinate depths of the stations together with the vertical overburden in metres of water equivalent and muon energy cutoffs are given in the inset. The ground level is at a coordinate depth of 7 metres. The compass points refer to a mine coordinate system.
Fig. 4.16: Muon cutoff energies at zenith angles 30 degrees (upper panel) and zenith angles 45 degrees (lower panel) as a function of azimuthal angles as viewed from the location of different detector stations (colours). The compass points refer to a mine coordinate system.
4.7 Summary

The EMMA array consists of eleven detector stations, three central detector stations with the tracking detectors and the high-granularity scintillator setup and eight outer detector stations with the drift chambers and LSTs.

The tracking detectors of the central stations are used for shower arrival direction determination and muon number estimation. As of 2014, the tracking detectors have been taking data for EAS analysis. A study about the muon multiplicity spectra recorded by the tracking detectors is given in the next chapter.

The high-granularity scintillator setup is designed to improve the muon multiplicity measurements especially at high multiplicities. It will also bring useful redundancy to the shower arrival direction reconstruction by the array and provides a start time for drift chambers. As of 2014, the SC16 scintillators are taking underground test data.

The outer stations of the EMMA array consist of drift chambers and LSTs. Their designed functionality is to record the muon densities at the shower tails.
5. EMMA commissioning study

The chapter describes the latest analysis of EAS data recorded by the EMMA array. The analysis confirms that the experiment is taking data efficiently as planned and that the data set is according to the EAS physics expectations. It also demonstrates the methods of event selection and tracking efficiency correction. Analysis results of measured track multiplicity spectra are given. The shape of the recorded multiplicity spectrum indicates that the simplest model of a knee-like spectrum with a pure proton composition cannot explain the data and that further analysis of the spectrum is required. The current systematical uncertainties of multiplicity measurements are commented on.

5.1 Measurement configuration and data taking

In the year 2013, underground data taking with the array took place for nearly 95% of the time. Data taking with the purpose of EAS analysis took place from February 2013 to the end of September 2013. These data were taken solely with drift chambers, and with five stations: with three tracking detectors in the three central stations, and with two outer station prototypes. A top view of the detector stations is shown in Figure 5.1.

The data analysed here were taken over the time interval from May 2013 to September 2013. The analysis consists of the data recorded with the three tracking detectors. The selected set for the analysis is 2352 hours (98 days) and consists of multiple runs, i.e. the data set is not continuous. The data set is chosen on the basis that all three tracking detectors were in stable operation. The recorded data are checked for consistency in the measured single-muon, multiple-muon and station-coincidence rates.

The trigger is generated from the anode signals of the chambers. In the data taking, trigger inputs were generated from each station as a triple-coincidence between any chambers from the three separate detector planes. The final trigger was OR of the station trigger inputs. This trigger logic was used to guarantee efficient recording for the muon events. DAQ and trigger electronics situate inside the stations and signals are exchanged between the stations via cable canals running in the cavern. High voltage is supplied separately within the stations and
5.2 Event reconstruction

The recorded events are reconstructed with the ETANA program (EMMA Tracking Analysis [21]). The event reconstruction procedure consists of hit position, track and shower arrival direction determinations.

The geometry of a tracking detector and a schematic diagram of track reconstruction are shown in Figure 5.2. The hit positions, i.e. the charged particle coordinates within the chambers, in an event are determined based on the recorded anode, near and far signals. An initial estimate of the shower arrival direction is constructed on the basis of the direction, which maximizes the number of parallel tracks passing through the three detector planes. Track searching is then committed based on the initial estimate of shower arrival direction and also tracks, which pass through only two detector planes are reconstructed. Tracks are selected based on a parallelity criterion $\delta \theta < 3^\circ$, where $\delta \theta$ is the angle between the track and the shower direction. A reconstructed track may consist of 2-6 hits. Not all hits are associated with tracks and therefore the hits are classified into two types, hits on-track and hits off-track. The final shower arrival direction is determined as an average of the track arrival directions. The hits, tracks and shower direction are reconstructed independently in each of the three tracking detectors.

5.3 Data set

A data set of 2352 hours (98 days) is selected to the analysis.

Two reconstruction variables are calculated (see [65]) for each multi-track event.
Fig. 5.2: The configuration of drift chambers in a central station (C, F and G). A tracking detector consists of three detector planes (top, middle, bottom), separated vertically by 1125 mm, and each consisting of two layers of chambers (X- and Y-layers). Each plane contains 35 chambers, 15 in Y-layer and 20 in X-layer. A detector consists of 105 chambers in total. The dimensions of the detector are $4220 \times 3650 \times 2250$ mm$^3$. The horizontal bars denote the positions and sizes of the chambers (width 200 mm and height 20 mm). The lengths of the chambers are 3650 mm. The two parallel lines depict the muon tracks, which are reconstructed using on-track hits (solid squares). Events contain also off-track hits, i.e. hits unassociated with tracks (open squares).

The parameter $\langle R \rangle$, called 'bundle size', is calculated based on the spatial track distribution within a tracking detector. The bundle size is defined as $\langle R \rangle = \Sigma_{i=1}^{N_{\text{track}}} r_i / N_{\text{track}}$, where $N_{\text{track}}$ is the number of tracks, and $r_i$ is the distance between the individual track and the mass centre of all the tracks in the event. This distance is calculated on the detector plane and the values of the bundle sizes are specific for a given detector geometry. Another parameter $P$ describes the 'purity' of the reconstruction and is defined as the ratio between the number of all reconstructed hits and the number of reconstructed tracks: $P = N_{\text{hit}} / N_{\text{track}}$. The value of $P$ is expected to be larger for events with accompanying muon-induced secondaries than for 'pure' muon events. $\langle R \rangle$ and $P$ are calculated for each tracking detector separately.

Figure 5.3 shows the distribution of all events in the $(\langle R \rangle, P)$-plane with reconstructed multiplicity $N_{\text{track}} > 4$ for Stations C, F and G. Two distinct regions can be seen in the $(\langle R \rangle, P)$-plane. EAS events have typical values of $\langle R \rangle \sim 1.0 - 1.8$ and $P \sim 6$. Events with smaller values of $\langle R \rangle = 0.1 - 1.0$ and values of $P = 2 - 20$ are interpreted as being single-muon-induced cascade events. All stations record the cascade events (the highest contribution peaking at $(0.4, 7.0)$) and the EAS
component (the bulge at $\langle R \rangle \sim 1.3$) as expected. Simulation expectations of $\langle R \rangle$ and $P$ are presented in Section 5.4.

Figure 5.4 shows 'cascade' and 'bundle' rates over the selected data set. The average cascade rates are approximately 1400, 1400 and 1100 per day for Stations C, F and G respectively. The corresponding bundle rates are some 3% of the cascade rates: approximately 40, 40 and 30 per day. Station G rates are lower than the rates of the other stations because Station G had the largest number of chambers disconnected from HV, namely 22 out of 105, where as C has 14/105 and F 9/105. The rate changes of the order of 10% are due to chamber efficiencies which anticorrelate with air pressure and evolve over time. The development of algorithms for correcting the efficiency changes over time is ongoing [56].

Figure 5.5 shows the coincidence rates between Stations C, F and G. A co-incidence between the stations is defined for events where at least one track is reconstructed in two or three of the stations. The average two-fold coincidence rates between Stations C and F, F and G, and G and C are approximately 850, 750 and 650 per day, where as the rate of three-fold coincidences is approximately 100 per day.

The figures above demonstrate that the tracking detectors are recording data as expected and that these data are consistent over time and between the different stations.
Fig. 5.4: The cascade (black) and bundle (red) rates (events per day) for Stations C, F and G for the data set of 2352 hours. The cascade rates are for events with $N_{\text{track}} > 4$ and the bundle rates are for events with $N_{\text{track}} > 4$ and $P < 1.375 \cdot (\langle R \rangle - 48 \text{ cm})$. The cascade rates are rescaled by a factor of 0.1 for clarity.

Fig. 5.5: The track coincidence rates (events per day) between the stations for the data set of 2352 hours. The rates are given for events with at least one track in two stations (two-fold, the stations indicated by the titles) and at least one track in all three stations (three-fold, CFG).
5.4 Simulations and efficiency correction

A CORSIKA-QGSJET01C simulation is performed in order to compare the measured track multiplicities with expectations from cosmic-ray spectrum and EAS physics model. An important aspect in the simulation is the determination of the detector response, i.e. the relation between the simulated muon numbers and the number of reconstructed tracks in data. In this section, the simulated and measured reconstruction variables are first compared in order to qualify the simulation. Second, an efficiency correction scheme is introduced. It is a method with which the reconstruction efficiencies are deducted from the data and the simulated muon multiplicities corrected in accordance with the efficiencies.

Samples of discrete energy primaries are simulated. The shower arrival angles are randomized according to isotropy and flat detector geometry for zenith angles $\theta < 35$ degrees. The effect of rock is approximated by applying a muon energy cut-off of $E = 45/\cos(\theta)$ GeV. The muon propagation through rock is thus simplified and effects relating to the muon scattering (direction changes and spatial displacement in relation to unscattered muon track) or muon-induced electron production are not included. The shower axis positions are distributed uniformly around the array in an area, which ensures correct yields for the muon multiplicities larger than 4, i.e. circular area of radius up to 100 m. The geometry of the tracking detectors and their locations are as in the experiment. In total, 14 samples (5000 to 200 showers per fixed energy) of different discrete energies are simulated. The simulated energies are $10^i$ PeV, where $i = -0.8, -0.6, \ldots, 1.8$ (an energy interval from 0.2 to 60 PeV). The chosen upper energy limit underestimates the yields for $N_{\text{track}} = 30$ by approximately 10% and for $N_{\text{track}} = 20$ by approximately 5%. For multiplicities less than 20 the value is less. (The estimates are based on an investigation of relative yields from different energy intervals.) The simulation is performed for proton and iron primaries. To generate sufficient statistics, each shower is sampled 50 to 200 times with different shower axis positions. The simulation statistics is 3 to 15 times the data for the energy region most relevant for the yields $N_{\text{track}} > 4$. For the highest energies, the simulation statistics is more than 100 times the data. In order to simulate a realistic cosmic-ray spectrum, each simulated discrete energy $10^i$ PeV is chosen to represent the contribution from an energy interval from $10^{i-0.1}$ to $10^{i+0.1}$ PeV. The contributions from the different energy intervals are weighted according to a realistic energy dependence of the flux. Overall normalization is based on data taking time and an empirical value of the cosmic-ray integrated intensity ($J(E > 10^{15} \text{eV}) = 1.6 \cdot 10^{-10}/\text{cm}^2\ \text{s sr}$). The method is documented in more detail in [61].

A simple detector model is applied to estimate the effect of the chamber efficiencies on the recorded muon numbers in EAS. The inputs to the code are the tracking detector geometry and the efficiency of hit detection, which is given separately for the 1st, 2nd, 3rd and $\geq 4$th hits within a chamber. The hit ordering is based on the hit distance from the anode wire. An event consists of a number of muon trajectories as given by the CORSIKA simulation, and the crossing points of muon trajectories with individual chambers are determined. These crossing points are referred to as 'hits'. Each hit 'survives' according to a survival probability, i.e. efficiency, which depends on the number of hits in the chamber. A track is
counted if a muon trajectory has generated 2 or more survived hits (neglecting trajectories with only 2 hits in the same detector plane). This resembles the track definitions used in ETANA. The output of simulated hits and tracks is compared with ETANA-reconstructed data.

**Qualification of the simulation**

The simulations are qualified by comparing them with data. The data are first selected to the analysis by three conditions: i) the reconstructed shower zenith angle is less than 35 degrees, ii) at least one track is reconstructed in each station (three-fold coincidence) and iii) the reconstructed shower directions from each three stations are parallel within 10 degrees. This selection is used to minimize the cascade contribution in the data. In the upcoming qualification plots the used simulation scenario is a pure proton knee-like spectrum with chosen chamber efficiencies: 1st-hit efficiency 0.84, 2nd-hit efficiency 0.40, 3rd-hit efficiency 0.20 and ≥4-hit efficiency 0.05. The simulated $P$ and $\langle R \rangle$ are given for a proton simulation, but they are insensitive to the primary mass. The data are for Station F.

The dependence of $\langle R \rangle$ on track multiplicity is shown in Figure 5.6 for data and simulation. The width of the $\langle R \rangle$-distribution gets smaller as multiplicities get higher, a fact arising from the definition of $\langle R \rangle$: as the number of tracks increases so do the fluctuations in $\langle R \rangle$ lessen. This behavior is well reproduced in the data except the cascade part, which is evident in the data ($\langle R \rangle \lesssim 0.6$) and which extends only up to multiplicity of about 20 while the EAS events yield up to approximately 40 tracks.

The $(\langle R \rangle, P)$-distribution of events surviving the selection is shown in Figure 5.7 and is compared with a simulation expectation. The figure is generated for track multiplicities $N_{\text{track}} > 4$. The distribution in $\langle R \rangle$ is well matching with the simulation expectation, expect that in the data some 4% of events have $\langle R \rangle$ -values less than expected ($\langle R \rangle \lesssim 0.6$), see also projection in Figure 5.8. (The cascade contribution is reduced by a factor of a thousand in comparison to the case of no coincidence requirement between the stations, compare with Figure 5.3.) Besides the excess of low-$\langle R \rangle$ events, the peak value of $\langle R \rangle$ is slightly less in the data than in the simulation (gaussian fits of the peaks have the means $\mu_{\text{data}} = 1.30 \pm 0.01$ m and $\mu_{\text{sim}} = 1.34 \pm 0.01$ m). The values of $P$ are on the average higher in the data set than in the simulation and distribution width is larger, as shown in the projection in Figure 5.8. These differences are understood by the fact that the used simulation method does not take into account the contribution of muon-induced secondaries (electrons) in the overall hit production in the events. Geant4-simulation of electron contribution was studied earlier and was estimated to be approximately 20% – 50% of all the hits in an event [21, 66]. These numbers fit well to the ratio of average $P$ between the data and the simulation, which is 1.5, the averages being $\langle P \rangle = 6.5$ (data) and $\langle P \rangle = 4.3$ (sim). According to [21, 66], only 5 – 10% of all hits, which are associated with a track, are caused by muon-induced secondaries and therefore this discrepancy is interpreted being non-detrimental for the application of the simulation.

Although the simulation fails to reproduce the hit number dependence $P$, the distribution of the average numbers of on-track hits per track is well reproduced,
Fig. 5.6: The distribution of events in $\langle R \rangle$ as a function of the track multiplicity $N_{\text{track}}$ for Station F. Upper panel: Data. Lower panel: Simulation. The number of events per bin are indicated by the colours.
Fig. 5.7: The ⟨⟨R⟩, P⟩-distribution for Station F for events with \( N_{\text{track}} > 4 \). Upper panel: Data (1192 events in total). Lower panel: Simulation. The colours indicate the number of events per bin. Other stations yield similar distributions.
Fig. 5.8: Upper panel: The distribution of events in $\langle R \rangle$. Data are denoted by the symbols and simulation by solid lines. Lower panel: The distribution of events in $P$. The distributions are shown for Station F for $N_{\text{track}} > 4$. Other stations yield similar distributions. The simulated distributions are normalized to the total number of counts in data.
Fig. 5.9: The distribution of events as a function of average number of on-track hits per track in data and simulation. Data (symbols) are compared with a proton simulation (solid line). The distributions are shown for a multiplicity interval 5-14 and for Station F. The selected data includes 1091 events, an integral to which the simulation curve is normalized.

see Figure 5.9. The numbers of on-track hits can be used as a measure of the detection efficiencies. This will subsequently allow the deduction of the relation between the number of tracks and number of simulated muons, which is presented next.

**Efficiency correction**

In order to determine the expected relation between the muon numbers and the reconstructed track numbers, the average number of on-track hits per track as a function of track multiplicity is investigated. The idea is to compare the output of the reconstructed number of on-track hits with that given by simulations with different efficiency inputs.

The multiplicity dependence of the average number of on-track hits per track averaged over all events with the same multiplicity for three different efficiency scenarios are shown in the left panel of Figure 5.10 and are compared with data. The data samples contain approximately 7600 events each. In efficiency scenario 1, the efficiencies for the 1st, 2nd, 3rd and ≥4th are 0.82, 0.40, 0.20, 0.05 (for C), 0.84, 0.40, 0.20, 0.05 (for F) and 0.76, 0.40, 0.20, 0.05 (for G). Scenario 2 underestimates the number of on-track hits and the corresponding efficiencies are 0.78, 0.30, 0.10, 0.00 (for C), 0.80, 0.30, 0.10, 0.00 (for F) and 0.72, 0.30, 0.10, 0.00 (for G). Scenario 3 is an overestimation with the efficiencies of 0.90, 0.50, 0.30, 0.10
(for C), 0.92, 0.50, 0.30, 0.10 (for F) and 0.84, 0.50, 0.30, 0.10 (for G). The three efficiency scenarios are used in order to reveal the sensitivity of the relation between the simulated muons and the reconstructed tracks on the selected efficiencies.

Stations C and F have similar efficiencies while the 1st hit efficiency of G is the lowest as was expected from previous investigations of the rates. The 2nd and 3rd hit efficiencies are within preliminary results of multihit-efficiency measurements, i.e. 0.1 – 0.5 [56]. The last hit efficiency given is only tentative. The simulation describes the overall features in the data very well, efficiency scenario 1 being the best description. As the track multiplicity rises so does the number of 2nd, 3rd etc. hits, for which the efficiencies are lower than for the 1st hit. This leads to the decrease in the average of the number of on-track hits as a function of track multiplicity. The drop in the value of the average number of hits after the first bin arises from a tracking condition: reconstructed events must contain at least one track which passes through three detector planes. The efficiency model matches the overall behavior of the data well in a multiplicity interval 1 – 20, but for multiplicities larger than 20 the data set has insufficient statistics and therefore the simulated curves are only indicative examples. The maximum reconstructed track number is highly sensitive on the values of the efficiencies, being approximately 45 for the scenario 2, 55 for the scenario 1 and 70 for the scenario 3.

The corresponding response curves, i.e. the relation between the number of simulated muons and the number of reconstructed tracks, as given by the simulation are shown in the right panel of Figure 5.10. The average number of reconstructed tracks is plotted as a function of a reference multiplicity, which is defined as the number of simulated muons, which pass through the middle detector plane of a tracking detector. For the scenario 1, a saturation limit defined as $1.5 \cdot N_{\text{track}}' = N_{\text{muons}}$ is $N_{\text{track}}' \approx 32$ for Stations C and F, and $N_{\text{track}}' \approx 27$ for Station G. These limits vary depending on the efficiency scenario, but in a multiplicity interval up to $N_{\text{tracks}} = 20$, which is also the number of chambers in the X-layer of a detector plane, the average number of tracks per given reference multiplicity differ for the different efficiency scenarios only by less than 10%, and the determination of the response is the most accurate in this multiplicity region. The response curves are shown for fixed energy proton showers (25 PeV), but the response curves are insensitive to the energy and mass of the primary, only the maximum reference multiplicity varies as their function.

The qualification of the simulation described above indicates that the data and simulations have robust conformity and that the overall dependence between the reconstructed tracks and simulated muons is well understood. However, the detailed roles of electron contribution in the hit production and event reconstruction were not taken into account. The uncertainties arising from these simplifications are assumed to be negligible in the multiplicity interval 1 – 20, which is well below the interpreted saturation region affecting the multiplicities above 30. Efficiency determination was done on an average basis over the chambers and time. The role of these dependencies is beyond the scope of the thesis.
Fig. 5.10: Left panel: The average number of on-track hits per reconstructed track as a function of reconstructed track multiplicity for Stations C, F and G. The values are the averages for all events of a given multiplicity. The data are denoted by black symbols and the three simulation options are shown as coloured symbols. The calculated errors are standard errors. The data statistics are of the orders of 1000 for bins 1-4, 100 for bins 5-10, 10 for bins 11-22 and 1 for the highest multiplicities. Bins with only 1 count are shown as open circles. Right panel: The average number of tracks recorded in a tracking detector as a function of the reference multiplicity (number of muons passing the middle detector plane). The example is shown for proton-initiated showers with $E = 25$ PeV and are for the three different efficiency options depicted in Left panel. Upper panel is for Station C, middle panel for Station F and lower panel for Station G.
5.5 Shower direction reconstruction performance

Figure 5.11 shows the difference spectra of reconstructed shower arrival directions from two stations. The spectra contain two-fold coincidence events, i.e. events with at least one reconstructed track in two stations, are shown for coincidences CF, CG and FG, and are for zenith angle cut $\theta < 35^\circ$. The numbers of these coincidences are 69515, 53577 and 60639, respectively, and the distributions are similar in all three cases. It is evident that most of the events are reconstructed as having parallel arrival directions, the modes of the distributions are $1^\circ$. For example, 91% of the events measured simultaneously by Stations C and F are parallel within 10 degrees. The corresponding fractions for CG and FG are 90% and 91%. For cuts $N_{\text{track}}>4$ for C and $N_{\text{track}}>0$ for F, which leaves approximately 2000 events, the shape of the distribution is similar to the one shown in Figure 5.11, but the corresponding fraction of parallel ($\delta\theta < 10^\circ$) shower directions is 83%. The corresponding value for both CG and FG is 82%.

![Fig. 5.11: The spectra of the difference of reconstructed shower arrival directions for station pairs, i.e. at least one track in Stations C and F (blue solid), in Stations C and G (red dashed) and in Stations F and G (dot-dashed green). The spectra include showers with reconstructed zenith angles $\theta < 35$ degrees (the direction determined by the station given first in each pair).](image)

In the case of three-fold coincidences, that is for the events where all three stations record a multiplicity $N_{\text{track}}>0$, there are approximately 8000 events and
two different fractions are calculated. Given an event in Station C with zenith angle $\theta < 35^\circ$, the fraction of events where the shower direction reconstructed by either one of Stations F or G is within $10^\circ$ with respect to C is 96% (called as 2/3 parallelity criterion). The fraction of events where the shower directions reconstructed by Stations F and G are both within $10^\circ$ with respect to C is 87% (3/3 parallelity). The corresponding numbers for Stations F and G are 96%, 87%, and 95%, 86%, respectively. For an example cut on multiplicity, $N_{\text{track}} > 4$, which leaves approximately 1000 events, the corresponding numbers are 96% and 85% for C, 95% and 83% for F, and 94% and 86% for G. The fraction of showers surviving the selections are sensitive to the multiplicity cut as shown in the case of Station F in Figure 5.12.

Fig. 5.12: The fraction of events passing the 2/3 (left panel) and 3/3 (right panel) parallelity criterion as a function of the reconstructed track multiplicity. The data (symbols) are for Station F and the errors are according to Clopper-Pearson intervals [67]. The dotted lines denote fitted linear trends with slopes $= (-1.7 \pm 1.1) \cdot 10^{-3}$ for 2/3 and $= (-6.4 \pm 1.6) \cdot 10^{-3}$ for 3/3.

The differences of reconstructed shower arrival directions indicate that the shower direction reconstruction in ETANA is performing well. More exact investigations and interpretations of the shower arrival direction reconstruction uncertainties and efficiencies are beyond this work.
5.6 Multiplicity spectra

In this section, the track multiplicity spectra of each Station (C, F and G) are compared with simulation expectations. The comparisons are made for shower zenith angles $\theta < 35^\circ$ and for data taking time of 98 days. The cut $\theta < 35^\circ$ is used to minimize effects from an uniformity of rock overburden and possible shower reconstruction failures. The comparisons are shown for station multiplicities $N_{\text{track}} > 4$ and for two groups of event selections: two-fold and three-fold coincidences. Descriptions of these selections are given together with a simulation description, after which the multiplicity spectra are compared and discussed.

Two-fold coincidence selections

Two-fold coincidences are events, where at least one track is recorded in two different tracking detectors. Events are accepted for a given station if the reconstructed shower zenith angle is less than 35 degrees and if at least one track is reconstructed in any of the two other stations. The shower arrival directions between the two stations coincide within 10 degrees for approximately 83% of the events. The $\langle (R), P \rangle$-distribution of these events is shown in Figure 5.13. About 20% of the events have $\langle R \rangle$-values less than expected from EAS simulations, which only includes the muons. The events are cascade-like events and therefore, for these events, the determination of the muon multiplicity is expected to fail. The cascade events are rejected by $\langle R \rangle > 0.6$.

![Fig. 5.13: The distribution of events with $N_{\text{track}} > 4$ and $\theta < 35^\circ$ in the ($\langle R \rangle, P$)-plane for Station F after requirements of a two-fold coincidence and a parallel shower arrival direction ($< 10^\circ$). The selection leaves in total 2930 events. The colours indicate the number of events per bin. Other stations yield similar distributions.](image-url)
Three selections will be shown in the multiplicity spectrum comparisons. The first selection is the number of tracks in a given station after requiring a two-fold coincidence. The second selection is the number of tracks in a given station after requiring a two-fold coincidence and a parallel ($\delta \theta < 10^\circ$) shower direction from another station. The third selection is the number of tracks in a given station after requiring a two-fold coincidence, parallelity and applying the cascade cut $\langle R \rangle > 0.6$.

**Three-fold coincidence selections**

Three-fold coincidence events are such that at least 1 track is reconstructed in each of the three tracking detectors. After the requirement, only 6% of the events are cascades, i.e. $\langle R \rangle < 0.6$, and 4% after subsequent parallelity requirements as was shown in Figure 5.7.

Three selections will be shown in the multiplicity spectrum comparisons. The first selection is the number of tracks in a given station after requiring a three-fold coincidence. The second selection is the number of tracks in a given station after a three-fold coincidence and subsequent cuts: parallel ($\delta \theta < 10^\circ$) shower direction at least from one other station (2/3 parallelity, 96% survival) and a cascade cut (94% survival). The third selection is the number of tracks in a given station after a three-fold coincidence and subsequent cuts: parallel shower direction from both two other stations (3/3 parallelity, 85% survival) and a cascade cut (96% survival).

**Simulated multiplicities**

The simulated multiplicity spectra correspond to two different cosmic-ray spectra: a knee-like spectrum with a pure proton composition ($\gamma = -2.7; E < 3.3$ PeV and $\gamma = -3.1; E \geq 3.3$ PeV) and a pure iron spectrum with a constant spectral index $\gamma = -2.7$. The two selected cases are the lower and upper limit of multiplicity spectrum expected from a two-component description of cosmic-ray spectrum. The normalization of the simulated spectra is based on the data taking time of 98 days and a value of the integrated intensity ($J(E > 10^{15}$ eV) = $1.6 \cdot 10^{-10}$/cm$^2$ s sr). The track multiplicities do not directly translate to primary energies, since the air shower axis position can be anywhere around the array and therefore air shower tails contribute to the spectrum. Different energy intervals contribute up to different multiplicities, the maximum rising as a function of the energy. As an example, the contributions from energies below the knee ($E < 3.3$ PeV) and above are shown in Figure 5.14. The track multiplicity region 5 – 11 is dominated by a contribution from energies below the knee whereas for the multiplicities larger than 11 the dominant contribution is from the energies above the knee.

**Spectrum comparisons**

Figure 5.15 shows the comparisons between the simulation and the measured data for Station C. The comparisons are shown for track multiplicities $N_{\text{track}} > 4$. The two-fold coincidence selections are shown in the upper panel. It is evident that the data are more consistent with a proton simulation than that of iron. The selection of two-fold coincidences only lies above the others, while the parallelity cut ($\delta \theta < 10^\circ$) lessens the numbers by approximately 17%. The cascade cut $\langle R \rangle > 0.6$
Fig. 5.14: The simulated track multiplicity spectrum for the energy regions below the knee $E < 3.3\text{ PeV}$ (red line) and above the knee $E \geq 3.3\text{ PeV}$ (blue line) assuming a pure proton composition with a knee-like spectrum (Station C, eff. scenario 1).

...affects mostly the lowest multiplicity region $N_{\text{track}} = 5 - 15$, lessening the event numbers by $10 - 20\%$. For $N_{\text{track}} > 18$ the cascade cut plays no role. The data lie on the proton simulation at the multiplicity 5 and rises above the proton simulation up to the multiplicity 20. For the highest multiplicities ($N_{\text{track}} > 20$), both the statistics and the increasing uncertainty in the exact efficiency correction hinders the interpretation. The data lie closer to the proton simulation, but since this region is already in the saturation domain of the tracking detector, an interpretation of a proton dominant composition at the highest multiplicities can not be done. The differences between the different data sets reflect the current systematical uncertainties of the event selection. The efficiency scenarios 1-3 illustrate the role of the efficiency correction. The differences between the event numbers from scenarios 2 and 3 in comparison to the scenario 1 are within $\pm 10\%$ for the multiplicity region $5 - 20$, above which the uncertainty of the efficiency correction increases.

Three-fold coincidences are shown in the lower panel of Figure 5.15. After the requirement of the third coincidence, the number of events in the multiplicity region $N_{\text{track}} \lesssim 20$ are lower than in the two-fold coincidence case, the effect being approximately $45\%$ at the multiplicity of 5. Here the data lie below the proton simulation by approximately $30\%$ at the multiplicity of 5 and rise on and above the proton simulation up to the multiplicity 20. At the highest multiplicities, the simulations and data match with the case of two-fold coincidences. Of the different selections of three-fold coincidences, the cascade cuts play only a marginal
role, while the uncertainty of the event numbers on the used parallelity criteria is approximately 15%. Overall the case of three-fold coincidences is similar to that of the two-fold case, especially if the selection of three-fold coincidence and a $2/3$-parallelity criterion is compared with the selection of two-fold coincidence, parallelity and cascade cut: the data lie below the proton simulation at the multiplicity of 5 and rises above it in the range up to the multiplicity 20.

In Figures 5.16–5.17 the spectra are shown for Stations F and G. The results and interpretations of the measured spectrum are alike that of Station C. For Station G, the rise of data points in the multiplicity interval 5 – 20 with respect to the proton simulation is not as evident as for C and F. It is assumed that this difference may be due to higher number of chambers not in operation in Station G than in the others, which can bring unquantified uncertainty in to the efficiency correction currently based on averages.
Fig. 5.15: The measured (symbols) and simulated (lines) track multiplicity spectra for Station C. Upper panel: the case of two-fold coincidences. Data: gray crosses = the number of two-fold coincident events, open circles = the number of events after a parallelity requirement, filled circles = the number of events after a subsequent cascade cut. Simulation: red lines represent a pure proton composition with a knee-like spectrum, blue lines a pure iron composition with a constant spectral index (see text). The simulated curves include the efficiency correction: solid line = efficiency scenario 1, dotted line = scenario 2, dashed line = scenario 3. Lower panel: the case of three-fold coincidences. Data: gray crosses = the number of three-fold coincident events, open circles = the number of events after a parallelity requirement in any of the two other stations and a cascade cut, filled circles = the number of events after a parallelity requirement in both two other stations and a cascade cut. Simulated cosmic-ray spectra are as in the upper panel. The symbols overlap especially in the region of the highest multiplicities. Errors are Gehler’s approximation of 1-sigma Poisson asymmetric errors [68].
Fig. 5.16: The measured (symbols) and simulated (lines) track multiplicity spectra for Station F. See the caption of Figure 5.15 for explanations.
Fig. 5.17: The measured (symbols) and simulated (lines) track multiplicity spectra for Station G. See the caption of Figure 5.15 for explanations.
5.7 Hypothesis tests on the shape of the spectrum

The shape of the measured multiplicity spectrum is studied in this section. For
the study a normalized ratio \( Y = N/N_p \) is defined, where \( N \) is the normalized
number of events such that \( N = N_p \) for the track multiplicity 5 and \( N_p \) is the
simulated number of events for pure proton knee-like spectrum. Figure 5.18 shows
the normalized ratios of data, proton and iron simulations in the multiplicity range
5 – 20. In this multiplicity range, the systematical uncertainty related to the
saturation effects is the least. The data are the sum of Stations C and F. These
two stations were selected to the sum since they have more chambers in operation
and their efficiencies are more stable than in Station G.

From Figure 5.18 it is evident that the normalized ratio \( Y \) rises above the
pure proton knee-like spectrum. To estimate the significance of the rise, a \( p \)-value
test is committed by fitting the simulation shape (a constant function \( Y = C \))
with the data and calculating the \( \chi^2 \)-value and the corresponding \( p \)-values for the
case of 15 degrees of freedom. The test indicates that the shape of the recorded
multiplicity spectrum is not consistent with a pure proton knee-like spectrum with
very high significance (\( \chi^2_{\text{two-fold}} = 64.38 \) and \( \chi^2_{\text{three-fold}} = 61.07 \) correspond to \( p \)-values < 0.0001 for both two-fold and three-fold coincidence cases). Similarly,
the data are tested for consistency with a proton simulation of constant spectral
index \( \gamma = -2.7 \) by calculating the normalized ratio \( Y' = N/N_{p'} \), where \( N_{p'} \) is the
number of events of the proton simulation with \( \gamma = -2.7 \), and fitting a constant
function \( Y' = C \) with the data. The test indicates that the shape of the recorded
multiplicity spectrum is consistent with a pure proton spectrum with constant
spectral index (\( \chi^2_{\text{two-fold}} = 12.06 \) and \( \chi^2_{\text{three-fold}} = 18.92 \) correspond to \( p \)-values
0.63 and 0.22, respectively).

The shape of the measured multiplicity spectrum may also be interpreted as
being due to an increase in the proportion of a heavy component in the spectrum.
Figure 5.19 shows the normalized ratios of data, proton, iron and a composition-
mixture simulation in the multiplicity range of 5 – 20. The mixture model is
constructed as a weighted sum of iron spectrum (\( \gamma = -2.7 \)) and proton knee-like
spectrum, where the weights depend on energy. A linear decrease of proton content
from 1.0 to 0.4 and a respective increase in the iron content from 0.0 to 0.6 as a
function of the logarithm of energy in the energy interval of 4 – 80 PeV is used.
The shape of the recorded multiplicity spectrum is consistent with the mixture
model (\( \chi^2_{\text{two-fold}} = 12.06 \) and \( \chi^2_{\text{three-fold}} = 16.54 \) correspond to \( p \)-values 0.67 and
0.35, respectively).
Fig. 5.18: The measured (symbols) and simulated (lines) normalized ratio $Y$ as a function of track multiplicity (see text). Upper panel: Two-fold coincidence with a requirement of a parallel shower direction and a cascade cut. Lower panel: Three-fold coincidence with a requirement of a parallel shower direction from one station (2/3 parallelity) and a cascade cut. The simulations are for the efficiency scenario 1 described in Section 5.4 and are listed in the legend.
Fig. 5.19: The measured (symbols) and simulated (lines) normalized ratio $Y$ as a function of track multiplicity (see text). Upper panel: Two-fold coincidence with a requirement of a parallel shower direction and a cascade cut. Lower panel: Three-fold coincidence with a requirement of a parallel shower direction from one station (2/3 parallelity) and a cascade cut. The simulations shown are for the efficiency scenario 1 described in Section 5.4 and are listed in the legend.
5.8 Conclusions

The analysis of the data recorded in the year 2013 shows that the experiment is recording data in an efficient and rigorous way. The EAS events are selected to the analysis by applying coincidence requirements between the stations and by further filtering the cascade contribution with a method based on reconstruction variables $\langle R \rangle$ and $P$. The relation between the number of reconstructed tracks in a tracking detector and the number of simulated muons is determined by an efficiency model, where the efficiencies of hit detection are set to match the dependence of the number of on-track hits on the reconstructed track multiplicity.

The multiplicity spectrum is analysed in the track multiplicity interval 5 – 20, which is below the interpreted saturation region of a tracking detector and corresponds to primary cosmic rays of the energies below and above the knee. The current systematical uncertainty of the absolute intensity of the multi-muon events is about 20%. The uncertainty arises from the uncertainties in the shower direction reconstruction and the cascade contribution. The studied shape of the multiplicity spectrum is not consistent with the simplest model of a pure proton knee-like spectrum, but it is consistent with an increase in the proportion of heavy nuclei as a function of rising energy above the knee energy of 4 PeV as was demonstrated by a simple two-component model of cosmic-ray spectrum.

To study the muon multiplicity spectrum further, ten-fold data taking time is required to guarantee sufficient statistics also for the highest multiplicities. This corresponds to three years of running time. Besides taking more data, the analysis of the muon multiplicity spectrum is to be extended to a wider multiplicity range, also accommodating the multiplicities $N_{\text{track}} \gtrsim 30$. This will be achieved after data taking with the high-granularity scintillator setup especially designed for the purpose of the measurement of the highest multiplicity events.
6. Discussion and outlook

The methods and results presented in the work are discussed in this chapter.

**Shower reconstruction**

The latest simulations of shower reconstruction in EMMA are presented in Section 3.6. The shower reconstruction is based on a parametrization of the muon lateral density distribution and a fit algorithm for determining the shower axis position, the shower core density $\rho(1)$ and the shape parameter $R_0$. The function in Eq. 3.1 has been used in the analysis of shower reconstruction, but different functions can also be used. Further work to be trialled for, include the optimization of the used fit algorithm. In the shower reconstructions presented only one fit was used in reconstructing the shower axis position, the core density and the shape parameter. Additional work on the optimization of shower axis reconstruction should include a study of different fit methods trying to find iterative approaches for the reconstruction employing several fits. The development of the shower reconstruction algorithm should aim to establish the optimal methodology and selection cuts for the shower reconstruction. Such methodology should aim to lessen the deviation of the reconstruction bias factors from the ideal and aim to establish better energy and composition indicators for the experiment (such as linear combinations of $R_0$ and $\rho(1)$). The efficiencies of muon detection in the detector stations should be included into the simulations.

**Unfolding of composition**

In the strategy for the reconstruction of the spectrum and composition, the distribution of the events in the $(\rho(1), R_0)$-plane is used. An example of unfolding is presented in Section 3.5 and demonstrates the potential of such a strategy. It has been shown that the distributions of proton and iron showers separate in the $(\rho(1), R_0)$-plane making it possible to reconstruct the composition in a two-component description of cosmic-ray spectrum, but for instance in the unfolding analysis of KASCADE data, five different element groups were used. The muon lateral density distributions of primaries with masses between proton and iron are between that given for proton and iron and correspondingly the showers of these
intermediate masses are expected to group between proton and iron showers in the \((\rho(1), R_0)\)-plane. Further work is required to study whether additional elements besides proton and iron could be used in the unfolding analysis of EMMA data.

**Event reconstruction in tracking detectors**

The commissioning analysis shown in Chapter 5 indicates that the track reconstruction in the tracking detectors is saturating for track multiplicities \(N_{\text{track}} \gtrsim 30\). Further understanding about the saturation of the tracking detector is to be acquired after a successful determination of the multi-hit efficiencies of the drift chambers by comparing the drift chamber and scintillator data.

In the EMMA commissioning study, the data is reconstructed with ETANA analysis program. The algorithm of ETANA program is optimally designed to work with scintillator start time, which is not yet utilized in the data taking of 2013. This may lead to hit position uncertainties up to 5 cm in the direction of the drift for hits with no counterpart in an overlapping chamber and further optimization is required in the delay line direction as well. Further checks are to be done about the relative alignments between the detector stations currently estimated to be known within about 1°.

The shower arrival direction difference spectra shown in Section 5.5 include all events with \(N_{\text{track}} > 0\). Further studies on the multiplicity dependence of shower arrival direction reconstruction accuracy and resolution are required. In this work, the average of all track directions is used to estimate the shower arrival direction. Better estimates can be achieved by selecting only the most parallel tracks to the average.

**Event selection**

The basis of the event selection for EAS analysis described in Chapter 5 is well understood. The application of event classification in the \(\langle R \rangle, P\) plane and use of coincidences between the stations allows a clean selection of the EAS events from all the recorded events. Further work remains in establishing the absolute intensity values for the multi-muon events. A careful study on the shower arrival direction reconstruction efficiencies and direction difference spectra is required.

**Efficiency correction**

The method of efficiency correction presented in Section 5.4 demonstrates a straight way for the correction between the simulated muon numbers and the recorded track numbers. The contribution of muon-induced secondaries is omitted in the simulation and the efficiency correction relies on the use of average efficiency parameters. In reality some chambers are disconnected from HV and have different efficiencies, and the efficiencies may evolve over time. The effect of these spatial and temporal efficiency variations should be evaluated by testing the sensitivity of the efficiency correction with different efficiency schemes that include the spatial and temporal variations or by using the exact efficiencies determined from data in to the simulations. It will also be useful to validate the efficiency correction method with Geant4 simulations of ETANA, which should include the muon-induced secondary production together with a detector model of sufficient details.
Muon multiplicity spectrum

The $p$-value tests presented in Section 5.7 indicate that the shape of the measured muon multiplicity spectrum is not consistent with a pure proton knee-like spectrum and that the shape can be explained by a model of increasing mass composition in the knee region of cosmic rays in accordance with many other experimental results presented in Chapter 2. There is ambiguity in the interpretation however, since a pure proton spectrum of constant spectral index $-2.7$ may explain the data as well. This ambiguity can be resolved by extending the analysis to higher multiplicities. These sort of ambiguities are intrinsic in the analysis of the muon multiplicity spectrum, but are not expected to be present in the composition analysis strategy utilizing the muon densities recorded in several stations, the shower reconstruction and energy and mass indicators.

The simulations and hadronic interaction models

The simulation used in the analysis of Chapter 5 is executed for the purpose of a successful qualitative interpretation of the data. In order to analyse the data further, a preferred scenario will be to simulate a continuous spectrum choosing the lower and upper energy limits such that the muon multiplicities $1 - 4$ will also be incorporated into the analysis and that the high-energy limit will be extended further than the used $10^{1.8}$ PeV.

The hadronic interaction model QGSJET-01C has been used throughout the simulations executed for the design and commissioning of EMMA. This model is a pre-LHC model and results of LHC have been incorporated into more recent interaction models used in EAS simulations, which are discussed for example in [69]. One example of a post-LHC model is the QGSJETII-04 model. To estimate the influence of the latest model to the analysis presented in Chapter 5, samples of 100 EAS are simulated with QGSJETII-04 in the zenith angle interval $0 - 35$ degrees and the yields of the muons with energies larger than $45/\cos \theta$ GeV are counted. Simulations with QGSJETII-04 result in larger muon yields than with QGSJET-01C: approximately $8 \pm 3\%$ and $11 \pm 3\%$ larger muon yields for 4 and 40 PeV proton showers, and $7 \pm 1\%$ and $9 \pm 1\%$ for 4 and 40 PeV iron showers. For continuing the analysis of multiplicity spectrum, it will be preferable to incorporate the latest model tuned with LHC data into the analysis.

High muon-multiplicity events

With the data recorded by the tracking detectors of the central stations, EMMA is not able to confirm the existence of the highest muon-multiplicity bundles reported by the ALEPH, DELPHI and ALICE experiments presented in Section 2.3. It is concluded that this is expected since the tracking detector saturates at the multiplicities $N_{\text{track}} \gtrsim 30$. The EMMA experiment will use the high-granularity scintillator setup in order to study the highest multiplicity events with data taking time of years surpassing the data taking time of the CERN experiments.

The simulation work presented in Section 4.3.2 shows that the secondaries accompanying the muons contribute on average approximately $20\%$ in the number of fired scintillators in an event. This simulation work is to be continued together
with scintillator measurements to further understand the resolution of muon number determination in the scintillator setup. Spatial clustering of fired scintillators in these events should be studied in order to find optimal muon number estimators. The muon numbers measured with the scintillators are to be cross-checked with the multiplicities recorded by the tracking detector for further improving the understanding of the muon number measurement in the central stations.

The scintillator setup of the array is expected to saturate only at the muon densities of approximately $64/m^2$ which correspond to primary energies of about 100 PeV. A setup of 24 SC16s ($A = 6$ m$^2$) started to take data in Station G in the summer of 2014.

\textit{Prospects}

The next step in the development of the experiment should be to design and construct a surface array above the EMMA underground array. A surface array would measure the shower axis position and the electron numbers in coincidence with the high-energy muon densities. With a measurement of two shower components, array’s performance in the measurement of cosmic-ray spectrum and composition would improve. Work towards an optimal design for the surface array of EMMA should be pursued for.
7. Summary of thesis

The work describes the design, construction and commissioning of the Experiment with MultiMuon Array (EMMA) introduced in Chapter 1. The experiment is an underground array of eleven detector stations at the depth of 80 m (and 40 m) in the Pyhäsalmi mine, in Pyhäjärvi, Finland. The experiment measures the lateral density distribution of muons in Extensive Air Showers (EAS) with an energy cutoff $E > 45$ GeV ($E > 22$ GeV).

In Chapter 2, the scientific background to EMMA is presented. The cosmic-ray energy spectrum and elemental composition are introduced and the distinct features of the spectrum, as measured with several experiments utilizing the measurement of different components of EAS, are presented. At the knee energy of about 4 PeV, the spectrum steepens and the average mass of elemental composition is observed to increase. In order to further measure the details of the spectrum and composition at the knee region new approaches and new experiments are needed. Anomalously high muon multiplicities have been measured by experiments at shallow depths underground.

In Chapter 3, the design calculations of the EMMA experiment are given. The EMMA experiment is designed and optimized to measure the cosmic-ray composition at the knee region by applying shower reconstruction and composition unfolding techniques. The experiment has an acceptance of at least 300 m$^2$ sr and about 3000 EAS above the energy of 4 PeV, 500 above 10 PeV and 30 above 40 PeV will be reconstructed with a data taking time of three years. The reconstruction of cosmic-ray composition will be limited by a threshold energy of shower reconstruction around 1–4 PeV.

In Chapter 4, the realization of the EMMA experiment is introduced. The EMMA array consists of eleven detector stations: three central stations containing the tracking detectors and the high-granularity scintillator setup, and eight outer detector stations. The tracking detectors of the central stations are used for shower arrival direction determination and muon number estimation. The high-granularity scintillator setup is designed to improve the muon number measurements especially at high multiplicities. The designed functionality of the outer detector stations is to record the muon numbers at the tails of the shower lateral density distribution.
In Chapter 5, the EMMA commissioning study is presented. The analysis of the data recorded by array’s three tracking detectors in 2013 shows that the experiment is recording data in an efficient and rigorous way. The EAS events are selected to the analysis by applying coincidence requirements between the stations and by further filtering the cascade contribution with a method based on reconstruction variables $\langle R \rangle$ and $P$. The relation between the number of reconstructed tracks in a tracking detector and the number of simulated muons is determined by an efficiency model, where the efficiencies of hit detection are set to match the dependence of the number of on-track hits on the reconstructed track multiplicity. The shape of the muon multiplicity spectrum in the multiplicity range $5 - 20$ is not consistent with the simplest model of a pure proton knee-like spectrum, but it is consistent with an increase in the proportion of heavy nuclei in the energies above the knee as was demonstrated with a two-component model of the cosmic-ray spectrum.

The results and methods of the work are discussed in Chapter 6. It is important to continue the spectral studies with longer data taking times and with the high-granularity scintillator setup.
List of author’s publications and presentations

List of most relevant publications related to the thesis


**List of other publications related to cosmic-ray physics**


**List of publications related to large-scale neutrino detectors**


**List of publications related to ALICE**

The author has contributed to the commissioning, data analysis and monitoring of the T0 detector of the ALICE experiment by 13 visits each lasting between 1 week to 2 months to CERN during the years 2008–2014. This has entitled the author to participate in several publications as a member of the ALICE Collaboration:


ALICE Collaboration: Centrality Dependence of Charged Particle Production at Large Transverse Momentum in Pb-Pb Collisions at $\sqrt{s_{NN}} = 2.76$ TeV, *Phys.*
ALICE Collaboration: Centrality dependence of Pion, Kaon, and Proton Production in Pb-Pb Collisions at $\sqrt{s_{NN}} = 2.76$ TeV, Phys. Rev. C 88 (2013) 044910


ALICE Collaboration: Centrality determination of Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV in the ALICE experiment, Phys. Rev. C 88 (2013) 044909


ALICE Collaboration: Charge separation relative to the reaction plane in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, Phys. Rev. Lett. 110 (2013) 012301


ALICE Collaboration: Long-range angular correlations on the near and away side in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, Phys. Lett. B 719 (2013) 29-41


ALICE Collaboration: Mid-rapidity anti-baryon to baryon ratios in pp collisions at $\sqrt{s} = 0.9, 2.76$ and 7 TeV measured by ALICE, Eur. Phys. J. C 73 (2013) 2496

ALICE Collaboration: Net-Charge Fluctuations in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, Phys. Rev. Lett. 110 (2013) 152301

ALICE Collaboration: Performance of the ALICE V0 system, JINST 8 (2013) P10016

ALICE Collaboration: Pseudorapidity density of charged particles in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, Phys. Rev. Lett. 110 (2013) 032301
ALICE Collaboration: Transverse Momentum Distribution and Nuclear Modification Factor of Charged Particles in p-Pb Collisions at $\sqrt{s_{NN}} = 5.02$ TeV, Phys. Rev. Lett. 110 (2013) 082302


ALICE Collaboration: $K_{0s} - K(0s)$ correlations in $pp$ collisions at $\sqrt{s} = 7$ TeV from the LHC ALICE experiment, Phys. Lett. B 717 (2012) 151-161

ALICE Collaboration: Production of Muons from Heavy Flavor Decays at Forward Rapidity in $pp$ and Pb-Pb Collisions at $\sqrt{s_{NN}} = 2.76$ TeV, Phys. Rev. Lett. 109 (2012) 112301


ALICE Collaboration: Neutral pion and $\eta$ meson production in proton-proton collisions at $\sqrt{s} = 0.9$ TeV and 7 TeV, Phys. Lett. B 717 (2012) 162-172


ALICE Collaboration: Measurement of charm production at central rapidity in proton-proton collisions at $\sqrt{s} = 2.76$ TeV, JHEP 1207 (2012) 191

ALICE Collaboration: Transverse sphericity of primary charged particles in minimum bias proton-proton collisions at $\sqrt{s} = 0.9, 2.76$ and 7 TeV, Eur. Phys. J. C72 (2012) 2124


ALICE Collaboration: Suppression of high transverse momentum D mesons in central Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, JHEP 1209 (2012) 112


ALICE Collaboration: Measurement of Event Background Fluctuations for Charged Particle Jet Reconstruction in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, JHEP 1203 (2012) 053


ALICE Collaboration: Underlying Event measurements in $pp$ collisions at $\sqrt{s} = 0.9$ and 7 TeV with the ALICE experiment at the LHC, JHEP 1207 (2012) 116


ALICE Collaboration: Femtoscopy of $pp$ collisions at $\sqrt{s} = 0.9$ and 7 TeV at the LHC with two-pion Bose-Einstein correlations, Phys. Rev. D84 (2011) 112004


ALICE Collaboration: Suppression of Charged Particle Production at Large Transverse Momentum in Central Pb-Pb Collisions at $\sqrt{s_{NN}} = 2.76$ TeV, Phys. Lett B 696 (2011) 30-39


ALICE Collaboration: Charged-particle multiplicity density at mid-rapidity in


ALICE Collaboration: Midrapidity antiproton-to-proton ratio in pp collisions at $\sqrt{s} = 0.9$ and 7 TeV measured by the ALICE experiment, *Phys. Rev. Lett.* 105 (2010) 072002


List of presentations given by the author since 2006

'The Muon Multiplicity Analysis with EMMA Tracking Stations', EMMA Collaboration Meeting, Pyhäjärvi Finland, 8.5.2014


'Progress in data analysis’, EMMA Scientific Advisory Board meeting, Pyhäjärvi, Finland, 15.-17.5.2013

'Analysis of Runs 04 and 105’, EMMA-Days Pyhäjärvi, Finland, 23.11.2012

'EAS selection in the EMMA underground array’ (poster), ECRS2012, Moscow, Russia, 5.7.2012

'About delay matching of calibration and underground data, multiplicity data 2012 and EMMA-site rock composition’, EMMA-Days Pyhäjärvi, Finland, 26.4.2012


'Muon multiplicity measurements in EMMA’, Particle Physics Day, Helsinki, Finland, 28.10.2011

'EMMA - the experiment and analysis of muon multiplicity events’, EMMA+ALICE Cosmics meeting, Pyhäjärvi, Finland, 31.8.2011

'Progresses on analysis of muon multiplicity distribution’ and 'Scintillator performance in EMMA’, EMMA Scientific Advisory Board meeting, Pyhäjärvi, Finland, 4.-5.5.2011

'Event classification in EMMA’ (poster), Physics Days 2011, Helsinki, Finland, 30.3.2011

'Beam as seen by T0’, ALICE Heavy Ion First Physics, CERN, Geneva, Switzerland, 18.11.2010

'Analysis of multitrack events’, EMMA/LAGUNA-FI meeting, Pyhäjärvi, Finland, 2.11.2010

'Developments in shower reconstruction and composition analysis for CARPET-3 EAS array’, European Cosmic Ray Symposium, Turku, Finland, 4.8.2010

'Timing measurement with SC16 scintillator detectors’, EMMA/LAGUNA-FI meeting, Pyhäjärvi, Finland, 23.6.2010
'Physics topics for EMMA’ and 'Cosmic ray experiment EMMA - Underground muon tracking unit’ (poster), Physics Days, Jyväskylä, Finland, 11.-13.3.2010

'Listing some of the physics topics for EMMA’, EMMA/LAGUNA-FI meeting, Pyhäjärvi, Finland, 5.2.2010

'EMMA layout design and shower reconstruction’ and 'Analysis of muon intensity on ground’, EMMA Scientific Advisory Board meeting, Pyhäjärvi, Finland, 28.9-29.9.2009

'Introduction to ALICE cosmics: a look at cosmic run’ (poster), Physics Days, Espoo, Finland, 12.-14.3.2009


'CARPET EAS array: Analysis of average lateral density distribution of electrons’, EMMA-Days, Pyhäjärvi, Finland, 6.10.2008

'Status of the Muon Underground-Experiment EMMA in the Pyhäsalmi Mine, Finland’, Karlsruhe Research Center, Germany, 9.6.2008

'Status of the Underground Cosmic-Ray Experiment EMMA’, Physics Days, Turku, Finland, 27.-29.3.2008

'EMMA: Simulations of physics performance' and 'On underground muon multiplicity spectra’, EMMA Collaboration Meeting, Pyhäjärvi, Finland, 17.5.2007

'EMMA-experiment: an idea to measure cosmic-ray composition’, XIV-th International School Particles and Cosmology, Baksan, Russia, 21.4.2007

'Measuring cosmic-ray composition with the EMMA-experiment’, Particle Physics Day, Helsinki, Finland, 24.11.2006

'Simulations and analysis of the EMMA-experiment’ (poster), Physics Days, Tampere, Finland, 9.-11.3.2006
References

[8] E.S. Seo, Astropart. Phys. 39-40 (2012) 76-87 (Figure reprinted with permission from Elsevier.)
[42] V. Avati et al., Astropart. Phys. 19 (2003) 513-323 (Figure reprinted with permission from Elsevier.)
[43] J. Abdallah et al., Astropart. Phys. 28 (2007) 273-286 (Figure reprinted with permission from Elsevier.)
[44] B. Alessandro, “ALICE: Study of the cosmic multiple muon events: results and perspectives”, talk given at ‘Results and prospects of forward physics at the LHC: Implications for the study of diffraction, cosmic ray interactions, and more’, CERN, Switzerland, 17.2.2013
[46] V. Avati et al., CERN/2001-003 SPSC/P231
[52] F. James, CERN Program Library Long Writeup D506, CERN 1994
[56] P. Kuusiniemi et al., EMMA Instrument paper, to be published
[58] ”User Manual for set of scientific measurement equipment of the scientific setup for cosmic rays (SSCR-SSE)”, internal (2010)


[66] T. Räihä, Private communication

[67] ROOT Manual, TEfficiency:ClopperPearson

[68] Gehler’s approximation taken from
