

**UNIVERSIDADE DE SÃO PAULO  
INSTITUTO DE FÍSICA DE SÃO CARLOS**

**Raul Ribeiro Prado**

**Modelo para teses e dissertações em  $\text{\LaTeX}$  utilizando a  
classe USPSC para o IFSC**

**São Carlos**

**2017**



**Raul Ribeiro Prado**

**Modelo para teses e dissertações em  $\text{\LaTeX}$  utilizando a  
classe USPSC para o IFSC**

Tese apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para obtenção do título de Doutor em Ciências.

Área de concentração: Física Aplicada

Supervisor: Prof. Dr. Luiz Vitor de Souza Filho

**Versão original**

**São Carlos  
2017**

Folha de aprovação em conformidade  
com o padrão definido  
pela Unidade.

No presente modelo consta como  
folhadeaprovacao.pdf





# Acknowledgements





*“O estudo, a busca da verdade e da beleza são domínios  
em que nos é consentido sermos crianças por toda a vida.”*

*Albert Einstein*



# Abstract

PRADO, R. R. **Model for theses and dissertations in  $\text{\LaTeX}$  using the USPSC class to the IFSC.** 2017. 62p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2017.

This is the english abstract.

**Keywords:** .



# Resumo

PRADO, R. R. **Modelo para teses e dissertações em  $\text{\LaTeX}$  utilizando a classe USPSC para o IFSC**. 2017. 62p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2017.

**Palavras-chave:** .



# List of Figures





# List of Tables

Table 1 – . . . . .	54
---------------------	----



# Contents

<b>1</b>	<b>Introduction</b>	<b>19</b>
<b>2</b>	<b>Extensive air showers</b>	<b>21</b>
2.1	EM component and the $X_{\max}$	21
2.2	Hadronic component and the $N_{\mu}$	21
2.3	Observables	21
2.3.1	$X_{\max}$	21
2.3.2	$N_{\mu}$	21
2.3.3	$X_{\max}^{\mu}$	21
<b>3</b>	<b>Pierre Auger Observatory</b>	<b>23</b>
<b>4</b>	<b>NA61/SHINE experiment</b>	<b>25</b>
<b>5</b>	<b>Interpretation of measurements of the number of muons in extensive air shower experiments</b>	<b>27</b>
<b>6</b>	<b>A new air-shower observable to constrain hadronic interaction models</b>	<b>41</b>
<b>7</b>	<b>Hadron production in pion-carbon interactions</b>	<b>53</b>
7.1	Dataset and simulations	53
7.2	Upstream and event selections	54
7.3	Track selection	54
7.4	$V^0$ selection	54
7.5	Phase space binning	55
7.6	$dE/dx$ analysis	55
7.6.1	$dE/dx$ measurements	56
7.6.2	$dE/dx$ model	57
7.6.3	$dE/dx$ fit strategy	58
7.7	$V^0$ analysis	58
7.7.1	Signal extraction	58
7.8	Corrections	58
7.9	Spectra	58
7.9.1	Statistical uncertainties	58
7.9.2	Systematic uncertainties	58
7.10	Results	58
7.11	Summary and conclusions	58
<b>8</b>	<b>Conclusions</b>	<b>59</b>
	<b>Bibliography</b>	<b>61</b>



# 1 Introduction



## 2 Extensive air showers

### 2.1 EM component and the $X_{\max}$

### 2.2 Hadronic component and the $N_{\mu}$

[1–5]

average pt against spectra of muons at the ground

### 2.3 Observables

2.3.1  $X_{\max}$

2.3.2  $N_{\mu}$

2.3.3  $X_{\max}^{\mu}$





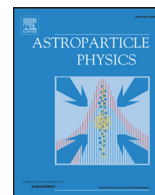
### 3 Pierre Auger Observatory



## 4 NA61/SHINE experiment



## 5 Interpretation of measurements of the number of muons in extensive air shower experiments



# Interpretation of measurements of the number of muons in extensive air shower experiments



Raul R. Prado<sup>a,\*</sup>, Ruben Conceição<sup>b</sup>, Mário Pimenta<sup>b</sup>, Vitor de Souza<sup>a</sup>

<sup>a</sup> Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, Brazil

<sup>b</sup> Laboratório de Instrumentação e Física Experimental de Partículas, Lisbon, Portugal

## ARTICLE INFO

### Article history:

Received 25 January 2016

Revised 8 June 2016

Accepted 19 July 2016

Available online 20 July 2016

### Keywords:

Ultra high energy cosmic rays

Muons

Composition

## ABSTRACT

In this paper we analyze the energy evolution of the muon content of air showers between  $10^{18.4}$  and  $10^{19.6}$  eV to be able to determine the most likely mass composition scenario from future number of muons measurements. The energy and primary mass evolution of the number of muons is studied based on the Heitler–Matthews model and Monte Carlo simulation of the air shower. A simple model to describe the evolution of the first and second moments of number of muons distributions is proposed and validated. An analysis approach based on the comparison between this model's predictions and data to discriminate among a set of composition scenarios is presented and tested with simulations. It is shown that the composition scenarios can be potentially discriminated under the conditions imposed by the method. The discrimination power of the proposed analysis is stable under systematic changes of the absolute number of muons from model predictions and on the scale of the reconstructed energy.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

The energy spectrum of ultra high energy cosmic rays (UHECRs) has been measured recently with high precision and two major features were confirmed. The ankle ( $\log(E/\text{eV}) \sim 18.7$ ) and the flux suppression ( $\log(E/\text{eV}) \sim 19.5$ ) have been undoubtedly established by HiRes [1], the Pierre Auger Observatory [2,3] and Telescope Array [4]. However, the astrophysical interpretation of these structures cannot be inferred with complete certainty mainly because of the lack of knowledge on the UHECR composition at these energies. In a light abundance scenario, the ankle could be interpreted as the modulation resulting from the particle interaction with radiation backgrounds [5,6]. On the other hand, it could also be explained as the transition from galactic to extra-galactic cosmic rays [7]. The flux suppression can be equally well described by the energy losses of extra-galactic particles due to interactions with CMB photons [8] or by the maximum reachable energy of the astrophysical acceleration mechanisms in nearby sources [9]. In each one of these astrophysical scenarios, the energy evolution of the UHECR composition is significantly different.

The UHECR measurements are done indirectly through the detection of extensive air showers. Therefore, the determination of

the composition depends strongly on the data analysis capability to correlate the measured properties of the shower to the primary particle type. This correlation is achieved using air shower simulations. However, intrinsic fluctuations of the showers and uncertainties in the high energy hadronic interaction models for energies above  $10^{17}$  eV prevent us from a definitive conclusion about the primary particle type for each event. Statistical analysis and evolution trends [10,11] are used to minimize the fluctuation effects, nevertheless an unique interpretation of the data is not possible because of the hadronic interaction model uncertainties. Currently, the most reliable observable to investigate composition at higher energies is  $X_{\text{max}}$ , the atmospheric depth at which the shower reaches the maximum number of particles [12]. A second very powerful observable sensitive to primary particle mass is the number of muons ( $N_{\mu}$ ) in the showers. However, the lack of knowledge of the high energy hadronic interactions and the systematic uncertainties in the energy determination limit the interpretation of  $N_{\mu}$  data in terms of composition in a more severe way than they do for  $X_{\text{max}}$ . There are several indications that the current most often used hadronic interaction models fail at predicting the muonic component features of air showers [13,14]. Moreover, as  $N_{\mu}$  scales directly with shower energy, the systematic uncertainty in energy reconstruction (typically  $\sim 10 - 20\%$ ) represents also a difficult challenge to overcome in the interpretation of the  $N_{\mu}$  data. As a consequence, it is not straightforward to envisage a data analysis procedure that extracts the mass abundance from the  $N_{\mu}$  data.

\* Corresponding author.

E-mail address: [raul.prado@usp.br](mailto:raul.prado@usp.br) (R. R. Prado).

In this paper we propose a new approach to interpret  $N_\mu$  data which accommodates the systematic uncertainties of the high energy hadronic interaction models and of the energy reconstruction. The analysis proposed here is based on the energy evolution of the first ( $\langle \log_{10} N_\mu \rangle$ ) and second ( $\sigma[\log_{10} N_\mu]$ ) moments of the  $\log_{10} N_\mu$  distribution. There are two central features of the proposed procedure: (a) a simplified model to describe the energy and mass evolution of  $\langle \log_{10} N_\mu \rangle$  and  $\sigma[\log_{10} N_\mu]$  which minimizes the hadronic interaction model dependencies, and (b) a comparison between the predictions of this model for a set of given composition scenarios and the data integrated in energy to maximize the discrimination power.

First in Section 2 we propose a simplified model to describe the energy and mass evolution of  $\langle \log_{10} N_\mu \rangle$  and  $\sigma[\log_{10} N_\mu]$ . We argue that to a very good approximation only two parameters ( $a$  and  $b$ ) summarize all uncertainties of the currently used high energy hadronic interaction models. This simplification of the description of  $\langle \log_{10} N_\mu \rangle$  and  $\sigma[\log_{10} N_\mu]$  with energy and mass is an important step in the analysis procedure because it minimizes the dependencies on hadronic interaction models in the interpretation of the data. In Section 3.1 we use shower simulations to study the energy and mass evolution of  $\langle \log_{10} N_\mu \rangle$  and  $\sigma[\log_{10} N_\mu]$  and to validate the model proposed in Section 2. We also introduce in Section 3.1 the algorithm developed to build the large set of simulations used in this paper. This simulation process is complemented in Appendix A.

In Section 4 we introduce a set of six benchmark composition scenarios defined by the percentage of proton, helium, nitrogen and iron nuclei as a function of energy. Four composition scenarios are astrophysical motivated (based in Refs. [6–9,15]) and two were derived from the  $X_{\max}$  measurements performed by the Pierre Auger Collaboration (based on Ref. [16]). By using simulations we also study the energy evolution of  $\langle \log_{10} N_\mu \rangle$  and  $\sigma[\log_{10} N_\mu]$  for each one of these scenarios and evaluate the effects of the uncertainties on the energy scale and on the absolute  $N_\mu$  due to the misprediction by the hadronic interaction models.

In Section 5 we show how the model proposed in Section 2 can be used to discriminate between these representative composition scenarios. The comparison of the model predictions for the composition scenarios with the data in an energy range is the important step of the analysis procedure proposed here because it maximizes the discrimination power allowing us to identify the most likely scenario that generated a set of  $N_\mu$  data. This comparison is done by the traditional  $\chi^2$ , which assumes the minimal value for the composition scenario which best describes the data. We use simulations to test our approach and show that it is possible to achieve a good discrimination between the chosen scenarios supposing a realistic case with the statistic to be collected during 3 years of data taking with the Pierre Auger Observatory Upgrade – AugerPrime. We also show that the systematic uncertainties in the energy reconstruction and on the absolute scale of the number of muons do not mix the composition scenarios. Hence we conclude in Section 6 that by using only the energy evolution of  $\langle \log_{10} N_\mu \rangle$  and  $\sigma[\log_{10} N_\mu]$  it would be possible to identify, by comparing the composition scenarios to the data, the scenario which best describes the measurements of  $N_\mu$ .

## 2. A model for the energy and mass evolution of $\log_{10} N_\mu$ moments

In this section we present a model to describe the energy and primary mass evolution of the  $\log_{10} N_\mu$  first and second moments. The Heitler–Matthews model [17] is a semi-empirical description of the shower development which describes the dependencies of

the mean  $N_\mu$  as

$$\langle N_\mu \rangle_A = A^{1-\beta} N_\mu^p \quad (1)$$

and

$$\langle N_\mu \rangle_E = \left( \frac{E}{\zeta_c^\pi} \right)^\beta, \quad (2)$$

where  $N_\mu^p$  is the number of muons in a proton shower and  $\zeta_c^\pi$  is the pion critical energy, assumed to be equal to 20 GeV in [17].  $\beta$  is often taken to be constant because its value is shown to vary in a small interval from 0.85 to 0.92 [17,18].

Both equations define a clear linear relation of  $\langle \log_{10} N_\mu \rangle$  with energy and mass that can be summarized as

$$\langle \log_{10} N_\mu \rangle_{E,A} = a + D_A \cdot \ln(A) + D_E \cdot (\log_{10} E - 19.0), \quad (3)$$

where  $D_E = \beta \simeq 0.85 - 0.92$ ,  $D_A = (1 - \beta) \cdot \log_{10} e \simeq 0.434 \cdot (1 - \beta) \simeq 0.0347 - 0.0651$ , and the energy  $E$  is given in eV. Because of our lack of knowledge of the hadronic interactions at the highest energies, the value of  $a$  is highly model dependent and presents a large variability. It can be written as  $a = \log_{10}(N_\mu^p) - \beta \log_{10}(\zeta_c^\pi)$  and varies approximately from 6.5 to 8.0, depending on the hadronic interaction model. These values of  $a$  were obtained using the simulations described in Section 3.

In addition to the  $\langle \log_{10} N_\mu \rangle$ , the  $\sigma[\log_{10} N_\mu]$  could also be modeled by the same approach. However, no analytic model has been proposed to describe the shower-to-shower fluctuations and our study relies on simulations to propose a similar description of  $\sigma[\log_{10} N_\mu]$  evolution with energy and mass. We propose that the  $\sigma[\log_{10} N_\mu]$  can be described as

$$\sigma[\log_{10} N_\mu]_A = \sigma[\log_{10} N_\mu]_{Fe} + b \cdot [\ln(A) - \ln(56)]^2, \quad (4)$$

where  $\sigma[\log_{10} N_\mu]_{Fe}$  is the  $\sigma[\log_{10} N_\mu]$  for iron nucleus initiated showers. Two main assumptions were used in this proposal: a) for a fixed primary (A), the  $\sigma[\log_{10} N_\mu]$  does not depend on energy and b) a quadratic dependency of  $\sigma[\log_{10} N_\mu]$  with  $\ln(A)$ . These assumptions are justified in Section 3 via Monte Carlo simulation of the air shower.

The description of the  $\sigma[\log_{10} N_\mu]$  is analogous to the deduction of  $\langle \log_{10} N_\mu \rangle$  using the Heitler–Matthews models in the following way. We will show in Section 3 that, for the purposes of this paper's analysis,  $\sigma[\log_{10} N_\mu]_{Fe}$  can be taken to be constant, in other words, the small model dependence of  $\sigma[\log_{10} N_\mu]_{Fe}$  can be ignored. On the other hand,  $b$  changes significantly with the hadronic interaction model, which reflects the theoretical uncertainties concerning the muonic component description.

Eqs. (3) and (4) summarize the first step of this paper. These equations offer a simple, but good description of the two first moments of the  $\log_{10} N_\mu$  distribution with energy and mass. The uncertainties due to hadronic interaction model descriptions are only significant for two parameters,  $a$  and  $b$ , while for the further parameters there is a good agreement between their predictions. The quality of the description given by Eqs. (3) and (4) is going to be numerically studied in the next section.

For a mixture of primaries in which each primary particle type,  $i$ , has mass  $A_i$  and contributes to the total flux with a fraction given by  $f_i$ , we can show that  $\langle \log_{10} N_\mu \rangle$  and  $\sigma[\log_{10} N_\mu]$  of the mixture ( $mix$ ) can be calculated as follows

$$\langle \log_{10} N_\mu \rangle_{mix} = \sum_i f_i \cdot \langle \log_{10} N_\mu \rangle_{A_i} \quad (5)$$

$$\langle \log_{10} N_\mu \rangle_{mix} = a + D_A \cdot \langle \ln(A) \rangle_{mix} + D_E \cdot (\log_{10} E - 19.0),$$

and

$$\sigma^2[\log_{10} N_\mu]_{\text{mix}} = \sum_i f_i \cdot \left[ (\langle \log_{10} N_\mu \rangle_{A_i} - \langle \log_{10} N_\mu \rangle_{\text{mix}})^2 + \sigma^2[\log_{10} N_\mu]_{A_i} \right]. \quad (6)$$

Using Eq. (3) we can write

$$\sigma^2[\log_{10} N_\mu]_{\text{mix}} = \sum_i f_i \cdot \left[ D_A^2 \cdot (\ln(A_i) - \langle \ln(A) \rangle_{\text{mix}})^2 + \sigma^2[\log_{10} N_\mu]_{A_i} \right]. \quad (7)$$

Note that  $\sigma^2[\log_{10} N_\mu]_{\text{mix}}$  does not depend on  $a$ . The dependence on  $b$  is implicit in the  $\sigma^2[\log_{10} N_\mu]_{A_i}$  term.

### 3. Simulation studies of $\log_{10} N_\mu$ moments

In this section we briefly describe the procedure adopted to produce simulated  $\log_{10} N_\mu$  distributions that are extensively employed in the following sections of this paper. The present discussion is complemented by Appendix A where more details about the simulations are given. Furthermore, in this section we also use the simulated showers to validate the  $\log_{10} N_\mu$  moment descriptions proposed in Section 2 and to study the energy evolution of  $\log_{10} N_\mu$  moments for a set of mass composition scenarios.

#### 3.1. Simulation technique

In our analysis we aim to assess the number of muons measured in UHECR experiments. A combination of detector technology, observatory altitude, spatial configuration of the detectors and analysis procedures determines the lateral distance range and the energy threshold of detectable muons. To avoid saturation of the detectors (close to the shower axis) and large statistical fluctuations (far from the shower axis), a fiducial lateral distance range is commonly defined to get the lateral distance function integrated. Therefore, the measured number of muons ( $N_\mu^{\text{meas}}$ ) is not the total number of muons at the ground but only a sample of them above an energy threshold and within a distance range.

In this paper  $N_\mu^{\text{meas}}$  is defined as the number of muons with energy above 0.2 GeV reaching the ground (1400 m above sea level, the Auger mean altitude) at a distance between 500 m and 2000 m from the shower axis. This choice is motivated by the design of the main current high energy cosmic ray experiments, for example, the Pierre Auger Observatory [19] and Telescope Array [20].

The muons spatial and energy distributions at the ground can be evaluated by CORSIKA [21] (version 7.4000), which is a full Monte Carlo code able to perform 3D shower simulations.  $N_\mu^{\text{meas}}$  could be determined by CORSIKA, in despite of its high computational cost [22]. CONEX [23] (version 2r4.37) is a very fast hybrid simulation code which combines full Monte Carlo with solutions of one-dimensional cascade equations. From CONEX simulations it is possible to determine the total number of muons at the ground above 1 GeV ( $N_\mu^{\text{tot}}$ ).

$N_\mu^{\text{tot}}$  and  $N_\mu^{\text{meas}}$  can be simultaneously obtained from full simulated showers (CORSIKA), allowing us to parametrize the relation between them. We propose the following parametrization:

$$N_\mu^{\text{meas}} = R(E, X_{\text{max}}) \cdot N_\mu^{\text{tot}}, \quad (8)$$

where the conversion factor  $R$  should be determined for each primary and depends on the energy and  $X_{\text{max}}$ . The parametrization of  $R(E, X_{\text{max}})$  is explored in detail in Appendix A. The  $X_{\text{max}}$  dependence of the factor  $R$  ensures that the parametrization takes into account the shower-to-shower fluctuations due to the variance of the first interaction depth. Furthermore, the most relevant physical

processes responsible for muons production in showers are reliably reproduced by the CONEX simulations, and consequently they should also be represented in  $N_\mu^{\text{meas}}$ . As shown in Appendix A, the  $N_\mu^{\text{meas}}$  distributions obtained based on the proposed parametrization are in very good agreement with the ones obtained from full Monte Carlo simulation.

The parametrization was done only for shower at  $38^\circ$  zenith angle. The zenith angle dependence can be taken into account by simulating other primaries with the corresponding arrival direction and by dividing the data in zenith angle intervals.

#### 3.2. Simulating $\log_{10} N_\mu$ moments

We generated 60,000 CONEX (version 2r4.37) showers with energies between  $10^{18.4}$  and  $10^{19.6}$  eV, for four primaries (proton, helium, nitrogen and iron) and two hadronic interaction models (EPOS-LHC [24] and QGSJetII-04 [25]). The showers are distributed uniformly in  $\log_{10}(E)$  and the zenith angle is fixed at  $38^\circ$ . From the  $R(E, X_{\text{max}})$  parametrization of Appendix A, the CONEX showers were converted into a set of  $N_\mu^{\text{meas}}$ .

Fig. 1 shows the evolution of  $\langle \log_{10} N_\mu^{\text{meas}} \rangle$  with the primary mass for three energy intervals. Lines are the result of a linear fit which demonstrates the dependence of  $\langle \log_{10} N_\mu^{\text{meas}} \rangle$  with mass as proposed in Eq. (3). The fits resulted in  $D_A \simeq 0.034 - 0.037$  and  $a \simeq 6.64 - 7.70$ , with errors from the fit less than 0.0005 and 0.005, respectively.

The energy evolution of  $\langle \log_{10} N_\mu^{\text{meas}} \rangle$  is shown in Fig. 2, where one can note the linear behavior as proposed in Eq. (3). The fits resulted in  $D_E \simeq 0.915 - 0.928$ , with errors from the fit less than 0.0003. The energy evolution of  $\sigma[\log_{10} N_\mu^{\text{meas}}]$  is shown in Fig. 3. Note the flatness of  $\sigma[\log_{10} N_\mu^{\text{meas}}]$  and the coincidence of the  $\sigma[\log_{10} N_\mu^{\text{meas}}]$  constant value of iron initiated showers for both hadronic interaction models. These figures validate both assumptions made in Section 2 concerning  $\sigma[\log_{10} N_\mu^{\text{meas}}]$ .

The primary mass dependence of  $\sigma[\log_{10} N_\mu^{\text{meas}}]$  can be seen in Fig. 4 for three energy intervals and both hadronic interaction models. The dashed lines are the quadratic curves shown in Eq. (4) fitted to the points. The fits resulted in  $\sigma[\log_{10} N_\mu]_{\text{Fe}}$  being indeed nearly constant, varying from 0.0258 to 0.0275, with errors from the fits less than 0.003. The fit also resulted in  $b = 0.0024 \pm 0.0002$  for QGSJetII-04 and  $b = 0.0033 \pm 0.0003$  for EPOS-LHC. The simulations shown in this section confirmed all the assumptions made in Section 2.

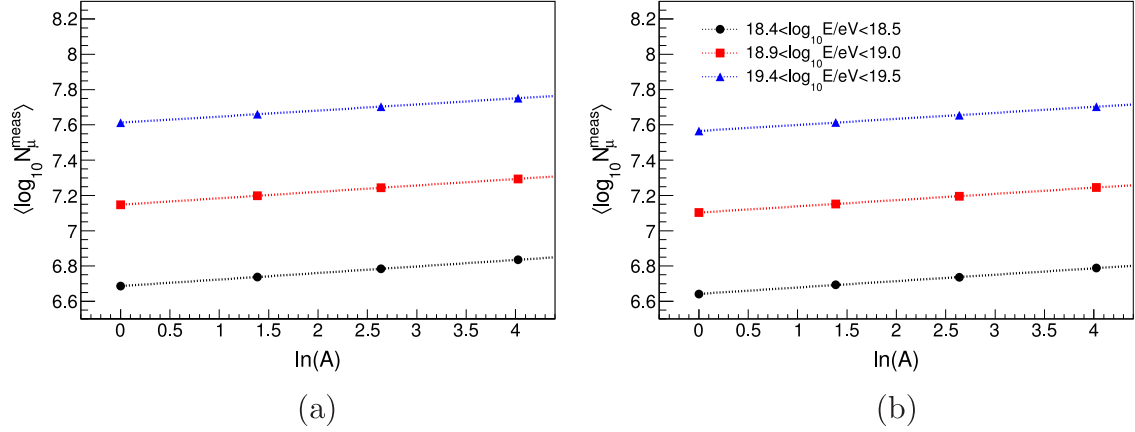
### 4. Mass composition scenarios and the energy evolution of the $\log_{10} N_\mu$ moments

In this section we simulate the energy evolution of  $\log_{10} N_\mu^{\text{meas}}$  moments for six mass composition scenarios, which are defined by setting the fractions  $f_i(E)$  of the total flux corresponding to each particle with mass  $A_i$ . Given  $f_i(E)$  and  $A_i$  we can calculate  $\langle \log_{10} N_\mu^{\text{meas}} \rangle$  and  $\sigma[\log_{10} N_\mu^{\text{meas}}]$  as a function of energy using the procedure described in Sections 2 and 3.

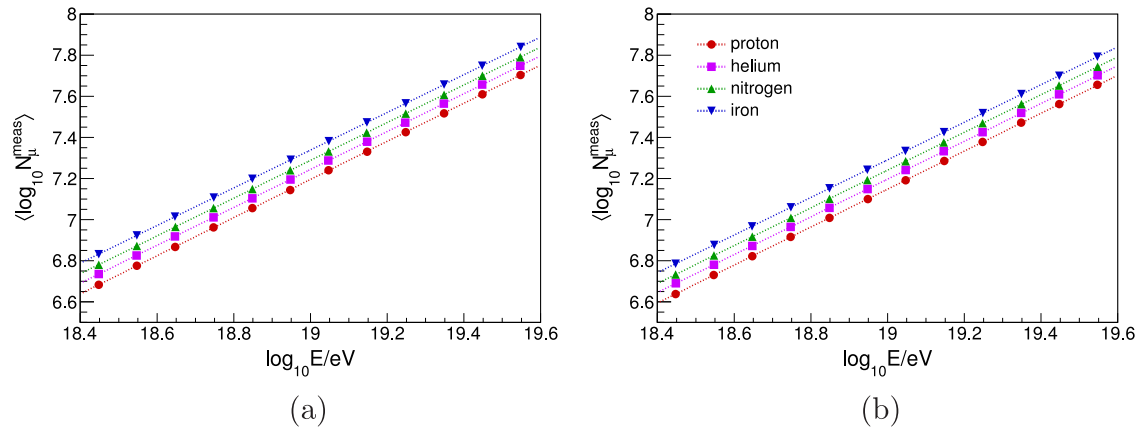
The mass composition scenarios we used are divided in two groups. The first one includes the astrophysical motivated scenarios, which are labeled by the letter A. The second group includes two scenarios obtained from the  $X_{\text{max}}$  distributions fit performed by the Pierre Auger Collaboration [11,16] and they are labeled by the letter X. Below, we present a brief description of the composition scenarios, which can be skipped by the reader that is familiar with the subject.

**Scenario A1:** This scenario proposes a pure proton flux. It was the first model proposed to explain the dip in the energy spectrum as the effect of pair-production in the propagation of the UHECR. This model is described in Refs. [5,8] and was also explored in Refs. [6,15] (labeled as Model B in Ref. [15]).

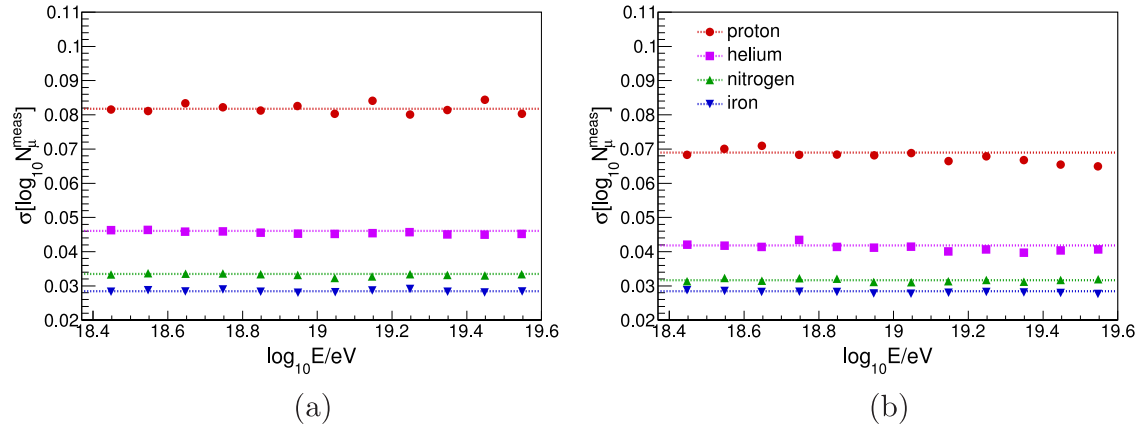




**Fig. 1.**  $\langle \log_{10} N_{\mu}^{\text{meas}} \rangle$  as a function of  $\ln(A)$  for both hadronic interaction models, (a) EPOS-LHC and (b) QGSJetII-04 and three energy intervals. The dotted lines are the results of the linear fit, represented in Eq. (3). The statistical error bars are smaller than the markers.



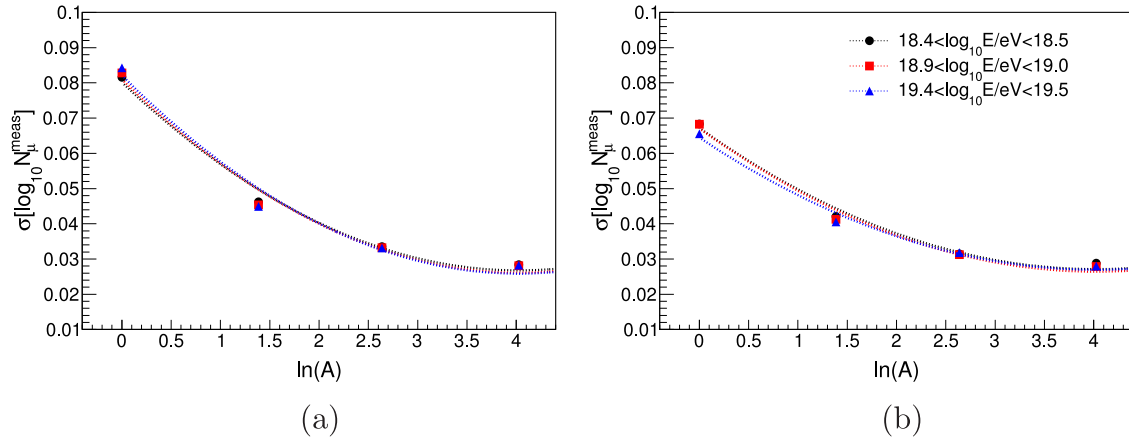
**Fig. 2.**  $\langle \log_{10} N_{\mu}^{\text{meas}} \rangle$  as a function of  $\log_{10}(E)$  for both hadronic interaction models, (a) EPOS-LHC and (b) QGSJetII-04, and four primaries (proton, helium, nitrogen and iron). The dotted lines are the results of the linear fit, represented in Eq. (3). The statistical error bars are smaller than the markers.



**Fig. 3.**  $\sigma[\log_{10} N_{\mu}^{\text{meas}}]$  as a function of  $\log_{10}(E)$  for both hadronic interaction models, (a) EPOS-LHC and (b) QGSJetII-04, and four primaries (proton, helium, nitrogen and iron). The dotted lines are the results of the fit of a constant energy function. The statistical error bars are smaller than the markers.

**Scenario A2:** This scenario assumes a mixed source composition with abundances similar to the data at lower energies. It was proposed by Allard et al. (labeled as Model A in Ref. [15]). In this model the ankle is explained as the transition in the predominance of the flux from the galactic to the extra-galactic component. The abundances are originally given for five groups of nuclei, however, in this paper the fluxes of the two heaviest groups were summed into the iron component.

**Scenario A3:** Biermann and de Souza [9] have proposed a model in which the observed cosmic ray energy spectrum from  $10^{15.0}$  to  $3 \times 10^{20.0}$  eV is explained by the galactic and only one extra-galactic source, the radio galaxy Cen A. In this model the element abundances from extra-galactic origin are similar to the galactic ones, but shifted up in energy because of the relativistic shock in the jet emanating from the active black hole. The abundances are originally given for six groups of nuclei, however, in this paper the flux of



**Fig. 4.**  $\sigma[\log_{10} N_{\mu}^{\text{meas}}]$  as a function of  $\ln(A)$  for both hadronic interaction models, (a) EPOS-LHC and (b) QGSJetII-04, and three energy intervals. The dotted lines are the results of the fit of Eq. (4). The statistical error bars are smaller than the markers.

the element group Ne–S was summed into the nitrogen flux and the flux of the Cl–Mn group was summed into the iron group flux.

**Scenario A4:** The model proposed by Globus et al. [7] describes the whole cosmic ray spectrum by superposing a rigidity dependent galactic component and a generic extra-galactic component. This model gives an adequate description of the energy spectrum and the moments of the  $X_{\text{max}}$  distribution measured by the Pierre Auger Observatory.

**Scenario X1:** It has been shown by the Pierre Auger Collaboration that the measured  $X_{\text{max}}$  distributions can be well described by a combination of four components [11,16]. By fitting the  $X_{\text{max}}$  simulated distributions to the data, the abundances of the separate components were obtained as a function of energy. This scenario is based on the abundances obtained by using the hadronic interaction model QGSJetII-04. However, the abundances obtained with Sibyll2.1 are also very close to the one we used. In order to minimize point-to-point fluctuations, we used here a smooth curve fitted to the fractions obtained in the Auger analysis [16].

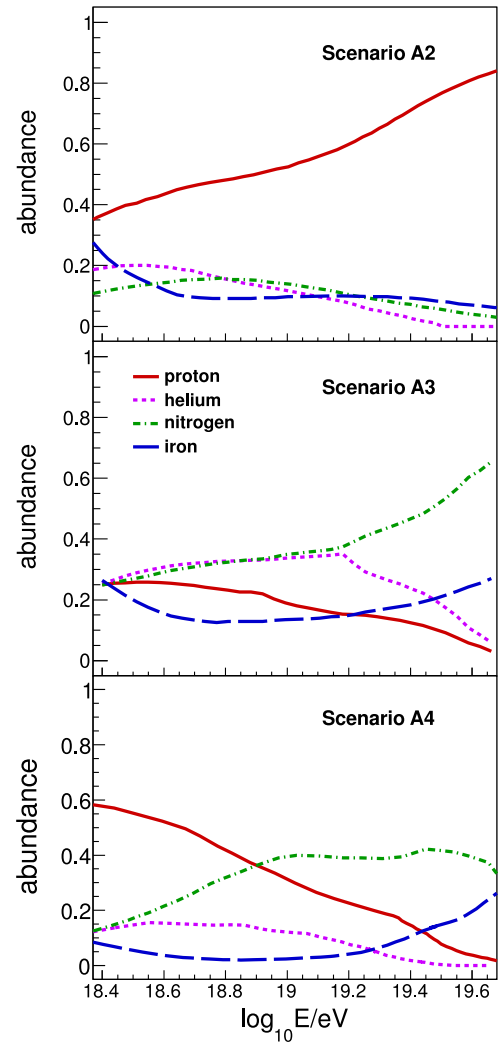
**Scenario X2:** This scenario was obtained by fitting  $X_{\text{max}}$  distributions measured by the Pierre Auger Observatory using showers simulated with the EPOS-LHC hadronic interaction model. The procedure is the same as the one adopted for Scenario X1.

The merging of components done for models A2 and A3 is necessary to allow us to use the parametrization elaborated in Section 3. Since we present in this paper only the analysis procedure, verified with simulations, this choice has no limiting consequence. Besides that, the systematic uncertainties of the abundances obtained from the scenarios are also going to be neglected here. Figs. 5 and 6 show the abundances for each scenario in the energy range from  $10^{18.4}$  to  $10^{19.6}$  eV as explained above. Scenario A1 is not shown because it assumes a 100% proton flux.

Fig. 7 shows the energy evolution of the  $\langle \log_{10} N_{\mu}^{\text{meas}} \rangle$  and  $\sigma[\log_{10} N_{\mu}^{\text{meas}}]$  for all mass composition scenarios. The error bars correspond to the one sigma fluctuation of the mean value considering the statistics from 3 years of AugerPrime data (3000 km<sup>2</sup> of muon detectors). The all particle flux was taken from Ref. [3]. Fig. 7a shows the mean normalized to the proton simulation for better visualization.

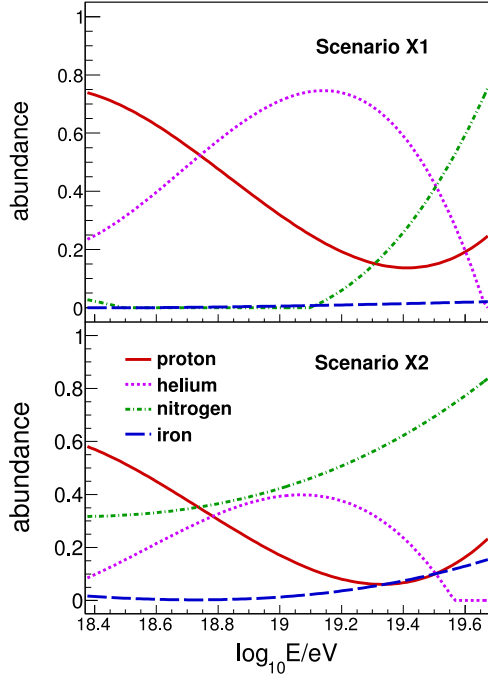
## 5. Discrimination between mass composition scenarios

Given the theoretical uncertainties on the  $N_{\mu}$  predictions and the systematic uncertainties on the energy reconstruction, the



**Fig. 5.** The composition component abundances as a function of  $\log_{10}(E)$  for the mass composition scenarios A2, A3 and A4 (see text).

question we would like to answer in this section is how it is possible to discriminate between the mass composition scenarios shown above using the evolution of the  $\log_{10} N_{\mu}^{\text{meas}}$  moments with energy. Examining Fig. 7 it might seem easy to differentiate



**Fig. 6.** The composition component abundances as a function of  $\log_{10}(E)$  for the mass composition scenarios X1 and X2 (see text).

the scenarios by using the absolute value or the evolution of the  $\log_{10} N_{\mu}^{\text{meas}}$  moments with energy. However, if we include in this figure the uncertainties in the hadronic interaction model and systematic in energy reconstruction the interpretation of the data is not straightforward.

Fig. 8 shows how the uncertainties on the hadronic interaction model predictions and on the energy reconstruction influence the interpretation of the  $\langle \log_{10} N_{\mu}^{\text{meas}} \rangle$  in terms of composition. We show in this figure the extreme composition scenarios (A1 and A3), since the other four scenarios lie within them. In Fig. 8a we calculate  $\langle \log_{10} N_{\mu}^{\text{meas}} \rangle$  for scenarios A1 and A3 adding arbitrarily 20% more muons to the simulation predictions to mimic the theoretical uncertainties in the hadronic interaction model predictions [14]. Even the extreme models A1 and A3 would overlap if the uncertainty is considered. In Fig. 8b we calculate  $\langle \log_{10} N_{\mu}^{\text{meas}} \rangle$  for scenarios A1 and A3 and changed the simulated energy by  $\pm 15\%$  in order to evaluate the effect of the systematic uncertainty in the energy reconstruction. Once more it is clear that even the extreme

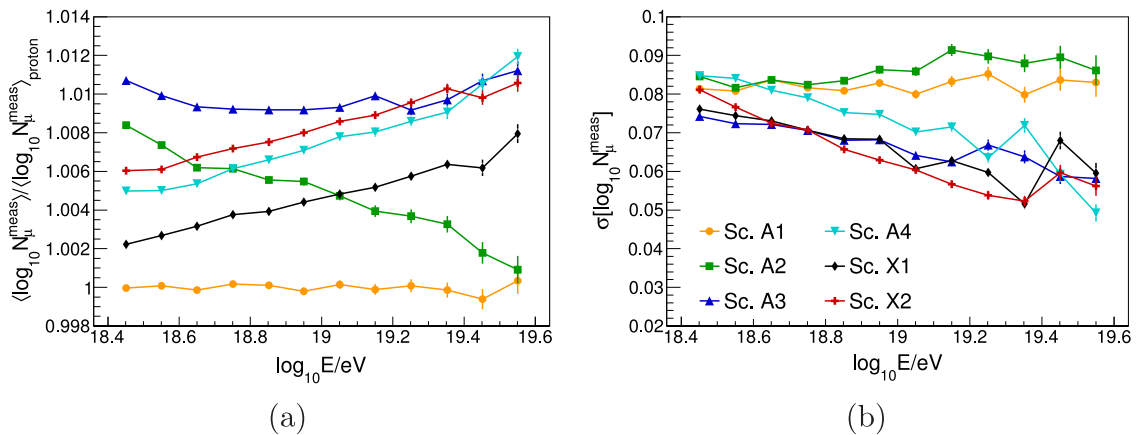
scenario cases cannot be distinguished anymore. Moreover a combination of both uncertainties in the  $N_{\mu}$  predictions and energy applied to this analysis would make the discrimination between the scenarios even harder. The conclusion is clear: the measurement of  $N_{\mu}$  does not lead to a straightforward interpretation of the data in terms of composition if all the uncertainties are considered.

It is worthwhile to remember here how the interpretation of the  $X_{\text{max}}$  measurement is done. The Pierre Auger Collaboration, for example, fits  $f_i$  to the measured  $X_{\text{max}}$  distribution in bins of energy [16]. The calculation of  $f_i$  depends on simulation and therefore on the hadronic interaction model. However, because the electromagnetic cascade of the shower dominates the determination of the  $X_{\text{max}}$  position, the discrepancy between the hadronic interaction model  $X_{\text{max}}$  predictions is minimized. The difference in  $\langle X_{\text{max}} \rangle$  is at most 20 g/cm<sup>2</sup> and in  $\sigma[X_{\text{max}}]$  is 6 g/cm<sup>2</sup> for the most often used hadronic interaction models (EPOS-LHC, Sibyll2.1 and QGSJetII-04) [11]. Given the small differences in the predictions of  $X_{\text{max}}$  and its consistency with data, the fit of  $f_i$  leads to acceptable differences in the calculation of  $f_i$  for different hadronic interaction models and then to mass composition scenarios which are physically consistent.

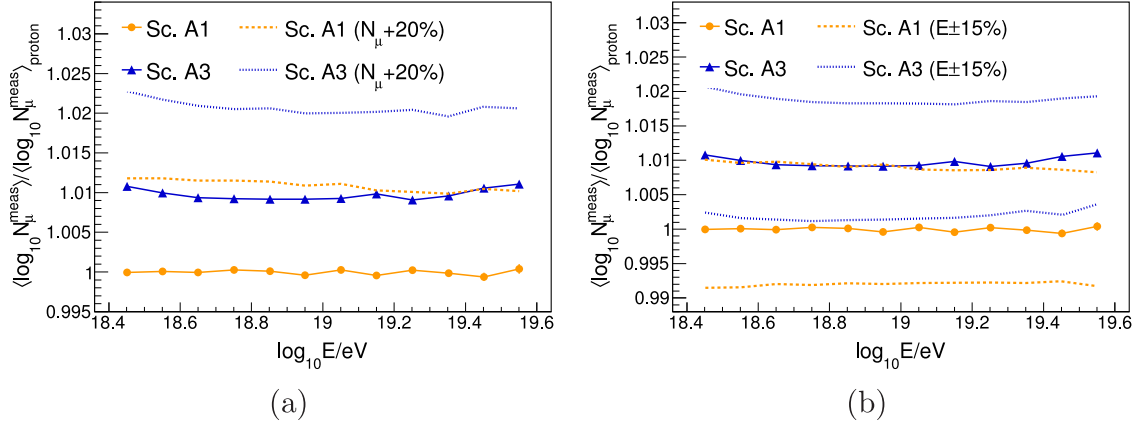
Unfortunately, the same procedure cannot be applied to  $N_{\mu}$  because of the discrepancies between the hadronic interaction model predictions and the inconsistency between simulations and data. It is known that the simulations are off by at least 20% in the calculation of  $N_{\mu}$  [13,14]. A fit of  $f_i$  based on the  $N_{\mu}$  distribution would lead to non-physical results. Therefore we propose an alternative analysis to discriminate between composition scenarios. The idea is to fix  $f_i$ , choosing a mass composition scenario, and fit the data with the energy evolution of  $\log_{10} N_{\mu}^{\text{meas}}$  moments to search for the scenarios which better describe the data.

If the composition ( $f_i$ ) were known by an independent measurement, this procedure would allow us to calculate  $a$  and  $b$  and constrain the hadronic interaction models by limiting fundamental properties of the interactions. This hypothesis needs to be explored further by using the results from the  $X_{\text{max}}$  measurement to fix  $f_i$ .

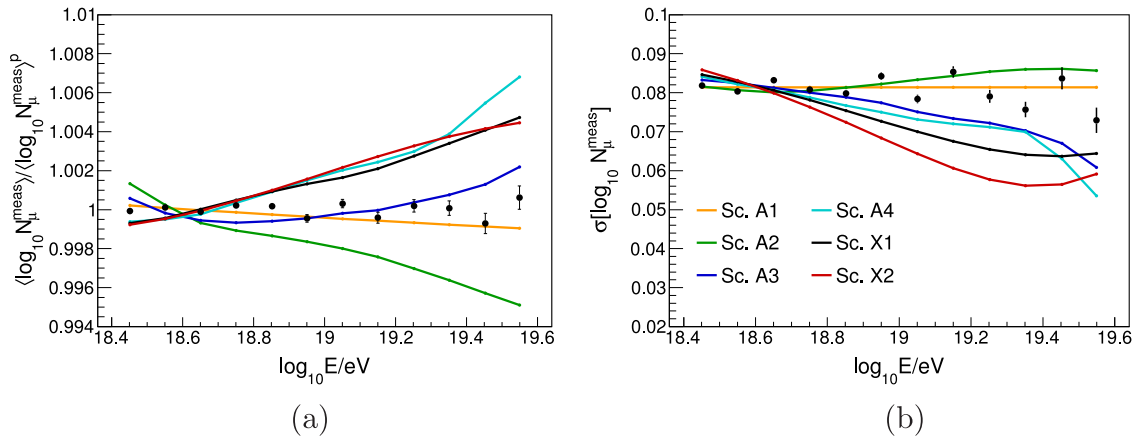
We propose here a procedure that allows a statistically robust test of composition scenarios against data. The method starts by using the model proposed in Section 2 to predict the energy evolution of  $\langle \log_{10} N_{\mu}^{\text{meas}} \rangle$  and  $\sigma[\log_{10} N_{\mu}^{\text{meas}}]$  for a given composition scenario. Here, all the parameters of the model are fixed, except  $a$  and  $b$ . The next step is to compare these predictions with data and find the values of  $a$  and  $b$  which make the model most similar to the data. This can be done by a  $\chi^2$  minimization. The minimal values of  $\chi^2$  determine which scenario best describes the data. Since  $a$  and  $b$  take all the hadronic interaction model dependence, the



**Fig. 7.** Energy evolution of (a)  $\langle \log_{10} N_{\mu}^{\text{meas}} \rangle$  and (b)  $\sigma[\log_{10} N_{\mu}^{\text{meas}}]$  for the six composition scenarios described in the text. The values of  $\langle \log_{10} N_{\mu}^{\text{meas}} \rangle$  are divided by the corresponding value of pure proton composition for better visualization. The hadronic interaction model used was EPOS-LHC.



**Fig. 8.** Energy evolution of  $\langle \log_{10} N_{\mu}^{\text{meas}} \rangle$  for 2 mass composition scenarios, A1 and A3. The dashed lines show the effects of (a) an increase of 20% in the  $N_{\mu}$  and (b) a variation of  $\pm 15\%$  in energy. The values of  $\langle \log_{10} N_{\mu}^{\text{meas}} \rangle$  are divided by the corresponding value of pure proton composition for better visualization. The hadronic interaction model used was EPOS-LHC.



**Fig. 9.** Black dots show the simulated (a)  $\langle \log_{10} N_{\mu}^{\text{meas}} \rangle$  and (b)  $\sigma[\log_{10} N_{\mu}^{\text{meas}}]$  using scenario A1 as the true one. The colored lines show the results of the fit for each one of the test scenarios. The values of  $\langle \log_{10} N_{\mu}^{\text{meas}} \rangle$  are divided by the corresponding value of pure proton composition for better visualization.

composition scenario can be tested independently of hadronic interaction model limitations.

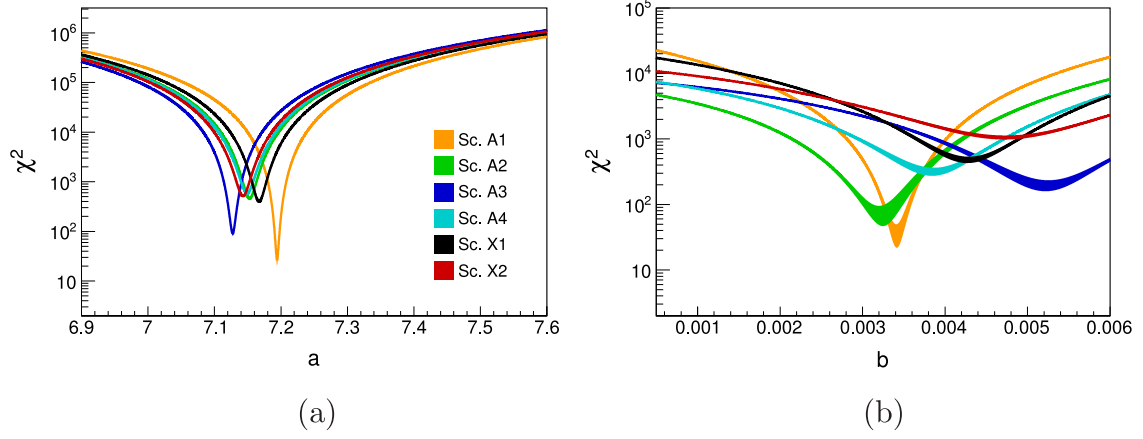
We explored this analysis proposal by choosing a composition scenario as if it would represent the true measurement, and we name it *true scenario*. We generate the  $\log_{10} N_{\mu}$  moments as a function of energy for the *true scenario* using the simulation described in Section 3 in order to emulate the real data. Here, the energy bins are defined by 12 intervals of width  $\Delta \log_{10}(E/\text{eV}) = 0.1$ , from  $10^{18.4}$  to  $10^{19.6}$  eV. This choice is mainly motivated by the Auger experimental acceptance, which reaches a 100% efficient trigger probability around  $10^{18.4}$  eV [26]. The number of events in each bin is determined by considering 3 years of data taken by the full array of Auger, following the energy spectrum of Ref.[3].

It is important to note that the simulations used here do not take into account any detector effects or zenith angle dependence. Although in this paper we do not intend to approach these issues because the focus here are on the general aspects of the analysis, it is clear that in practical applications of the method one should deal with these experimental difficulties. The detector effects, like resolution and limited acceptance, could be addressed by unfolding or unbiasing techniques once the detector response is well known. One example of these process is the Auger analysis of  $X_{\text{max}}$  moments [11]. The zenith angle dependence could be addressed in a conservative approach by dividing the data in zenith angle intervals or by correcting the data using a *constant intensity cut* (CIC) method [27–29]. This later class of method has been successfully used, for example, to determine the shower size

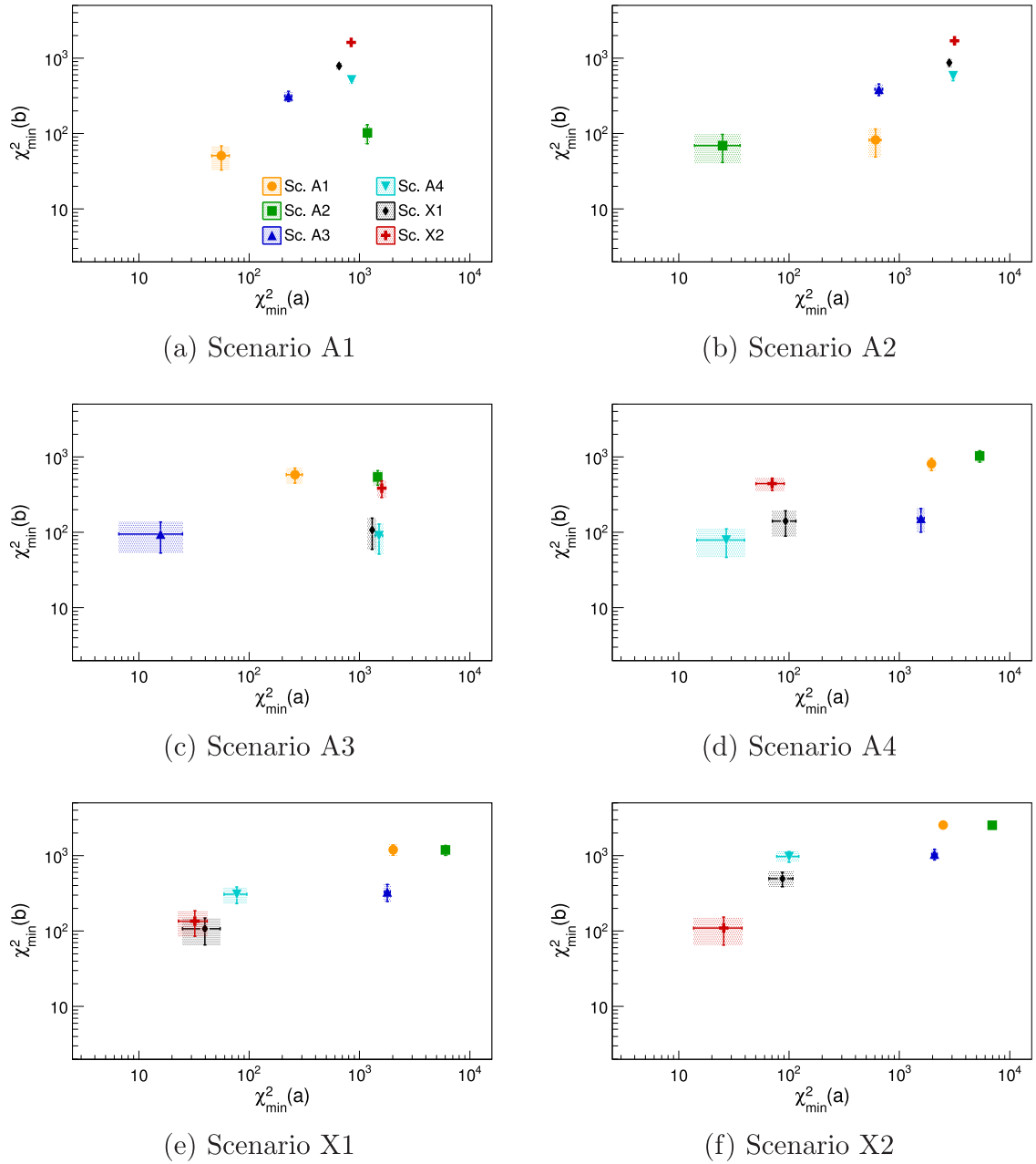
parameter by Pierre Auger [30] and KASCADE-Grande Collaboration [31] and to correct the  $N_{\mu}$  parameter by KASCADE-Grande Collaboration [32]. The systematics uncertainties from these procedure are usually small ( $<10\%$ ) and should be taken into account in a realistic approach of our method.

In next step, we perform a  $\chi^2$  fit using the model described by Eqs. (5) and (7), with  $a$  and  $b$  as free parameters of the fit, for all the composition scenarios. The scenarios which are not the true one are named *test scenarios*. Fig. 9 shows one realization of these fits in which scenario A1 was used as the *true scenario* to generate the black dots. The fit of the  $\langle \log_{10} N_{\mu}^{\text{meas}} \rangle$  with energy (Fig. 9) sets the best value of  $a$  and the minimal value of  $\chi^2(a)$ . The fit of the  $\sigma[\log_{10} N_{\mu}^{\text{meas}}]$  with energy (Fig. 9) sets the best value of  $b$  and the minimal value of  $\chi^2(b)$ . In all fits,  $D_E = 0.920$ ,  $D_A = 0.0354$  and  $\sigma[\log_{10} N_{\mu}]_{\text{Fe}} = 0.0265$ . Each line in Fig. 9 is the fit of one out of the six composition scenarios. We compared all scenarios (*test scenarios*) to the *true scenario*.

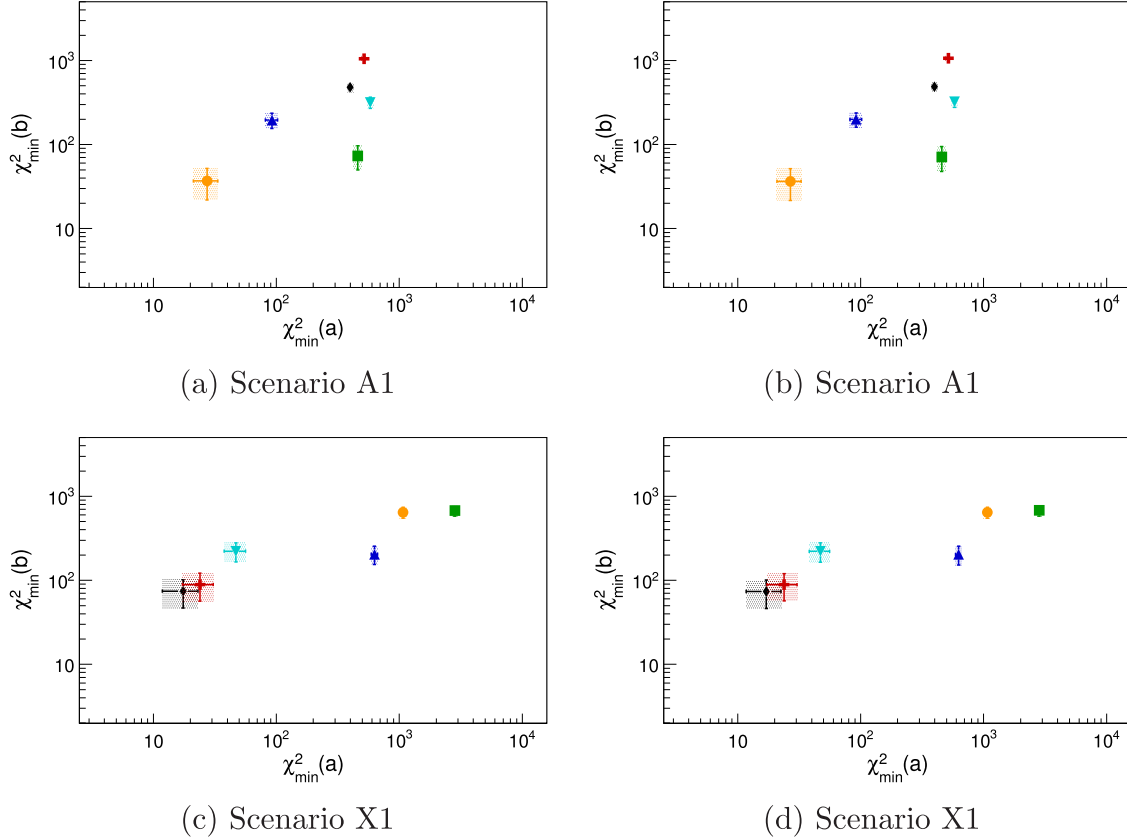
A simple  $\chi^2$  comparison finds the *test scenarios* which best fit the data generated with the *true scenario*. The average value of  $\chi^2$  as a function of the fitted parameters  $a$  and  $b$  is shown in Fig. 10 for a set of 500 realizations. In this case, it is clear that the scenario A1 better describes the  $\langle \log_{10} N_{\mu}^{\text{meas}} \rangle$  and  $\sigma[\log_{10} N_{\mu}^{\text{meas}}]$  evolution with energy because of the smaller values of  $\chi_{\text{min}}^2(a)$  and  $\chi_{\text{min}}^2(b)$ . Fig. 11 shows the plots of  $\chi_{\text{min}}^2(a)$  vs  $\chi_{\text{min}}^2(b)$  for all six scenarios as the *true scenarios*. The error bars represent one standard deviation around the mean for 500 realizations. All scenarios,



**Fig. 10.**  $\chi^2$  as a function of the parameters  $a$  (a) and  $b$  (b) for all the scenarios. The scenario A1 is the *true scenario*. The colored bands represent one standard deviation around the mean for a set of 500 realizations.



**Fig. 11.**  $\chi_{\min}^2(a)$  vs  $\chi_{\min}^2(b)$  for all composition scenarios.



**Fig. 12.**  $\chi^2_{\min}(a)$  vs  $\chi^2_{\min}(b)$  for the true scenario A1 with (a)  $\alpha_{N_\mu} = 1.3$  and (b)  $\alpha_{N_\mu} = 1.6$  and for the true scenario X1 with (c)  $\alpha_{N_\mu} = 1.3$  and (d)  $\alpha_{N_\mu} = 1.6$  (see text).

except in X1 case, can be discriminated by the smallest  $\chi^2_{\min}(a)$  and  $\chi^2_{\min}(b)$ . In other words, the *true scenario* is the one with the  $\chi^2_{\min}(a) - \chi^2_{\min}(b)$  point closer to the left-down corner. Note that only  $\chi^2_{\min}(a)$  or only  $\chi^2_{\min}(b)$  cannot alone discriminate most of the scenarios. In the case of X1 as *true scenario* one can see that, even if it is not possible to discriminate scenario X1 and X2, it is still possible to discriminate the  $X_{\max}$  scenarios from the astrophysical ones.

### 5.1. Sensitivity to the systematic uncertainties on energy scale and absolute number of muons

As mentioned above, the greatest obstacles in interpreting  $N_\mu$  data currently are the systematic uncertainties in the theoretical description and reconstruction of air showers. In this section we demonstrate that the procedure proposed in the previous section to discriminate between composition scenarios is stable under systematic changes of absolute  $N_\mu$  prediction and of energy scale.

The systematic uncertainties in  $N_\mu$  scale were tested by applying the rescaling factor  $\alpha_{N_\mu}$  in  $N_\mu^{\text{meas}}$  generated by simulations. The examination of Eq. (3) shows that a rescaling factor on  $N_\mu$  corresponds to an additive term in  $a$ . This is how the systematic effect on  $N_\mu^{\text{meas}}$  is incorporated in the simple description of  $\langle \log_{10} N_\mu^{\text{meas}} \rangle$ . Eq. (4) shows that  $b$  does not depend on  $\alpha_{N_\mu}$ . If only  $\langle \log_{10} N_\mu^{\text{meas}} \rangle$  changes by an additive term when  $\alpha_{N_\mu}$  is applied, it is clear that  $\chi^2_{\min}(a)$  and  $\chi^2_{\min}(b)$  are independent of  $\alpha_{N_\mu}$ . Fig. 12 shows the values of  $\chi^2_{\min}(a)$  and  $\chi^2_{\min}(b)$  for true scenarios A1 and X1 and  $\alpha_{N_\mu} = 1.3$  and 1.6, as examples. One can see that the values of  $\chi^2_{\min}(a)$  and  $\chi^2_{\min}(b)$  are stable under systematic changes in  $N_\mu^{\text{meas}}$ .

The energy scale effect was tested by including a rescaling factor  $\alpha_E$  in simulated energy of each shower. The same analysis of Eqs. (3) and (4) reveals that  $a$  and  $b$  accommodate the sys-

tematic effects in energy as additive terms. All additive terms are canceled in the  $\chi^2$  comparison resulting in the independence of the conclusions under systematic effects. In Fig. 13 we show the  $\chi^2_{\min}(a)$  and  $\chi^2_{\min}(b)$  for true scenarios A1 and X1 and for  $\alpha_E = 0.85$  and 1.15, which represents a systematic uncertainties of 15% in energy. Again, it can be observed that the values of  $\chi^2_{\min}(a)$  and  $\chi^2_{\min}(b)$  would not lead to a different conclusion, and therefore, the results of the method would be stable under energy shifts.

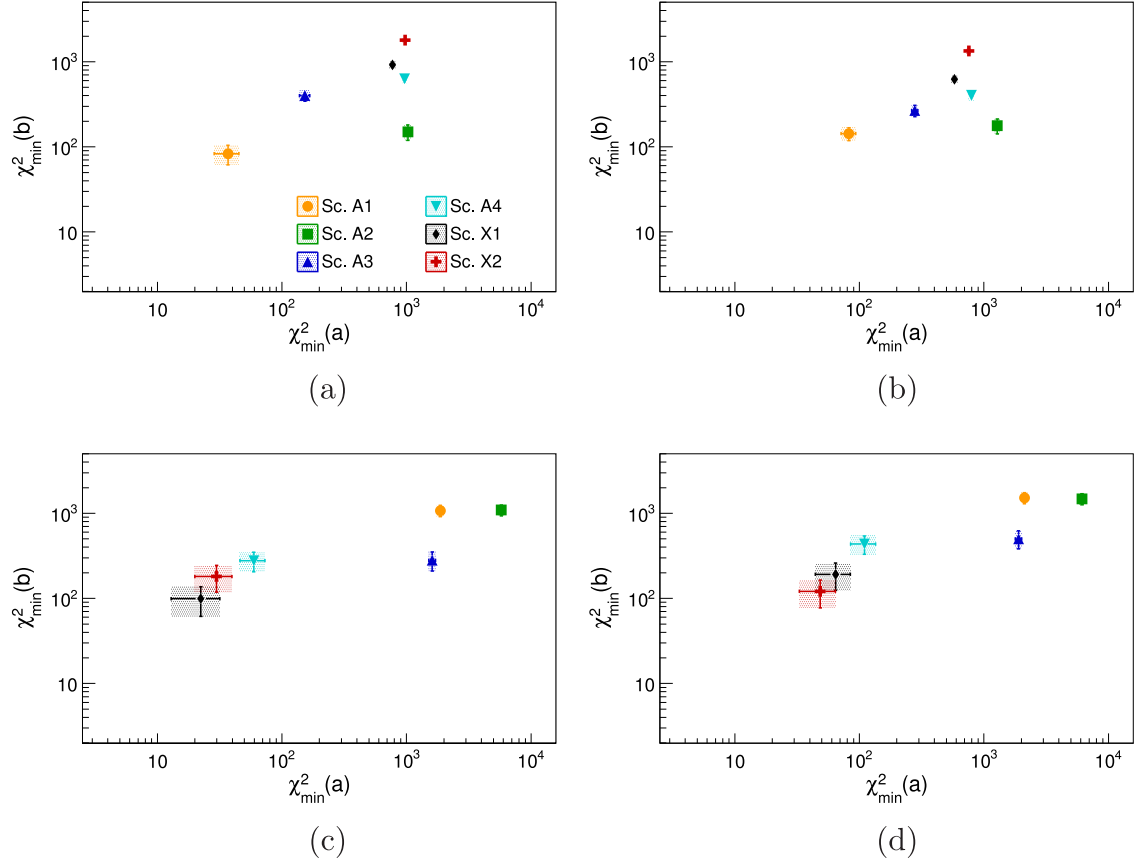
## 6. Conclusion

In this paper, we have analyzed the muon content of air showers, proposed a parametrization of the first two moments of the number of muons with energy and primary particle mass and showed how the measured mean and  $\sigma$  of  $\log_{10} N_\mu$  can be used to discriminate between composition scenarios.

We proposed a model to describe  $\langle \log_{10} N_\mu^{\text{meas}} \rangle$  and  $\sigma[\log_{10} N_\mu^{\text{meas}}]$  as a function of energy and primary particle mass ( $A$ ). This model was conceived to keep the most relevant hadronic interaction uncertainties concentrated in only two parameters ( $a$  and  $b$ ). We have validated the model with Monte Carlo simulation of the air shower and its capability to describe the  $\log_{10} N_\mu^{\text{meas}}$  moments was proven.

Six composition scenarios were considered. The particle flux predicted by these scenarios was transformed into the corresponding  $\langle \log_{10} N_\mu^{\text{meas}} \rangle$  and  $\sigma[\log_{10} N_\mu^{\text{meas}}]$  evolution with energy. The  $\langle \log_{10} N_\mu^{\text{meas}} \rangle$  and  $\sigma[\log_{10} N_\mu^{\text{meas}}]$  evolution with energy was fitted using the proposed parametrization. A comparison of the  $\langle \log_{10} N_\mu^{\text{meas}} \rangle$  and  $\sigma[\log_{10} N_\mu^{\text{meas}}]$  model using a simple  $\chi^2$  test allows the discrimination between the scenarios. The discrimination is effective even considering the systematic uncertainties on the  $N_\mu$  prediction and on energy scale uncertainty.





**Fig. 13.**  $\chi^2_{\min}(a)$  vs  $\chi^2_{\min}(b)$  for the true scenario A1 with (a)  $\alpha_E = 0.85$  and (b)  $\alpha_E = 1.15$  and for the true scenario X1 with (c)  $\alpha_E = 0.85$  and (d)  $\alpha_E = 1.15$  (see text).

The effect of the systematic in the  $N_\mu$  number and energy reconstruction was studied for constant values of the uncertainty with energy. This choice is justified by the narrow energy interval used in the analysis. Abrupt changes of the systematic uncertainties with energy could change the conclusion drawn here since  $\log_{10} N_\mu^{\text{meas}}$  is assumed to be fixed in the proposed model.

The upgrade of Telescope Array [33] and Pierre Auger Observatory [34], to be constructed in the next few years, will for the first time allow precise measurements of the muon component of air showers for energies above  $10^{18}$  eV. This will open up a new window of analyses and tests in astroparticle physics. Once data is acquired, the parametrization proposed here could be tested and if proven to be right, the analysis method proposed in Section 5 could be used to find the most probable composition scenario in the energy range from  $10^{18.4}$  to  $10^{19.6}$  eV.

### Acknowledgments

We thank Stéphane Coutu and Carola Dobrigkeit for the review of the manuscript on behalf of the Pierre Auger Collaboration. RRP thanks the financial support given by FAPESP (2014/10460-1). RC and MP gratefully acknowledge the financial support by Fundação para a Ciência e Tecnologia (SFRH/BPD/73270/2010) and OE, FCT-Portugal, CERN/FIS-NUC/0038/2015. VdS thanks the support of the Brazilian population via CNPq and FAPESP (2010/07359-6, 2014/19946-4).

### Appendix A. Parametrization of $R(E, X_{\max})$

As mentioned in Section 3, the parametrization of  $R(E, X_{\max})$  can be done by means of full air shower simulations. In this paper, we choose CORSIKA [21] as the full Monte Carlo code.

First, a set of 200 CORSIKA (version 7.4000) showers for each primary (proton, helium, nitrogen and iron), fixed energy ( $E = 10^{18.5}, 10^{19.0}, 10^{19.5}$  eV) and high energy hadronic interaction model (EPOS-LHC [24] and QGSjetII-04 [25]) were generated. The low energy hadronic interaction model is Fluka [35] in all cases. Furthermore, the zenith angle was set fixed to  $38^\circ$  for all showers. This is clearly a simplification, and then it should be stressed that, in a more realistic analysis, the  $N_\mu$  zenith angle dependence must be treated.

In Fig. A.1 the factor  $R$  is shown as a function of  $X_{\max}$  for all primaries, one primary energy,  $10^{19.0}$  eV, and one hadronic interaction model, EPOS-LHC. The observed behavior suggests a linear parametrization of the form,

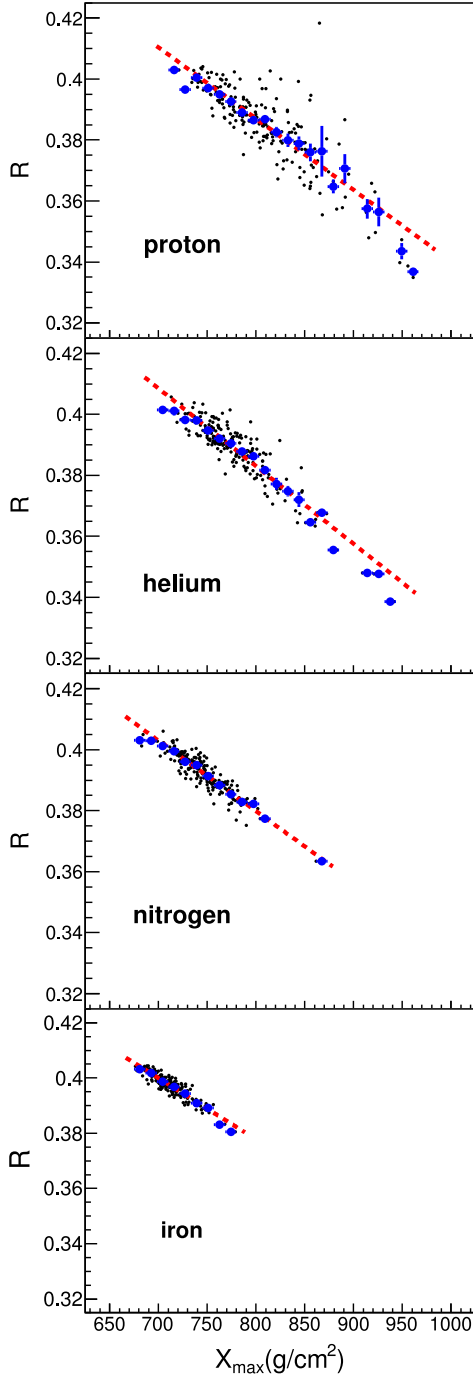
$$R(E, X_{\max}) = p_1(E) \cdot X_{\max} + p_0(E), \quad (\text{A.1})$$

where  $p_0$  and  $p_1$  have to be also parametrized as a function of energy. The values of  $p_0$  and  $p_1$  for each primary and energy were determined from the linear fit, shown in Fig. A.1 by a dotted red line.

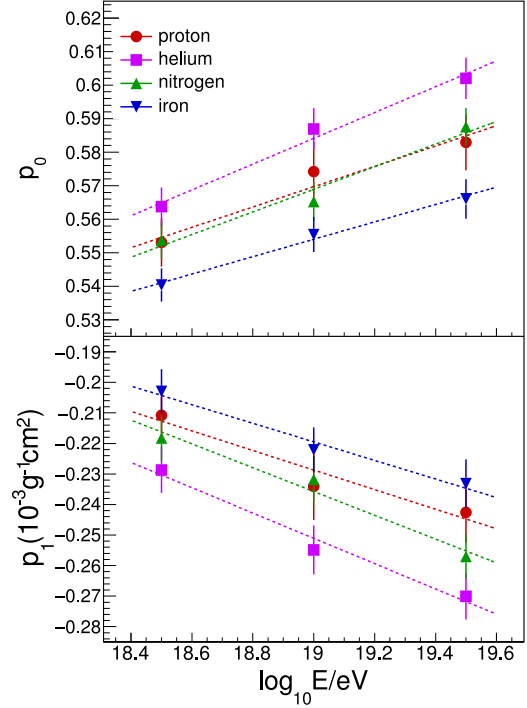
In Fig. A.2  $p_0$  and  $p_1$  are shown as a function of logarithmic energy for all primaries and the hadronic interaction model EPOS-LHC, as an example. Again, we are able to perform a linear parametrization of  $p_0$  and  $p_1$  as a function of  $\log_{10}(E)$ , in the form

$$\begin{aligned} p_0(E) &= \alpha_0 \cdot \log_{10}(E/\text{eV}) + \beta_0, \\ p_1(E) &= \alpha_1 \cdot \log_{10}(E/\text{eV}) + \beta_1. \end{aligned} \quad (\text{A.2})$$

The dotted lines in Fig. A.2 show the function from Eq. (A.2) fitted to the point obtained from CORSIKA simulations. The values of  $\alpha_0$ ,  $\beta_0$ ,  $\alpha_1$  and  $\beta_1$  for all primaries and hadronic interaction models are presented in Table A.1.



**Fig. A.1.** Conversion factor  $R$  as a function of  $X_{\max}$  for  $E = 10^{19.0}$  eV showers with EPOS-LHC as hadronic interaction model. The black dots are individual showers, the blue circles are the profile and dotted red line is the linearfit. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. A.2.**  $p_0$  and  $p_1$  as a function of  $\log_{10}(E)$  (see text) for EPOS-LHC as the hadronic interaction model. The dashed lines are the linear fits represented in Eq. (A.2).

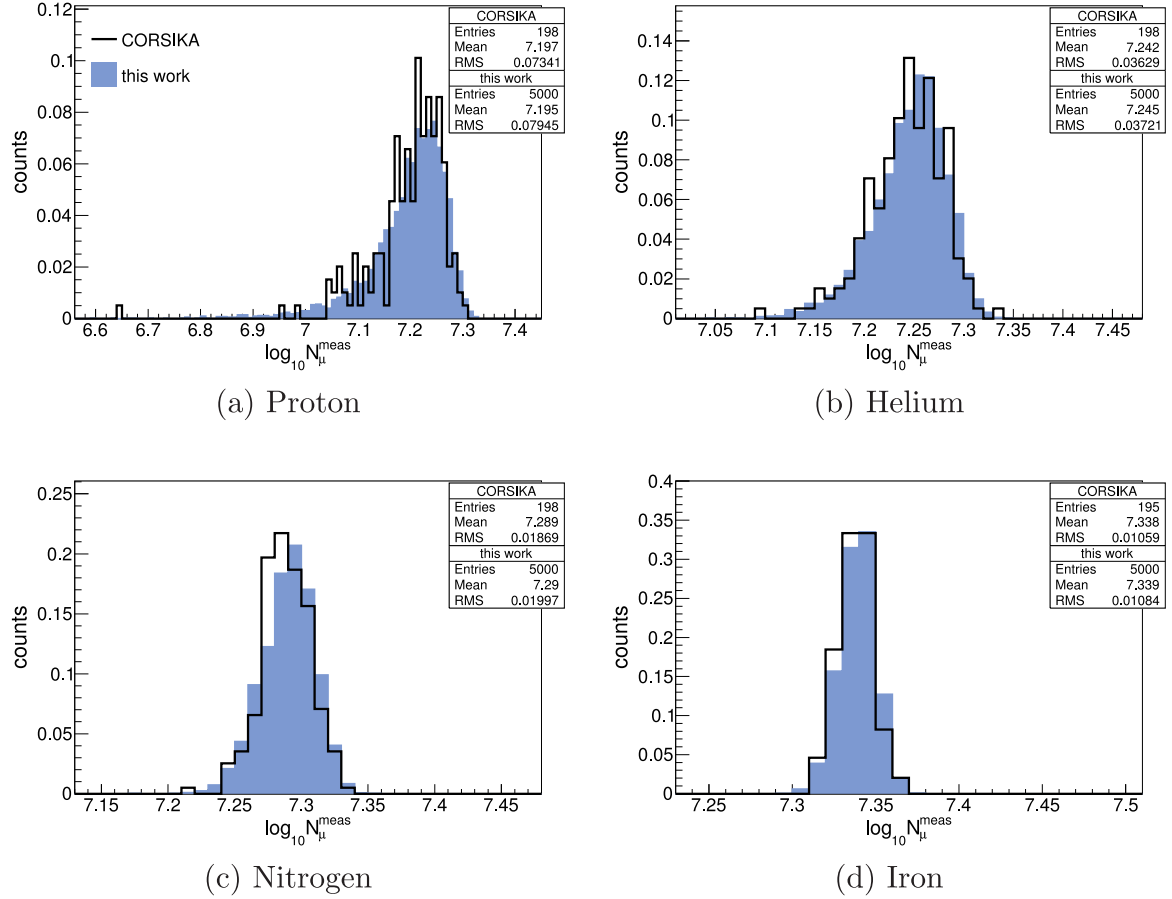
A comparison between  $\log_{10} N_{\mu}^{\text{meas}}$  distributions achieved directly from CORSIKA showers and from the simulated method described in this work can be seen in Fig. A.3. The energy is fixed at  $E = 10^{19.0}$  eV and the hadronic interaction model is EPOS-LHC. The discrepancies in the mean values and in the  $\sigma$  of the distributions are less than 2% and 5% respectively, for any combination of primary and hadronic interaction model. Indeed, considering the differences between the approaches assumed by both software, CORSIKA and CONEX, these observed discrepancies are really satisfactory. Among the several physical effects which are treated differently we can highlight the lack of geomagnetic field and muon multiple scattering in CONEX.

Although there is a small discrepancy between our method's and CORSIKA's  $\log_{10} N_{\mu}^{\text{meas}}$  distributions, we do not expect to find any loss in the development of this paper. This can be assured because the method proposed here is not dependent on the comparison between our simulations and any other set of simulated showers.

**Table A1**  
Fitted parameters  $\alpha_0$ ,  $\beta_0$ ,  $\alpha_1$  and  $\beta_1$  given in Eq. (A.2).

Had. int. model	Primary	$\alpha_0$	$\beta_0$	$\alpha_1(10^{-5} \text{g}^{-1} \text{cm}^2)$	$\beta_1(10^{-5} \text{g}^{-1} \text{cm}^2)$
EPOS-LHC	proton	0.0303	-0.00635	-3.20	38.0
	helium	0.0385	-0.147	-4.14	53.6
	nitrogen	0.0337	-0.0721	-3.90	50.5
	iron	0.0259	0.0613	-3.04	35.8
QGSJetII-04	proton	0.0216	0.130	-2.04	17.3
	helium	0.0143	0.264	-1.20	1.35
	nitrogen	0.0167	0.209	-1.60	9.94
	iron	0.0245	0.0427	-2.99	38.5





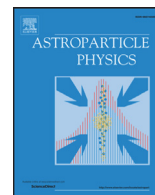
**Fig. A.3.** Comparison between the normalized  $\log_{10} N_{\mu}^{\text{meas}}$  distributions generated by CORSIKA and by this work's algorithm. The energy is  $E = 10^{19.0}$  eV and the hadronic interaction model is EPOS-LHC. The primary particles are (a) proton, (b) helium, (c) nitrogen and (d) iron.

## References

- [1] R. Abbasi (The HiRes Collaboration), Measurement of the flux of ultra high energy cosmic rays by the stereo technique, *Astropart. Phys.* 32 (1) (2010) 53–60. [0904.4500](#), doi:[10.1016/j.astropartphys.2009.06.001](#).
- [2] J. Abraham (The Pierre Auger Collaboration), Measurement of the energy spectrum of cosmic rays above  $10^{18}$  eV using the Pierre Auger Observatory, *Phys. Lett. B* 685 (4–5) (2010) 239–246. [1002.1975](#), doi:[10.1016/j.physletb.2010.02.013](#).
- [3] A. Schulz, Measurement of the Energy Spectrum of Cosmic Rays above  $3 \times 10^{17}$  eV with the Pierre Auger Observatory, in: *Proceedings of the 33rd ICRC*, 2013.
- [4] T. Abu-Zayyad (The Telescope Array Collaboration), The cosmic-ray energy spectrum observed with the surface detector of the telescope array experiment, *Astrophys. J.* 768 (1) (2013) L1, doi:[10.1088/2041-8205/768/1/L1](#).
- [5] D. Allard, et al., UHE nuclei propagation and the interpretation of the ankle in the cosmic-ray spectrum, *Astron. Astrophys.* 443 (2005) L29–L32. [astro-ph/0505566](#), doi:[10.1051/0004-6361:200500199](#).
- [6] V. Berezhinsky, A. Gazizov, S. Grigorieva, On astrophysical solution to ultrahigh energy cosmic rays, *Phys. Rev. D* 74 (4) (2006) 043005, doi:[10.1103/PhysRevD.74.043005](#).
- [7] N. Globus, D. Allard, E. Parizot, A complete model of the CR spectrum and composition across the Galactic to Extragalactic transition, *Phys. Rev. D* 92 (2015) 021302(R). [1505.01377](#), doi:[10.1103/PhysRevD.92.021302](#).
- [8] V. Berezhinsky, A. Gazizov, S. Grigorieva, Dip in UHECR spectrum as signature of proton interaction with CMB, *Phys. Lett. B* 612 (3–4) (2005) 147–153, doi:[10.1016/j.physletb.2005.02.058](#).
- [9] P.L. Biermann, V. de Souza, Centaurus a: the extragalactic source of cosmic rays with energies above the knee, *Astrophys. J.* 746 (1) (2012) 72. [1106.0625](#), doi:[10.1088/0004-637X/746/1/72](#).
- [10] J. Abraham (The Pierre Auger Collaboration), *Phys. Rev. Lett.* 104 (9) (2010) 091101, doi:[10.1103/PhysRevLett.104.091101](#).
- [11] A. Aab (The Pierre Auger Collaboration), Depth of maximum of air-shower profiles at the Pierre Auger Observatory. I. measurements at energies above  $10^{17.8}$  eV, *Phys. Rev. D* 90 (12) (2014) 122005, doi:[10.1103/PhysRevD.90.122005](#).
- [12] K.-H. Kampert, M. Unger, Measurements of the cosmic ray composition with air shower experiments, *Astropart. Phys.* 35 (10) (2012) 660–678. [1201.0018](#), doi:[10.1016/j.astropartphys.2012.02.004](#).
- [13] G.R. Farrar, The muon content of hybrid events recorded at the Pierre Auger Observatory, in: *Proceedings of the 33rd ICRC*, 2013.
- [14] A. Aab (The Pierre Auger Collaboration), Muons in air showers at the Pierre Auger Observatory: mean number in highly inclined events, *Phys. Rev. D* 91 (2015) 032003, doi:[10.1103/PhysRevD.91.032003](#).
- [15] D. Allard, E. Parizot, A.V. Olinto, On the transition from galactic to extragalactic cosmic-rays: spectral and composition features from two opposite scenarios, *Astropart. Phys.* 27 (1) (2007) 61–75. [astro-ph/0512345](#), doi:[10.1016/j.astropartphys.2006.09.006](#).
- [16] A. Aab (The Pierre Auger Collaboration), Depth of maximum of air-shower profiles at the Pierre Auger Observatory. II. composition implications, *Phys. Rev. D* 90 (12) (2014) 122006, doi:[10.1103/PhysRevD.90.122006](#).
- [17] J. Matthews, A Heitler model of extensive air showers, *Astropart. Phys.* 22 (5–6) (2005) 387–397, doi:[10.1016/j.astropartphys.2004.09.003](#).
- [18] J. Alvarez-Muñiz, et al., Hybrid simulations of extensive air showers, *Phys. Rev. D* 66 (3) (2002) 033011. [astro-ph/0205302](#), doi:[10.1103/PhysRevD.66.033011](#).
- [19] A. Aab (The Pierre Auger Collaboration), *Nucl. Instrum. Meth. A* 798 (2015) 172–213, doi:[10.1016/j.nima.2015.06.058](#).
- [20] H. Kawai (The Telescope Array Collaboration), Telescope Array Experiment, *Nucl. Phys. B Proc. Suppl.* 175–176 (2008) 221–226, doi:[10.1016/j.nuclphysbps.2007.11.002](#).
- [21] D. Heck, et al., CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers, Technical Report, Forschungszentrum Karlsruhe GmbH, Karlsruhe, 1998.
- [22] J.A. Ortiz, G. Medina-Tanco, V. de Souza, Longitudinal development of extensive air showers: hybrid code SENECA and full Monte Carlo, *Astropart. Phys.* 23 (5) (2005) 463–476. [astro-ph/0411421](#), doi:[10.1016/j.astropartphys.2005.02.007](#).
- [23] T. Bergmann, et al., One-dimensional hybrid approach to extensive air shower simulation, *Astropart. Phys.* 26 (6) (2007) 420–432, doi:[10.1016/j.astropartphys.2006.08.005](#).
- [24] K. Werner, T. Pierog, Extended air shower simulations using EPOS, in: *AIP Conference Proceedings*, vol. 928, AIP, 2007, pp. 111–117, doi:[10.1063/1.2775903](#).
- [25] S. Ostapchenko, Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: QGSJET-II model, *Phys. Rev. D* 83 (1) (2011) 014018. [1010.1869](#), doi:[10.1103/PhysRevD.83.014018](#).
- [26] J. Abraham (The Pierre Auger Collaboration), Trigger and aperture of the surface detector array of the Pierre Auger Observatory, *Nucl. Instrum. Meth. A* 613 (1) (2010) 29–39, doi:[10.1016/j.nima.2009.11.018](#).

- [27] J. Hersil, et al., Observations of extensive air showers near the maximum of their longitudinal development, *Phys. Rev. Lett.* 6 (1) (1961) 22–23, doi:[10.1103/PhysRevLett.6.22](https://doi.org/10.1103/PhysRevLett.6.22).
- [28] D.M. Edge, et al., The cosmic ray spectrum at energies above  $10^{17}$  eV, *J. Phys. A* 6 (1973) 1612, doi:[10.1088/0305-4470/6/10/019](https://doi.org/10.1088/0305-4470/6/10/019).
- [29] M. Nagano, et al., Energy spectrum of primary cosmic rays between  $10^{14.5}$  and  $10^{18}$  eV, *J. Phys. G* 10 (1984) 1295–1310, doi:[10.1088/0954-3899/18/2/022](https://doi.org/10.1088/0954-3899/18/2/022).
- [30] C. Di Giulio, Energy calibration of data recorded with the surface detectors of the Pierre Auger Observatory, in: *Proceedings of the 31st ICRC, Łódź, Poland, 2009*.
- [31] W.D. Apel (The KASCADE-Grande Collaboration), Cosmic ray energy reconstruction from the S(500) observable recorded in the KASCADE-grande air shower experiment, *Astropart. Phys.* 77 (2016) 21–31, doi:[10.1016/j.astropartphys.2015.12.002](https://doi.org/10.1016/j.astropartphys.2015.12.002).
- [32] J.C. Arteaga-Velazquez, The energy spectrum of primary cosmic rays reconstructed with the KASCADE-grande muon data, in: *Proceedings of the 31st ICRC, Łódź, Poland, 2009*.
- [33] H. Sagawa, Telescope array extension: TA x 4, in: *Proceedings of the 34th ICRC, 2015*.
- [34] R. Engel, Upgrade of the Pierre Auger Observatory (AugerPrime), in: *Proceedings of the 34th ICRC, 2015*.
- [35] G. Battistoni, et al., The FLUKA code: description and benchmarking, in: *AIP Conference Proceedings*, vol. 896, 2007, pp. 31–49.

## 6 A new air-shower observable to constrain hadronic interaction models



# A new air-shower observable to constrain hadronic interaction models

Raul R. Prado\*, Vitor de Souza

Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, Brazil

## ARTICLE INFO

### Article history:

Received 24 January 2017

Revised 26 April 2017

Accepted 19 May 2017

Available online 30 May 2017

### Keywords:

Extensive air showers

Muons

Hadronic interaction models

## ABSTRACT

The energy spectrum of muons at ground level in air showers is studied and a new observable is proposed to constrain hadronic interaction models used in air shower simulations. An asymmetric Gaussian function is proposed to describe the muon ground energy spectrum and its parameters are studied regarding primary particle, energy and hadronic interaction models. Based on two realistic measurements of the muon density at a given distance from the shower axis, a new observable ( $r_\mu$ ) is defined. Considering realistic values of detector resolutions and number of measured events, it is also shown  $r_\mu$  can be successfully used to constrain low and high energy hadronic interaction models. The study is focused in the energy range between  $10^{17.5}$  and  $10^{18.0}$  eV because of the importance of this interval for particle physics and astrophysical models. The constraining power of the new observable is shown to be large within current experimental capabilities.

© 2017 Elsevier B.V. All rights reserved.

## 1. Introduction

The interaction of ultra high energy cosmic rays with atmospheric nuclei allows us to access hadronic interactions at energies much beyond the reach of man-made accelerators. However several properties of the cosmic ray phenomena and of the detection techniques impose limitations on the available information to study the highest energetic interactions. Experiments are only able to measure the air shower produced by the interaction of the cosmic particles with the nucleus of atoms in the atmosphere. The study of elementary properties of particle physics is done by relating global shower parameters to the properties of the hadronic interactions.

The depth at which the shower reaches its maximum number of particles ( $X_{\max}$ ) and the muon component are the most used shower features related to properties of particle interactions. At the same time, they are also strongly sensitive to the type of the shower primary particle [1]. Given that the primary particle type cannot be determined on event-by-event basis, different primary compositions must be considered in the study of the interaction properties. Usually, the large possibility of primary particles, from proton to iron nuclei, makes it impossible to disentangle the mass of the cosmic particle from particle interaction properties.

This paper proposes an analysis procedure to constrain hadronic interaction models by the measurements of the muon density by two different detectors. The idea proposed here is to compare the

predictions of the models to measurements based on the muon density at ground level. It will be shown in Section 4 that the muon ground energy spectrum has valuable information, which allow to discriminate among different hadronic interaction models.

Hadronic interaction models used in Monte Carlo simulations of air showers are limited to phenomenological approaches tested and tuned to collider data [2]. Three hadronic models for high energy interaction ( $>80$  GeV) and three hadronic models for low energy interaction ( $<80$  GeV) were investigated. A new and realistic air-shower observable based on two measurements of the muon density by different detectors is proposed in this paper and it is demonstrated to be powerful enough to constrain the hadronic interactions models.

The analysis procedure proposed here does not aim to infer properties of the particle interactions, instead, the new proposed parameter has power to test the predictions of the hadronic interaction models within realistic experimental conditions. By comparing the prediction of the models to future measurements of the new parameter, it will be possible to select the best model and guide the way towards a better understanding of the underlying assumptions taken by that model.

The work is based on Monte Carlo simulations of air showers. The new proposed parameter is studied in the energy range from  $10^{17.5}$  to  $10^{18.0}$  eV. The energy range is limited in order to minimize the effect of systematic uncertainties in energy reconstruction and also because of the importance of this range, which is the transition energy range from collider data to the extrapolation domain. It is shown that the new parameter can be used to test hadronic interaction models without knowledge on the primary composition.

\* Corresponding author.

E-mail address: [raul.prado@usp.br](mailto:raul.prado@usp.br) (R.R. Prado).

The paper is organized as follows: Section 2 describes the simulations, Section 3 shows a study on the energy spectra of muons at the ground level, Section 4 proposes the new parameter and studies its behavior with relation to primary mass, energy and detector properties, Section 5 quantifies the constraining power of the new parameter and Section 6 concludes the work.

## 2. Simulations

Extensive air showers were simulated using the Monte Carlo code CORSIKA v7.500 [3]. The applied thinning factor was  $10^{-2}$  for the electromagnetic component and  $10^{-4}$  for the hadronic component, with the maximum weight for any particle set to 100. It was verified that the choice of this thinning configuration does not result in a bias in the present analysis. The energy of the primary particles was sampled continuously between  $10^{17.5}$  and  $10^{18.0}$  eV, following a power law energy spectrum with index  $-3$ . The arrival directions were sampled following a uniform distribution in solid angle, up to a maximum zenith angle of  $60^\circ$ . Three primaries were simulated: proton, nitrogen nucleus and iron nucleus. The minimum energy of particles simulated in air showers were set to 0.3 GeV for hadrons and muons, and 0.003 GeV for electrons and photons. Three hadronic interaction models were used for high energies interactions ( $>80$  GeV): QGSJetII-04 [4], EPOS-LHC [5,6] and Sibyll2.3 [7], and three for low energies ( $<80$  GeV): FLUKA [8], GHEISHA [9] and UrQMD [10]. For each selected combination of hadronic interaction model and primaries particle, 1200 showers were generated. The ground altitude (1400 m above sea level) was chosen to be the one at the Pierre Auger Observatory.

The energy spectrum of muons at ground level is the shower feature to be considered in this paper. The energy spectra were built by collecting from the simulated air showers all muons reaching ground in a lateral distance between 425 and 475 m from the shower axis. Although a full detector reconstruction is not applied, in Section 4.1 the detector effects are taken into account by artificial smearing the shower observables around its simulated values. Detector thresholds and geometry are considered as well.

To sample showers at the same stage of development, the commonly used  $DX$  parameter is used.  $DX$  is the slant atmospheric depth between the shower maximum and the ground. It is given by

$$DX = \frac{X_{\text{gr,vert}}}{\cos \theta} - X_{\text{max}}, \quad (1)$$

where  $X_{\text{gr,vert}}$  is the vertical slant depth of the ground,<sup>1</sup>  $\theta$  is the zenith angle of the shower axis and  $X_{\text{max}}$  is the depth in which the shower reaches its maximum. For each simulated event, the  $\theta$  corresponds to the true zenith angle as set in the input of the simulation and  $X_{\text{max}}$  was taken directly from the Gaisser–Hillas function fitted to the longitudinal energy deposit profile.

## 3. Characterization of the muon ground energy spectrum

In this section, the simulations described in Section 2 are used to characterize the energy spectrum of muons at ground level and to study its relations with primary mass and hadronic interaction models.

Fig. 1 shows examples of the ground energy spectrum of muons for six simulated events, which differ by the primary particle and the shower geometry. The low and high energy hadronic interaction models are FLUKA and QGSJetII-04, respectively. Examples have been selected to illustrate the general shape of the muon spectrum for different primary particles and extreme values of  $DX$ . The left-hand column shows deep showers, with relatively small

$DX$  values, while the right-hand column shows shallow showers, with relatively high values of  $DX$ . The normalization and the mean of the distributions are clearly different but the overall shape is very similar.

As shown by the red lines in Fig. 1, the ground energy spectrum of muons is well described by the following asymmetric Gaussian function

$$\frac{dN_\mu}{dx} = \begin{cases} N_0 \exp \left[ -\frac{1}{2} \left( \frac{x-\eta}{\sigma} \right)^2 \right], & \text{if } x < \eta \\ N_0 \exp \left[ -\frac{1}{2} \left( \frac{x-\eta}{\alpha\sigma} \right)^2 \right], & \text{if } x > \eta \end{cases} \quad (2)$$

where  $x = \log_{10}(E/\text{GeV})$ .  $N_0$  is the normalization parameter and it is correlated with the total number of muons in the shower.  $\eta$  is the mode of the energy distribution, and it is strongly correlated with the average energy of muons reaching ground between 425 and 475 m distance from the shower axis.  $\sigma$  and  $\alpha$  give the width of the distribution,  $\alpha$  being the parameter that measures the degree of asymmetry of the distribution.

Muon energy spectra of all simulated showers were fit by the function presented in Eq. (2), in which  $N_0$ ,  $\eta$ ,  $\sigma$  and  $\alpha$  were taken as free parameters. The fitting was performed using a binned maximum likelihood method with Poissonian probability distribution functions.

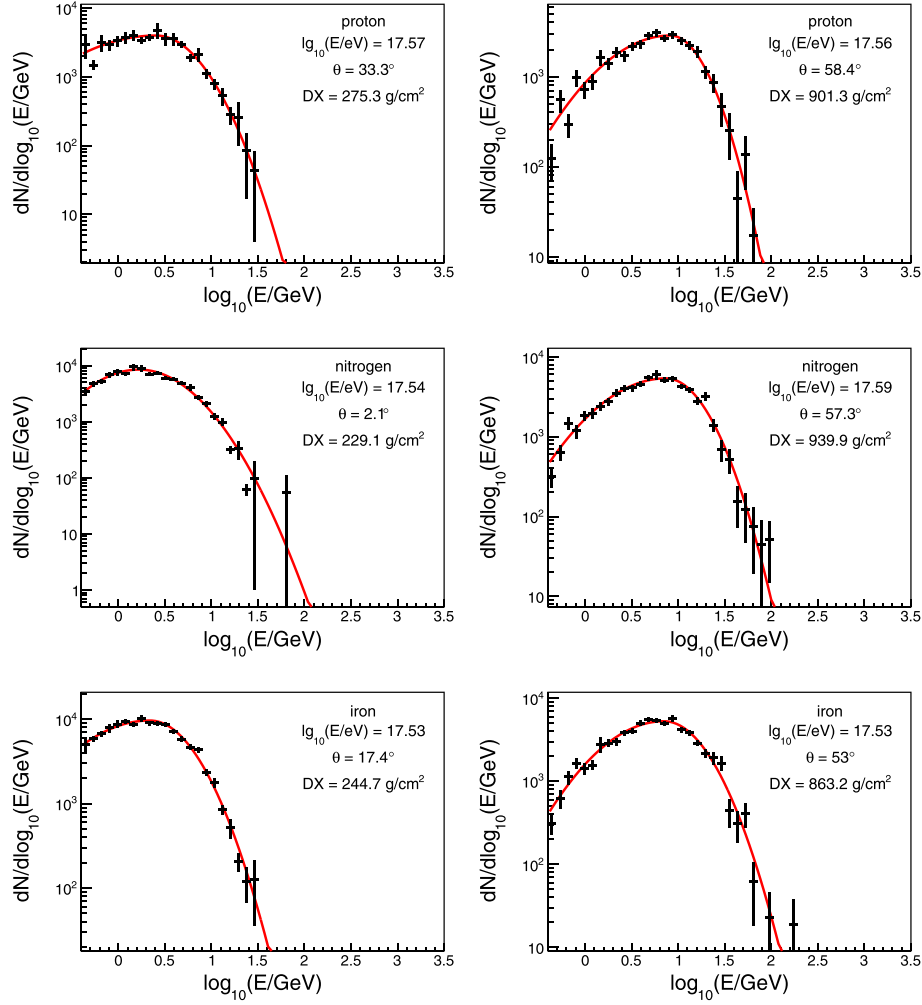
Since the ground energy spectrum of muons is well described by the asymmetric Gaussian function, one may study its dependencies on the energy, primary particle and hadronic interaction models through the evolution of the parameters  $N_0$ ,  $\eta$ ,  $\sigma$  and  $\alpha$  with  $DX$ . Fig. 2 shows the  $DX$  evolution of the four parameters for three different energy intervals, Fig. 3 for three primary particles and Fig. 4 for five combinations of high and low energy hadronic interaction models.

It is clear from Fig. 2 that the normalization ( $N_0$ ) is the only property of ground muon energy spectrum which shows a significant dependence on the primary energy. As expected,  $N_0$  also depends strongly on the primary particle and hadronic interaction model (see Figs. 3 and 4), which reflects the very known behavior of number of muons in air showers. Concerning the peak position of the distributions,  $\eta$ , a strong evolution with  $DX$  is observed, revealing the shift of the average energy of muons to higher values with increasing  $DX$ . Besides normalization and peak positions, changes on the overall shape of the energy spectrum of muons can be evaluated through the parameters  $\sigma$  and  $\alpha$ . These parameters clearly show a very weak evolution with  $DX$ , and a nearly null dependence on the primary energy and primary particle and a relatively very small dependence on the hadronic interaction models.

The analysis of Fig. 4 also repeats known though less popular lessons: the effect of low energy hadronic interaction models are as important as the high energy one regarding the description of the muonic component in air-shower simulations [11]. The differences of  $N_0$  between QGSJetII-04/FLUKA and QGSJetII-04/UrQMD or GHEISHA are of the same order or larger than the largest difference between the high energy interaction models. The average energy of muons, correlated with  $\eta$ , is larger for GHEISHA, followed by UrQMD and then FLUKA. The differences in  $\eta$  due to the low energy hadronic interaction models are as large as the differences due to the high energy hadronic interaction models. Regarding  $\sigma$  and  $\alpha$  parameters, one can see in Fig. 4 they show the same weak  $DX$  evolution for all the hadronic model combinations. Furthermore, the effect of low energy hadronic models is again clear as shown by the difference between FLUKA and GHEISHA, or UrQMD.

Finally, the analysis of Fig. 4 also opens new possibilities. It turns out that  $\eta$  is strongly sensitive to the hadronic interaction models, and at the same time, it does not show any significant dependence on the primary energy and on the primary particle. The lack of primary energy dependence is important experimen-

<sup>1</sup>  $X_{\text{gr,vert}} = 870 \text{ g/cm}^2$  for the Pierre Auger Observatory.



**Fig. 1.** Example of the ground muon energy spectrum for six simulated showers. Shower parameters are shown in each panel. The energy spectra were built by collecting from the air-shower simulation all muons hitting the ground in a lateral distance between 425 and 475 m from the shower axis. The low and high energy hadronic interaction models are FLUKA and QGSJetII-04, respectively. Red lines are the result of fitting Eq. (2) to the data points. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

tally to eliminate effects from the experimental energy scale, while the lack of primary particle dependence is an essential property to disentangle the hadronic interaction effects from the primary composition determination. The evolution of  $\eta$  with  $DX$  is simple (linear) and strong, which makes easy to study showers in different evolution stages.

All of these are indications that accessing experimentally the information carried by  $\eta$  could be successfully used to constrain the hadronic interaction models. In the next section a new observable which is strongly correlated to  $\eta$  is studied and its constraining power in realistic experimental conditions is tested.

#### 4. An observable to test hadronic interaction models

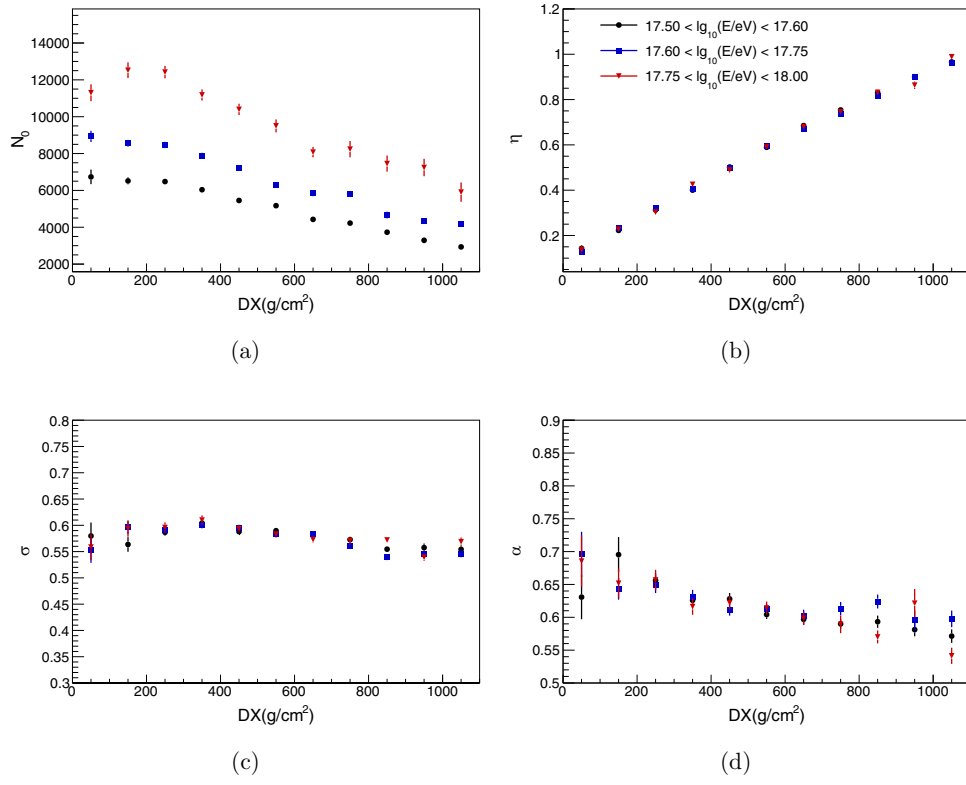
Accessing  $\eta$  directly is not possible for any running or planned UHECR experiment. Therefore, instead of proposing an unrealistic parameter, this study starts from a realistic experimental scenario and aims to find an observable which correlates to  $\eta$ . A generic experimental set-up is considered with two muon detector arrays with different amount of shielding leading to two energy thresholds. With such an experimental set-up, it is possible to measure the integral of the energy spectrum or the energy density of muons above two energy thresholds.

Using the simulations described in Section 2 the density of muons in the lateral distance from 425 to 475 m from the shower axis was calculated for a surface ( $S_{\mu}^{\text{sur}}$ ) and a buried ( $S_{\mu}^{\text{bur}}$ ) detectors and a new parameter is defined:

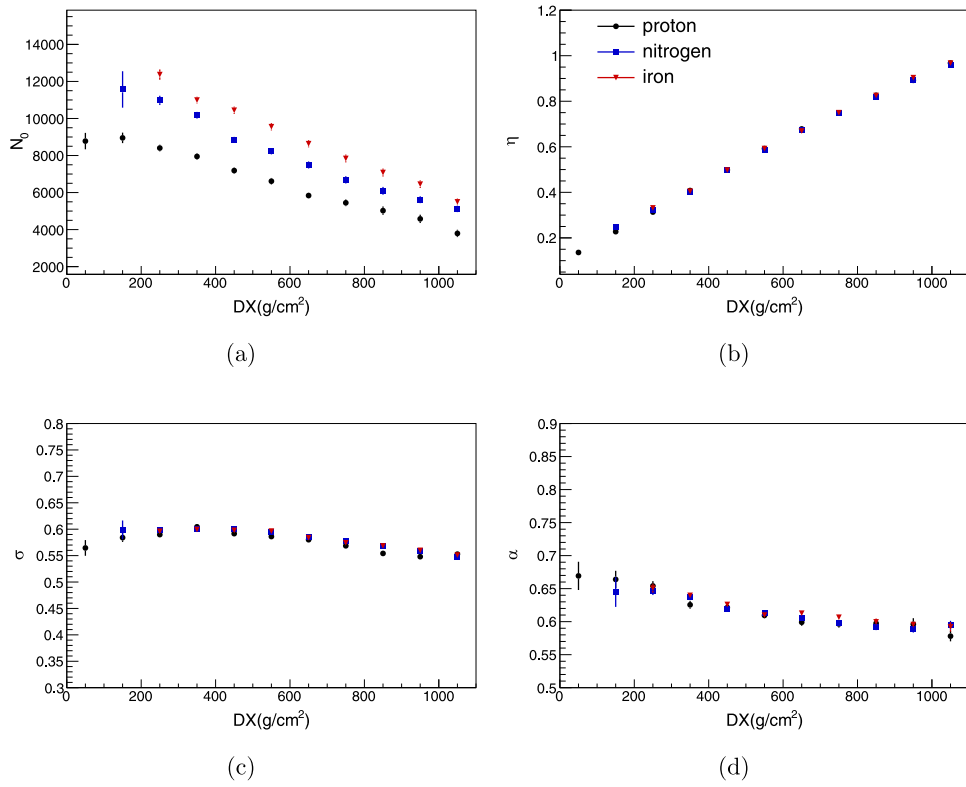
$$r_{\mu} = \sec^{\beta} \theta \frac{S_{\mu}^{\text{bur}}}{S_{\mu}^{\text{sur}}}, \quad (3)$$

where  $\theta$  is the zenith angle of the primary particle. The term  $\sec^{\beta} \theta$  compensates the zenith dependence on the energy threshold and the effective area of the buried detectors. The surface detector ( $S_{\mu}^{\text{sur}}$ ) is considered to have 10 m<sup>2</sup> and the buried detector ( $S_{\mu}^{\text{bur}}$ ) to have 30/cos $\theta$  m<sup>2</sup>, which are motivated by water-Cherenkov stations and flat buried scintillators respectively.  $\beta$  was determined to minimize the primary mass dependence of  $r_{\mu}$ . In Fig. 5 the  $r_{\mu}$  dependence on  $\beta$  is shown for one hadronic interaction models, which justify the choice of  $\beta = 0.6$  as the value that minimize the difference between primaries. The detector features considered is described in the next section.

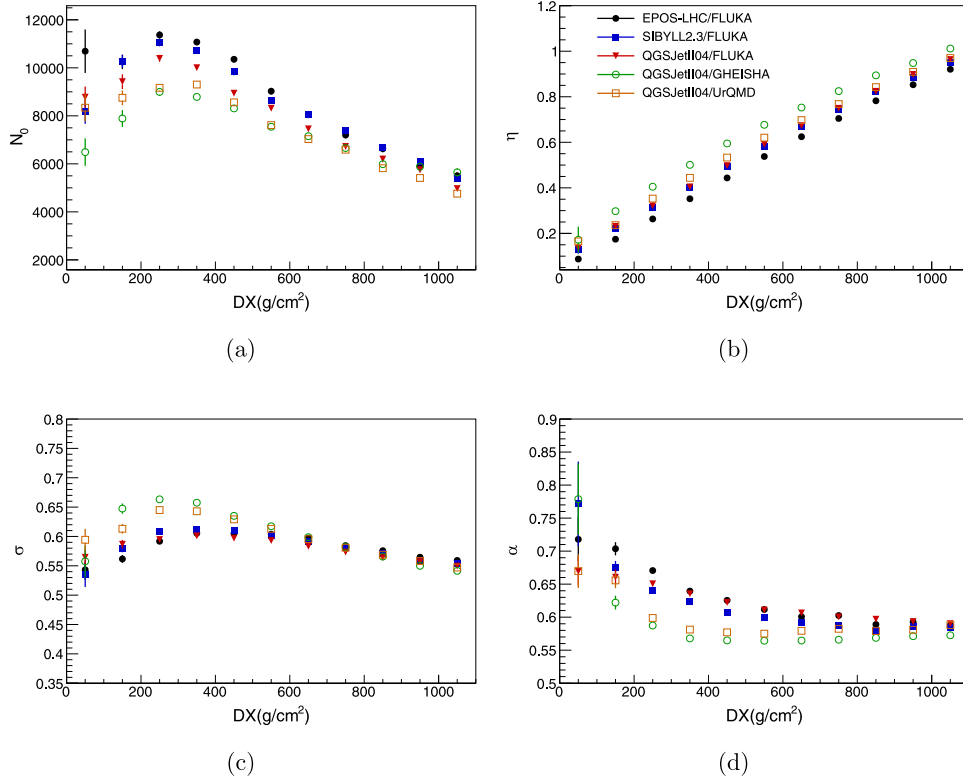
Figs. 6 and 7 show the distributions of  $r_{\mu}$  for several energy thresholds of the buried detector. Three cases are shown in which the vertical muon energy threshold of the buried detector ( $E_{\text{vert},\mu}^{\text{th}}$ ) is changed. The effective energy threshold for a muon with incident zenith angle  $\theta_{\mu}$  is  $E_{\text{vert},\mu}^{\text{th}}/\cos\theta_{\mu}$ . For the surface detectors, the energy threshold is kept fixed at 0.3 GeV. Fig. 6 compares the



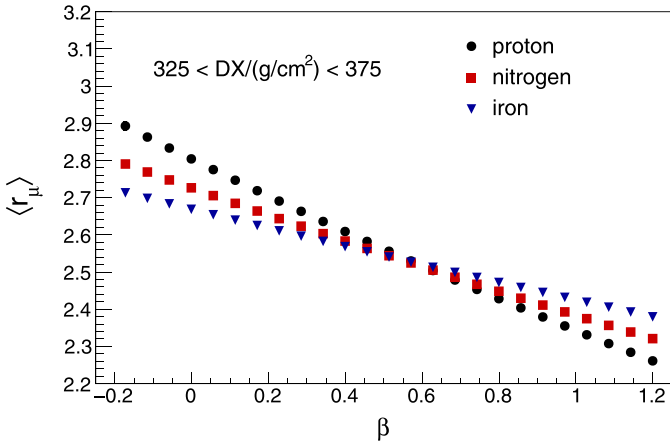
**Fig. 2.** (a)  $N_0$ , (b)  $\eta$ , (c)  $\sigma$  and (d)  $\alpha$  as a function of  $DX$  for three energy intervals. Same number of p, N and Fe showers are considered. The hadronic interaction models used are FLUKA and QGSjetII-04.



**Fig. 3.** (a)  $N_0$ , (b)  $\eta$ , (c)  $\sigma$  and (d)  $\alpha$  as a function of  $DX$  for three primary particles. The energy interval for all primaries is  $17.5 < \lg_{10}(E/\text{eV}) < 18.0$ . The hadronic interaction models used are FLUKA and QGSjetII-04.



**Fig. 4.** (a)  $N_0$ , (b)  $\eta$ , (c)  $\sigma$  and (d)  $\alpha$  as a function of  $DX$  for five combinations of hadronic interaction models. The energy interval for all primaries is  $17.5 < \log_{10}(E/\text{eV}) < 18.0$ . Same number of p, N and Fe showers are considered.

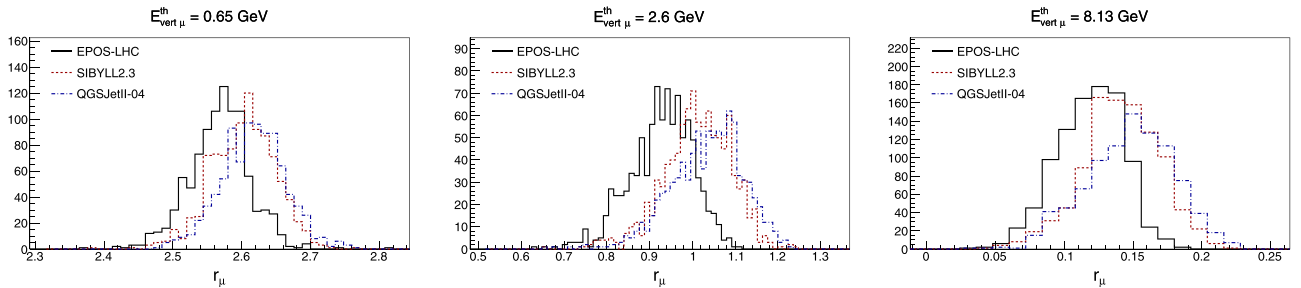


**Fig. 5.**  $\langle r_\mu \rangle$  as a function of  $\beta$  for one small  $DX$  interval and for three different primaries.  $\beta = 0.6$  was chosen to minimize the dependence of  $r_\mu$  with primary particle.

distributions for different high energy hadronic interaction models and Fig. 7 for different low energy hadronic interaction models. The different degrees of separation of the hadronic interaction models with different energy threshold is clear, pointing to the possibility to constrain the hadronic interaction models using  $r_\mu$ . In the next sections, the constraining power of  $r_\mu$  is estimated and its correlation with  $\eta$  is shown.

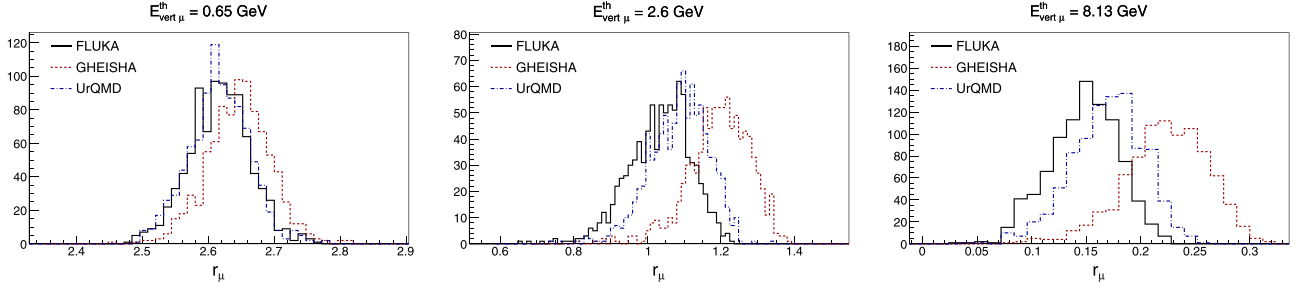
#### 4.1. $r_\mu$ Determination including detector characteristics

In this section, the effect of detector geometry, muon energy thresholds, detector resolutions and systematic uncertainties on  $r_\mu$  are studied. To include the detector features in the proposed analysis, the general characteristics of the muon detectors of the Pierre Auger Observatory are considered. A surface detector with 10 m<sup>2</sup> of effective area is considered to measure muons with energies above 0.3 GeV. Its general characteristics are inspired in the AugerPrime [12–14] design of water-Cherenkov stations with plastic scintillators on top. A second detector is considered based on

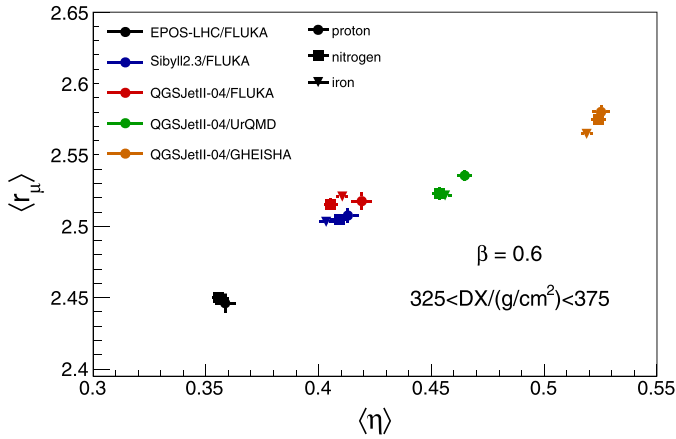


**Fig. 6.**  $r_\mu$  distributions of the three high energy hadronic interaction models. Each panel shows the distributions for a given energy thresholds of the buried detector,  $E_{\text{vert},\mu}^{\text{th}} = 0.65, 2.6, 8.13$  GeV from left to right. Low energy hadronic interaction model is FLUKA.

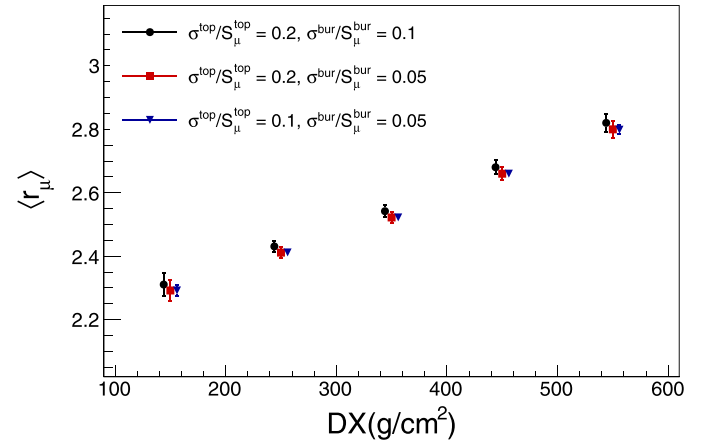




**Fig. 7.**  $r_\mu$  distributions of the three low energy hadronic interaction models. Each panel shows the distributions for a given energy thresholds of the buried detector,  $E_{\text{vert},\mu}^{\text{th}} = 0.65, 2.6, 8.13$  GeV from left to right. High energy hadronic interaction model is QGSJetII-04.



**Fig. 8.** Mean  $r_\mu$  as a function of  $\langle \eta \rangle$  for  $325 < DX/(g/cm^2) < 375$ .  $\beta$  was set to 0.6. See Eqs. (2) and (3) for definitions of the parameters. Detectors effective area and threshold were considered according to Section 4.1. Five combinations of hadronic interaction models is shown for three primary particles. A nearly linear correlation is seen and a clear separation of many hadronic interaction models is visible.



**Fig. 9.**  $\langle r_\mu \rangle$  as a function of  $DX$  for different values of  $S_\mu^{\text{sur}}$  and  $S_\mu^{\text{bur}}$  resolution. The mean value  $\langle r_\mu \rangle$  is calculated over 2000 realizations in which the detector resolution was applied to each simulated shower. Points were artificially shifted in  $DX$  for clarity.

the general characteristics of the Auger AMIGA detector [15,16]. Those are 30 m<sup>2</sup> flat scintillator detectors buried 2.5 m below the ground. The muon energy threshold for the buried detectors depends on the incident zenith angle of the particle and it is given by  $E_{\text{th}} = \beta \rho h / \cos \theta_\mu$ , where  $\beta = 1.808$  MeV cm<sup>2</sup> g<sup>-1</sup> is the fractional energy loss per depth of standard rock,  $\rho = 1.8$  g cm<sup>-3</sup> is the soil density,  $h = 2.5$  m is the vertical depth of the detectors and  $\theta_\mu$  is the zenith incidence angle of the muon. Because of its flatness, the effective collection area of the detector decrease by a factor  $\cos \theta$ , where  $\theta$  is the zenith angle of the shower. The reconstruction of the muon density by an AMIGA-like detector has been shown to be satisfactorily possible for events with  $\theta < 45^\circ$  [17–19].

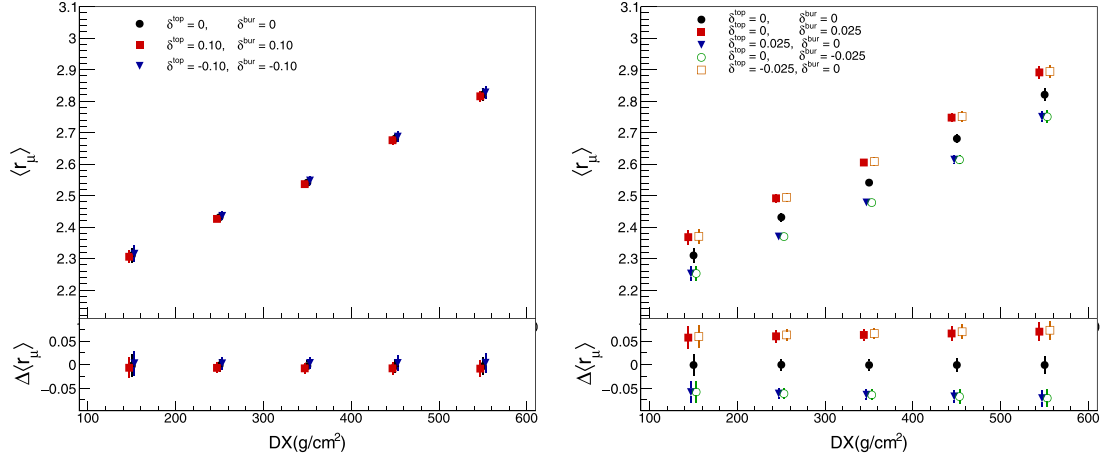
From now on, the results were obtained by considering the detector features shown above to calculate the muon density from the simulated air showers explained in Section 2. In this way, the most important properties of the detectors are considered without the need to perform a full detector simulation.

Fig. 8 shows the relation between the average value of  $r_\mu$  ( $\langle r_\mu \rangle$ ) and the average value of  $\eta$  ( $\langle \eta \rangle$ ) for showers with  $325 < DX/(g/cm^2) < 375$  according to the detector properties explained above. Five combinations of low and high energy hadronic interaction models are shown in different colors and three primaries are shown in different marker styles. From Fig. 8, one can see a nearly linear relation between  $\langle r_\mu \rangle$  and  $\langle \eta \rangle$ . Furthermore, the relatively large separation between the hadronic interaction model combinations by  $\langle r_\mu \rangle$  is clear, which shows  $r_\mu$  is a good observable to constrain hadronic interaction properties.

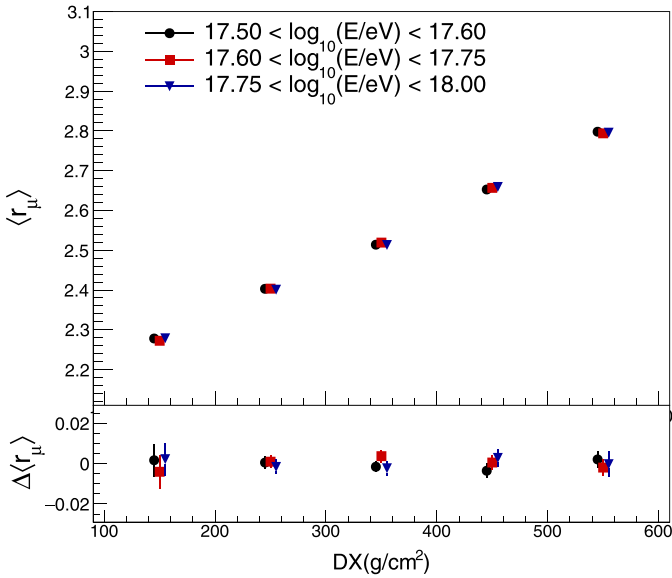
The resolutions on the muon density reconstruction were taken into account by applying a Gaussian smearing on the true signal

obtained from the simulations.  $S_\mu^{\text{sur}}$  resolution was considered to vary in the range from 10% to 20% [12,14] and  $S_\mu^{\text{bur}}$  resolution in the range from 5% to 10% [18,19]. Fig. 9 shows the effect of the resolution in the calculation of  $\langle r_\mu \rangle$ . The Gaussian smearing were performed 2000 times and the  $\langle r_\mu \rangle$  is shown as a function of  $DX$ . The standard deviation of  $r_\mu$  distributions is shown by the error bars. Three cases are shown in which the resolution on both  $S_\mu^{\text{sur}}$  and  $S_\mu^{\text{bur}}$  vary. The systematic effect of the detector resolution in the determination of  $\langle r_\mu \rangle$  is smaller than 5%.

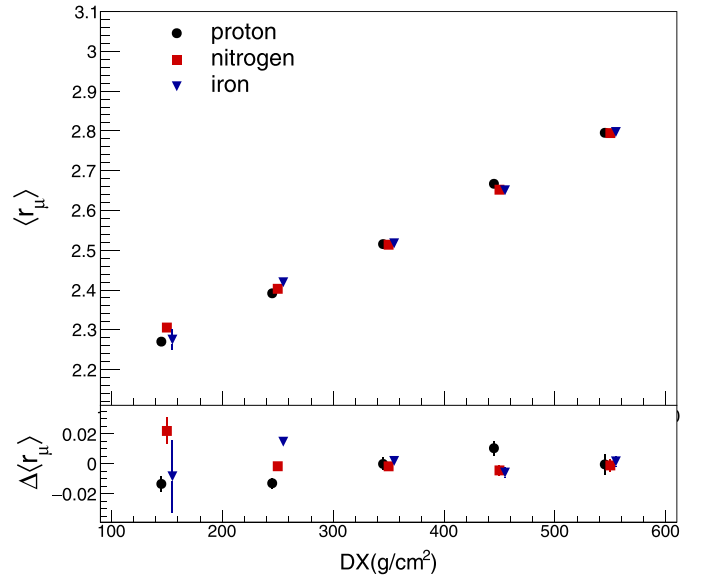
Besides the experimental resolutions, systematic uncertainties on  $S_\mu^{\text{sur}}$  and  $S_\mu^{\text{bur}}$  can be originated from the detection and reconstruction procedures. Typically the most significant systematic uncertainty on muon density measurements are due to systematic uncertainties on the shower energy determination, which affects both  $S_\mu^{\text{sur}}$  and  $S_\mu^{\text{bur}}$  in the same magnitude. To evaluate the effect of systematic uncertainties on  $r_\mu$ , the simulated  $S_\mu^{\text{sur}}$  and  $S_\mu^{\text{bur}}$  were shifted artificially and the resulting  $r_\mu$  were calculated. First it was considered the shifts on  $S_\mu^{\text{sur}}$  and  $S_\mu^{\text{bur}}$  are totally correlated, which means the same magnitude and direction. This case represents the energy reconstruction uncertainty effect. Fig. 10 shows the  $r_\mu$  as a function of  $DX$ , for one hadronic interaction models combination, for different cases in which  $S_\mu^{\text{sur}}$  and  $S_\mu^{\text{bur}}$  were shifted by a factor  $1 + \delta^{\text{top}}$  and  $1 + \delta^{\text{bur}}$  respectively. In Fig. 10 left panel it is shown the effect of a 10% shift on both muon density at the same direction. Clearly, correlated systematic uncertainties on  $S_\mu^{\text{sur}}$  and  $S_\mu^{\text{bur}}$  have an insignificant effect on  $r_\mu$ . In Fig. 10 right panel it is shown the effect of systematic shifts of 2.5% on  $S_\mu^{\text{sur}}$  and  $S_\mu^{\text{bur}}$  in opposite directions. The magnitude of the  $\langle r_\mu \rangle$  deviation is of order 0.05.



**Fig. 10.**  $\langle r_\mu \rangle$  as a function of  $DX$  for different combinations of the systematic uncertainties on  $S_\mu^{\text{sur}}$  and  $S_\mu^{\text{bur}}$ .  $\delta^{\text{top}}$  and  $\delta^{\text{bur}}$  are defined in Section 4.1. All simulated primary particles are included (p, N, Fe). Bottom panel shows the difference with relation to the average value. The hadronic interaction model combination used is QGSJetII-04/FLUKA. Points were artificially shifted in  $DX$  for clarity.



**Fig. 11.**  $\langle r_\mu \rangle$  as a function of  $DX$  for three energy intervals. All simulated primary particles are included (p, N, Fe). Bottom panel shows the difference with relation to the average value. The hadronic interaction model combination used is QGSJetII-04/FLUKA. Points were artificially shifted in  $DX$  for clarity.



**Fig. 12.**  $\langle r_\mu \rangle$  as a function of  $DX$  for three primary particles (pr, N, Fe). All simulated energies are included. Bottom panel shows the difference with relation to the average value. The hadronic interaction model combination used is QGSJetII-04/FLUKA. Points were artificially shifted in  $DX$  for clarity.

The consequences of this deviation on the separation between different hadronic interaction models is discussed in Section 5.

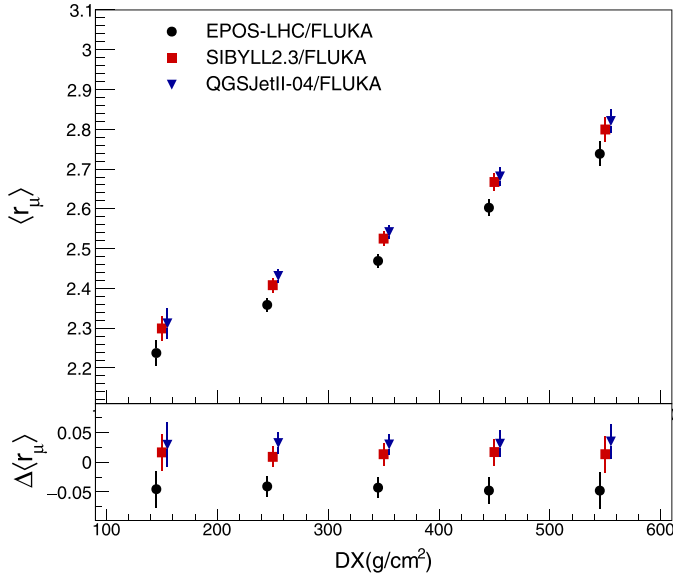
#### 4.2. $r_\mu$ Dependence on primary mass and energy

For the following analysis, the  $DX$  range was defined to preserve a good statistics in all  $DX$  bins and it goes from 100 to 600 g/cm<sup>2</sup> divided in 5 bins of 100 g/cm<sup>2</sup>. The upper bound of 600 g/cm<sup>2</sup> is highly influenced by the shower zenith angle limitation at  $\theta < 45^\circ$  due to the buried detector features.

Fig. 11 shows the evolution of  $\langle r_\mu \rangle$  as a function of  $DX$  for different energy ranges. All simulated primary particles are include (p, N and Fe). Hadronic interaction model combination here is QGSJetII-04/FLUKA. A better visualization of the differences in  $\langle r_\mu \rangle$  is seen in the bottom panel of Fig. 11, where  $\Delta r_\mu$  are the differences with relation to the average value of  $\langle r_\mu \rangle$  for the three energy ranges considered. The differences of  $\langle r_\mu \rangle$  in all energy intervals is smaller than 1% for the entire  $DX$  range. The energy independence of  $\langle r_\mu \rangle$

is expected, since  $S_\mu^{\text{sur}}$  and  $S_\mu^{\text{bur}}$  evolve similarly with energy in the range from  $10^{17.5}$  to  $10^{18.0}$  eV. The lack of energy dependence is an advantage because it allows the analysis of events in a large energy interval, increasing significantly the available statistics. Furthermore, it also contribute to diminishing any effect due to the experimental energy scale.

The primary mass dependence of the  $\langle r_\mu \rangle$  is shown in Fig. 12. The hadronic interaction model combination shown is again QGSJetII-04/FLUKA. Very similar results were obtained for all combinations of models. The bottom panel shows the  $\Delta\langle r_\mu \rangle$ . The observed primary mass dependence is below 2%. The dependence of  $\langle r_\mu \rangle$  with the primary particle was minimized by choosing  $\beta = 0.6$ . The lack of  $\langle r_\mu \rangle$  dependence on the primary particle is a great advantage because it disentangles the study of the hadronic interaction properties from the determination of the primary particle type.



**Fig. 13.**  $\langle r_\mu \rangle$  as a function of  $DX$  for three high energy hadronic interaction models. All simulated energies are included. All simulated primary particles are included (p, N, Fe). Bottom panel shows the difference with relation to the average value. Low energy hadronic interaction model is FLUKA. The detector resolution is set to 20% for  $S_\mu^{\text{sur}}$  and 10% for  $S_\mu^{\text{bur}}$ . Points were artificially shifted in  $DX$  for clarity.

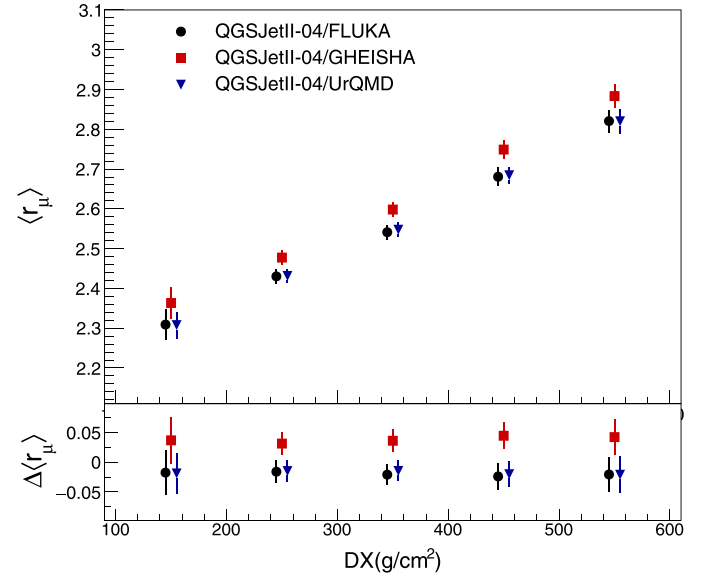
## 5. Results: constraining hadronic interaction models with $r_\mu$ parameter

In this section, the capacity to constrain hadronic interaction models by measuring  $r_\mu$  is demonstrated by studying the  $DX$  evolution of  $\langle r_\mu \rangle$  for different combinations of low and high energy hadronic models. Detector resolution is taken into account as explained above. The effect of a limited number of events is considered here. The total number of simulated air showers used (3500) is approximately the number of hybrid events to be measured by the Pierre Auger Observatory infill array in 2 years of operation. The infill array consists of 750 m spaced water-Cherenkov stations spread over an area of 23.5 km<sup>2</sup>. In this same area the AMIGA-Grande and AugerPrime muon detectors are going to be installed. Considering the duty cycle of the fluorescence detectors to be 14%, the total exposure of this experimental set-up is 4.32 km<sup>2</sup>.sr.yr. Taking into account the cosmic-rays flux between 10<sup>17.5</sup> and 10<sup>18.0</sup> eV, the expected number of events to be measured per year is 1806.

Figs. 13 and 14 show the  $\langle r_\mu \rangle$  as a function of  $DX$  for different combinations of hadronic interaction models. In Fig. 13 the three high energy hadronic models are shown in combination with one low energy hadronic interaction model: FLUKA. In Fig. 14 the three low energy hadronic models are shown in combination with one high energy hadronic interaction model: QGSJetII-04. The bottom panels show the  $\Delta\langle r_\mu \rangle$ . The worst case for the detector resolution was considered: 10% for  $S_\mu^{\text{bur}}$  and 20% for  $S_\mu^{\text{sur}}$ . Figs. 13 and 14 show that even with a relatively poor detector resolution a clear separation between hadronic interaction models is achieved. In Fig. 13 it is observed that EPOS-LHC can be distinguished from QGSJetII-04 and Sibyll2.3, while in Fig. 14 it is seen that GHEISHA can be distinguished from UrQMD and FLUKA.

To better quantify the discriminating power of  $\langle r_\mu \rangle$ , the commonly used Merit Factor can be used. It is defined as:

$$\text{Merit Factor} = \frac{|\langle r_\mu \rangle_a - \langle r_\mu \rangle_b|}{\sqrt{\sigma_a^2 + \sigma_b^2}}, \quad (4)$$



**Fig. 14.**  $\langle r_\mu \rangle$  as a function of  $DX$  for three low energy hadronic interaction models. All simulated energies are included. All simulated primary particles are included (p, N, Fe). Bottom panel shows the difference with relation to the average value. High energy hadronic interaction model is QGSJetII-04. The detector resolution is set to 20% for  $S_\mu^{\text{sur}}$  and 10% for  $S_\mu^{\text{bur}}$ . Points were artificially shifted in  $DX$  for clarity.

where  $a$  and  $b$  refer to any two hadronic interaction model combination and the  $\sigma$ 's are the standard deviations of  $\langle r_\mu \rangle$ .

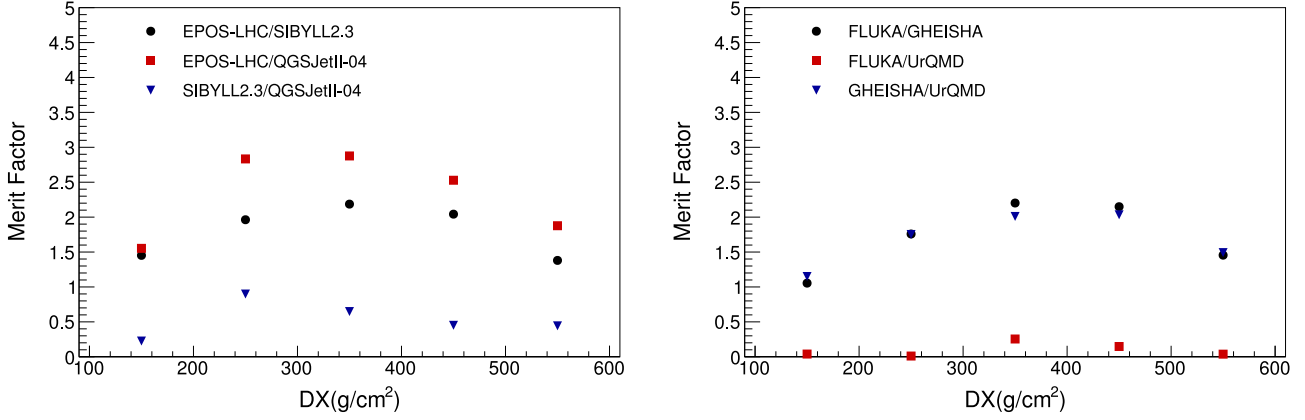
Fig. 15 shows the Merit Factor as a function of  $DX$  for the same experimental resolutions and statistics described above. The left-hand panel refers to hadronic model combinations with different high energy models, and on the right-hand panel with different low energy models. The best Merit Factor is 3.0 between EPOS-LHC and QGSJetII-04 and 2.3 between FLUKA and GHEISHA.

Figs. 16 and 17 show the Merit Factor as a function of the number of events and detector resolution for one particular  $DX$  bin: 300 <  $DX$  (g/cm<sup>2</sup>) < 400. The resolutions on  $S_\mu^{\text{sur}}$  and  $S_\mu^{\text{bur}}$  were re-scaled by a common factor  $f$ , being that  $\sigma_{\text{bur}} = 0.1fS_\mu^{\text{bur}}$  and  $\sigma_{\text{top}} = 0.2fS_\mu^{\text{top}}$ . The effect of the number of events was calculated by re-scaling the standard deviation of the muon densities by a factor  $\sqrt{N_{\text{sim}}/N}$ , where  $N_{\text{sim}}$  is the number of simulated showers and  $N$  is the number of showers in each case.

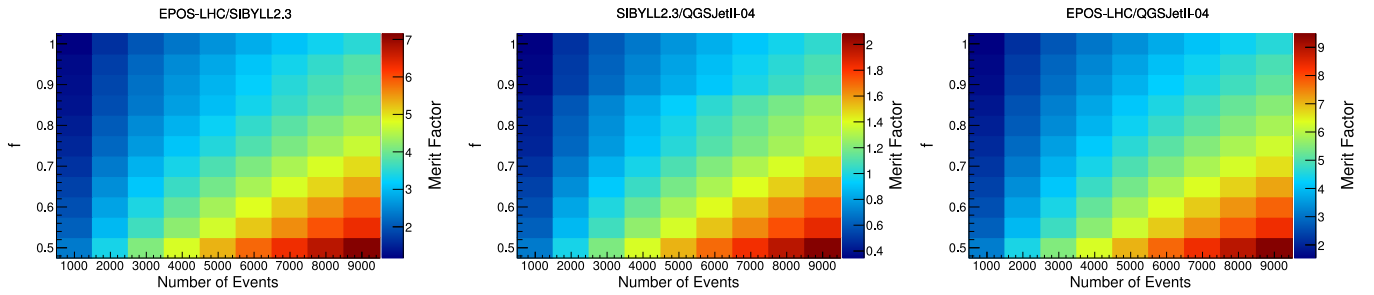
Fig. 16 shows that it is possible to reach large Merit Factor values (>5) for the separation between EPOS-LHC and QGSJetII-04/Sibyll2.3 using a reasonably small number of events (<6000) considering realistic detector resolutions ( $\sigma_{\text{bur}}/S_\mu^{\text{bur}} < 0.06$  and  $\sigma_{\text{top}}/S_\mu^{\text{top}} < 0.13$ ). On the other hand, the separation between Sibyll2.3 and QGSJetII-04 is small for any resolutions and number of events, which is expected because of their similar values of  $\eta$ . The same conclusions can be drawn about the low energy hadronic interactions models from Fig. 17.  $\langle r_\mu \rangle$  provides a very good separation between FLUKA and GHEISHA/UrQMD, but the separation power is limited for GHEISHA and UrQMD.

## 6. Conclusions

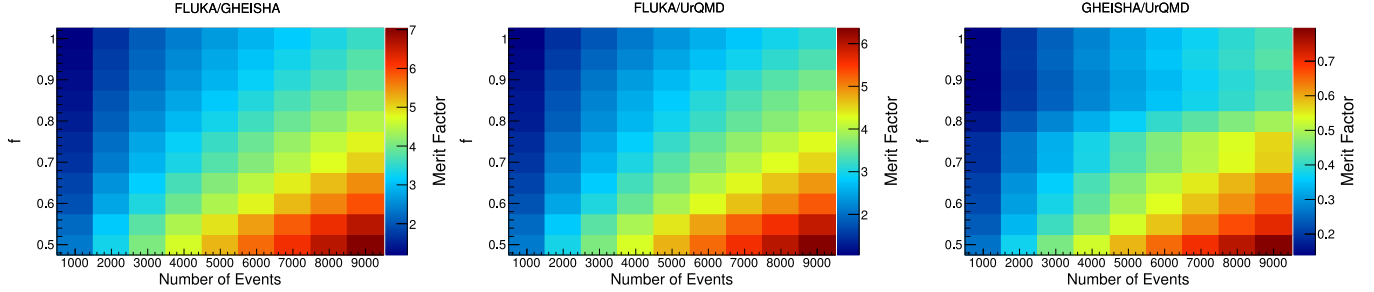
This paper studies the ground muon energy spectrum of air showers and proposes an analysis procedure to constrain hadronic interaction models used in Monte Carlo simulations. In Section 3, it was shown that the energy distribution of muons at ground level can be well described by an asymmetric Gaussian function with four parameters. The study of the evolution of the four parameters with  $DX$  concluded that the overall shape of the energy spectrum of muons does not depend strongly on the combination of low and



**Fig. 15.** Merit Factor calculated for two different hadronic interaction model combination as a function of  $DX$ . Left panel shows the cases in which the high energy hadronic interaction models are different and the low energy one is the same, FLUKA. The legend indicates the two hadronic models considered to calculate the Merit Factor. Right panel shows the cases in which the low energy hadronic interaction models are different and the high energy one is the same, QGSJetII-04.



**Fig. 16.** Merit Factor as a function of the number of events and detector resolution. Detector resolution are defined as  $\sigma_{bur} = 0.1fS_{\mu}^{bur}$  and  $\sigma_{top} = 0.2fS_{\mu}^{sur}$ . The Merit Factor is given in the color scale. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 17.** Merit Factor as a function of the number of events and detector resolution. Detector resolution are defined as  $\sigma_{bur} = 0.1fS_{\mu}^{bur}$  and  $\sigma_{top} = 0.2fS_{\mu}^{sur}$ . The Merit Factor is given in the color scale. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

high energy hadronic interactions models. Also, it was verified that the average muon energy, or the parameter  $\eta$ , is sensitive to the model combination and presents a strong evolution with  $DX$ .

However,  $\eta$  is not an easy parameter to be measured. Therefore in Section 4 a new and experimentally motivated parameter ( $r_{\mu}$ ) is proposed and its correlation with  $\eta$  is shown.  $r_{\mu}$  dependencies on the primary mass and energy were proven to be insignificant which allows on to disentangle the composition and the hadronic interaction studies as well as minimize the effect of systematic uncertainties in the energy reconstruction.

The general properties of the current and proposed muon detectors of the Pierre Auger Observatory are considered to study  $r_{\mu}$  under realistic experimental limitations. Section 5 shows that the discrimination power of  $r_{\mu}$  is significantly large. EPOS-LHC can be separated from Sibyll2.3 and QGSJetII-04 with large Merit Factor ( $>5$ ) using a reasonably small number of events ( $<6000$ ). Sibyll2.3 and QGSJetII-04 show similar average muon energy at ground and therefore can, in the best case, be discriminated with Merit Factor  $\sim 2$  with about 9000 events. As for the low energy hadronic

models, FLUKA can be separated from GHEISHA and UrQMD with large Merit Factor ( $>5$ ) using a reasonably small number of events ( $<6000$ ). GHEISHA and UrQMD can, in the best case, be discriminated with Merit Factor  $\sim 0.8$  with about 9000 events. It was shown that correlated systematic uncertainties on  $S_{\mu}^{sur}$  and  $S_{\mu}^{bur}$  have insignificant influence on  $\langle r_{\mu} \rangle$ . Uncorrelated opposite systematics of order of 2.5% can generate a systematic uncertainty on  $\langle r_{\mu} \rangle$  of the same order of the separation between the hadronic interaction models. This means that for a realistic application of  $r_{\mu}$  analysis, the systematic uncertainties on the muon densities measured by the two detectors have to be correlated (more realistic case). If the systematic uncertainties of the two detectors are in opposite directions (less realistic case) they have to be smaller than 2.5% in order to keep the discrimination power of the hadronic interaction models.

It was also shown that  $r_{\mu}$  is a very robust parameter which can be used irrespective of ignorance of the primary mass composition to test the hadronic interaction models. Constraints imposed by an analysis based on  $r_{\mu}$  can in a short period of time contribute to the

solution of the know problems with muon production in extensive air shower [20,21].

## Acknowledgements

The authors would like to thank Roger Clay for the review of the manuscript on behalf of the Pierre Auger Collaboration. RRP thanks the financial support given by FAPESP (2014/10460-1). VdS thanks the support of the Brazilian population via CNPq and FAPESP (2015/15897-1, 2014/19946-4).

## References

- [1] K.-H. Kampert, M. Unger, Measurements of the cosmic ray composition with air shower experiments, *Astropart. Phys.* 35 (10) (2012) 660–678, doi:[10.1016/j.astropartphys.2012.02.004](https://doi.org/10.1016/j.astropartphys.2012.02.004).
- [2] R. Engel, D. Heck, T. Pierog, Extensive air showers and hadronic interactions at high energy, *Annu. Rev. Nucl. Part. Sci.* 61 (1) (2011) 467–489, doi:[10.1146/annurev.nucl.012809.104544](https://doi.org/10.1146/annurev.nucl.012809.104544).
- [3] D. Heck, et al., *CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers*, Technical Report FZKA 6019, Forschungszentrum Karlsruhe GmbH, Karlsruhe, 1998.
- [4] S. Ostapchenko, Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: QGSJET-II model, *Phys. Rev. D* 83 (1) (2011) 014018, doi:[10.1103/PhysRevD.83.014018](https://doi.org/10.1103/PhysRevD.83.014018).
- [5] K. Werner, T. Pierog, Extended air shower simulations using EPOS, in: AIP Conference Proceedings, 928, AIP, 2007, pp. 111–117, doi:[10.1063/1.2775903](https://doi.org/10.1063/1.2775903).
- [6] T. Pierog, I. Karpenko, J.M. Katzy, E. Yatsenko, K. Werner, EPOS LHC: Test of collective hadronization with data measured at the CERN large hadron collider, *Phys. Rev. C* 92 (2015) 034906, doi:[10.1103/PhysRevC.92.034906](https://doi.org/10.1103/PhysRevC.92.034906).
- [7] F. Riehn, R. Engel, A. Fedynitch, T.K. Gaisser, T. Stanev, A new version of the event generator Sibyll, in: *Proceedings of the 34th International Cosmic Ray Conference (ICRC)*, The Hague, Netherlands, 2015, p. 558.
- [8] A. Ferrari, P.R. Sala, A. Fasso, J. Ranft, FLUKA: A Multi-Particle Transport Code, Technical Report CERN-2005-10 (2005), INFN/TC\_05/11, SLAC-R-773,
- [9] H.C. Fesefeldt, *Simulation of Hadronic Showers, Physics and Applications*, Technical Report PITHA 85-02, RWTH, Aachen, Germany, 1985.
- [10] M. Bleicher, E. Zabrodin, C. Spieles, S.A. Bass, C. Ernst, S. Soff, L. Bravina, M. Belkacem, H. Weber, H. Stöcker, W. Greiner, Relativistic hadron-hadron collisions in the ultra-relativistic quantum molecular dynamics model, *J. Phys. G* 25 (1999) 1859, doi:[10.1088/0954-3899/25/9/308](https://doi.org/10.1088/0954-3899/25/9/308).
- [11] C. Meurer, J. Bluemer, R. Engel, A. Haungs, M. Roth, Muon production in extensive air showers and its relation to hadronic interactions, *Czech. J. Phys.* 56 (1) (2006) A211–A219, doi:[10.1007/s10582-006-0156-9](https://doi.org/10.1007/s10582-006-0156-9).
- [12] R. Engel, for the Pierre Auger Collaboration, upgrade of the Pierre Auger Observatory (AugerPrime), in: *Proceedings of the 34th International Cosmic Ray Conference (ICRC)*, The Hague, Netherlands, 2015, p. 686.
- [13] A. Aab, et al. (The Pierre Auger Collaboration), The Pierre Auger Observatory Upgrade - Preliminary Design Report (2016).
- [14] J.G. Gonzalez, R. Engel, M. Roth, Mass composition sensitivity of combined arrays of water cherenkov and scintillation detectors in the EeV range, *Astropart. Phys.* 74 (2016) 37–46, doi:[10.1016/j.astropartphys.2015.09.003](https://doi.org/10.1016/j.astropartphys.2015.09.003).
- [15] F. Sanchez, For the Pierre Auger Collaboration, in: *Proceedings of the 32nd International Cosmic Rays Conference (ICRC)*, Beijing, China, 2011.
- [16] A. Aab, et al., (The Pierre Auger Collaboration), prototype muon detectors for the AMIGA component of the pierre auger observatory, *J. Instrum.* 11 (02) (2016) P02012, doi:[10.1088/1748-0221/11/02/P02012](https://doi.org/10.1088/1748-0221/11/02/P02012).
- [17] A.D. Supanitsky, A. Etchegoyen, I. Alekotte, M.G. Berisso, M.C. Medina, Underground muon counters as a tool for composition analyses, *Astropart. Phys.* 29 (2008) 461–470, doi:[10.1016/j.astropartphys.2008.05.003](https://doi.org/10.1016/j.astropartphys.2008.05.003).
- [18] D. Ravnani, A.D. Supanitsky, A new method for reconstructing the muon lateral distribution with an array of segmented counters, *Astropart. Phys.* 65 (2015) 1–10, doi:[10.1016/j.astropartphys.2014.11.007](https://doi.org/10.1016/j.astropartphys.2014.11.007).
- [19] D. Ravnani, A.D. Supanitsky, D. Melo, Reconstruction of air shower muon densities using segmented counters with time resolution, *Astropart. Phys.* 82 (2016) 108–116, doi:[10.1016/j.astropartphys.2016.06.001](https://doi.org/10.1016/j.astropartphys.2016.06.001).
- [20] A. Aab, et al., (The Pierre Auger Collaboration), muons in air showers at the pierre auger observatory: mean number in highly inclined events, *Phys. Rev. D* 91 (2015) 032003, doi:[10.1103/PhysRevD.91.032003](https://doi.org/10.1103/PhysRevD.91.032003).
- [21] A. Aab, et al., (The Pierre Auger Collaboration), Testing hadronic interactions at ultrahigh energies with air showers measured by the pierre auger observatory, *Phys. Rev. Lett.* 117 (19) (2016) 192001, doi:[10.1103/PhysRevLett.117.192001](https://doi.org/10.1103/PhysRevLett.117.192001).



# 7 Hadron production in pion-carbon interactions

## (Introduction)

### 7.1 Dataset and simulations

#### (DONE)

The  $\pi^-$ -C data were collected by NA61/SHINE in 2009 at two beam energies: 158 and 350 GeV/c. The  $\pi^-$  beam was a secondary one produced by the collisions of a 400 GeV/c proton beam against a 10 cm long beryllium target. The carbon target consisted of an isotropic graphite plate with 2 cm thickness along the beam axis. For more details about the  $\pi^-$ -C dataset see Ref. [6].

Two trigger modes are relevant for the present analysis: the beam and interaction trigger, which are denominated by T1 and T2 respectively. The definition of T1 is  $S1 \wedge S2 \wedge \overline{V0} \wedge \overline{V1} \wedge \overline{V1'}$  and T2 is  $T1 \wedge \overline{S4}$ . While the T1 assures that a beam particle reached the target position, the T2 is suppose to eliminate events in which a beam particle crossed the target without interacting. Because of the position of the S4 detector, it can also be reached by high energy particles produced by the inelastic interaction at the target, causing the removal of events which are desirable for the analysis. It was verified that the rate of this events is very small and they do not produce a significant bias on our results. For more details about the trigger modes see Ref. [7]. The standard calibration algorithm applied to NA61/SHINE data is described in Ref. [8].

To estimate and remove from the particle spectra the contribution of interactions that do not occur at the target, a set of data were also taken with the target removed. The amount of target removed data is approximately 10% of the total data taken. In ?? we describe the procedure for the target removed subtraction.

The Monte Carlo simulation sets were generated by first generating the primary interactions using hadronic interaction models and then by passing the produced particles through a detailed detector simulation based on GEANT3 package [9]. Three hadronic interaction models were used: EPOS 1.99 [10], DPMJET 3.06 [11] and QGSJET II-04 [12]. For each beam energy and hadronic interaction model, a simulation set was produced with approximately the same event number as the datasets. Both the data and the simulations were reconstructed by the standard NA61/SHINE reconstruction chain [13].



	158 GeV/c	350 GeV/c
Data	3.46 M	3.04 M
EPOS 1.99	3.71 M	3.12 M
DPMJET 3.06	3.92 M	3.47 M
QGSJET II-04	3.71 M	3.06 M

Table 1 –

## 7.2 Upstream and event selections

The upstream and event selection applied in this analysis are described in Ref. [6]. It is important to mention the cut on the vertex position of the main vertex. For this analysis, the events in which the fitted Z position of the main vertex is more than 17 cm distant from the nominal target position (-580 cm) were removed. The motivation for that is to reduce the number of out of target interactions.

In Tab. 1 we show the number of events available for each data and simulation set after the event selection.

In addition to the standard dataset, a set of data were recorded with the carbon target removed. This dataset is used to estimate the contribution of the interactions which do not occur in the target. The two dataset will be refereed as target inserted and target removed along the text.

## 7.3 Track selection

(rewrite)

The following selection criteria were applied to assure the selection of well reconstructed tracks:

- (i) The fitted track has to be inside the geometrical acceptance of the detector.
- (ii) The total number of cluster on the track should be greater than or equal to 25.
- (iii) The sum of clusters on both VTPCs has to be greater than or equal to 12, or the number of cluster on the GTPC has to be greater than or equal to 6.
- (iv) The distance between the track extrapolated to the interaction plane and the interaction point (impact parameter) should be smaller than 4 cm in the both horizontal and vertical plane.

## 7.4 $V^0$ selection

The following criteria were applied to select the  $V^0$ 's:



1. The found vertex has to be identified as a  $V^0$  type vertex
2. The number of daughter tracks has to be equal to 2
3. Both daughter tracks have to be opposite charges
4. The total number of cluster has to be greater than 30 for both tracks
5. At least one track has to have more than 15 clusters in the VTPCs

## 7.5 Phase space binning

(rewrite)

For the present analysis, the phase space were divided in bins of  $p_{\text{and}} p_{\text{T}}$ . In ?? we show the  $(p, p_{\text{T}})$  configuration used in this analysis.

The phase space was split in bins of  $p_{\text{and}} p_{\text{T}}$ . In ?? we show the binning configuration, which is the same for  $\Lambda$  and  $\bar{\Lambda}$  but is different for  $K_S^0$ .

## 7.6 $dE/dx$ analysis

(rewrite)

In this section we present the analysis and the results of the spectra of  $\pi^\pm$ ,  $K^\pm$  and  $p(\bar{p})$ . The analysis is done in track basis and it starts with the track selection that is shown in ?. Second step, after splitting the measured tracks in bins of the phase space, is to perform the particle identification analysis aiming to determine the number of tracks which correspond to each particle type. This is done here by using the energy deposited by the tracks in the TPCs, the  $dE/dx$ , which is briefly explored in ?.

The particle identification (??) is done statistically by fitting the  $dE/dx$  distributions with a combination of particle types. Because of the complicated dependence of the  $dE/dx$  on the particle momentum and properties of the measured track (like number of clusters), the  $dE/dx$  fit turns to be a complex step, requiring the development of a suitable  $dE/dx$  model and an efficient  $dE/dx$  fit strategy. In addition to that, the overlapping of the  $dE/dx$  distributions of different particles in certain regions of momentum makes the  $dE/dx$  fit even more complicated for these regions of the phase space. Because of these difficulties, the particle identification step turns to be the most challenging one and several original ideas were proposed and applied to overcome them.

The following steps are the Monte Carlo correction (??) and the computation of the spectra (??). Finally the derived spectra are shown in ?.

In general, the  $dE/dx$  measurements can provide a very precise particle identification in a statistical basis. For a given momentum interval, the  $dE/dx$  distributions

originated by different charged particles can be well separated by their mean value and shape. Applying a proper model to describe the  $dE/dx$  distribution, it is possible to perform a fitting of the measured distributions and determine the fractions of tracks associated to each particle. This procedure can only fail for a restricted region of momentum in which the average of the energy deposit function of two particles are too close to each other. In the next section, the particle identification procedure will be described.

#### 7.6.1 $dE/dx$ measurements

**(rewrite)**

The  $dE/dx$  associated to each track is defined as the energy lost by the charged particle per unit of length. In NA61/SHINE the  $dE/dx$  is measured by the TPCs, which collect the number of freed electrons from the ionization of the gas by the passage of the charged particles. The determination of the  $dE/dx$  from the signal recorded at the TPCs requires a complex and detailed procedure, which has been very well established by the NA49 and NA61/SHINE experiment along the last decades. Since the detailed description of this procedure is out of the scope of this text, only the general idea and the most important aspects will be presented in the next paragraphs. More complete and detailed approaches can be found in Refs. [14].

Several processes can contribute to the energy loss of charged particle due to its interaction with atoms of the gas in the TPCs, being the emission of electrons by ionization the most relevant one. The electrons emitted are drifted through the chamber and collected in the readout pads, which records the signal as ADC charges. A set of consecutive charges defines a cluster. The 3-dimensional position of the cluster is determined by the position and time distributions in which the charges reaches the readout pad. This position gives the crossing point of the particle track inside the TPCs.

The total charge measured in each cluster is related to the  $dE/dx$  of each track. However, numerous detector effects have to be corrected at the cluster level before grouping the cluster in one unique  $dE/dx$  value. The simplest correction accounts for the geometrical differences due to the incident angle of the track in the pad and the pad widths. More complicated corrections account for differences in the electronic gain and gas pressure/temperature of the pads, differences in the sector gains and losses of electrons during the drift in the chambers and in the readout pad. A detailed description of the correction procedure can be found in Ref. [15].

The track  $dE/dx$  is then determined by the combination of the corrected charges in all clusters. The well known Landau-like shape of the energy loss probability distribution makes the simplest approach, based on the average over all clusters, not suitable. Because of the long tail of the probability distribution, the average and the variance of the measured charges are not well behaved for typical number of clusters ( $\sim 20-150$ ). To overcome this

issue and obtain a satisfactory  $dE/dx$  resolution, the method of the truncated mean is applied, in which only a subset of the clusters is selected to compute the average. The selected clusters are defined by ordering the values of the charge and the selecting the ones inside a given percentage interval. For the NA61/SHINE experiment, it was found the optimal interval being the smallest 50% of the clusters [14].

### 7.6.2 $dE/dx$ model

**(rewrite)**

To perform the particle identification by fitting the measured  $dE/dx$  distribution, a model that describe the  $dE/dx$  distributions of different particle types as a function of their momentum  $p$  is required. Once there is no universal choice of this model, several different alternatives have been used in previous analysis. Although the model chosen here is based on previous studies developed for NA49 and NA61/SHINE experiment, it contain particular features which were found to be the most suitable for the present analysis.

First, the notation adopted in this text has to be presented for clarification. The particle types are represented by the index  $i$ , and it can assume one of the five particle types treated here,  $i = e, \pi, K, p, d$ . The charges are represented by the index  $j$ , being that  $j = +$  or  $j = -$ . Also, the number of cluster measured in a track is represented by  $n_{cl}$  and the  $dE/dx$  is replaced by  $\varepsilon$  for simplicity.

Because the  $dE/dx$  is obtained by averaging the measured charge over a certain number of cluster, it is natural to assume that the shape of the  $\varepsilon$  distribution depends on the  $n_{cl}$ . To be more precise, the  $\varepsilon$  resolution should be larger for smaller  $n_{cl}$  and vice-versa. Additionally, it is obviously expected the mean of the distribution to change with the momentum of the particle  $p$  and the particle type.

Since the shape of the  $\varepsilon$  distribution, for a given  $n_{cl}$  and  $p$ , can be well described by an asymmetric Gaussian function, the probability density function of  $\varepsilon$  for a particle type  $i$  and charge  $j$  is written as

$$f_{i,j}(\varepsilon|p, n_{cl}) = \frac{1}{\sqrt{2\pi}\sigma_{i,j}} \exp \left[ -\frac{1}{2} \left( \frac{\varepsilon - \mu_{i,j}}{\delta \sigma_{i,j}} \right)^2 \right], \quad (7.1)$$

with

$$\delta = \begin{cases} 1 - d, & \varepsilon \leq \mu_{i,j} \\ 1 + d, & \varepsilon > \mu_{i,j}, \end{cases} \quad (7.2)$$

where the parameter  $\mu$  is the mode of the distribution,  $\sigma$  is the resolution and  $d$  is the asymmetry parameter. The  $p$  and  $n_{cl}$  dependence is implicit on the parameters  $\mu$  and  $\sigma$ , as will be explained next. The mode  $\mu$  is related to the mean of the distribution,  $\langle \varepsilon \rangle$ , by

$$\mu_{i,j} = \langle \varepsilon \rangle_{i,j} - \frac{\sigma_{i,j}}{\sqrt{2\pi}} \left[ (1 + d)^2 - (1 - d)^2 \right]. \quad (7.3)$$

The evolution of  $\langle \varepsilon \rangle$  is expected to follow a Bethe-Bloch-like function. In this model, a reference  $\langle \varepsilon \rangle$  (p) curve is defined by a data-based parametrization using a generic function which is a variation of the Bethe-Bloch function. The reference value of  $\langle \varepsilon \rangle$  for a given  $p$  is denoted as  $\langle \varepsilon \rangle^{\text{BB}}$ . To account for deviations from the reference  $\langle \varepsilon \rangle$ , the present model includes a set of parameters called *calibration constants*, which are denoted by  $X$ . These parameters act as logarithmic shifts of the  $\langle \varepsilon \rangle$  around  $\langle \varepsilon \rangle^{\text{BB}}$  and they can in principle be applied to each particle and charge separately. To reduce the complexity of the model, it is assumed here one global calibration constant for each charge that follows the  $\langle \varepsilon \rangle$  of the  $\pi$  distribution and individual calibration constants for the other particles, but being common for both charges. In the end, the  $\langle \varepsilon \rangle$  for a given particle type  $i$  and charge  $j$  is given by

$$\langle \varepsilon \rangle_{i,j} = \begin{cases} \langle \varepsilon \rangle_i^{\text{BB}} e^{X_i^j} & (i = \pi) \\ \langle \varepsilon \rangle_i^{\text{BB}} e^{X_\pi^j} e^{X_i^j} & (i \neq \pi). \end{cases} \quad (7.4)$$

In total, 6 calibration constants are included in the model:  $X_\pi^+$ ,  $X_\pi^-$ ,  $X_e$ ,  $X_K$ ,  $X_p$  and  $X_d$ .

The dependence of the resolution  $\sigma$  on  $n_{\text{cl}}$  is assumed to be of the form  $\sigma \sim 1/\sqrt{n_{\text{cl}}}$ . Besides that,  $\sigma$  is assumed to depend on the  $\langle \varepsilon \rangle$  by a power law relation and a normalization parameter for each charge is also included ( $\sigma_0^j$ ). The final expression for the resolution is,

$$\sigma_{i,j} = \frac{\sigma_0^j}{\sqrt{n_{\text{cl}}}} \langle \varepsilon \rangle_{i,j}^\alpha, \quad (7.5)$$

in which 3 more parameters are included:  $\sigma_0^+$ ,  $\sigma_0^-$  and  $\alpha$ .

By combining the ????????? with the ??, we obtain the probability density function of  $\varepsilon$  for each particle  $i$  and charge  $j$ . Besides the 6 calibration constants, the model includes 4 *shape parameters*:  $\sigma_0^+$ ,  $\sigma_0^-$ ,  $\alpha$  and  $d$ . Altogether there are 10 parameters that can be set free to fit the model to the measured  $\varepsilon$  distributions.

### 7.6.3 $dE/dx$ fit strategy

## 7.7 $V^0$ analysis

### 7.7.1 Signal extraction

## 7.8 Corrections

## 7.9 Spectra

### 7.9.1 Statistical uncertainties

### 7.9.2 Systematic uncertainties

## 7.10 Results

## 7.11 Summary and conclusions

## 8 Conclusions



# Bibliography

- 1 PIEROG, T. Connecting accelerator experiments and cosmic ray showers. **EPJ Web Conf.**, v. 53, p. 01004, 2013.
- 2 KNAPP, J.; HECK, D.; SCIUTTO, S. J.; DOVA, M. T.; RISSE, M. Extensive air shower simulations at the highest energies. **Astropart. Phys.**, v. 19, p. 77–99, 2003.
- 3 MEURER, C.; BLUEMER, J.; ENGEL, R.; HAUNGS, A.; ROTH, M. Muon production in extensive air showers and its relation to hadronic interactions. **Czech. J. Phys.**, v. 56, p. A211, 2006.
- 4 PIEROG, T. Open issues in hadronic interactions for air showers. **EPJ Web Conf.**, v. 145, p. 18002, 2017.
- 5 ESPADANAL, J.; CAZON, L.; CONCEIÇÃO, R. Sensitivity of EAS measurements to the energy spectrum of muons. **Astropart. Phys.**, v. 86, p. 32–40, 2017.
- 6 ADUSZKIEWICZ, A. et al. Measurement of meson resonance production in  $\pi^- + C$  interactions at SPS energies. **Eur. Phys. J.**, C77, n. 9, p. 626, 2017.
- 7 RUPRECHT, M. **Measurements of the Spectrum of Charged Hadrons in  $\pi^- + C$  Interactions with the NA61/SHINE Experiment**. 2012. Tese (Doutorado) — Institute of Experimental Particle Physics, Karlsruhe Institute of Technology, 2012.
- 8 ABGRALL, N. et al. Calibration and analysis of the 2007 data. 2008.
- 9 BRUN, R.; BRUYANT, F.; CARMINATI, F.; GIANI, S.; MAIRE, M.; MCPHERSON, A.; PATRICK, G.; URBAN, L. GEANT Detector Description and Simulation Tool. 1994.
- 10 PIEROG, T.; WERNER, K. Muon Production in Extended Air Shower Simulations. **Phys. Rev. Lett.**, v. 101, p. 171101, 2008.
- 11 ROESLER, S.; ENGEL, R.; RANFT, J. The Monte Carlo event generator DPMJET-III. In: **Advanced Monte Carlo for radiation physics, particle transport simulation and applications. Proceedings, Conference, MC2000, Lisbon, Portugal, October 23-26, 2000**. [s.n.], 2000. p. 1033–1038. Available from Internet: <http://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=SLAC-PUB-8740>.
- 12 OSTAPCHENKO, S. Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: I. QGSJET-II model. **Phys. Rev.**, D83, p. 014018, 2011.
- 13 ABGRALL, N. et al. Measurements of Cross Sections and Charged Pion Spectra in Proton-Carbon Interactions at 31 GeV/c. **Phys. Rev.**, C84, p. 034604, 2011.
- 14 VERES, G. I. **Baryon Momentum Transfer in Hadronic and Nuclear Collisions at the CERN NA49 Experiment**. 2001. Tese (Doutorado) — Department of Atomic Physics, Eotvos Lorand University, 2001.

15 MARCINEK, A.  $\phi$  meson production in proton-proton collisions in the NA61/SHINE experiment at CERN SPS. 2016. Tese (Doutorado) — Institute of Physics, Jagiellonian University, 2016.