PHYSICAL REVIEW D **109,** 043003 (2024)

Simple and less hadronic-model-dependent method for cosmic-ray composition determination

A. C. Cobos[*]

Instituto de Tecnologías en Detección y Astropartículas Mendoza (CNEA, CONICET, UNSAM),
CNEA Regional Cuyo, Godoy Cruz, Mendoza, Argentina

A. D. Supanitsky and A. Etchegoyen

Instituto de Tecnologías en Detección y Astropartículas (CNEA, CONICET, UNSAM),
Centro Atómico Constituyentes, San Martín, Buenos Aires, Argentina

(Received 7 November 2023; accepted 19 December 2023; published 5 February 2024)

The origin and nature of ultrahigh-energy cosmic rays remains a mystery despite the great progress that has been made in recent years mainly due to the observations performed by the Pierre Auger Observatory and Telescope Array. In particular, the composition information of the cosmic rays as a function of the energy appears to be crucial for the understanding of their origin. The best observables for primary mass composition are the muon content of extensive air showers and the atmospheric depth of the shower maximum development. In this work we present a novel method which allows us to use any type of classification method to perform mass composition analyses based on the number of muons. The method also provides the information of the mean values of this observable corresponding to different primary particles. The analyses are based on numerical simulations of the showers but including experimental uncertainties in the reconstruction of the energy and in the measurement of the number of muons. We also study the impact of the use of different high-energy hadronic interaction models in the composition analyses performed. The biases introduced by the use of different high-energy hadronic interaction models to analyze the data obtained with the presented method are considerably reduced compared with those obtained with classic statistical methods that use a single physical observable.

DOI: 10.1103/PhysRevD.109.043003

I. INTRODUCTION

The cosmic-ray energy spectrum extends from below 10^9 to above 10^{20} eV. It can be approximated by a broken power law with some spectral features: the knee at a few 10^{15} eV [1–5], a second knee at $\sim 10^{17}$ eV [6], the ankle at $\sim 5 \times 10^{18}$ eV [7] and a suppression at $\sim 4 \times 10^{19}$ eV [7,8]. Different experimental techniques have been used for the observation of cosmic rays depending on the energy range under consideration. Due to their low flux at energies $\gtrsim 10^{15}$ eV, their detection can only be achieved by measuring extensive air showers (EASs), cascades of billions of secondary particles resulting from the interaction of the primary cosmic rays with molecules of Earth's atmosphere. The EASs present three main components: the electromagnetic one which is formed by electrons, positrons and gamma rays, the muonic one which is formed by muons and antimuons, and the hadronic one which is formed by high-energy nucleons and other hadrons. The EASs consist of a core of high-energy hadrons that continuously feed the electromagnetic component, mainly through the decay of neutral pions and eta particles. The muons are generated by the decay of lower-energy charged pions and kaons.

The Pierre Auger Observatory [9] constructed in the province of Mendoza, Argentina is at present the largest observatory ever built for measuring ultrahigh-energy cosmic rays (UHECRs, with energies $\gtrsim 10^{18}$ eV). This observatory combines arrays of surface detectors (water-Cherenkov tanks) with fluorescence telescopes. Surface detectors allow one to reconstruct the lateral development of the showers by detecting secondary particles that reach the ground. Fluorescence telescopes are used to study the longitudinal development of the showers. The combination of the two techniques into a hybrid observatory maximizes the precision in the reconstruction of the EAS properties and minimizes systematic errors. The Telescope Array project [10], located in Utah, USA, is also a hybrid detector that combines arrays of surface detectors with fluorescence telescopes, allowing also the reconstruction of the EAS properties with high precision. In this case the surface detectors are composed of scintillator detection devices housed inside metal clad containers.

The cosmic-ray origin still remains a mystery despite great theoretical and experimental efforts made in recent

---

[*] agustin.cobos@iteda.cnea.gov.ar

years. Recent results suggest that the UHECR flux is composed predominantly of hadronic primary particles [11–13]. As charged particles, they suffer deflections in cosmic magnetic fields and, then, their arrival directions do not point back directly to their sources. Therefore, an indirect search for their origin is necessary. In particular, composition information appears to be crucial to find the transition between the galactic and extragalactic components of the cosmic rays and to elucidate the origin of the suppression at the highest energies [14–17].

The two best indicators of primary mass composition are the atmospheric depth corresponding to the maximum shower development, X_{max}, and the muon content of the shower [17–20]. Primary mass composition analyses can only be performed by comparing experimental data with EAS simulations. These simulations are subject to large systematic uncertainties because they are based on high-energy hadronic interaction models (HEHIMs) that extrapolate low-energy accelerator data to the highest energies. The most used HEHIMs in the literature have been recently updated by using data taken by the Large Hadron Collider (LHC). These models (EPOS-LHC [21], SIBYLL2.3c [22] and QGSJetII-04 [23]) are called post-LHC models.

Machine learning research studies algorithms and modeling tools used to infer patterns from data. The better the algorithm, the more accurate decisions and predictions will become as it processes more data. In recent years, machine learning has been used extensively in the physical sciences, including particle and astroparticle physics [24–26]. For the particular case of ultrahigh-energy cosmic-ray physics, the performance of deep neural networks has been studied to estimate X_{max} and the muon content of EASs from the signals in the water-Cherenkov detectors of the Pierre Auger Observatory [27–29]. As mentioned before, the main limitation for the composition determination is the systematic uncertainties introduced by the use of different HEHIMs. For this reason the machine learning techniques are not used directly to determine the composition from data at the detector level. Currently, these techniques are used to determine parameters that are sensitive to primary mass and then the composition can be inferred in subsequent analyses.

In this work, we develop a method to perform primary mass composition analyses that allows us to significantly reduce the systematic uncertainties introduced by the use of HEHIMs to analyze the experimental data. This new method makes use of the information provided by any type of classification algorithm. In particular, the classifier used in this work is a simple neural network. The parameter sensitive to primary mass is the number of muons, focusing on the relation between the mean values corresponding to different HEHIMs. The analysis is performed for binary mixtures of different theoretical values of proton abundance. The studies are performed by using numerical simulations, which include experimental uncertainties in the reconstruction of the energy and in the number of muons. It is worth mentioning that, although the analyses are performed for binary mixtures, the method could be extended to mixtures of more than two types of particles.

II. ANALYSIS

A. Development of the method

In this subsection we present a new method to determine the primary mass composition of cosmic rays by using a neural network. In a first stage we define and analyze different variables obtained with the neural network in a straightforward way in order to have a conceptual scaffolding, then, we describe the method developed. As mentioned in the previous section, the composition analysis is carried out based on the number of observable muons and the simplified case in which there are just two nuclear species, A_0 and A_1, considered. The analysis is also carried out for a fixed primary energy; therefore, all shower variables and distribution functions defined hereafter will be referred to as this fixed parameter. The dependence on this parameter will be omitted for compactness of the notation. A detailed description of the construction of the shower variables and the functions defined here can be found in Sec. II B. A Multilayer perceptron neural network is used with one neuron in the input/output layers. Two hidden layers are considered with 20 neurons for each one. The output function is sigmoid type, where an output $= 0$ (1) is considered if the sigmoid output is ≤ 0.5 (>0.5).

Let be $N_\mu(A_0, \mathrm{HM}_{\mathrm{ref}})$ and $N_\mu(A_1, \mathrm{HM}_{\mathrm{ref}})$ the total number of muons that reach the ground due to primaries A_0 and A_1, corresponding to a reference HEHIM, $\mathrm{HM}_{\mathrm{ref}}$. The neural network is subjected to the training process in order to provide an output $= 0$ or 1 when their input variables correspond to $N_\mu(A_0, \mathrm{HM}_{\mathrm{ref}})$ or $N_\mu(A_1, \mathrm{HM}_{\mathrm{ref}})$, respectively. In other words, we train the neural network so that it can recognize the primaries A_0 (output $= 0$) and A_1 (output $= 1$) from their N_μ values.

Let be $N_\mu(c_{A_0}, \mathrm{HM})$ a muon number sample of size N of a given HEHIM, HM, corresponding to a binary mixture of primaries A_0 and A_1 with respective abundances c_{A_0} and c_{A_1}; $c_{A_0} + c_{A_1} = 1$. The estimator of the abundance c_{A_0}, $c^{\mathrm{nn}}_{A_0}(\mathrm{HM})$, obtained from the output of the neural network when this sample is used, is given by the ratio between the number of output $= 0$ and the total number of events, N. Figure 1 (top) shows the mean values of the proton abundance estimator, $\langle c^{\mathrm{nn}}_{\mathrm{p}} \rangle$, obtained directly with the neural network for different values of c_{p} corresponding to binary mixtures of proton and iron ($A_0 = \mathrm{p}$, $A_1 = \mathrm{Fe}$). The primary energy corresponds to $E = 10^{18}$ eV and the size of the samples is $N = 5000$ (see Sec. II B for more details). Experimental uncertainties have been included. For each value of c_{p} and for each HEHIM, 100 independent samples were generated to obtain the $\langle c^{\mathrm{nn}}_{\mathrm{p}} \rangle$ values. As mentioned in the previous section, the HEHIMs used are EPOS-LHC [21],

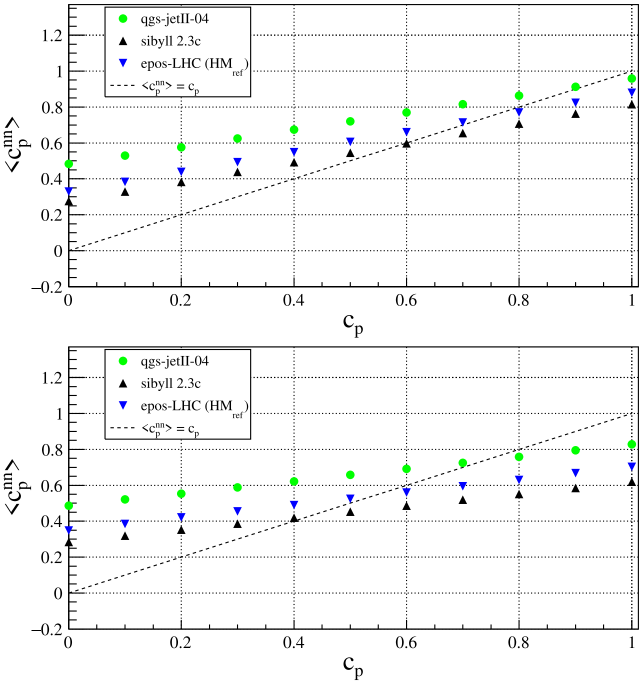

FIG. 1. Mean value of proton abundance estimator, $\langle c_{\mathrm{p}}^{\mathrm{nn}} \rangle$, as a function of c_{p} obtained directly with the neural network used in this work for proton-iron (top) and proton-nitrogen (bottom) binary mixtures. The primary energy is $E = 10^{18}$ eV. The size of the samples is $N = 5000$. For each value of c_{p} and for each HEHIM, 100 independent samples were generated taking into account experimental uncertainties (see Sec. II B for more details). The HEHIMs used are EPOS-LHC [21], SIBYLL2.3c [22] and QGSJetII-04 [23]. EPOS-LHC is used as the reference model ($\mathrm{HM_{ref}}$). The error bars are smaller than the marker size. The dash line corresponds to $\langle c_{\mathrm{p}}^{\mathrm{nn}} \rangle = c_{\mathrm{p}}$.

SIBYLL2.3c [22] and QGSJetII-04 [23]. EPOS-LHC is used as the reference model ($\mathrm{HM_{ref}}$). Figure 1 (bottom) shows $\langle c_{\mathrm{p}}^{\mathrm{nn}} \rangle$ values for the same cases of Fig. 1 (top) but for proton and nitrogen mixtures ($A_1 = \mathrm{nit}$).

Note in Fig. 1 that the error bars are smaller than the marker size since, for the same primary abundance, the standard deviation of the number of outputs $= 0$ is very small even considering experimental uncertainties. However, the neural network performance is relatively poor and hence the misclassification rate is relatively high. This can be seen from the fact that the estimator $c_{\mathrm{p}}^{\mathrm{nn}}$ produces biases even if $\mathrm{HM} = \mathrm{HM_{ref}}$. This can be explained by computing, after the training process, the efficiencies (also called true positive rate) $\varepsilon_{A_i} = \varepsilon_{A_i}(\mathrm{HM_{ref}})$, defined only for $\mathrm{HM_{ref}}$ as the quotient between the obtained number of output $= i$ due to primaries of type A_i and the total number of primaries of type A_i used as input of the neural network when this number tends to infinity; $i = 0, 1$. Then, given a sample $N_\mu(c_{A_0}, \mathrm{HM_{ref}})$, the normalized numbers of output $= i$ due to primaries of type A_j, $n_{A_j}(i) = N_{A_j}(i)/N$, can be expressed by the following equations:

$$n_{A_0}(0) = c_{A_0}\varepsilon_{A_0}, \tag{1}$$

$$n_{A_0}(1) = c_{A_0}(1 - \varepsilon_{A_0}), \tag{2}$$

$$n_{A_1}(1) = c_{A_1}\varepsilon_{A_1}, \tag{3}$$

$$n_{A_1}(0) = c_{A_1}(1 - \varepsilon_{A_1}) \tag{4}$$

and the estimator $c_{A_0}^{\mathrm{nn}}(\mathrm{HM_{ref}})$ will be expressed by the normalized number of output $= 0$, i.e. by the sum of Eqs. (1) and (4)

$$c_{A_0}^{\mathrm{nn}}(\mathrm{HM_{ref}}) = c_{A_0}\varepsilon_{A_0} + c_{A_1}(1 - \varepsilon_{A_1}). \tag{5}$$

Taking into account that c_{A_0} does not depend on ε_{A_0} or ε_{A_1}, from Eq. (5) one can see that $c_{A_0}^{\mathrm{nn}}(\mathrm{HM_{ref}})$ is equal to c_{A_0} only if $\varepsilon_{A_0} = \varepsilon_{A_1} = 1$.

The effect of assuming a given HEHIM as the true one ($\mathrm{HM_{ref}}$) whereas the real one is another (HM) can be studied by replacing $c_{A_0}^{\mathrm{nn}}(\mathrm{HM_{ref}})$ with $c_{A_0}^{\mathrm{nn}}(\mathrm{HM})$ in Eq. (5). In this way, the test samples are generated with HM but the neural network is trained—and the ε_{A_i}'s are calculated—with $\mathrm{HM_{ref}}$. Taking into account that $c_{A_1} = 1 - c_{A_0}$ and solving for c_{A_0} from Eq. (5) with the mentioned replacement, the estimator $c_{A_0}^{\mathrm{nn}\varepsilon} = c_{A_0}^{\mathrm{nn}\varepsilon}(c_{A_0}^{\mathrm{nn}}, \varepsilon_{A_0}, \varepsilon_{A_1})$ can be defined as

$$c_{A_0}^{\mathrm{nn}\varepsilon} = \frac{c_{A_0}^{\mathrm{nn}}(\mathrm{HM}) + \varepsilon_{A_1} - 1}{\varepsilon_{A_0} + \varepsilon_{A_1} - 1}. \tag{6}$$

Figure 2 shows the mean values $\langle c_{\mathrm{p}}^{\mathrm{nn}\varepsilon} \rangle$ as a function of c_{p} obtained from Eq. (6) for the same samples of Fig. 1. It can be seen that $c_{A_0}^{\mathrm{nn}\varepsilon}$ does not produce biases when $\mathrm{HM} = \mathrm{HM_{ref}}$. By comparing the top and bottom panels of Fig. 2, one can see that the magnitudes of the biases corresponding to the proton-nitrogen mixtures are greater than those of the proton-iron mixtures; this is due to the decrease in the merit factor between the primary distributions of the former mixtures [see Eq. (A1) of Appendix A]. For the same reason, the magnitudes of the biases for the proton-helium mixtures are even larger (not shown). It is worth mentioning that the absolute values of the biases obtained with Eq. (6) are of similar order of magnitude to those obtained with other methods that use X_{max} or the number of muons as observables (see for instance Refs. [20,30–32]).

In order to introduce the method developed in this work, we first characterize the N_μ distributions corresponding to different primary particles. The produced number of muons increases with a small power of the mass number and almost linearly with the primary energy. This behavior can be explained in terms of the Heitler-Matthews model of hadronic air showers [33], which predicts that the mean value of the total number of muons produced in a shower is

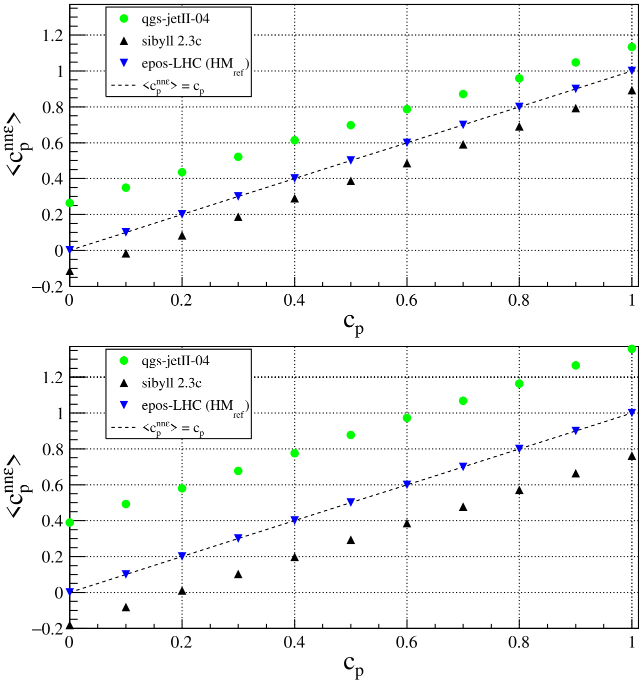

FIG. 2. Mean value of proton abundance estimator, $\langle c_{\mathrm{p}}^{\mathrm{nn}\varepsilon} \rangle$, as a function of c_{p} obtained from Eq. (6) for proton-iron (top) and proton-nitrogen (bottom) binary mixtures corresponding to the same cases of Fig. 1. The error bars are smaller than the marker size. The dash line corresponds to $\langle c_{\mathrm{p}}^{\mathrm{nn}\varepsilon} \rangle = c_{\mathrm{p}}$.

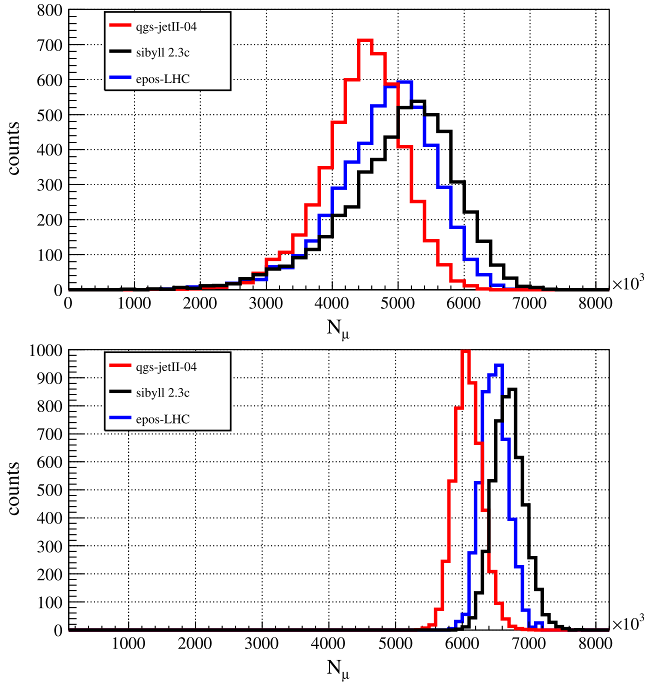

FIG. 3. N_μ histograms corresponding to the HEHIMs used for proton (top) and iron (bottom). The number of entries of each histogram is 5000.

$\langle N_\mu \rangle(E, A) = A[E/(A\xi_c)]^\beta$, where ξ_c is the critical energy at which charged pions decay into muons. This prediction relies on the superposition model in which a primary nucleus of mass number A and energy E is considered as the A independent nucleus of energy E/A each. The Heitler-Matthews model predicts quite well the functional dependence of the mean number of muons at the shower maximum development with A and E. Simulations with post-LHC HEHIMs show that $\beta \simeq 0.915$–0.928 [34,35]. Figure 3 shows the N_μ histograms obtained from simulations corresponding to different HEHIMs for proton (top) and iron (bottom). The mean values of N_μ between SIBYLL2.3c and EPOS-LHC are related by a factor ∼1.043 and ∼1.033 for proton and iron, respectively. For the case of QGSJetII-04 and EPOS-LHC, these factors are ∼0.930 and ∼0.939. Doing the same analysis with any primary A_i, one can state that the numerical factor between the N_μ mean values of two HEHIMs is almost independent of the primary type, with differences less than 1%. The same conclusions can be made for the numerical relationship between the N_μ standard deviations of two HEHIMs. Moreover, these conclusions are not affected when experimental uncertainties are taken into account since they are estimated, for any primary A_i, as proportional to the mean values of $N_\mu(A_i)$ [36,37]. Figure 4 shows the N_μ histograms for the same cases as Fig. 3 but with the N_μ values of SIBYLL2.3c and QGSJetII-04 divided by the aforementioned factors. As expected, a better match is seen between the histograms than those shown in Fig. 3.

The method presented in this work is based on one assumption:

(I) The number of muons corresponding to a HEHIM differs by a scale factor from that of another HEHIM.

According to this assumption, the random variables of two hadronic models HM and $\mathrm{HM_{ref}}$, corresponding to any primary A_i, will satisfy $N_\mu(\mathrm{HM}, A_i) = (1 + \gamma_o) \times N_\mu(\mathrm{HM_{ref}}, A_i)$, where $1 + \gamma_o$ is the scale factor, being $\gamma_o = \gamma_o(\mathrm{HM}, \mathrm{HM_{ref}}) > -1$. Note that from assumption (I) the value of β is the same for any HEHIM and, for a given primary, the shape of the N_μ distribution is the same regardless of the HEHIM considered.

After the training process, we now use as inputs the $N_\mu(A_0, \mathrm{HM_{ref}})$ and $N_\mu(A_1, \mathrm{HM_{ref}})$ values but multiplied by different quantities $1 + \gamma \geq 0$, in order to obtain the efficiencies $\varepsilon_{A_0}(\gamma)$ and $\varepsilon_{A_1}(\gamma)$ (see Sec. II B for more details). Note that $\varepsilon_{A_0}(\gamma = 0)$ and $\varepsilon_{A_1}(\gamma = 0)$ correspond to the efficiencies previously defined, i.e. when the $N_\mu(A_0)$ and $N_\mu(A_1)$ variables of $\mathrm{HM_{ref}}$ are used.

Figure 5 (top) shows the $\varepsilon_{\mathrm{p}}(\gamma)$ and $\varepsilon_{\mathrm{Fe}}(\gamma)$ values as functions of γ after training the neural network with the $N_\mu(\mathrm{p}, \mathrm{HM_{ref}})$ and $N_\mu(\mathrm{Fe}, \mathrm{HM_{ref}})$ ($A_0 = \mathrm{p}$, $A_1 = \mathrm{Fe}$). Note that as the γ value increases (decreases), $\varepsilon_{\mathrm{p}}(\gamma)$ decreases (increases) approaching 0 (1). In the opposite direction, $\varepsilon_{\mathrm{Fe}}(\gamma)$ increases (decreases) approaching 1 (0) as the values of γ increase (decrease). The same behavior is observed when the neural network is trained with other types of

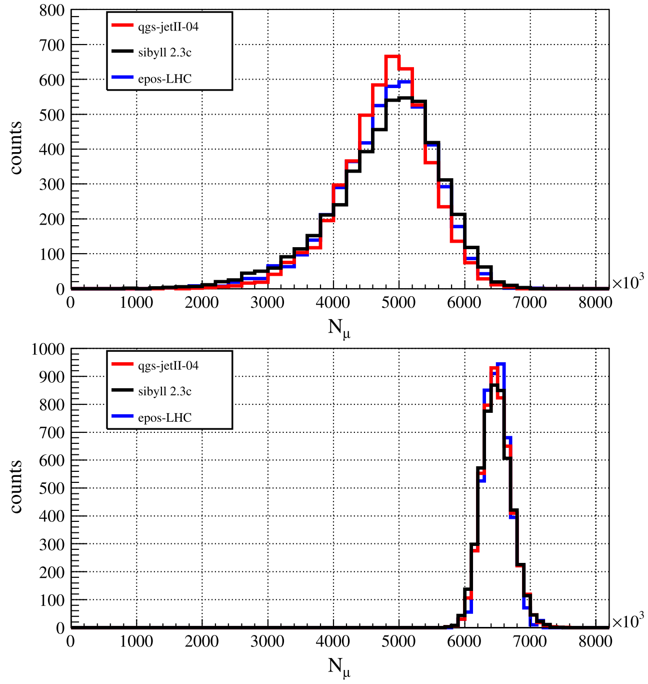

FIG. 4. N_μ histograms for the same cases of Fig. 3 but with the N_μ values of SIBYLL2.3c and QGSJetII-04 multiplied by the corresponding scaled factor (see text).

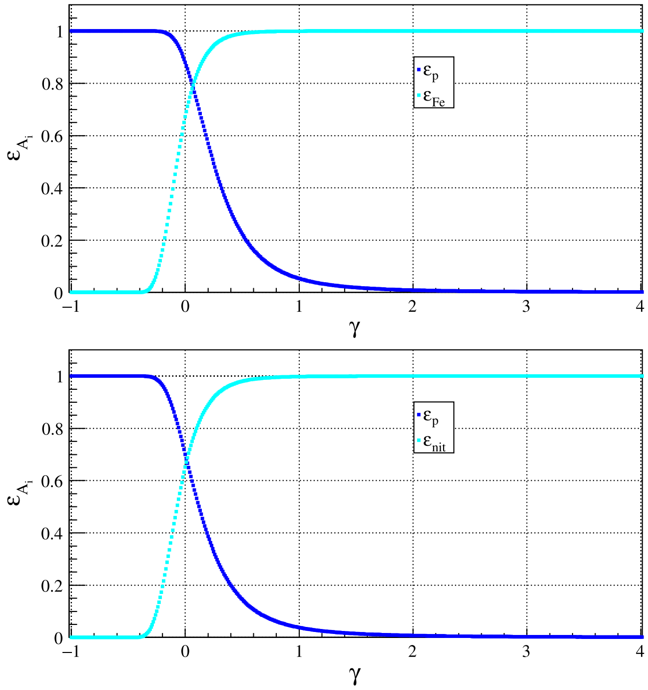

FIG. 5. Top: efficiency values, $\varepsilon_\mathrm{p}(\gamma)$ and $\varepsilon_\mathrm{Fe}(\gamma)$, as functions of γ after training the neural network with the $N_\mu(\mathrm{p},\mathrm{HM}_\mathrm{ref})$ and $N_\mu(\mathrm{Fe},\mathrm{HM}_\mathrm{ref})$ values corresponding to proton and iron ($A_0 = \mathrm{p}$, $A_1 = \mathrm{Fe}$). Bottom: $\varepsilon_\mathrm{p}(\gamma)$ and $\varepsilon_\mathrm{nit}(\gamma)$ values as functions of γ after training the neural network with the $N_\mu(\mathrm{p},\mathrm{HM}_\mathrm{ref})$ and $N_\mu(\mathrm{nit},\mathrm{HM}_\mathrm{ref})$ corresponding to proton and nitrogen ($A_0 = \mathrm{p}$, $A_1 = \mathrm{nit}$).

primary particles as can be seen from Fig. 5 (bottom) where $\varepsilon_\mathrm{p}(\gamma)$ and $\varepsilon_\mathrm{nit}(\gamma)$, corresponding to proton and nitrogen ($A_1 = \mathrm{nit}$) are shown.

In a similar way, we use as inputs the $N_\mu(c_{A_0},\mathrm{HM})$ values of the sample test mixture multiplied by different amounts $1+\gamma$, i.e. $(1+\gamma)N_\mu(c_{A_0},\mathrm{HM})$, obtaining the neural network estimator, $c^\mathrm{nn}_{A_0}(\mathrm{HM},\gamma)$ (see Sec. II B for more details). Figures 6–8 show, respectively, the mean values, $\langle c^\mathrm{nn}_\mathrm{p}\rangle(\mathrm{HM},\gamma)$ delivered by the neural network for EPOS-LHC, SIBYLL2.3c and QGSJetII-04 corresponding to proton-iron (top) and proton-nitrogen (bottom) mixtures for the same independent samples used in Fig. 1. The error bars are included but are smaller than the marker size in most of the cases. From Figs. 6–8 one can see that the $c^\mathrm{nn}_\mathrm{p}(\gamma)$'s have similar behaviors for different HEHIMs, increasing (decreasing) their values up to 1 (down to 0) as γ decreases (increases). By comparing the top and bottom panels of Figs. 6–8 one can also verify that the variations in the shapes of the $c^\mathrm{nn}_\mathrm{p}(\mathrm{HM},\gamma)$ curves and the separations between them for different values of c_p are more accentuated in the proton-iron mixture than in the proton-nitrogen mixture; this is due to the increase in the merit factor between the primary distributions of the former mixture.

According to assumption (I) and Eq. (5), a general expression for the $c^\mathrm{nn}_{A_0}(\mathrm{HM},\gamma)$ estimator provided by the neural network can be established,

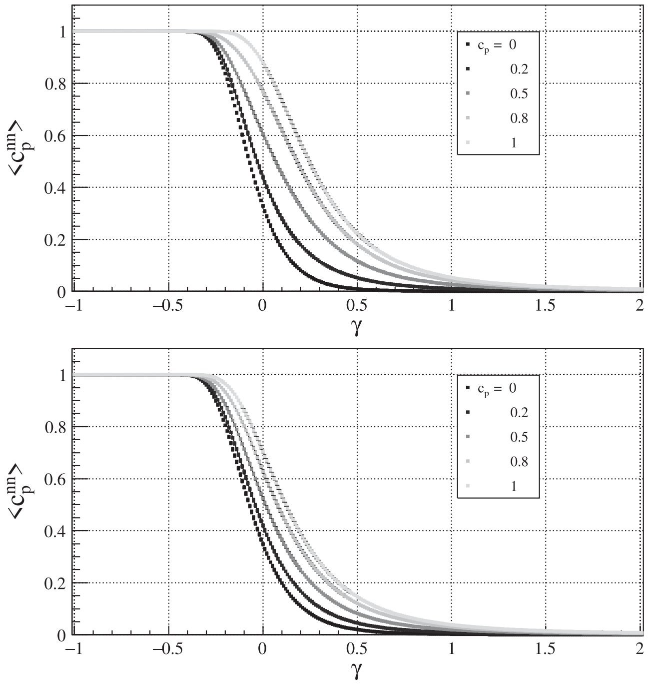

FIG. 6. Mean values, $\langle c^\mathrm{nn}_\mathrm{p}\rangle(\mathrm{HM},\gamma)$, as a function of γ provided by the neural network for different values of c_p corresponding to proton-iron (top) and proton-nitrogen (bottom) binary mixtures. The error bars are smaller than the marker size in most of the cases. The HEHIM used corresponds to HM = EPOS-LHC with the same mixture samples used in Fig. 1 (see Sec. II B for more details).

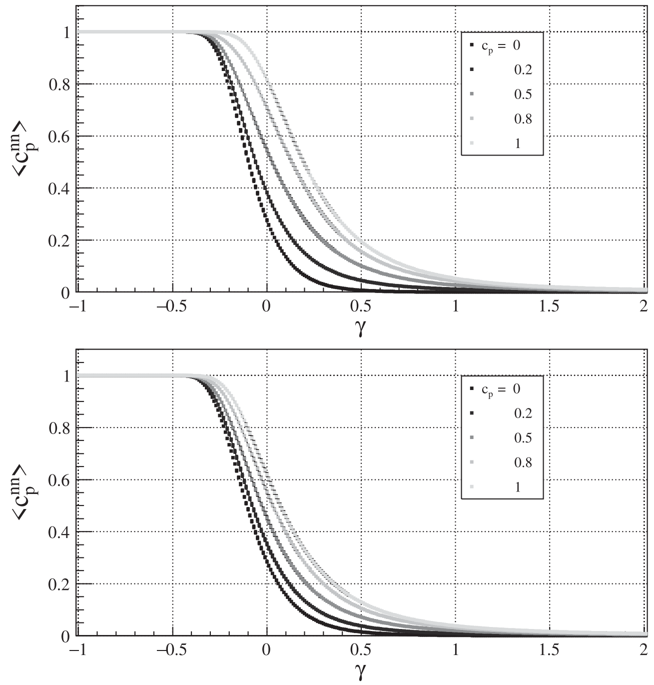

FIG. 7. Mean values, $\langle c_{\mathrm{p}}^{\mathrm{nn}}\rangle(\mathrm{HM},\gamma)$, as a function of γ provided by the neural network for different values of c_{p} corresponding to proton-iron (top) and proton-nitrogen (bottom) binary mixtures. The error bars are smaller than the marker size in most of the cases. The HEHIM used corresponds to HM = SIBYLL2.3c with the same mixture samples used in Fig. 1 (see Sec. II B for more details).

$$c_{A_0}^{\mathrm{nn}}(\mathrm{HM},\gamma)=F(c_{A_0},\gamma+\gamma_o), \tag{7}$$

where

$$F(c_{A_0},\gamma+\gamma_o)=-\varepsilon_{A_1}(\gamma+\gamma_o)+1 \\ +c_{A_0}[\varepsilon_{A_0}(\gamma+\gamma_o)+\varepsilon_{A_1}(\gamma+\gamma_o)-1]. \tag{8}$$

From the behavior of $c_{A_0}^{\mathrm{nn}}(\mathrm{HM},\gamma)$ observed in Figs. 6–8 and taking into account assumption (I), it can be shown that (see Appendix B) if L is a value of γ large enough to ensure that $c_{A_0}^{\mathrm{nn}}(\mathrm{HM},\gamma)\cong 0$ for any HEHIM, HM, then

$$\gamma_o(c_{A_0})=\int_{-1}^{L}[F(c_{A_0},\gamma)-c_{A_0}^{\mathrm{nn}}(\mathrm{HM},\gamma)]d\gamma, \tag{9}$$

where $F(c_{A_0},\gamma)$ corresponds to $F(c_{A_0},\gamma+\gamma_o)$ when $\gamma_o=0$. Therefore, from Eq. (9) one can express the $\gamma_o(c_{A_0})$ value as a function of c_{A_0}. Finally, the estimator, $\hat{c}_{A_0}$, is defined and obtained as the one that minimizes

$$S(c_{A_0})=S_1(c_{A_0})+S_2(c_{A_0}), \tag{10}$$

being

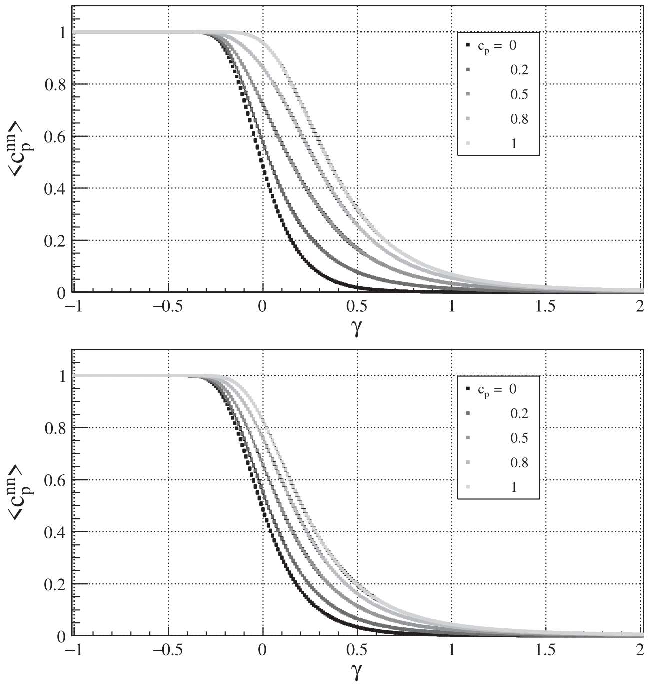

FIG. 8. Mean values, $\langle c_{\mathrm{p}}^{\mathrm{nn}}\rangle(\mathrm{HM},\gamma)$, as a function of γ provided by the neural network for different values of c_{p} corresponding to proton-iron (top) and proton-nitrogen (bottom) binary mixtures. The error bars are smaller than the marker size in most of the cases. The HEHIM used corresponds to HM = QGSJetII-04 with the same mixture samples used in Fig. 1 (see Sec. II B for more details).

$$S_1(c_{A_0})=\left|\sum_{k=1}^{n}[c_{A_0}^{\mathrm{nn}}(\mathrm{HM},\gamma_k)-F(c_{A_0},\gamma_k+\gamma_o(c_{A_0}))]\right| \tag{11}$$

and

$$S_2(c_{A_0})=\left(\frac{1}{n}\sum_{k=1}^{n}[c_{A_0}^{\mathrm{nn}}(\mathrm{HM},\gamma_k)-F(c_{A_0},\gamma_k+\gamma_o(c_{A_0}))]^2\right)^{1/2}, \tag{12}$$

where $\gamma_1=-1$ and $\gamma_n=L$.

Note that the method developed in this work allows us to obtain the estimator $\hat{\gamma}_o=\hat{\gamma}_o(\mathrm{HM},\mathrm{HM}_{\mathrm{ref}})$ associated with the relation between the mean values of $N_\mu(A_i)$ corresponding to HM and $\mathrm{HM}_{\mathrm{ref}}$. Note that once $\hat{c}_{A_0}$ is found, the corresponding estimator $\hat{\gamma}_o$ is obtained by replacing in Eq. (9) c_{A_0} with $\hat{c}_{A_0}$ and then, solving for γ_o. We define the percentage bias $\delta N_\mu^{A_i}[\%]$ as

$$\delta N_\mu^{A_i}=\left[\frac{\langle\hat{N}_\mu\rangle(A_i,\mathrm{HM})}{\langle N_\mu\rangle(A_i,\mathrm{HM})}-1\right]100\%, \tag{13}$$

where the estimator of $N_\mu(A_i,\mathrm{HM})$ is given by $\hat{N}_\mu(A_i,\mathrm{HM})=(1+\hat{\gamma}_o)\langle N_\mu\rangle(A_i,\mathrm{HM}_{\mathrm{ref}})$.

B. Numerical approach

In order to obtain the N_μ distributions considered in this work, different EAS simulations were performed. The shower library used is generated with CONEX2r6.40 [38–40]. The HEHIMs considered are EPOS-LHC [21], SIBYLL2.3c [22] and QGSJetII-04 [23]. EPOS-LHC is used as the reference model ($\mathrm{HM_{ref}}$). The showers are simulated for primary energy $E = 10^{18}$ eV. For each HEHIM, a set of 5000 EASs are generated for proton, helium, nitrogen and iron. From each set of EASs obtained, a histogram of each $N_\mu(A_i)$ distribution is constructed. N_μ distributions of proton ($A_0 = \mathrm{p}$), helium ($A_1 = \mathrm{He}$), nitrogen ($A_1 = \mathrm{nit}$) and iron ($A_1 = \mathrm{Fe}$) are considered. Only muons with energies greater than 1 GeV are taken into account. The ground level is set at the sea level (CONEX2r6.40 default ground level); however, the surface detector energy resolution of Auger is used (see text below).

The simulations for the training process are performed by using the ROOT package [41]. In this work a multilayer perceptron (mlp) neural network is used. The TMultiLayerPerceptron ROOT class with one neuron in the input/output layers and 20 neurons for each of the two hidden layers considered is used. The output function is sigmoid type, where an output $= 0$ (1) is considered if the sigmoid output is ≤ 0.5 (>0.5). The mlp is subjected to the training process so that it provides an output $= 0$ or 1 when we use as input random values of N_μ generated from the histogram of $\mathrm{HM_{ref}}$ corresponding to A_0 or A_1, respectively. The number of these random $N_\mu(A_i, \mathrm{HM_{ref}})$ values generated for each type of primary A_i during the learning process is equal to 10^6.

To obtain the efficiencies $\varepsilon_{A_i}(\gamma)$, new $N_\mu(A_i, \mathrm{HM_{ref}})$ values for each primary A_i were generated. In order to include experimental uncertainties each generated value is then randomly fluctuated. For this, it is assumed that measurements of primary energy and of the number of muons follow Gaussian distributions centered at the true value, with widths given by the detector resolution. An average relative energy resolution of 13.1% is considered. This value corresponds to the surface detector's energy resolution of Auger for $E = 10^{18}$ eV [42], which treats the energy bin independently and also takes account of the effect of the migration of events between bins [36]. The relative errors of the reconstructed number of muons are extracted from Ref. [37] which correspond to $\simeq 12.5$ and 9.5% for proton and iron, respectively. For the case of helium and nitrogen, the same relative error value as for proton is used in order to be conservative. Each value is then multiplied by a certain amount $1 + \gamma$. The $(1 + \gamma) N_\mu(A_i, \mathrm{HM_{ref}})$ obtained is used as input of the mlp. This process is repeated 10^6 times for each primary type A_i and each value of γ used. The corresponding efficiencies $\varepsilon_{A_i}(\gamma)$ are finally calculated as the quotient between the obtained number of output = i and the total number of $(1 + \gamma) N_\mu(A_i, \mathrm{HM_{ref}})$, i.e. 10^6, used as input of the mlp. The process described in this paragraph to obtain the $\varepsilon_{A_i}(\gamma)$ is done for γ values ranging from -1 to 10 in steps of 0.01 as can be partially seen in Fig. 5.

The simulations performed for the estimator of the proton abundance obtained by the mlp, $c_\mathrm{p}^\mathrm{nn}(\mathrm{HM}, \gamma)$, are also performed by using the ROOT package. The values of c_p considered range between 0.0 and 1.0 in steps of 0.1. For each value of c_p, the number of events, N_p, due to proton-induced air showers is obtained by sampling a binomial distribution function, $B(N, c_\mathrm{p})$, where $N = 5000$ is the total number of events. The number of events due to the other primary A_1 (iron, nitrogen or helium) is calculated as $N_{A_1} = N - N_\mathrm{p}$. In order to analyze the impact of the differences between HEHIMs on composition analyses, EPOS-LHC, SIBYLL2.3c and QGSJetII-04 are used to generate the event samples from their $N_\mu(A_i)$ histograms and including also experimental uncertainties in the measurement of N_μ as described before. The values of the sample $N_\mu(c_\mathrm{p}, \mathrm{HM})$ are multiplied by a certain amount $1 + \gamma$ and used as inputs of the mlp. The estimator $c_\mathrm{p}^\mathrm{nn}(\mathrm{HM}, \gamma)$ obtained by the mlp is finally calculated as the ratio between the number of outputs $= 0$ and the total number of outputs (i.e. N) corresponding to each sample. The $c_\mathrm{p}^\mathrm{nn}(\mathrm{HM}, \gamma)$ is computed for γ values ranging from -1 to 20 in steps of 0.01. For each value of c_p, 100 independent samples (i.e. the number of iterations) are generated to obtain the distribution of $c_\mathrm{p}^\mathrm{nn}(\mathrm{HM}, \gamma)$ corresponding to EPOS-LHC, SIBYLL2.3c and QGSJetII-04. The mean value of the $c_\mathrm{p}^\mathrm{nn}(\mathrm{HM}, \gamma)$, $\langle c_\mathrm{p}^\mathrm{nn} \rangle(\mathrm{HM}, \gamma)$, and its standard deviation are also calculated as can be seen in Figs. 6–8. Once the $\varepsilon_{A_0}(\gamma)$, $\varepsilon_{A_1}(\gamma)$ and $c_\mathrm{p}^\mathrm{nn}(\mathrm{HM}, \gamma)$ functions are obtained, $\hat{c}_\mathrm{p}$ are $\hat{\gamma}_0$ are computed as described in Sec. II A. For the integral of Eq. (9) a limit $L = 20$ is used. The values of c_p used to minimize Eq. (10) range from -0.03 to 1.03.

III. RESULTS AND DISCUSSION

A. Results of the $\hat{c}_\mathrm{p}$ estimator

The top and middle panels of Fig. 9 show $\langle \hat{c}_\mathrm{p} \rangle$ as a function of c_p for the proton-iron and proton-nitrogen mixtures, respectively. One can see that the method does not produce biases when $\mathrm{HM} = \mathrm{HM_{ref}}$. For $\mathrm{HM} \neq \mathrm{HM_{ref}}$, the absolute values of the biases are less than 0.05 in most cases, showing a marked improvement compared to the results shown in Fig. 2 obtained with Eq. (6). The cost to pay with respect to Eq. (6) is an increase in the error bars in some cases. The bottom panel of Fig. 9 shows $\hat{c}_\mathrm{p}$ as a function of c_p for proton-helium mixture samples. One can see again the absence of biases for the $\mathrm{HM} = \mathrm{HM_{ref}}$ case. From Fig. 9 it can be surprisingly seen that the magnitudes of the biases for different types of mixtures are smaller when the merit factor corresponding to their primary components decreases. This means, for example, that the

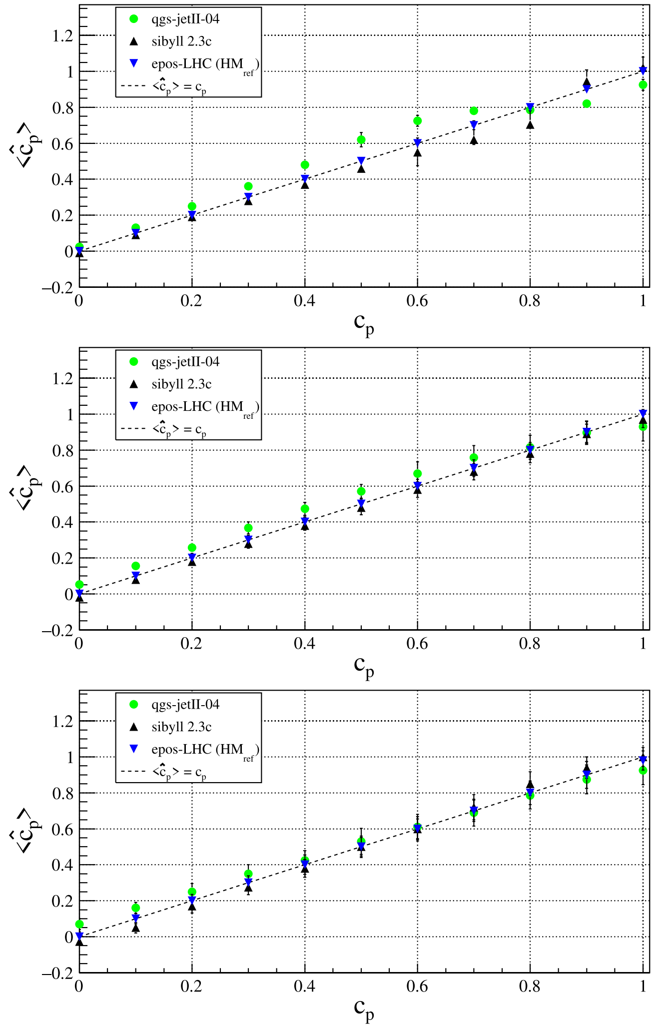

FIG. 9. Mean value, $\langle\hat{c}_\text{p}\rangle$, of the proton abundance estimator as a function of c_p computed with the method presented in this work for proton-iron (top), proton-nitrogen (middle) and proton-helium (bottom) mixture samples. The dash line corresponds to $\langle\hat{c}_\text{p}\rangle = c_\text{p}$.

improvement obtained for the proton-helium mixtures is greater than that obtained for the proton-iron mixtures. It is concluded that the method presented in this work is slightly dependent on the HEHIM used for composition determination, implying an improvement compared to classical methods where the X_max or the number of muons are used as the physical observable (see for instance Refs. [20,30]). This improvement is quantified in the following subsection.

B. Results for the estimator of $\langle N_\mu\rangle$

As mentioned in Sec. II A, the method presented in this work not only provides the estimators of the primary abundances but also the estimators of the $\langle N_\mu\rangle(A_i)$ values. Figures 10–12 show the biases δN_μ^p (top) and $\delta N_\mu^{A_1}$

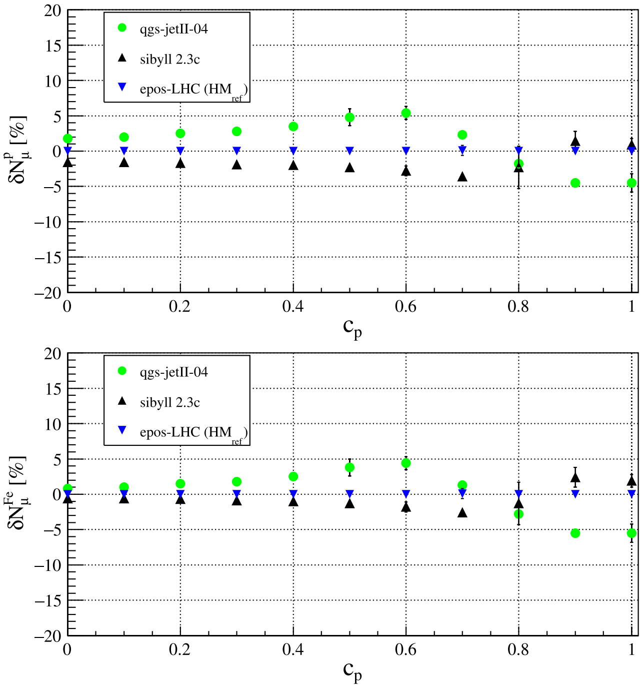

FIG. 10. Biases δN_μ^p (top) and δN_μ^Fe (bottom) as functions of c_p obtained from Eq. (13) corresponding to proton-iron mixtures for the three HEHIMs used in this work.

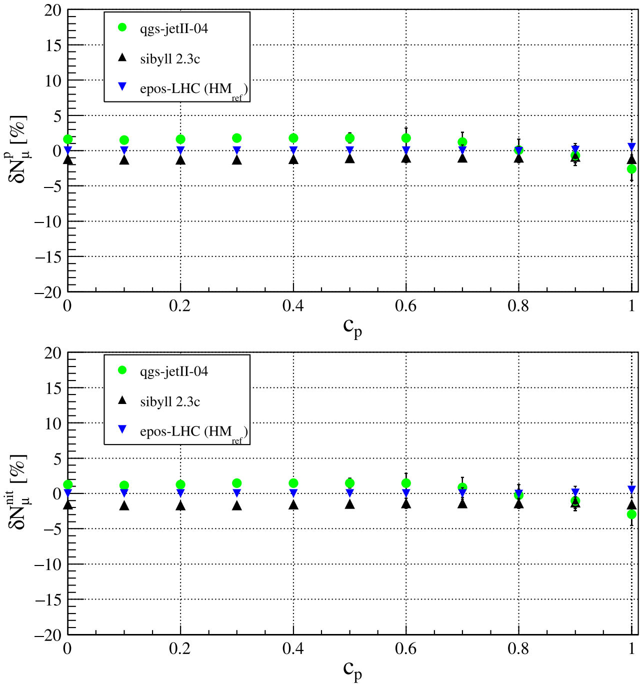

FIG. 11. Biases δN_μ^p (top) and δN_μ^nit (bottom) as functions of c_p obtained from Eq. (13) corresponding to proton-nitrogen mixtures for the three HEHIMs used in this work.

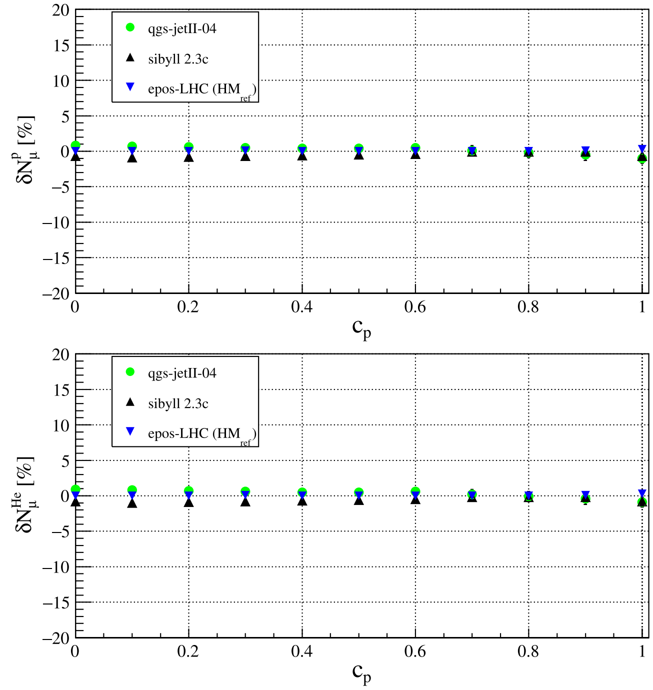

FIG. 12. Biases $\delta N_{\mu}^{\mathrm{p}}$ (top) and $\delta N_{\mu}^{\mathrm{He}}$ (bottom) as functions of c_{p} obtained from Eq. (13) corresponding to proton-helium mixtures for the three HEHIMs used in this work.

(bottom) as functions of c_{p} corresponding to proton-iron, proton-nitrogen and proton-helium mixtures, respectively. According to the definition of $\delta N_{\mu}^{A_i}$ given in Eq. (13), the differences between the estimated and the true number of muons are $\lesssim 2\%$ in most cases, being negligible when $\mathrm{HM} = \mathrm{HM}_{\mathrm{ref}}$. It can also be seen that the results obtained for the proton-nitrogen mixtures are better than those obtained for the proton-iron mixtures. And the results of the proton-helium mixtures are in turn better than those of the proton-nitrogen mixtures, with differences between the estimated and the true number of muons smaller than 1%. It is worth mentioning that the implication of the biases shown in this subsection on the primary abundance shown in Fig. 9 is in good agreement with the results obtained in Ref. [20] where a study of the proton abundance biases as a function of the relative differences between the number of muons was done—also for binary mixture samples—by using a maximum likelihood method. One could therefore conclude that the percentage biases in the number of muons obtained with the presented method are $\lesssim 2\%$ or, in other words, the primary abundance biases obtained in this work are comparable to the biases that would be obtained with classical methods (for example, maximum likelihood) in a hypothetical situation in which the differences between the number of muons of different HEHIMs were $\lesssim 2\%$.

C. Discussion

As commented in Sec. I, machine learning has been used in the physical sciences, including particle and astroparticle physics. The large amount of events collected by the experiments of these fields has allowed the exploitation of machine learning techniques that require vast amounts of data to build and train a model. As an example, in Ref. [27] it was presented a method to reconstruct X_{max} from the signal traces induced in the water-Cherenkov detectors. This method is based on a deep neural network, whose architecture and optimization was specifically developed for the measurement conditions at the Pierre Auger Observatory. In this case each signal trace is characterized by many machine-learned features. The network parameters for forming these observables to characterize the signal traces and for combining all available information are adjusted in a training procedure. In a similar way, the work of Ref. [28] presented a machine learning method to extract the muon signal with the time integral of the muon trace. Generalizing this idea, in Ref. [29] a deep learning algorithm was used for the estimation of the muon contribution to the signal recorded in each time trace bin.

In this work, instead of developing a suitable architecture for the machine learning process, we have presented a method that allows us to extract information about the composition of UHECR with a simple classifier. Here, a multilayer perceptron neural network was used. The method is based on one assumption, called assumption (I), that consists of linking the N_{μ} values of two HEHIMs by a single scale factor, γ_o, regardless of the primary particle in question. We have shown that the method has a low dependency on the HEHIM used as an experimental sample despite the poor performance of the neural network observed in Fig. 1. This was possible thanks to the degree of accuracy of assumption (I) discussed and verified in Sec. II A. Hence, the method has also the ability of obtaining automatically the estimator of the mean value of N_{μ} corresponding to any HEHIM. It should be note that Eq. (9) constitutes an important tool since it simplifies the computational architecture—and reduces computation time—expressing γ_o as a function of c_{A_0}, disregarding a sweep on the γ_o variable. It is worth mentioning that the approach presented in this work uses machine learning but instead of the event-by-event technique a finite sample of N_{μ} is finally needed, just like with classical statistical methods. As can be seen from Eqs. (9)–(12) and from the technique described in Sec. II A to obtain the functions $c_{A_0}^{\mathrm{nn}}(\mathrm{HM}, \gamma)$, the method only needs a relatively low mathematical level since the main mathematical power lies in the optimization associated with the learning and training processes of the neural network.

The method has achieved satisfactory results despite using only one physical observable, N_{μ}. The hybrid feature of the Pierre Auger Observatory could improve the performance of the classifier by using as inputs more than one type of variable, as for example, the signal traces induced in the water-Cherenkov detectors. Furthermore, this type of variable which is very close to the detector output is particularly appropriate to exploit the capabilities of the neural networks. However, exploring cosmic-ray composition results with these multidimensional input spaces would need a suitable and complex architecture adapted for the measurement conditions at the Pierre Auger Observatory, not to mention the need to introduce new assumptions, similar to assumption (I), so that the results are less dependent on the HEHIM used as experimental sample.

Although the method should be extended to mixtures of more than two types of particles, the binary case could be applied in regions of the cosmic-ray energy spectrum where the abundances of a third and fourth type of primary appear to be negligible or null [30]. It is worth mentioning that different mass composition methods that seem to have a reduced dependence on the assumed HEHIM have been developed in recent years. In Ref. [43], a method based on parametrizations of the $X_{\max}$ distributions obtained from simulations is presented in which the normalization levels of the mean value and the standard deviation of $X_{\max}$ are determined from experimental data. In this way, the influence of HEHIMs on composition analyses is reduced. In Ref. [44], a method based on the correlation between $X_{\max}$ and the number of muons in air showers is described. The purpose of this method is to determine whether the mass composition is pure or mixed. A similar method is used by Auger to study the composition in the ankle region [45]. The results obtained in this case are robust with respect to the details of the hadronic interactions and to experimental systematic uncertainties. Therefore, it is expected that a combination of the method presented in this work with the other method based on the $X_{\max}$ observable can be developed in order to further reduce the dependence of the composition analyses on HEHIMs. Following the same goal, a promising alternative would be the development of a composition method, similar to the one presented here, but using as input the observable $X_{\max}$ and some assumption "analogous to assumption (I)" that links the $X_{\max}$ of different HEHIMs. Once the robustness of the method has been verified, the next step would be its generalization by using both observables, N_μ and $X_{\max}$, with the corresponding assumptions (probably correlated) in order to obtain even less hadronic-model-dependent results than those obtained in the present work. These topics are under current research.

IV. CONCLUSIONS

In this work, we have presented a method to perform mass composition analyses based on the N_μ observable. The method processes the information provided by a simple neural network (or from any type of classification method) to obtain the abundance estimators of different primary particles, providing at the same time, the estimators of the mean values of their N_μ distributions. It is worth mentioning that the approach presented only needs a relatively low mathematical level and instead of using the event-by-event technique, a finite sample of N_μ is needed, just like with classical statistical methods.

We have also studied the impact of the use of different high-energy hadronic interaction models in the composition analyses. The biases found are smaller than those obtained with classic statistical methods that use a single physical observable.

ACKNOWLEDGMENTS

The authors thank Armando di Matteo and Roger Clay for a careful reading of the manuscript and to the members of the Pierre Auger Collaboration for useful discussions.

APPENDIX A: MERIT FACTOR

The discrimination power of a given mass sensitive parameter, q, can be assessed by the commonly used merit factor, which is defined as

$$\mathrm{MF}(q) = \frac{\langle q\rangle_{A_1} - \langle q\rangle_{A_0}}{\sqrt{\mathrm{Var}[q]_{A_1} + \mathrm{Var}[q]_{A_0}}}, \tag{A1}$$

where $\mathrm{Var}[q]_A$ is the variance of parameter q for the primary type A. By using N_μ as the q parameter in Eq. (A1) for $E = 10^{18}$ eV one can verify from simulations that the merit factors between proton and iron are ≃2.2, 1.9 and 2.4 for EPOS-LHC, SIBYLL2.3c and QGSJetII-04, respectively. For proton and nitrogen (proton and helium) the merit factors decrease to ≃1.3, 1.1 and 1.4 (≃0.58, 0.54 and 0.64) for the respective HEHIMs mentioned.

APPENDIX B: PROOF OF EQ. (9)

In this appendix the proof of Eq. (9) is given under the assumption (I). Figure 5 can be seen as a representative example of how the efficiencies $\varepsilon_{A_0}(\gamma)$ and $\varepsilon_{A_1}(\gamma)$ behave as functions of γ. Let L be any value of γ large enough to ensure that for any HEHIM, HM, it is true that

$$\varepsilon_{A_0}(L) \cong \varepsilon_{A_0}(L+\gamma_o) \cong 0, \tag{B1}$$

and

$$\varepsilon_{A_1}(L) \cong \varepsilon_{A_1}(L+\gamma_o) \cong 1. \tag{B2}$$

Then, it can be easily seen that

$$\int_{-1}^{L} \varepsilon_{A_0}(\gamma+\gamma_o)d\gamma = \int_{-1}^{L} \varepsilon_{A_0}(\gamma)d\gamma - \gamma_o \tag{B3}$$

and

$$\int_{-1}^{L} \varepsilon_{A_1}(\gamma+\gamma_o)d\gamma = \int_{-1}^{L} \varepsilon_{A_1}(\gamma)d\gamma + \gamma_o. \tag{B4}$$

As stated in Eqs. (7) and (8) from assumption (I), the $c_{A_0}^{\mathrm{nn}}(\mathrm{HM},\gamma)$ estimator can be expressed as

$$c_{A_0}^{\mathrm{nn}}(\mathrm{HM},\gamma) = -\varepsilon_{A_1}(\gamma+\gamma_o) + 1 + c_{A_0}[\varepsilon_{A_0}(\gamma+\gamma_o) + \varepsilon_{A_1}(\gamma+\gamma_o) - 1]. \tag{B5}$$

Integrating Eq. (B5) from −1 to L, replacing then with Eqs. (B3) and (B4) and solving for γ_o, we finally have that

$$\gamma_o(c_{A_0}) = \int_{-1}^{L}[F(c_{A_0},\gamma) - c_{A_0}^{\mathrm{nn}}(\mathrm{HM},\gamma)]d\gamma, \tag{B6}$$

where

$$F(c_{A_0},\gamma) = -\varepsilon_{A_1}(\gamma) + 1 + c_{A_0}[\varepsilon_{A_0}(\gamma) + \varepsilon_{A_1}(\gamma) - 1]. \tag{B7}$$

[1] G. Kulikov and G. Khristiansen, J. Exp. Theor. Phys. **35**, 635 (1958).
[2] M. Aglietta *et al.*, Astropart. Phys. **20**, 641 (2004).
[3] T. Antoni *et al.*, Astropart. Phys. **24**, 1 (2005).
[4] W. Apel *et al.*, Astropart. Phys. **31**, 86 (2009).
[5] M. Amenomori *et al.*, Adv. Space Res. **47**, 629 (2011).
[6] W. Apel *et al.*, Phys. Rev. Lett. **107**, 171104 (2011).
[7] F. Fenu *et al.*, Proc. Sci., ICRC2017 (**2017**) 486, https://pos.sissa.it/301/486.
[8] D. Ivanov *et al.*, Proc. Sci., ICRC2017 (**2017**) 498, https://pos.sissa.it/301/498.
[9] A. Aab *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **798**, 172 (2015).
[10] M. Fukushima *et al.*, Prog. Theor. Phys. Suppl. **151**, 206 (2003).
[11] M. Risse *et al.*, Mod. Phys. Lett. A **22**, 749766 (2007).
[12] J. Abraham *et al.*, Astropart. Phys. **27**, 155168 (2007).
[13] J. Abraham *et al.*, Phys. Rev. Lett. **104**, 091101 (2010).
[14] G. Medina-Tanco *et al.*, in *Proceedings of the 30th ICRC, Merida, Mexico* (2007), p. 1101, http://particle.astro.ru.nl/pub/30ICRC5-1101.pdf.
[15] A. Supanitsky, A. Cobos, and A. Etchegoyen, Phys. Rev. D **98**, 103016 (2018).
[16] K. Kampert, Braz. J. Phys. **43**, 375 (2013).
[17] A. Supanitsky, Galaxies **10**, 75 (2022).
[18] K. Kampert and M. Unger, Astropart. Phys. **35**, 660 (2012).
[19] A. Supanitsky, A. Etchegoyen, G. Medinatanco, I. Allekotte, M. Gomezberisso, and M. Medina, Astropart. Phys. **29**, 461 (2008).
[20] A. Cobos, A. D. Supanitsky, and A. Etchegoyen, Phys. Rev. D **102**, 083022 (2020).
[21] T. Pierog, Iu. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner, Phys. Rev. C **92**, 034906 (2015).
[22] F. Riehn, H. P. Dembinski, R. Engel, A. Fedynitch, T. K. Gaisser, and T. Stanev, Proc. Sci., ICRC2017 (**2017**) 301 [arXiv:1709.07227].
[23] S. Ostapchenko, EPJ Web Conf. **52**, 02001 (2013).
[24] G. Carleo *et al.*, Rev. Mod. Phys. C **91**, 045002 (2019).
[25] A living review of machine learning for particle physics, https://iml-wg.github.io/HEPML-LivingReview/.
[26] A. Supanitsky, G. Medina-Tanco, and A. Etchegoyen, Astropart. Phys. **31**, 116 (2009).
[27] A. Aab *et al.* (Pierre Auger Collaboration), J. Instrum. **16**, P07019 (2021).
[28] A. Guillen, A. Bueno, J. M. Carceller, J. C. Martínez-Velázquez, G. Rubio, C. J. Todero Peixoto, and P. Sanchez-Lucas, Astropart. Phys. **111**, 12 (2019).
[29] A. Aab *et al.* (Pierre Auger Collaboration), J. Instrum. **16**, P07016 (2021).
[30] J. Bellido, Proc. Sci., ICRC2017 (**2017**) 506.
[31] P. Abreu *et al.* (Pierre Auger Collaboration), J. Cosmol. Astropart. Phys. 02 (2013) 026.
[32] A. Aab *et al.*, Phys. Rev. D **12**, 122006 (2014).
[33] J. Matthews, Astropart. Phys. **22**, 387 (2005).
[34] T. Pierog, Proc. Sci., ICRC2017 (**2017**) 1100.
[35] R. Prado, R. Conceição, M. Pimenta, and V. de Souza, Astropart. Phys. **83**, 40 (2016).
[36] A. Aab *et al.*, Phys. Rev. D **91**, 032003 (2015); **91**, 059901(E) (2015).
[37] A. Supanitsky, A. Etchegoyen, D. Melo, and F. Sanchez, Astropart. Phys. **68**, 7 (2015).
[38] T. Bergmann, R. Engel, D. Heck, N. Kalmykov, S. Ostapchenko, T. Pierog, T. Thouw, and K. Werner, Astropart. Phys. **26**, 420 (2007).
[39] T. Pierog *et al.*, Nucl. Phys. B, Proc. Suppl. **151**, 159 (2006).
[40] G. Bossard, H. J. Drescher, N. N. Kalmykov, S. Ostapchenko, A. I. Pavlov, T. Pierog, E. A. Vishnevskaya, and K. Werner, Phys. Rev. D **63**, 054030 (2001).
[41] https://root.cern.ch.
[42] A. Coleman, Proc. Sci., ICRC2019 (**2019**) 225.
[43] S. Blaess, J. A. Bellido, and B. R. Dawson, arXiv:1803.02520v1.
[44] P. Younk and M. Risse, Astropart. Phys. D **35**, 807 (2012).
[45] A. Aab *et al.*, Phys. Lett. B **762**, 288 (2016).