Hadronic Interactions Studies at the LHC

Tanguy Pierog¹-∗, Matias Perlin²-³, and Klaus Werner⁴

¹KIT, IAP, Postfach 3640, D-76021 Karlsruhe, Germany
²Instituto de Tecnologías en Detección y Astropartículas (CNEA, CONICET, UNSAM), Buenos Aires, Argentina
³Departamento de Física, FCEN, Universidad de Buenos Aires and IFIBA, CONICET, Buenos Aires, Argentina
⁴SUBATECH, University of Nantes – IN2P3/CNRS – IMT Atlantique, Nantes, France

Abstract. The deficit of muons in the simulation of extensive air showers is a long-standing problem and the origin of large uncertainties in the reconstruction of the mass of the high-energy primary cosmic rays. Hadronic interaction models re-tuned after early LHC data have a more consistent description of the muon content among them but still disagree with the data. Furthermore, some inconsistencies are also visible in the electromagnetic component. Ten years after the first LHC-tuned model release, much more detailed data are available both from LHC, SPS and hybrid air shower measurements, allowing for the understanding of some deficiencies in the model which could lead to a change in both $X_{\text{max}}$ and the muon production by air showers. A better treatment of the core-corona approach according to LHC data has important consequences on the muon production while an update of diffraction and nuclear fragmentation is changing the $X_{\text{max}}$ distribution. An updated version of EPOS LHC will be presented addressing the main issues of this model.

1 Introduction

Despite all the efforts made to take into account the first results of proton-proton collisions at the LHC in hadronic interaction models used for air-shower simulations, the observed number of muons, their height of production, or even the depth of shower maximum are still not reproduced consistently by the models [1]. Furthermore, the differences among model predictions introduce uncertainties in cosmic-ray data analysis, which are currently smaller than in the past but still exceed the experimental uncertainties in certain cases [2]. Nevertheless before claiming the need for “new physics”, it is important to guarantee that all the standard QCD physics is properly taken into account in these models. For that, it is necessary to go beyond the simplest observables which are usually used to test them. After ten years of running, the various LHC experiments provided a large amount of complex data to analyze and understand, in particular, thanks to the correlations between different observables, which are not yet fully investigated.

Among the hadronic interaction models used for air-shower analysis, only Eros LHC [3–6] includes all the features needed to have a detailed description of the correlation between various observables [1]. Indeed, the core-corona approach in this model, which allows the production of a collective hadronization phase, appears to be a key element to reproduce LHC data. Before LHC, it was usually accepted that hydrodynamical phase expansion, for instance due to the formation of a quark-gluon plasma, was possible only in central heavy-ion collisions. Proton-nucleus (pA) collisions were then used as a reference to probe the effect of such collective behavior (final state) but with some nuclear effect at the initial state level, while proton-proton (pp) interactions were free of any nuclear effects. With the LHC operated in pp, pPb and PbPb mode, it is now possible to compare high-multiplicity pp or pPb events with low-multiplicity PbPb events (which correspond to the same number of particles measured at mid-rapidity) and surprisingly, the very same phenomena are observed [7, 8] concerning the soft-particle production.

One of the most striking features observed in all systems is the long-range two-particle correlation and the evolution of the particle flow as described in Ref. [9]. In Ref. [10], the authors demonstrate how these data from the CMS Collaboration can be reproduced and explained using an approach combining standard perturbative calculations for initial conditions and hydrodynamical calculations for the final state interactions.

At the same time, the recent results compiled by the Working group on Hadronic Interactions and Shower Physics (WHISP) [11] clearly indicate that the discrepancy between the muon production in simulations and data gradually increases with energy. It is a strong indication of a different hadronization than the one used in the current hadronic models [12–17], including Eros LHC, which does not have enough core contribution according to data [7] published after the release of the model in 2012.

Furthermore, other studies showed some deficiencies of Eros LHC and other models not only for the number of muons but also for $X_{\text{max}}$ measurements [18]. Since this is the most important observable for the mass composition of cosmic rays, one should check the model against the
latest LHC data in terms of cross-section and multiplicity distributions which are key factors for the shower development [19]. This will be done in Section 2.

In Ref. [20], we presented a modified version of Eros LHC [4] based on Eros3 [3] to study the consequence of the extended range of collective hadronization on air-shower physics in this particular model with positive results on the number of muons. To have a more general approach, new results based on modified secondary particle spectra following the core-corona approach applicable to any model are now presented before a full implementation in a new version of the model. In Section 3, we will discuss the impact of collective hadronization on the total number of muons produced by air showers. In Section 4 the basic principles the core-corona approach will be presented and we will discuss the changes made in CONEX [21] to take this modifications into account and check the results.

Finally, in Section 5, we will discuss future developments.

2 Latest LHC data and $X_{\text{max}}$

Using the simple Heitler-Matthews model [22] or studying Monte-Carlo simulations [19], it is well known that the inelastic cross-section, the elasticity and the multiplicity of hadronic interactions, together with the mass and energy of the primary cosmic rays are the most important factors which determine the position of the maximum air shower development $X_{\text{max}}$. For Eros LHC, this was fixed using the early data of different LHC experiments at 7 TeV center-of-mass energy. Since then, new and more precise data have been released at 13 TeV or using lead as a projectile giving new constraints to the models which should be taken into account to give proper $X_{\text{max}}$ predictions. A new preliminary version of the model, called Eros LHC-R, is presented here.

2.1 pp cross-sections

The various pp cross-sections (inelastic, elastic and total) are key ingredients to tune the model parameters. As described in [4], the amplitude of the elementary scattering in Eros LHC can be used directly to compute these cross-sections via the profile function. Really recently, the ATLAS Collaboration, using its ALFA detector, published a very precise measurement at 13 TeV center-of-mass energy. Since then, new and more precise data have been released at 13 TeV or using lead as a projectile giving new constraints to the models which should be taken into account to give proper $X_{\text{max}}$ predictions. A new preliminary version of the model, called Eros LHC-R, is presented here.

2.2 Multiplicity and elasticity in pA

The elasticity is important for the shower development but is complicated to measure at LHC. There is no direct measurement, but some other distributions could be sensitive to the energy fraction of the leading particle. The average multiplicity of pp and pPb interactions, another important measurement for $X_{\text{max}}$, is well described by the hadronic models. In the meantime, the fluctuations of the multiplicity in pPb collision have been published and is, in fact, not so well reproduced by the two models which can run with lead projectile Eros LHC and QGSJetII.04. In fact, the latter is linked to the elasticity because the frequency of low multiplicity events is related to highly elastic events while the high multiplicity events are linked to low elasticity. So, a better description of the fluctuations of the multiplicity should improve the elasticity prediction.

In Fig. 2, the predictions for the multiplicity distribution of Pb-p interactions at 5.02 TeV from Eros LHC-R (dotted line), Eros LHC (full line) and QGSJetII.04 (dashed line) are compared to CMS data [30]. On the left-hand side, a linear scale for the y-axis is used to focus on the low multiplicity (high elasticity) events. Eros LHC clearly underestimates the very low multiplicity peak (related to diffraction, so with high elasticity), while QGSJetII.04 underestimates the intermediate range. On the right-hand side, a logarithmic scale is used to observe the tail at high multiplicity where the multiplicity is underestimated by Eros LHC and overestimated by QGSJetII.04. To correct Eros LHC, it is not possible to simply change some parameters. The lack of fluctuations at low and high multiplicity is due to the usage of fixed screening parameters which can reproduce the average behavior correctly but underestimate the natural fluctuations that are due to different configurations of elementary interaction reinteractions which are not taken into account explicitly. A contrario, in QGSJetII.04 these Pomeron-
Pomeron interactions are taken into account but neglecting the energy conservation at the amplitude level, leading to the overestimation of the high multiplicity events. In Eros LHC-R, a distribution of screening parameters has been introduced to take into account these fluctuations in an effective way with only one parameter to be tuned to these data, leading to a wider range of fluctuations. One important consequence is that more events have a low multiplicity because less nucleons are interacting leading to a larger elasticity.

2.3 Nuclear fragmentation

As stated in [31], the fluctuations of \( X_{\text{max}} \) for nuclear primaries are underestimated in Eros LHC because of an error in the generation of wounded nucleons which were produced first by the model itself, which is based on the list of nucleons, and then again by the subroutine creating the nuclear fragment from the list of the remaining nucleons, which includes some evaporation which was not necessary. Hence, the produced fragments are too light with too many free nucleons. This is now corrected in Eros LHC-R but the RMS of \( X_{\text{max}} \) for iron showers is increased by around 15%.

2.4 Hadronic interactions in Air and \( X_{\text{max}} \)

In order to understand the impact of these modifications on the \( X_{\text{max}} \) predictions by Eros LHC-R we can first check the p-air cross-section and elasticity. In Fig. 3, theinelastic p-air cross-section is shown on the left-hand side and the elasticity is on the right-hand side. The lines are for Eros LHC-R (dotted line), Eros LHC (full line), QGSJetII.04 (dashed line), and Sibyll 2.3d (dash-dotted line), and the points are data from [25]. As expected, Eros LHC-R has a cross-section about 10% lower than Eros LHC at high energy and an elasticity which is increased by around 15%.

These are not huge differences, and as shown in Fig. 4, it is only a 2% increase of \( X_{\text{max}} \) for proton-induced showers. Nevertheless this represents already around 15 g/cm\(^2\) which was the difference between Eros LHC (line with full stars) and Sibyll 2.3d (line with full triangles). So, in fact, Eros LHC-R (line with open circles) predicts around the same \( X_{\text{max}} \) values than Sibyll 2.3d, much deeper than the predictions from QGSJetII.04 (line with open squares).

In terms of mass composition, it means that the mean logarithmic mass deduced from \( X_{\text{max}} \) measurements using Eros LHC-R will also be around 15% larger, hence reducing the muon deficit in simulation compared to data. It will probably be very similar to Sibyll 2.3d, but these are still preliminary results.

3 Collective hadronization and air-shower physics

The dominant mechanism for the production of muons in air showers is via the decay of light charged mesons. The vast majority of mesons are produced at the end of the hadron cascade, after typically five to ten generations of hadronic interactions (depending on the energy and zenith angle of the cosmic ray). The energy carried by neutral pions, however, is directly fed into the electromagnetic shower component and is not available for further production of more mesons and subsequently muons. Thus, the energy carried by hadrons other than neutral pions is typically available to produce more hadrons and ultimately
Figure 3. Inelastic cross-section (left-hand side) and elasticity (energy fraction of the leading particle) (right-hand side) for p-air interactions as a function of center of mass energy. Simulations are done with EPOS LHC-R (dotted line), EPOS LHC (full line), QGSJetII-04 (dashed line), and SIBYLL 2.3d (dash-dotted line). Points are data from [25].

Figure 4. \(\langle X_{\text{max}} \rangle\) for proton- and iron-induced showers as a function of the primary energy. Predictions of different high-energy hadronic interaction models are presented with full lines for proton and dashed lines for iron with full triangles for SIBYLL 2.3d, open circles for EPOS LHC-R, open squares for QGSJetII.04, full stars for EPOS LHC. Refs. to the data can be found in [32] and [33].

muons in following interactions and decays. As explained in Ref. [12, 34], the ratio of the average electromagnetic to average hadronic energy, called \(R\), and its dependence on center-of-mass energy, are thus related to the muon abundance in air showers: if this energy ratio is smaller (larger), more (less) energy is available for the production of muons at the end of the hadronic cascade and ultimately more (fewer) muons are produced. In a simplified model where all particles have the same energy, we can define \(R = c/(1 - c)\) where \(c = N_{\pi^0}/N_{\text{all}}\) and \(N_{\text{all}}\) is the total multiplicity. Then it can even be demonstrated using the simple Heitler-Matthews model [22] that the exponent \(\beta\) of the energy dependence of the muon production is directly related to \(c\) and thus \(R\) as

\[
\beta = 1 + \frac{\ln(1 - c)}{\ln(N_{\text{all}})}.
\]

Since in a collective hadronization (or statistical model) the production of particles with higher mass (in particular, strange quarks) [35] is not suppressed as in a string hadronization, the fraction of secondary pions in the dense core is reduced because many other more massive hadrons and resonances are produced. This leads to
a lower ratio of the electromagnetic to hadronic energy density in particles produced from the core. As a consequence, such reduction of $R$ due to more collectivity in secondary particle production should increase the slope $\beta$ of the energy dependence of the muon production in air showers when compared to traditional hadronic models where only the string hadronization is taken into account.

### 4 Core-corona effect

The discussion in the previous section suggests that a change in $R$ is a potential way to reduce the discrepancy between measurements and air shower simulations. Nevertheless, $R$ is quite well constrained by theory as well as laboratory measurements and, thus, cannot be changed entirely arbitrarily as studied in another analysis [19]. In a naive model like Ref. [22], where only pions are considered as secondary particles, $R \approx 0.5$. In a more realistic approach based on string fragmentation, we have $R \approx 0.41$ to 0.45.

![Figure 5. Particle to pion ratio for the $\Omega$ baryon versus multiplicity at mid-rapidity, for different contributions (core (dash-dotted), corona (dotted), core+corona (dashed) and all (core+corona+hadronic gas) (full)) from the Eros simulations, for different systems (pp (thin), pPb (normal), PbPb (bold)). We also plot ALICE data from [7].](image)

But as shown in Ref. [7], particle ratios such as $K/\pi$, $p/\pi$ or $\Lambda/\pi$ change with increasing secondary-particle density, saturating to the value given by a thermal/statistical model with a freeze-out temperature of 156.5 MeV [36] yielding $R \approx 0.34$. Such a behavior can be explained in terms of a core-corona picture [35]. This approach has been used in the framework of realistic simulations [37], but also in simple model calculations [38–41]. The basic idea is that some fraction of the volume of an event (or even a fraction of events) behaves as a QGP and decays according to statistical hadronization (core), whereas the other part produces particles via string fragmentation (corona). The particle yield $N_i$ for particle species $i$ is then a sum of two contributions

$$N_i = \omega_{\text{core}} N_i^{\text{core}} + (1 - \omega_{\text{core}}) N_i^{\text{corona}},$$

where $N_i^{\text{core}}$ represents statistical (grand canonical) particle production, and $N_i^{\text{corona}}$ is the yield from string decay. Crucial is the core weight $\omega_{\text{core}}$. In order to explain LHC data [7], the weight $\omega_{\text{core}}$ needs to increase monotonically with the multiplicity, starting from zero for low-multiplicity pp scattering, up to 0.5 or more for very high multiplicity pp, reaching unity for central heavy-ion collisions (PbPb), as illustrated in Fig. 5 using the Eros3 model [3]. This process was already present in Eros LHC but in a very early stage without data to constrain the evolution of the core and corona ratio leading to a very small effect on the muon production. The development done for Eros3 can be used to improve Eros LHC but this requires non trivial changes which are not yet available. In this paper, we will still present the results of a more general study.

In order to know whether such a constrained value of $R$ could be low enough to increase the number of muons in air-shower simulations such that the data could be reproduced, a simplified core-corona approach can be used to at least set some realistic upper limit under the following assumptions:

- the fraction of core effectively increases with energy: the core-corona approach is originally a function of the multiplicity and independent of the collision energy for a given multiplicity. Since the average multiplicity increases with the energy, the average core fraction must increase with energy.
- only the change in hadronization is taken into account: collective effects in the core in principle include particle correlations and flow, but since the longitudinal particle

![Figure 6. Evolution of the ratio $R$ from pure corona to pure core in the case of a linear increase of core fraction as a function of the logarithm of the energy for three different models in proton-air interactions at mid-rapidity.](image)
momentum is dominating over the transverse momentum down to relatively low energy, these effects can be neglected in first order.

- no nuclear effect is introduced: the multiplicity increases with the mass of the projectile, so the core fraction would increase for higher primary mass. This is not taken into account due to technical limitations (minimizing the effect for higher primary mass but only for the first few hadronic generations).
- core-corona effect is applied on full phase space: core hadronization has been observed at mid-rapidity, but in order to see the maximal effect on air shower to set an upper limit, the modification should apply at larger rapidities.

So, in the following, we are going to employ a straightforward core-corona approach, based on Eq. (2), for any hadronic interaction model in CONEX air-shower simulations. The particle yield from the chosen interaction model is by definition considered to be the corona yield, whereas we use the standard statistical hadronization (also referred to as resonance gas) for the core part. So $\omega_{\text{core}} = 0$ would be the “normal” simulation with the default interaction model. Choosing $\omega_{\text{core}} > 0$, amounts to mixing the yields from the interaction model according to the core-corona superposition shown in Eq. (2) and changes the value of $R$ depicted in Fig. 6 for proton-air interactions for the three different hadronic interaction models Eros LHC, QGSJett99 and Sibyll 2.3d.

Technically, we directly modify individual particle ratios of the secondary particle spectra $dN_i/dE_j$ for particle species $i$ and energy bins $dE_j$, of hadronic interactions with air nuclei used by CONEX for numerical air-shower simulations based on cascade equations. Knowing the initial ratios $N^0/K^0, p/\pi^-, K^+/\pi^+, p/n, K^0/S^0$ (taking into account strange baryon decays) from a corona-type model and the value of the same ratios from the core model, we compute new spectra in which the particle yields include both, core and corona according to $\omega_{\text{core}}$. Since the hadronization mechanism can affect only newly produced particles, the properties of the leading particle should be preserved. To achieve that, the new particle yields are computed for all secondaries, but excluding the one corresponding to the respective projectile type, i.e. protons in proton-air, kaons in kaon-air interactions, and so on. The yield of the projectile-type particles is determined subsequently by exploiting energy conservation in all energy bins $dE_j$ summed over all secondary particle species $i$: the sum $\sum_i E_i dN_i/dE_j$ must be conserved. Since at high $x_p = E_i/E_{\text{lab}}$ only the projectile-type particles will have $dN_i/dE_j$ significantly different from zero (aka leading-particle effect), the resulting modified leading-particle type spectra at high $x_p$ follow the original distribution, and are only affected by the scaling procedure at lower values of $x_p$. Together, this assures that energy conservation as well as the total multiplicity are not affected, but only the particle ratios. More details will be given in a future publication.

We expect the core weight $\omega_{\text{core}}$ to increase with energy in a logarithmic way like the total multiplicity. Thus,

$$\omega_{\text{core}}(E_{\text{lab}}) = f_s F(E_{\text{lab}}; E_{\text{th}}, E_{\text{scale}})$$

with

$$F(E_{\text{lab}}; E_{\text{th}}, E_{\text{scale}}) = \frac{\log_{10}(E_{\text{lab}}/E_{\text{th}})}{\log_{10}(E_{\text{scale}}/E_{\text{th}})}$$

for $E_{\text{lab}} > E_{\text{th}}$.

To model this, starting already at fixed-target energies, $E_{\text{th}} = 100$ GeV. Different energy dependencies are explored by changing $E_{\text{scale}}$ from 100 GeV (corresponding to a step function), to $10^6$ GeV, and $10^{10}$ GeV. The $f_s$ scale is varied from 0.25, 0.5, 0.75 to 1.0; in addition, we enforce $F(E_{\text{lab}}; E_{\text{th}}, E_{\text{scale}}) = 1$ for all $E_{\text{lab}} \geq E_{\text{scale}}$. This yields the $\omega_{\text{core}}$ energy dependencies as depicted in Fig. 7.

All these scenarios have been used to simulate full air showers with CONEX, using cascade equations from the first interaction to the ground, for proton and iron primary particles at $E_0 = 10^{19}$ eV. In Fig. 8, the results are shown in the $X_{\text{max}}$-ln$N_p$ plane for QGSJett99.04. This typical example illustrates that it is well possible with modified hadronization in air-shower cascades to describe the data of the Pierre Auger Observatory. As expected, more core-like contributions are needed compared to what is currently provided by the models. This means, QGP-like effects also in light colliding systems and starting in central collisions at much lower center-of-mass energies may play a decisive role. It is worth noticing that if $X_{\text{max}}$ is shifted to the right in this plot, a smaller value of core fraction is needed to reduce the difference between the predicted and the measured number of muons.

5 Summary

The better description of the collective hadronization and, in particular, the fact that the core is produced earlier than predicted by Eros LHC can have very important consequences for the muon production in air showers. The effect of QGP using the standard Eros LHC was shown not to be significant. Indeed in this model, the QGP was produced only for very high-multiplicity events and at mid-rapidity.
which are both rare and not so important for air-shower development. Other studies using a QGP or alternative hadronization as a possible new source of muons were all based on changes under extreme conditions too [13, 43] or with extreme consequences not observed at the LHC [14]. As shown here and in Ref. [17, 44, 45], according to the most recent LHC results, the collective hadronization happens at a much lower multiplicity and, consequence, with effects at larger rapidities (lower particle densities than foreseen). In that case, the fact that much more particles coming from the hadronization of a QGP play a significant role in the air-shower development. A stronger effect could be observed in case of a nuclear projectile that could create a collective phase with a non-zero chemical potential, which could lead to an even stronger increase of the non-electromagnetic secondary particles [16]. The combination of a mild increase like observed in this study with a proton primary with an even stronger effect for a heavier projectile associated with a deeper \( X_{\text{max}} \) (thus a larger mean mass) could be the complete solution of the muon-deficit problem in air-shower simulations. The new Eros LHC-R taking into account the core-corona effect in a detailed enough way to reproduce all major effects observed at the LHC is still under development, but the preliminary results on \( X_{\text{max}} \) using up-to-date cross-section and multiplicity distributions have been presented and lead indeed to larger values of \( X_{\text{max}} \).

References


Figure 8. Comparison of different core-corona mixing scenarios, as described in the text, on air shower simulations at \( 10^{19} \) eV QGSJetII.04 in the \( X_{\text{max}} - \ln N_p \) plane. The solid lines represent changing the scale \( f_{\omega} \), while the dashed lines also indicate the effect of changing \( E_{\text{scale}} \). The default model corresponds to the corona-only simulations. The data point is from the Pierre Auger Observatory [42]. Each model line represents all values that can be obtained for any mixture of cosmic nuclei from proton (bottom right) to iron (top left).