

**Core-corona effect in hadron collisions and muon production in air showers**Sebastian Baur<sup>1</sup>, Hans Dembinski<sup>2,3</sup>, Matias Perlin<sup>1,4,5</sup>, Tanguy Pierog<sup>1,\*</sup>, Ralf Ulrich<sup>1</sup>, and Klaus Werner<sup>6</sup><sup>1</sup>*Institute for Astroparticle Physics, Karlsruhe Institute of Technology, Karlsruhe, Germany*<sup>2</sup>*Max Planck Institute for Nuclear Physics, Heidelberg, Germany*<sup>3</sup>*Department of Physics, TU Dortmund, Dortmund, Germany*<sup>4</sup>*Instituto de Tecnologías en Detección y Astropartículas (CNEA, CONICET, UNSAM), Buenos Aires, Argentina*<sup>5</sup>*Departamento de Física, FCEN, Universidad de Buenos Aires and IFIBA, CONICET, Buenos Aires, Argentina*<sup>6</sup>*SUBATECH, Nantes University—IN2P3/CNRS—IMT Atlantique, Nantes, France*

(Received 26 February 2019; revised 9 November 2022; accepted 21 April 2023; published 24 May 2023)

It is well known that the fraction of energy in a hadron collision going into electromagnetic particles (electrons and photons, including those from decays) has a large impact on the number of muons produced in air shower cascades. Recent measurements at the LHC confirm features that can be linked to a mixture of different underlying particle production mechanisms such as a collective statistical hadronization (core) in addition to the expected string fragmentation (corona). Since the two mechanisms have a different electromagnetic energy fraction, we present a possible connection between statistical hadronization in hadron collisions and muon production in air showers. Using a novel approach, we demonstrate that the core-corona effect as observed at the LHC can have a significant impact and should be properly taken into account before trying to find a more exotic solution for the lack of muon production in simulations of high energy cosmic rays.

DOI: [10.1103/PhysRevD.107.094031](https://doi.org/10.1103/PhysRevD.107.094031)**I. INTRODUCTION**

Cosmic ray particles reach Earth from galactic and extragalactic sources with enormous energies and produce huge particle cascades in the atmosphere. The resulting extensive air showers are measured with the aim to unveil the astrophysical nature and origin of high energy cosmic rays. The Pierre Auger Observatory [1,2] and the Telescope Array [3,3] are the largest contemporary experiments targeting the most energetic cosmic rays with energies beyond  $10^{18}$  eV.

Of particular interest is the cosmic ray mass composition, which may be proton to iron nuclei and is expected to carry a unique imprint of the physics at the sources. The mass composition as a function of the cosmic ray energy  $E_0$  is inferred from air shower observables, of which the most important ones are the depth of the shower maximum  $X_{\max}$  and the number of muons  $N_\mu$  [4]. The depth  $X_{\max}$  is the integrated matter density column that a shower traversed

until the maximum number of charged particles in the shower is reached. The number of muons is obtained by counting muons when the shower arrives at the ground. Experimentally the muon counting is limited to a radial range around the shower axis as well as to a minimal energy of muons.

To infer the cosmic ray mass composition from these observables, accurate predictions from air shower simulations are needed for cosmic rays with various primary masses. However, the Pierre Auger Observatory [5,6] and the Telescope Array [7] observed that the measured number of muons in air showers drastically exceeds expectations from model predictions at shower energies around and above  $10^{19}$  eV. A recent summary of muon measurements [8,9] shows that a consistent muon excess compared to simulation is seen by the majority of cosmic ray experiments over a very wide energy range. The discrepancy between results based on  $X_{\max}$  and  $N_\mu$  is currently preventing an unambiguous interpretation of air shower data in terms of mass composition.

The amount of energy ending up in electromagnetic particles in hadron collisions

$$R = \frac{E_{\text{em}}}{E_{\text{had}}}, \quad (1)$$

\*tanguy.pierog@kit.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.

where  $E_{\text{cm}}$  is the summed energy over all  $\gamma$  (mostly from  $\pi^0$  decay) as well as  $e^\pm$ , and  $E_{\text{had}}$  the summed energy of all hadrons, is one of the crucial parameters driving muon production in extensive air showers [10–12]. It is closely related to the way an excited partonic system hadronizes. In hadronic interaction models used to simulate air showers, the hadronization is mainly done using a string fragmentation model that was successfully developed to describe the hadron production in  $e^+e^-$  collisions, and low energy proton-proton collisions. In systems with higher energy densities, such as heavy ion collisions, a statistical hadronization of a fluid is expected where the production of heavy particles is favored, thus reducing the fraction of  $\pi^0$  compared to other types of particles and therefore giving a lower value of  $R$ . Since the 1980s, collective effects in the hadronic final state, such as flow [13–15] or strangeness enhancement [16–20], have been observed in heavy ion collisions (often referred to as large systems). Later on, similar effects have been predicted [21–26] for proton-proton collisions (also known as small systems) and were eventually discovered at the LHC [27] (see Refs. [28,29] for detailed reviews).

While a fluidlike behavior (referred to as collective effects in the following) is confirmed in both large and small systems, their origin is still unclear. In large systems the formation of a quark-gluon-plasma (QGP) is commonly assumed as a phase of parton matter where confinement is no longer required [30–32] and in particular since the Relativistic Heavy Ion Collider (RHIC) results [33–36]. This QGP will evolve according to the laws of hydrodynamics and eventually decay statistically. There are various expected consequences of such a scenario, such as long-range two-particle correlations, the so-called ridge phenomenon [27,37], jet quenching [38,39], flow [40], or enhanced production of strange hadrons [20]. It was initially a surprise when such effects were also discovered in small systems. While it was argued that also in central collisions of small systems the energy densities may be high enough to allow for the formation of a QGP [21] and that the formation of small QGP droplets is plausible [41,42], other recent studies have shown that collective effects can be achieved by alternative mechanisms such as microscopic effects in string fragmentation [43] or QCD interference [44]. The possibility to have both string fragmentation hadronization and statistical hadronization at the same time but in a different area of the collision, depending on the energy or particle densities, has been introduced in [45] to explain the peculiar evolution of collective effects as a function of collision centrality in heavy ion collisions. This is the so-called core-corona separation.

In this paper, we study the impact of collective effects (more precisely: statistical hadronization) in smaller systems than heavy ions and the resulting modified hadronization scheme in high-energy interactions on air showers. Air shower cascades are driven by collisions of hadrons and

light nuclei at ultrahigh energies. We show that the effect of statistical hadronization in collisions of hadrons and nuclei may have been underestimated so far in the understanding of muon production in air showers [46–49]. Statistical hadronization affects the energy fraction contained in electromagnetic versus hadronic particles,  $R$ , which has implications for the muon production in cosmic ray air showers.

In Sec. II, using a direct modification of hadronic properties, we first identify the most important quantity to modify the muon production in air showers, which appears to be  $R$ . Then, in Sec. III, the origin and behavior of the values of  $R$  are studied in different hadronic interaction models at the LHC. We introduce in Sec. IV a simplified version of the core-corona model in the CONEX framework to qualitatively predict the effect of a constrained change of  $R$  values. Finally, in Sec. V, we present the result on the muon production when statistical hadronization is present. Our goal is to demonstrate that the observed discrepancy between current air shower simulations with measurements can be potentially resolved with collective hadronization in agreement with existing collider data, and that further experimental and theoretical studies in this direction are justified [46,50]. The development of a full model able to reproduce all data from accelerator and cosmic experiments is a long process currently ongoing [51–53].

## II. THE MUON PROBLEM AND THE $R$ OBSERVABLE

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 to the electromagnetic shower component and is not available for further production of more mesons and subsequently muons. The energy carried by hadrons that are not neutral pions is, on the other hand, able to produce more hadrons and ultimately muons in following interactions and decays. Using a simple Heitler-type toy model [54,55] to describe air showers, the neutral pion fraction  $c = N_{\pi^0}/N_{\text{mult}}$ , defined as the number of neutral pions  $N_{\pi^0}$  divided by the total number of final-state particles  $N_{\text{mult}}$  (full phase space, all types of stable hadrons) in a collision, was found to have a strong impact on the muon number and in particular on the slope of the energy dependence of the muon production. Indeed in this model we get

$$N_\mu = \left( \frac{E_0}{E_{\text{dec}}} \right)^\beta \quad \text{with} \quad \beta = 1 + \frac{\ln(1-c)}{\ln N_{\text{mult}}}, \quad (2)$$

where  $E_0$  is the energy of the primary cosmic ray particle and  $E_{\text{dec}}$  is the typical energy at which mesons decay in the cascade. So the muon number  $N_\mu$  increases with

decreasing  $c$ , which is understandable since more hadrons are available to produce muons. A second quantity with an important impact on the muon number was identified to be the hadron multiplicity  $N_{\text{mult}}$ .

The value of  $c$  is very important for the muon production. Unfortunately, it is difficult to measure both  $N_{\pi^0}$  and  $N_{\text{mult}}$  experimentally (for example at the LHC) since neutral particles cannot be easily counted individually. In general, secondary particle identification is unavailable at large pseudorapidities  $\eta$  where the energy flow is large enough to become relevant for the air shower development. Hence, it is useful to explicitly study the ratio of the electromagnetic to the hadronic energy density  $R$  given by

$$R(\eta) = \frac{\langle dE_{\text{em}}/d\eta \rangle}{\langle dE_{\text{had}}/d\eta \rangle}, \quad (3)$$

which is sensitive to properties of the hadronization as shown in the next section. Here the energy densities  $\langle dE/d\eta \rangle$  are obtained by summing the energy of all final-state particles except for neutrinos in bins of  $\eta$  and averaging over a large number of collisions.

The neutral pion fraction  $c$  can be easily related to the energy ratio  $R$ , since both have very similar kinematic aspects of final state distributions. If all particles have the same energy such as in the generalized Heitler model, then we have simply  $R = c/(1 - c)$ . But  $R$  is experimentally much easier to measure, since, using a calorimeter, the signals deposited by electromagnetic particles and by hadrons are characteristically different. We compute a detailed conversion between  $R$  and  $c$  using the standalone EPOS LHC hadronic interaction model [56] to simulate fixed energy proton-proton collisions at various center-of-mass energies, and found that, for the relevant parameter range, a change of  $R$  by  $\Delta R$  affects  $c$  by  $\Delta c \approx 0.8 \cdot \Delta R$ , where  $R$  is computed by integrating Eq. (3) over all  $\eta$ . In Sec. III, we will study  $R$  for different models as a function of  $\eta$ , and at fixed  $\eta$  as a function of the charged particle density at central pseudorapidity  $dN_{\text{ch}}/d\eta|_{\eta=0}$ , which is determined as the average multiplicity within  $|\eta| < 0.5$ .

To distinguish the effect of  $R$  and the hadron multiplicity  $N_{\text{mult}}$  on the air shower properties, we will first use them as simple effective parameters, instead of hadronic observables, and study the correlation between the number of muons ( $\ln N_{\mu}$ ) and the depth of the shower maximum ( $X_{\text{max}}$ ). The influence on the main air shower observables (like  $X_{\text{max}}$  and  $\ln N_{\mu}$ ) of various effective parameters (like  $R$ ,  $c$ , or  $N_{\text{mult}}$ , named generically  $q$  in the following) in interaction models was investigated in a previous study [10] in which the behavior of hadronic interaction models in air shower simulations was modified during full air shower Monte Carlo simulations within the CONEX [57] framework. These changes were implemented such that the parameters  $q(E_{\text{lab}})$  of the hadronic event generators were

modified in an energy-dependent way by a simple correction function as

$$q(E_{\text{lab}}) \rightarrow q(E_{\text{lab}}) \times (1 + f_q F(E_{\text{lab}}; E_{\text{th}}, E_{\text{scale}})), \quad (4)$$

using the modification scale  $f_q$  and the energy-dependent factor

$$F(E_{\text{lab}}; E_{\text{th}}, E_{\text{scale}}) = \frac{\ln(E_{\text{lab}}/E_{\text{th}})}{\ln(E_{\text{scale}}/E_{\text{th}})} \quad \text{for } E_{\text{lab}} > E_{\text{th}}, \quad (5)$$

and representing the assumption that models are well constrained by accelerator data at lower energies than the energy threshold  $E_{\text{th}}$ , where  $F(E_{\text{lab}}) = 0$ , while they become logarithmically unconstrained going to higher energies. The parameter  $E_{\text{scale}}$  is the reference energy scale. Here, we will use  $E_{\text{scale}} = E_{\text{LHC}}^{\text{CR}} \approx s_{\text{LHC}}/(2m_p) \approx 90$  PeV, using an LHC center-of-mass energy of  $\sqrt{s_{\text{LHC}}} = 13$  TeV, which means that the factors  $f_q$  directly correspond to the prospective impact at LHC energies. Typical threshold values are  $E_{\text{th}} \approx s_{\text{Tevatron}}/(2m_p) \approx 1$  PeV, using the center-of-mass energy of the Tevatron accelerator that was the experiment with the largest energy at the time of the first study (highest energy to tune the hadronic interaction models). However, particle production in the important forward phase space is not well constrained by either Tevatron nor LHC data, allowing much lower values of  $E_{\text{th}}$  to be explored. It is a key point of the application of Eq. (5) inside CONEX that a significant fraction of the air shower cascade is consistently modified during full simulations. The individual particles produced after each high energy hadronic interactions are modified such that a global parameter  $q$  is changed according to Eq. (4), conserving the other ones and the total energy of each event. For instance, the multiplicity ( $N_{\text{mult}}$ ) of one event can be increased without changing the particle ratio ( $c$ ) and the elasticity (fraction of the total energy going to the projectile remnant), by duplicating particles of each particle type excluding the leading particle. All the technical details can be found in Appendix A of [10].

In the original study [10], each air shower observables were studied individually for a change of different parameters  $q$ . But here, we apply Eqs. (4), (5) to explore the impact of  $q = R$  or  $q = N_{\text{mult}}$  on  $X_{\text{max}}$  and  $\ln N_{\mu}$  simultaneously in full air shower simulations. The resulting correlated effect is shown in Fig. 1 as demonstrated for air showers at  $E_0 = 10^{19}$  eV using EPOS LHC in CONEX. Lines in this figure show all possible resulting mean values of  $X_{\text{max}}$  and  $\ln N_{\mu}$  for any mass composition of cosmic rays between pure proton (bottom right end of lines) and pure iron (top left end of lines). The resulting values of  $X_{\text{max}}$  and  $\ln N_{\mu}$  are located on a straight line because the mean values for both are linear functions of the mean-logarithmic mass of cosmic rays [58,59] given a fixed air shower energy. The line shape is universal, but its location and to a lesser degree

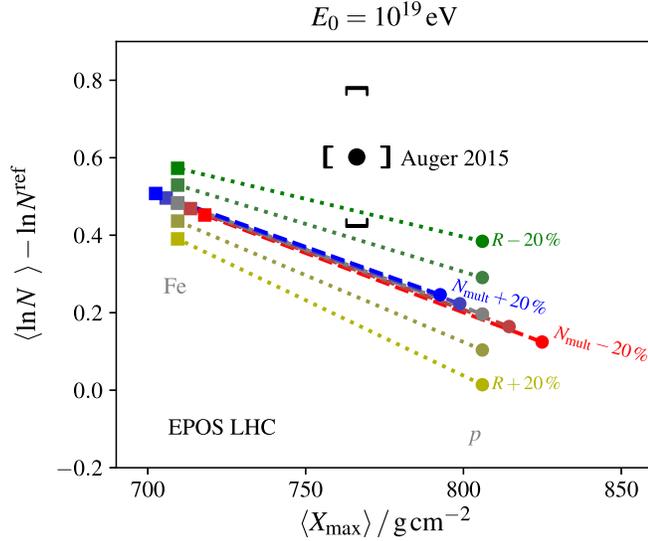


FIG. 1. Impact of the modification scales  $f_q$  (in %) of the hadron multiplicity  $N_{\text{mult}}$  (dashed lines) and the energy ratio  $R$  (dotted lines) in collisions at the LHC energy of  $\sqrt{s} = 13$  TeV on EPOS LHC predictions of the air shower observables  $X_{\text{max}}$  and  $\ln N_{\mu}$  in  $10^{19}$  eV air showers. The datum is from the Pierre Auger Observatory [5]. The model lines represent all values that can be obtained for any mixture of cosmic nuclei from proton (bottom right) to iron (top left).

the slope and length depend on the hadronic interaction model. Current hadronic interaction models predict lines that are too low compared to experimental data from air showers, as indicated by the vertical gap between the representative data point from the Pierre Auger Observatory [5] and the EPOS LHC line. This discrepancy is the expression of the muon problem outlined above.

When  $N_{\text{mult}}$  is modified the simulated line shifts along itself: the multiplicity has a correlated effect on  $X_{\text{max}}$  and  $\ln N_{\mu}$  that cannot close the gap to the data. However, modifications of  $R$  mainly affect the muon number and leave  $X_{\text{max}}$  unchanged, creating vertical shifts and tilts of the line in the plot. Thus, within the assumptions outlined here, we find that a decrease of  $R$  by  $f_q = -15\%$  at the LHC energy of  $\sqrt{s} = 13$  TeV would be sufficient to make the simulations compatible with the air shower data at  $10^{19}$  eV. These results have been cross-checked with alternate interaction models in the air shower simulations. There is a very good qualitative agreement in all cases.

Furthermore, it was established that the muon discrepancy in simulations increases smoothly with energy [8,60,61]. Thus, the slope of the energy dependence introduced in Eq. (2) is also affected, pointing to a too small value of  $\beta$ . This may be related to a too large  $\pi^0$  production leading to a too large value of  $R$  yet to be explained. We will explore the possible origin of a different  $R$  in the next section.

### III. MULTIPARTICLE PRODUCTION AND THE $R$ OBSERVABLE

The discussion in the previous section suggests that a change of  $R$  (or  $c$ , which is equivalent) is a potential way to reduce the muon content 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 as suggested by the studies of the previous Sec. II in an arbitrary way. In a naive model like Ref. [55], where only pions are considered as secondary particles,  $R = 0.5$ . In a more realistic approach based on string fragmentation we have  $R \approx 0.4$  (as seen in Fig. 3). But as shown in Ref. [20], 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 [62] yielding  $R \approx 0.34$ . Such a behavior can be explained in terms of a core-corona picture [45]. This approach has been used in the framework of full hydrodynamical simulations [51,63] but also in simple model calculations [64–67]. 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

$$N_i = \omega_{\text{core}} N_i^{\text{core}} + (1 - \omega_{\text{core}}) N_i^{\text{corona}}, \quad (6)$$

where  $N_i^{\text{core}}$  represents statistical (grand canonical) particle production,  $N_i^{\text{corona}}$  is the yield from string decay, and  $\omega_{\text{core}}$  is the core weight. In order to explain LHC data [20] 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 events, reaching unity for central heavy ion collisions (PbPb). A simplified version of the core-corona approach will be described in Sec. IV.

For an illustration we will use EPOS LHC as the baseline model to test sensitivity of  $R$  towards a QGP-like state. As alternative model we use PYTHIA8 [68,69], which provides entirely different (non-QGP-like) physics concepts for collectivity. EPOS LHC is a general purpose event generator widely used in high energy physics, and in particular also for heavy ion collisions. It includes the description of a QGP-like behavior in high energy collisions. PYTHIA8, on the other hand, is the reference model in high energy physics for proton-proton interactions. Both models generate a distribution of colored strings from the collision of a projectile and a target. Despite a very different underlying approach for the string generation [perturbative Quantum Chromo Dynamics (pQCD) factorization for PYTHIA8 and parton-based Gribov-Regge theory [70] for EPOS LHC], the string distributions are not very different, because they are

strongly constrained by the data on particle multiplicities. These strings can be hadronized directly in both generators using the Lund string model [71] in PYTHIA8, or the area law [70] in EPOS LHC—both cases are strongly constrained by LEP data. At low energy ( $\approx 10$  to  $100$  GeV) this is sufficient to successfully describe proton-proton interactions with good accuracy. Nevertheless, it turns out that at the LHC energies, additional physics mechanisms are needed to describe the observed particle correlations and abundances in the final state. In PYTHIA8, a modified color reconnection approach [72,73] or a “string shoving” mechanism [43] have been proposed to introduce collective effects such as a modified hadronization or particle correlations, similar to those obtained from a QGP. In EPOS LHC, on the other hand, the “core-corona” approach [45] is used as originally developed for heavy ion collisions. As already explained, the core amounts to regions with high string/energy densities, where strings are assumed to “melt” and produce matter that expands hydrodynamically and then decays statistically, whereas the corona represents particles from ordinary string fragmentation, which escape from the dense regions. While in EPOS3 [74] the hydrodynamic expansion is fully implemented and hadronization occurs on a freeze-out hypersurface, in EPOS LHC this expansion is mimicked by parametrizing the flow at hadronization. This has proven to well describe various collective observable phenomena [56]. Simulations of EPOS LHC are readily available via the CRMC software [75]. On generator level, we study particles with a lifetime  $c\tau > 1$  cm, which is a widely adopted definition of long-lived particles at the LHC [76].

In fact, in EPOS LHC final-state particles originate from three different production mechanisms: standard string fragmentation (corona), statistical decay of a fluid (core), and the decay of the beam remnants (as defined in this particular model [77]). While experimentally the origin of the production mechanism for a particle cannot be identified, individual production mechanisms can still be studied since they predominantly contribute to different regions of phase space. This is demonstrated in Fig. 2 (top), which shows the relative contribution of these mechanisms to the total energy density  $\langle dE/d\eta \rangle$  for minimum bias proton-proton collisions at a center-of-mass energy of 13 TeV. Three regions can be identified: The energy density at central pseudorapidities,  $|\eta| < 5$ , is dominated by particles originating in the dense core of the interaction; at intermediate rapidities,  $5 < |\eta| < 8$ , it is dominated by particles from string fragmentation; and at large rapidities,  $|\eta| > 8$ , by the fragmentation of beam remnants. Underlying differences in particle production, therefore, lead to varying observables as a function of pseudorapidity.

A corresponding effect is also observed as a function of the central charged particle multiplicity  $dN_{\text{ch}}/d\eta|_{\eta=0}$  (charged hadrons with  $c\tau > 1$  cm and no  $p_T$  cut). Final states with large particle multiplicity are known to be an effective proxy for pronounced statistical hadronization [27]. Therefore, at

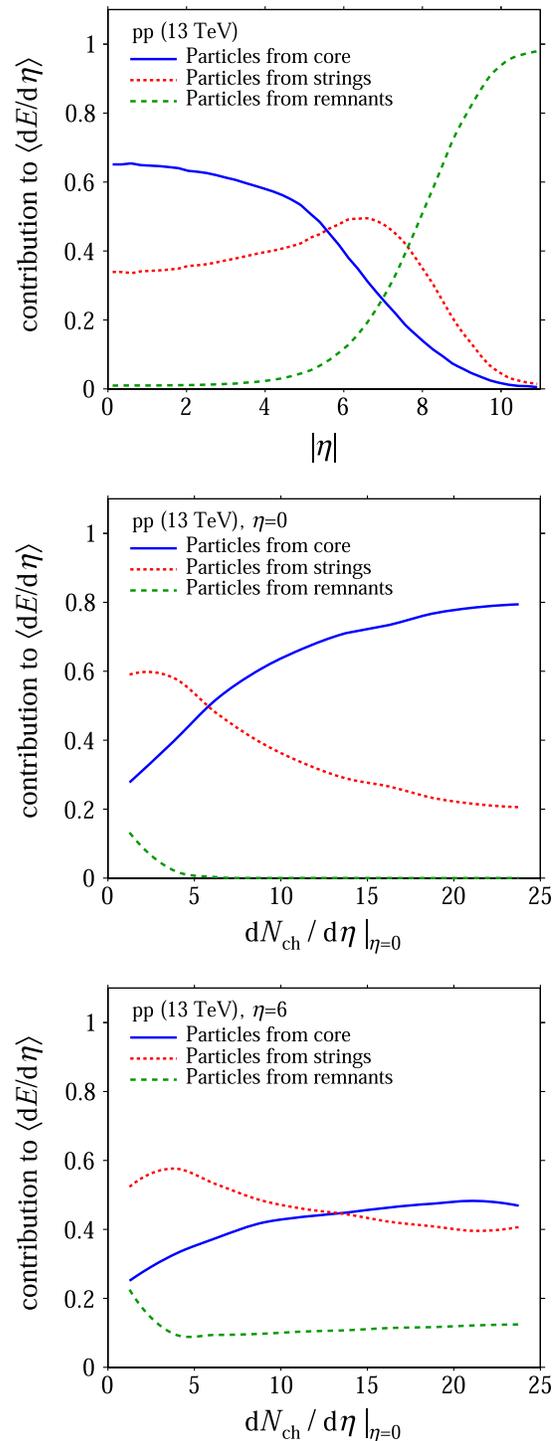


FIG. 2. Fractional contribution of particles originating from different production mechanisms to the total energy density  $dE/d\eta$  as predicted by EPOS LHC. The top figure shows the contribution as function of  $|\eta|$ ; the middle (bottom) figure shows the contributions at  $\eta = 0$  ( $\eta = 6$ ) as a function of the charged particle density at  $\eta = 0$ .

fixed pseudorapidity, the influence of the core increases as a function of particle multiplicity. This effect is expected to be most significant at  $|\eta| \approx 0$  since the relative contribution of

the core is largest. This is illustrated in the middle panel of Fig. 2 for  $\eta = 0$  and in the bottom panel for  $\eta = 6$ . It can be seen that the contribution of the core to the energy density at  $\eta = 0$  becomes dominant for pp collisions with more than  $\approx 7$  charged particles per unit of pseudorapidity, while at  $\eta = 6$  this transition is shifted to a larger value. Here the multiplicity at midrapidity should be seen as a proxy for the “centrality” of the collision. Thus very low multiplicity events are peripheral events, and, in particular, diffractive ones with particles coming only from the beam remnants. This explains the rise of the remnant contribution at very low  $dN_{\text{ch}}/d\eta|_{\eta=0}$ . In the EPOS LHC model, the core fraction will depend on the local particle density, which could come from the production of the secondary particles and which dominates at midrapidity. But this density could be modified by the density of partons in the beam remnants which depend on various assumptions. In other words, the  $\eta$  dependence of the core formation is very model dependent and very few data are currently available to test the core fraction in the forward phase space most relevant for the air shower development.

Using EPOS LHC we find that the fraction of secondary pions (both charged and neutrals) 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. Accordingly, this effect can be seen in the pseudorapidity-dependent ratio of the average electromagnetic to hadronic energy density  $R$  shown in the top panel of Fig. 3. At  $|\eta| \approx 0$ , hadron production is dominated by the core and therefore the value of  $R$  for EPOS LHC is as low as 0.34. As the contribution of the core to the total energy decreases with increasing pseudorapidity, also  $R$  increases and reaches a value of 0.4 at  $|\eta| \approx 7$  before it decreases rapidly due to the very low electromagnetic contribution in the beam remnants. In comparison, a flat ratio below  $|\eta| \approx 7$  is obtained when statistical hadronization is disabled in EPOS LHC (corona only). The data point shown in this figure at  $\eta \approx 6$  is derived from Ref. [78], where we have corrected the original values from detector level to generator level using the Rivet routines provided by the CMS collaboration [79,80]. The shaded region corresponds to the systematic uncertainties of the measurement. These data are consistent with all models within the experimental uncertainty; there is a slight tension with the PYTHIA8 simulations using the modified color reconnection approach [81]. Such data with smaller uncertainties, and measured over a wide range of  $\eta$ , have the potential to differentiate between some of the models. In particular, any slope observed in the region  $0 < |\eta| \lesssim 6$  would be a clear hint for a transition of several distinct hadronization mechanisms (i.e., the core-corona effect).

The ratio of the electromagnetic to hadronic energy density  $R$  at  $\eta = 0$  is shown as a function of the central multiplicity  $dN_{\text{ch}}/d\eta|_{\eta=0}$  in the middle panel of Fig. 3. It can be observed that  $R$  drops down to values of 0.3 when

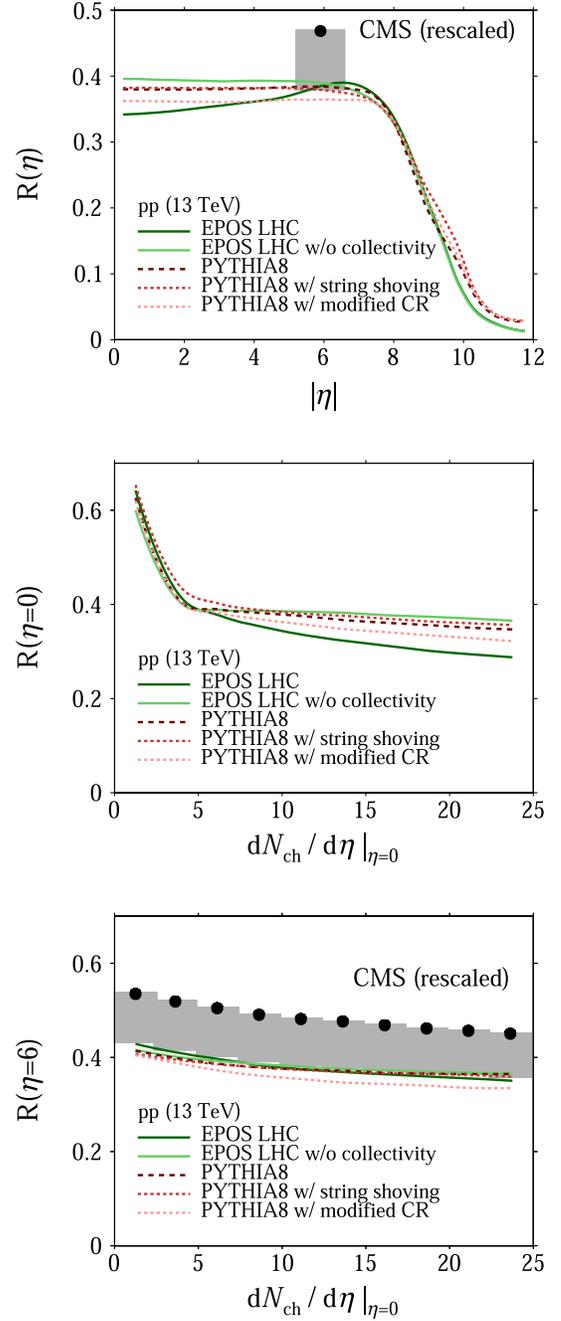


FIG. 3. Ratio of the average electromagnetic to hadronic energy densities  $R$  simulated for proton-proton collisions at 13 TeV with EPOS LHC (solid lines) with and without hydrodynamical treatment of the dense core, as well as PYTHIA8 (dashed lines) in the default configuration, with string shoving and with modified color reconnection (CR). The top figure shows  $R(\eta)$  as a function of  $|\eta|$ , the middle and bottom figure show  $R$  evaluated at  $\eta = 0$  and  $\eta = 6$  as a function of the central charged particle multiplicity. The asymmetric uncertainties of the CMS data are a feature of this measurement.

statistical hadronization is enabled in EPOS LHC while it reaches a constant plateau of 0.4 in the case of disabled statistical hadronization, which is similar to the PYTHIA8

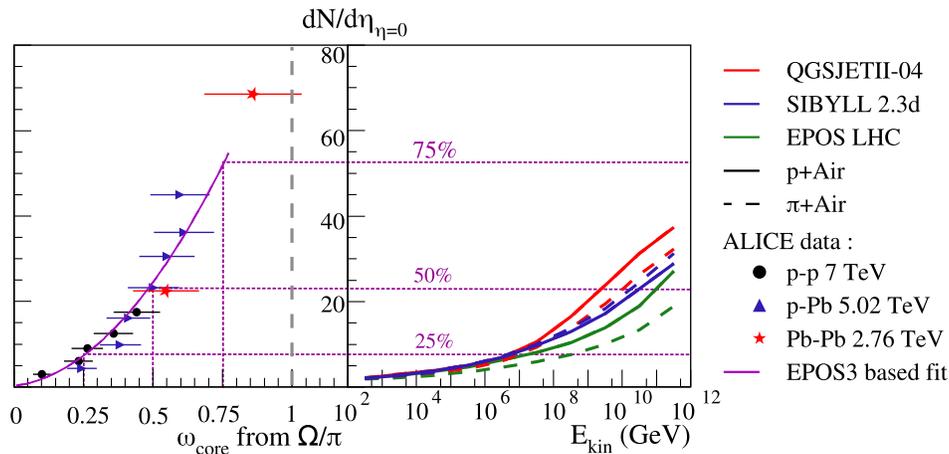


FIG. 4. Energy evolution of particle density at mid-rapidity for proton (solid lines) and pion (dashed lines) interaction with air for different hadronic interaction models (right-hand side) compared to  $\omega_{\text{core}}$  as extracted from the  $\Omega/\pi$  from [63] (left-hand side). Data points are from ALICE [20] recorded at pseudorapidity  $|\eta| < 0.5$ . See text for details.

predictions. The peak at very low multiplicity is due to low mass diffractive events where the  $\pi^0$  production is favored because of the lack of energy to produce heavier particles. At  $\eta = 6$ , it can be seen in the bottom panel of Fig. 3 how the different model predictions compare to the available CMS data (also from Ref. [78]). A slightly better description of the multiplicity dependence is achieved when collective effects are taken into account (both in EPOS LHC and PYTHIA8), but in the context of this paper, it is obvious that a more precise measurement at  $\eta \approx 6$  and more is of paramount importance for the muon problem in air shower physics. In particular using a light ion beam like oxygen to be as close as possible to the air target.

From this study, it is clear that the core fraction can have a complex behavior that depends on both multiplicity (on an event-by-event basis) and pseudorapidity and relies on various model assumptions. For instance, EPOS LHC was released after the first LHC data became available. At that time, only average values and the evolution of the mean transverse momentum as a function of the particle multiplicity were known precisely. The increase of multistrange baryon production with particle multiplicity was a prediction of the model, but—as shown in Ref. [20]—was only qualitatively correct. Effectively, the core is formed in EPOS LHC only at larger multiplicities compared to what is necessary to reproduce the data. Thus, it is expected that the density needed to produce the core is currently overestimated changing both  $\eta$  and multiplicity dependence [more core at lower multiplicity (and then energy) and larger rapidities].

Precise data on  $R$  versus multiplicity is needed to support (or reject) this hypothesis, but studies at the LHC in proton-proton and proton-nucleus collisions support such a scenario [20,27] at energy densities as reached by cosmic rays interacting with the atmosphere [47–49]. Studying LHC data at midrapidity, it is found that for events with

$\langle dN_{\text{ch}}/d\eta \rangle_{|\eta| < 0.5} \sim 10$  (corresponding to typical proton-air interactions as shown in Fig. 4 right-hand side)  $\omega_{\text{core}}$  is already  $\approx 30\%$  or more at higher energy. In fixed target experiments, where the full phase space is available for measurements, the  $\omega_{\text{core}}$  is below 5% which is less than the precision of the models. At RHIC (200–500 GeV), in the light of what is observed at the LHC, collective effects like flow or multiplicity-dependent strangeness enhancement are in fact observed in proton-proton [82] and proton-gold data [83–86].

In the next section, we will study the possible impact of the core production on the muon production in air showers.

#### IV. SIMPLIFIED CORE-CORONA APPROACH AS MOTIVATED BY LHC DATA

In order to investigate whether a constrained value of  $R$  could be low enough to increase the number of muons in air shower simulations, as it is required to describe observations, a simplified core-corona approach is used to estimate the potential consistency with LHC data under the following assumptions:

- (1) *The fraction of core  $\omega_{\text{core}}$  effectively increases logarithmically with energy.* The standard core-corona approach is fundamentally a function of the multiplicity and not primarily of the collision energy [20], but since the average multiplicity increase with the energy, the average core fraction  $\omega_{\text{core}}$  will also increase with energy. In Fig. 4, we show how the particle density at midrapidity can be converted into a evolution of  $\omega_{\text{core}}$  with energy. As the most sensitive example, the description of the  $\Omega/\pi$  ratio by a core-corona approach in EPOS3 [63] corresponds to a correlation between  $dN/d\eta_{\eta=0}$  and  $\omega_{\text{core}}$ : the common y axis denotes central multiplicity; the left plot shows how  $\omega_{\text{core}}$  smoothly changes

over various center-of-mass energies and colliding systems; the right plot illustrates  $\pi/p$ -air collisions as a function of lab energy.

- (2) *Only the change of hadronization is taken into account.* Collective effects in the core in principle include particle correlations and flow, but since particles with a large ratio of longitudinal to transverse momentum dominate most of the air shower development, these effects are expected to be negligible.
- (3) *Nuclear effect are not considered.* The multiplicity increases with the mass of the projectile, so the core fraction should also increase with the mass of the cosmic ray. For technical reasons this cannot be taken into account in our simulations—but since nuclear effects further enhance the proposed impact (mostly for the first interaction), our approach will yield conservative statements.
- (4) *The core-corona effect is applied equally to the full phase space,* while core hadronization has been experimentally established at midrapidity mainly. As discussed in the previous section, this is not currently excluded but ultimately needs more experimental data. Therefore, we simply assume that the modification equally affects particles emitted at all pseudorapidities (except for the leading particle).
- (5) *Similarly, the implementation of the core-corona effect is performed on all types of hadronic projectile* (nucleons, pions, and kaons). This can be justified by the simple argument that the increase of the multiplicity and thus of the energy density is mostly due to the increase of sea quarks and gluons interactions from the projectile (and target) particle, which are independent of the valence quarks. The latter change only the nature of the leading particles, which are unchanged in our simulations.

In the following, we are going to define a “simplified core-corona approach,” based on Eq. (6):

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

using CONEX air shower simulations, applied on different hadronic interaction models. The particle yield from the chosen model is by definition considered to represent the corona-type yield  $N_i^{\text{corona}}$ , whereas we use the standard statistical hadronization [also referred to as “resonance gas” or “grand canonical ensemble” (GCE)] for the core part. For the latter, in this study, one uses a model [62] that reproduces very well the particle ratios of central PbPb collisions at the LHC with the following parameters:  $T = 156.5$  MeV,  $\mu_B = 0.7$  MeV, and  $V = 5280$  fm<sup>3</sup>. By definition, this is the extreme case that we consider as the known observed limit for the core values and corresponding particle ratios. As a consequence, we impose as a constraint that the GCE parameters are fixed and only  $\omega_{\text{core}}$

changes with energy. So  $\omega_{\text{core}} = 0$  would be the “normal” simulation with the chosen hadronic interaction model. Choosing  $\omega_{\text{core}} > 0$  amounts to mixing the yields from the given model with the one from the GCE according to the core-corona superposition shown in Eq. (6). EPOS LHC is a particular case because it has already a core-corona model implemented, but with a quantitatively wrong value of  $\omega_{\text{core}}$  as shown in [20]. Hence the core production was switched of, to produced new spectra for CONEX that are based on string fragmentation (corona) only. SIBYLL 2.3d and QGSJETII.04 have by default corona component only.

The implementation of this simplified core-corona approach was performed by mean of CONEX taking advantage of its numerical air shower simulations based on cascade equations. This allow us to directly modify individual ratios of secondary particle species of the energy spectra  $dN_i/dE_j$ , for particle species  $i$  and energy bins  $dE_j$ , of hadronic interactions with air nuclei without modifying the hadronic model themselves (only the tabulated averaged spectra are modified). These energy spectra of secondary particles were calculated once for each full MC hadronic interaction model, for later use them as an input by CONEX for numerical air shower simulations. Knowing the initial ratios  $\pi^0/\pi^\pm$ ,  $p/\pi^\pm$ ,  $K^\pm/\pi^\pm$ ,  $p/n$ ,  $K^0/K^\pm$  of the particles used in CONEX from each corona-type model (i.e., current hadronic models) and the value of the corresponding ratios from the core model [62], we compute new spectra in which the particle yields include both core and corona according to  $\omega_{\text{core}}$ .

These ratios take into account the decay products of all other type of particles like strange baryons or mesons with a lifetime too short to propagate in the atmosphere. For leptonic decay product from  $\eta$  particle for instance, we approximate that it can be counted as  $\pi^0$ , since in the cascade equation it will directly contribute to the electromagnetic cascade and the possible change of particle spectra is not considered in this simplified approach (only the particle ratios). Neutrinos are not counted since they have no impact on the muon production in air showers.

Since the hadronization mechanism can affect only newly produced particles the properties of the leading particle must be preserved. To achieve that, the new particle yields are computed for all secondaries by applying a scaling procedure, but excluding the one corresponding to the respective projectile type, i.e., protons in proton-air and 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  $\sum_i E_j dN_i/dE_j$  must be conserved. Since at high  $x_F = E_j/E_{\text{lab}}$  only the projectile-type particles will have  $dN_i/dE_j$  significantly different from zero (also known as a leading-particle effect), the resulting modified leading-particle-type spectra at high  $x_F$  follow the original distribution and are only affected by the scaling procedure at

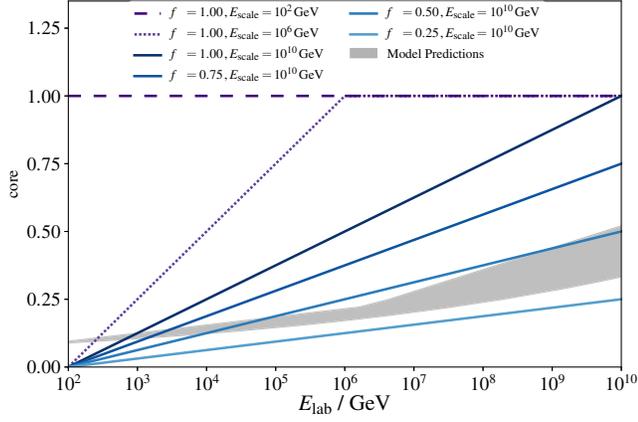


FIG. 5. Different energy evolutions probed for  $\omega_{\text{core}}$ . The solid lines represent changing the scale  $f_\omega$  of the effect, while the dashed lines also indicate the effect of changing  $E_{\text{scale}}$ . The shaded area is based on Fig. 4 motivate by ALICE data and current predictions on particle densities of the EPOS3 model.

lower values of  $x_F$ . Together, this assures us that energy conservation as well as the total multiplicity are not affected, while the species ratios are changed [87].

We assume the core weight  $\omega_{\text{core}}$  to increase with energy in a logarithmic way. Thus, we use

$$\omega_{\text{core}}(E_{\text{lab}}) = f_\omega F(E_{\text{lab}}; E_{\text{th}}, E_{\text{scale}}) \quad (7)$$

to model this [in analogy to Eq. (4)], 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  $f_\omega$  scale is varied from 0.25, 0.5, 0.75, to 1.0; in addition we limit the maximum of  $F(E_{\text{lab}}; E_{\text{th}}, E_{\text{scale}})$  to 1. This yields the  $\omega_{\text{core}}$  energy dependencies as depicted in Fig. 5. The comparison with the expected behavior estimated from central ( $\eta \sim 0$ ) data is overlaid as the shaded area (from Fig. 4) indicates that the logarithmic evolution is a reasonable simplified starting assumption, in particular at high energies where the effect is the strongest. Since the pseudorapidity dependency of  $\omega_{\text{core}}$  is, in fact, unknown and model dependent, it remains highly relevant to study various simplified scenarios beyond what is indicated by the data from central ( $|\eta| < 0.5$ ) particle production. All these scenarios have been simulated with CONEX, using cascade equations from the first interaction to the ground, for proton and iron primary particles at  $E_0 = 10^{19}$  eV. No explicit Monte Carlo method is used in order to obtain just the average response. Since only cascade equations are used, the superposition model is used to get the iron results as sum of 56 simultaneous proton showers, each with the energy reduced by a factor of 56.

To summarize this chapter: we introduced a “simplified core-corona approach,” where an event is composed of two

components, a core part and a corona part, with respective weights  $\omega_{\text{core}}$  and  $1 - \omega_{\text{core}}$ . The core part corresponds to “normal” particle production, as given by the underlying interaction model, whereas the core part corresponds to “statistical” or “thermal” particle production. The core weight is assumed to be energy dependent.

## V. IMPACT OF A SIMPLIFIED CORE-CORONA APPROACH ON AIR SHOWERS

In Fig. 6 the results are shown in the  $X_{\text{max}} - \ln N_\mu$  plane for three models EPOS LHC (top left), QGSJETII.04 [88,89] (top right), and SIBYLL 2.3d [90] (bottom left). These examples illustrate that it is possible to describe the data of the Pierre Auger Observatory with modified hadronization in air shower cascades, if a larger corelike contribution is considered compared to what is currently provided by the models. For this to work, however, QGP-like effects need to appear already in light colliding systems and start at comparably low center-of-mass energies. In the bottom-right panel, the three models are compared using the core fraction corresponding to the shaded area of Fig. 5. In the case of SIBYLL 2.3d, this scenario puts the dashed-blue line within one sigma (systematic) of the data point from the Pierre Auger Collaboration. In other words, the combination of a deeper  $X_{\text{max}}$  for this model and an increase by about 15% of the muon number thanks to the core-corona approach is enough to make the simulations compatible with the data point. For EPOS LHC and QGSJETII.04 a larger  $\omega_{\text{core}}$  is needed or an additional change in  $X_{\text{max}}$  to get a heavier composition. On the other hand, we can see that with the core-corona approach it is not possible to reach the nominal value of the data without extreme uncertainty, which are very unlikely even with the large uncertainty on the core fraction in forward particle production.

Furthermore, from Eq. (2) also a different energy evolution of the muon production follows. To study the effect of our core-corona model on the muon production as a function of the energy, we can compare the different scenarios with the compilation of data presented in Ref. [8] using the renormalized factor

$$z = \frac{\ln N_\mu - \ln N_\mu^{\text{p}}}{\ln N_\mu^{\text{Fe}} - \ln N_\mu^{\text{p}}}, \quad (8)$$

with  $N_\mu$  being any muon related experimental observable and  $\ln N_\mu^{\text{p}}$  and  $\ln N_\mu^{\text{Fe}}$  being the logarithm of the same observable simulated with proton and iron primaries, respectively, for a given reference hadronic interaction model. This allows a direct comparison between different experiments for various types of muon observables.

Considering the energy dependence of  $z$ , there is an implicit dependence on the cosmic-ray mass  $A$ , since  $\langle \ln A \rangle$  varies with energy. However, as expected from the Heitler

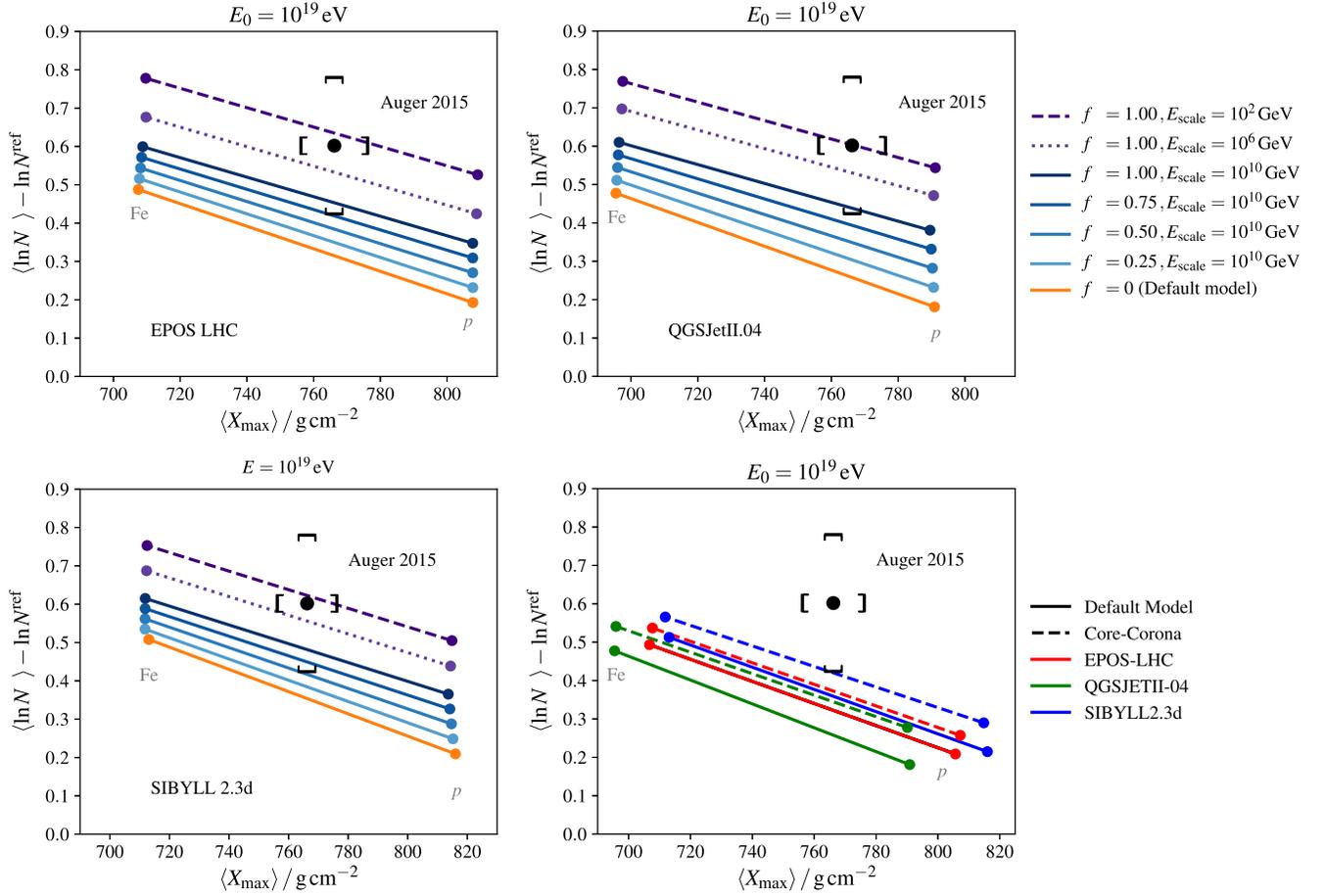


FIG. 6. The impact of different core-corona mixing scenarios on air shower simulations at  $10^{19}$  eV in the  $X_{\text{max}} - \ln N_{\mu}$  plane using EPOS LHC (top left), QGSJETII.04 (top right), SIBYLL 2.3d (bottom left), and the ALICE-motivated model corresponding to the shaded area in Fig. 5. 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 datum is from the Pierre Auger Observatory [5]. Each model line represents all values that can be obtained for any mixture of cosmic nuclei from proton (bottom right) to iron (top left). In the bottom-right panel, the full lines are the default model and the dashed line the ALICE-motivated core-corona mixing.

model formula, and even more importantly, verified via explicit simulations,  $z$  and  $\langle \ln A \rangle$  are related as  $z = a + b \langle \ln A \rangle$ , and from  $z(\text{pure Fe}) = 1$  and  $z(\text{pure p}) = 0$  we simply get  $a = 0$  and  $b = 1/\ln 56$ . This is very useful, since it means that the  $A$  dependence of  $z$  (called  $z_{\text{mass}}$ ) is given as

$$z_{\text{mass}} = \frac{\langle \ln A \rangle}{\ln 56}, \quad (9)$$

and the expectation of  $\Delta z = z - z_{\text{mass}}$  is zero for the case of full consistency between all experimental observables and the simulations based on a valid reference model. This means that, plotting  $\Delta z$  for experimental data, we should get zero if the reference model were perfect, whereas  $\Delta z > 0$  implies a muon deficit in the simulations. In this way we can visualize the energy dependence of the muon excess, corrected for mass dependencies. More details and references are given in Ref. [8].

As pointed out in Ref. [8], for all models the data have a positive  $\Delta z$  showing a significant logarithmic increase with the primary energy, indicating an increasing muon deficit in the simulations. In Fig. 7 the effect of the different energy evolution of  $\omega_{\text{core}}$  for EPOS LHC and QGSJETII.04 on  $\Delta z$  are shown. Here the new simulations are treated like data and the  $z$  factor is calculated using the original (quoted) models as a reference such that the new  $\Delta z$  can be compared to the data points directly. The positive  $\Delta z$  of the lines indicate a larger muon production when  $\omega_{\text{core}}$  increases and the positive slopes mean that the slope of the muon production as a function of the primary energy is larger when  $\omega_{\text{core}}$  increases. By including a consistent corelike hadronization, we thus reproduce the energy evolution as found in the data. This is even possible for values of  $\omega_{\text{core}}$  remaining well below 1. The ALICE-motivated [20] results corresponding to the shaded area in Fig. 5 are indicated by the red lines. As already seen in Fig. 6, for EPOS LHC and QGSJETII.04 the effect would

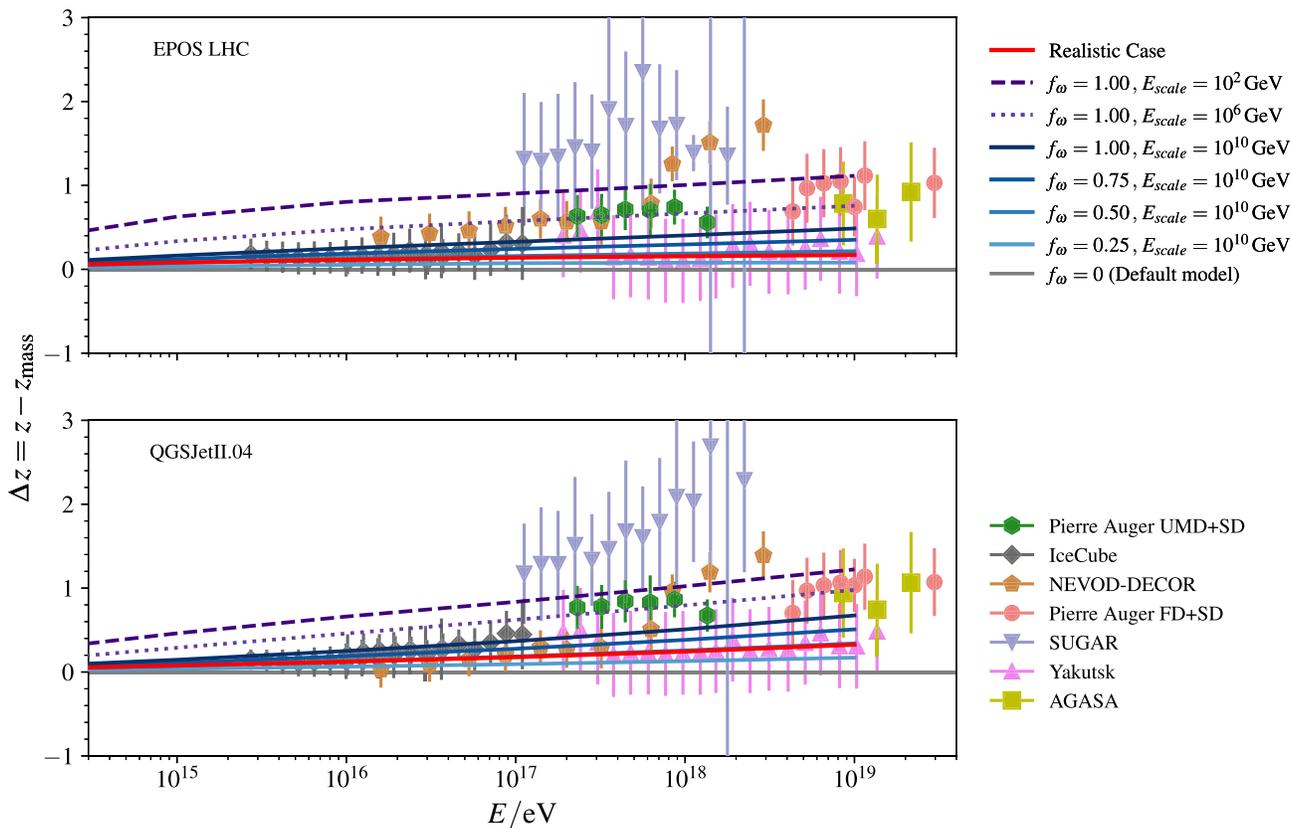


FIG. 7. Evolution of the mass corrected  $z$  factor,  $\Delta z = z - z_{\text{mass}}$ , as a function of the primary energy. The data are taken from Ref. [9], where the muon content measured by AGASA [91], IceCube [92], KASCADE-Grande [93], NEVOD-DECOR [94,95], Pierre Auger [96,97], SUGAR [98], and Yakutsk [99] are compared. Overlaid are predictions obtained from changing the scale  $f_\omega$  (solid lines) and  $E_{\text{scale}}$  (dashed and dotted lines) obtained with EPOS LHC (top) and QGSJETII.04 (bottom) air shower simulations.

not be large enough for the current mass composition of the models.

The possibility to see the effect of a core hadronization (QGP or similar more exotic phenomena) on air shower physics have already been studied in the literature [100–103]. Changes in the muon production because of a change of  $R$  under either extreme or exotic assumptions (which were not yet observed at the LHC) are usually assumed. Furthermore, it was shown that the production of a core only in very central, high-density, collisions is not sufficient to significantly change the muon numbers in air shower simulations [104].

In contrast to the new results presented here, in those previous studies the corelike production does not cover sufficient energy range in air showers to change the muon production significantly. We demonstrate that corelike effects potentially applied to small colliding systems, and for relatively low center-of-mass energies as studied here, have an important impact on muon production in air showers. Since our study is based on the simple assumption that the full phase space has a modified  $\pi^0$  ratio, it remains crucial for cosmic ray physics to conduct further dedicated measurements at the LHC to better understand  $\pi^0$  production relative to other particles. The phase space for the

formation of corelike effects is potentially significantly larger than previously studied, and in particular may extend towards larger rapidities.

## VI. SUMMARY

In the light of recent LHC data on soft QCD [20], it is clear that the current hadronic models used for air shower simulations are not up to date in terms of hadronization. We have shown that the muon production in air showers significantly depends on the ratio  $R = E_{\text{em}}/E_{\text{had}}$ , where  $E_{\text{em}}$  is the sum of energy in secondary  $\gamma$  (from  $\pi^0$ ) and  $e^\pm$ , while  $E_{\text{had}}$  is the sum of energy in hadrons in individual hadron collisions. We also showed that  $R$  itself depends on the hadronization mechanism. Thus, a change or transition in these mechanisms can help to explain the discrepancy between the observed number of muons in air showers by the Pierre Auger Observatory and the predictions based on current hadronic models. Since at the LHC, even in proton-proton interactions, one observes a transition from a string-type to a statistical-type hadronization at midrapidity, we used the particle ratios of the statistical model at all pseudorapidities (as an extreme case) to show that such

a hadronization scheme would in principle be sufficient to resolve the observed difference between air shower simulations and cosmic ray data. Experimental measurements at the LHC are currently compatible with this possibility if the mass composition predicted by the model is large enough (like in the SIBYLL 2.3d case). On the other hand, extreme scenarios where full statistical hadronization is reached at low energies [ $E_{\text{lab}} \sim \mathcal{O}(100 \text{ GeV})$ ] are already excluded by the slope of the energy dependence of air shower muon data. Furthermore, in future huge-aperture air shower experiments, the tail of the  $\ln N_\mu$  distribution could be used to indirectly measure the slope of the energy distribution of neutral pions far beyond the reach of the LHC [11,12].

Dedicated measurements at the LHC have now another complementary opportunity to study hadronization in hadronic collisions using the proposed  $R$  observable. This will provide new constraints on the extension of the phase space in which statistical hadronization contributes to final state particle distributions. In particular, since  $R$  can be measured

relatively easily over wide ranges of pseudorapidity—including the forward direction that is most important for air shower physics. Such measurements would complement the very important forward measurements done by the LHCf experiment on the neutral particles [105–107]. The core-corona effect might not be the full solution to the muon puzzle, but it should be considered thoroughly in the models before claiming the need for more exotic explanations. Measuring  $R$  at the LHC potentially has a significant impact on resolving the current mystery of muon production in cosmic ray induced extensive air showers. Thus, at last, one aspect to resolve the cosmic ray muon mystery is a better understanding of statistical hadronization in small collision systems at accelerators.

### ACKNOWLEDGMENTS

H. D. acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project No. 449728698.

- 
- [1] A. Aab *et al.* (Pierre Auger Collaboration), The Pierre Auger cosmic ray observatory, *Nucl. Instrum. Methods Phys. Res., Sect. A* **798**, 172 (2015).
  - [2] J. Abraham *et al.* (Pierre Auger Collaboration), The fluorescence detector of the Pierre Auger observatory, *Nucl. Instrum. Methods Phys. Res., Sect. A* **620**, 227 (2010).
  - [3] H. Tokuno *et al.*, New air fluorescence detectors employed in the Telescope Array experiment, *Nucl. Instrum. Methods Phys. Res., Sect. A* **676**, 54 (2012).
  - [4] K.-H. Kampert and M. Unger, Measurements of the cosmic ray composition with air shower experiments, *Astropart. Phys.* **35**, 660 (2012).
  - [5] A. Aab *et al.* (Pierre Auger Collaboration), Muons in air showers at the Pierre Auger Observatory: Mean number in highly inclined events, *Phys. Rev. D* **91**, 032003 (2015); **91**, 059901(E) (2015).
  - [6] A. Aab *et al.* (Pierre Auger Collaboration), Testing Hadronic Interactions at Ultrahigh Energies with Air Showers Measured by the Pierre Auger Observatory, *Phys. Rev. Lett.* **117**, 192001 (2016).
  - [7] R. U. Abbasi *et al.* (Telescope Array Collaboration), Study of muons from ultrahigh energy cosmic ray air showers measured with the Telescope Array experiment, *Phys. Rev. D* **98**, 022002 (2018).
  - [8] H. P. Dembinski *et al.* (EAS-MSU, IceCube, KASCADE-Grande, NEVOD-DECOR, Pierre Auger, SUGAR, Telescope Array, and Yakutsk EAS Array Collaborations), Report on tests and measurements of hadronic interaction properties with air showers, *EPJ Web Conf.* **210**, 02004 (2019).
  - [9] D. Soldin (EAS-MSU, IceCube, KASCADE-Grande, NEVOD-DECOR, Pierre Auger, SUGAR, Telescope Array, and Yakutsk EAS Array Collaborations), Update on the combined analysis of muon measurements from nine air shower experiments, *Proc. Sci. ICRC2021* (2021) 349.
  - [10] R. Ulrich, R. Engel, and M. Unger, Hadronic multiparticle production at ultra-high energies and extensive air showers, *Phys. Rev. D* **83**, 054026 (2011).
  - [11] L. Cazon, R. Conceição, and F. Riehn, Probing the energy spectrum of hadrons in proton air interactions at ultrahigh energies through the fluctuations of the muon content of extensive air showers, *Phys. Lett. B* **784**, 68 (2018).
  - [12] L. Cazon, R. Conceição, M. A. Martins, and F. Riehn, Probing the  $\pi^0$  spectrum at high- $x$  in proton-Air interactions at ultra-high energies, *EPJ Web Conf.* **210**, 02006 (2019).
  - [13] H. A. Gustafsson *et al.*, Collective Flow Observed in Relativistic Nuclear Collisions, *Phys. Rev. Lett.* **52**, 1590 (1984).
  - [14] J. Barrette *et al.* (E877 Collaboration), Observation of Anisotropic Event Shapes and Transverse Flow in Au + Au Collisions at AGS Energy, *Phys. Rev. Lett.* **73**, 2532 (1994).
  - [15] H. Appelshäuser *et al.* (NA49 Collaboration), Directed and Elliptic Flow in 158-GeV / Nucleon Pb + Pb Collisions, *Phys. Rev. Lett.* **80**, 4136 (1998).
  - [16] E. Andersen *et al.* (WA97 Collaboration), Strangeness enhancement at mid-rapidity in Pb Pb collisions at 158-A-GeV/c, *Phys. Lett. B* **449**, 401 (1999).
  - [17] F. Antinori *et al.* (NA57 Collaboration), Energy dependence of hyperon production in nucleus nucleus collisions at SPS, *Phys. Lett. B* **595**, 68 (2004).

- [18] B. I. Abelev *et al.* (STAR Collaboration), Enhanced strange baryon production in Au + Au collisions compared to p + p at  $s(\text{NN})^{1/2} = 200\text{-GeV}$ , *Phys. Rev. C* **77**, 044908 (2008).
- [19] B. B. Abelev *et al.* (ALICE Collaboration), Multi-strange baryon production at mid-rapidity in Pb-Pb collisions at  $\sqrt{s_{\text{NN}}} = 2.76\text{ TeV}$ , *Phys. Lett. B* **728**, 216 (2014); **734**, 409(E) (2014).
- [20] J. Adam *et al.* (ALICE Collaboration), Enhanced production of multi-strange hadrons in high-multiplicity proton-proton collisions, *Nat. Phys.* **13**, 535 (2017).
- [21] K. Werner, I. Karpenko, and T. Pierog, “Ridge” in Proton-Proton Scattering at 7 TeV, *Phys. Rev. Lett.* **106**, 122004 (2011).
- [22] P. Bozek, Elliptic flow in proton-proton collisions at  $\sqrt{s} = 7\text{ TeV}$ , *Eur. Phys. J. C* **71**, 1530 (2011).
- [23] D. d’Enterria, G. K. Eyyubova, V. L. Korotkikh, I. P. Lokhtin, S. V. Petrushanko, L. I. Sarycheva, and A. M. Snigirev, Estimates of hadron azimuthal anisotropy from multiparton interactions in proton-proton collisions at  $\sqrt{s} = 14\text{ TeV}$ , *Eur. Phys. J. C* **66**, 173 (2010).
- [24] S. K. Prasad, V. Roy, S. Chattopadhyay, and A. K. Chaudhuri, Elliptic flow ( $v_2$ ) in pp collisions at energies available at the CERN Large Hadron Collider: A hydrodynamical approach, *Phys. Rev. C* **82**, 024909 (2010).
- [25] G. Ortona, G. S. Denicol, P. Mota, and T. Kodama, Elliptic flow in high multiplicity proton-proton collisions at  $\sqrt{s} = 14\text{ TeV}$  as a signature of deconfinement and quantum energy density fluctuations, *arXiv:0911.5158*.
- [26] L. Cunqueiro, J. Dias de Deus, and C. Pajares, Nuclear like effects in proton-proton collisions at high energy, *Eur. Phys. J. C* **65**, 423 (2010).
- [27] V. Khachatryan *et al.* (CMS Collaboration), Observation of long-range near-side angular correlations in proton-proton collisions at the LHC, *J. High Energy Phys.* **09** (2010) 091.
- [28] K. Dusling, W. Li, and B. Schenke, Novel collective phenomena in high-energy proton-proton and proton-nucleus collisions, *Int. J. Mod. Phys. E* **25**, 1630002 (2016).
- [29] C. Loizides, Experimental overview on small collision systems at the LHC, *Nucl. Phys.* **A956**, 200 (2016).
- [30] E. V. Shuryak, Quantum chromodynamics and the theory of superdense matter, *Phys. Rep.* **61**, 71 (1980).
- [31] H. Stöcker and W. Greiner, High-energy heavy ion collisions: Probing the equation of state of highly excited hadronic matter, *Phys. Rep.* **137**, 277 (1986).
- [32] P. F. Kolb and U. W. Heinz, Hydrodynamic description of ultrarelativistic heavy ion collisions, in *Quark Gluon Plasma 3* (2003), pp. 634–714, *arXiv:nucl-th/0305084*.
- [33] J. Adams *et al.* (STAR Collaboration), Experimental and theoretical challenges in the search for the quark gluon plasma: The STAR Collaboration’s critical assessment of the evidence from RHIC collisions, *Nucl. Phys.* **A757**, 102 (2005).
- [34] K. Adcox *et al.* (PHENIX Collaboration), Formation of dense partonic matter in relativistic nucleus-nucleus collisions at RHIC: Experimental evaluation by the PHENIX collaboration, *Nucl. Phys.* **A757**, 184 (2005).
- [35] I. Arsene *et al.* (BRAHMS Collaboration), Quark gluon plasma and color glass condensate at RHIC? The Perspective from the BRAHMS experiment, *Nucl. Phys.* **A757**, 1 (2005).
- [36] B. B. Back *et al.*, The PHOBOS perspective on discoveries at RHIC, *Nucl. Phys.* **A757**, 28 (2005).
- [37] B. I. Abelev *et al.* (STAR Collaboration), Long range rapidity correlations and jet production in high energy nuclear collisions, *Phys. Rev. C* **80**, 064912 (2009).
- [38] G. Aad *et al.* (ATLAS Collaboration), Observation of a Centrality-Dependent Dijet Asymmetry in Lead-Lead Collisions at  $\sqrt{s_{\text{NN}}} = 2.77\text{ TeV}$  with the ATLAS Detector at the LHC, *Phys. Rev. Lett.* **105**, 252303 (2010).
- [39] S. Chatrchyan *et al.* (CMS Collaboration), Observation and studies of jet quenching in PbPb collisions at  $\sqrt{s_{\text{NN}}} = 2.76\text{ TeV}$ , *Phys. Rev. C* **84**, 024906 (2011).
- [40] S. Acharya *et al.* (ALICE Collaboration), Investigations of Anisotropic Flow Using Multiparticle Azimuthal Correlations in pp, p-Pb, Xe-Xe, and Pb-Pb Collisions at the LHC, *Phys. Rev. Lett.* **123**, 142301 (2019).
- [41] P. M. Chesler, Colliding Shock Waves and Hydrodynamics in Small Systems, *Phys. Rev. Lett.* **115**, 241602 (2015).
- [42] P. M. Chesler, How big are the smallest drops of quark-gluon plasma?, *J. High Energy Phys.* **03** (2016) 146.
- [43] C. Bierlich, G. Gustafson, and L. Lönnblad, Collectivity without plasma in hadronic collisions, *Phys. Lett. B* **779**, 58 (2018).
- [44] B. Blok, C. D. Jäkel, M. Strikman, and U. A. Wiedemann, Collectivity from interference, *J. High Energy Phys.* **12** (2017) 074.
- [45] K. Werner, Core-Corona Separation in Ultra-Relativistic Heavy Ion Collisions, *Phys. Rev. Lett.* **98**, 152301 (2007).
- [46] T. Pierog, B. Guiot, I. Karpenko, G. Sophys, M. Stefaniak, and K. Werner, EPOS 3 and air showers, *EPJ Web Conf.* **210**, 02008 (2019).
- [47] L. A. Anchordoqui, C. Garcia Canal, S. J. Scutto, and J. F. Soriano, Through the looking-glass with ALICE into the quark-gluon plasma: A new test for hadronic interaction models used in air shower simulations, *Phys. Lett. B* **810**, 135837 (2020).
- [48] L. A. Anchordoqui, C. G. Canal, F. Kling, S. J. Scutto, and J. F. Soriano, An explanation of the muon puzzle of ultrahigh-energy cosmic rays and the role of the Forward Physics Facility for model improvement, *J. High Energy Astrophys.* **34**, 19 (2022).
- [49] J. Manshanden, G. Sigl, and M. V. Garzelli, Modeling strangeness enhancements to resolve the muon excess in cosmic ray extensive air shower data, *J. Cosmol. Astropart. Phys.* **02** (2023) 017.
- [50] T. Pierog, S. Baur, H. Dembinski, R. Ulrich, and K. Werner, Collective hadronization and air showers: Can LHC data solve the muon puzzle ?, *Proc. Sci. ICRC2019* (2019) 387.
- [51] A. G. Knospe, C. Markert, K. Werner, J. Steinheimer, and M. Bleicher, Hadronic resonance production and interaction in p-Pb collisions at LHC energies in EPOS3, *Phys. Rev. C* **104**, 054907 (2021).
- [52] K. Werner, T. Pierog, B. Guiot, and J. Jahan, Recent developments in EPOS: Core–Corona effects in air showers?, *Phys. At. Nucl.* **84**, 1026 (2021).

- [53] K. Werner, On a deep connection between factorization and saturation: New insight into modeling high-energy proton-proton and nucleus-nucleus scattering in the EPOS4 framework, [arXiv:2301.12517](https://arxiv.org/abs/2301.12517).
- [54] T. Pierog and K. Werner, Muon Production in Extended Air Shower Simulations, *Phys. Rev. Lett.* **101**, 171101 (2008).
- [55] J. Matthews, A Heitler model of extensive air showers, *Astropart. Phys.* **22**, 387 (2005).
- [56] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner, EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider, *Phys. Rev. C* **92**, 034906 (2015).
- [57] T. Bergmann, R. Engel, D. Heck, N. N. Kalmykov, S. Ostapchenko, T. Pierog, T. Thouw, and K. Werner, One-dimensional hybrid approach to extensive air shower simulation, *Astropart. Phys.* **26**, 420 (2007).
- [58] P. Abreu *et al.* (Pierre Auger Collaboration), Interpretation of the depths of maximum of extensive air showers measured by the Pierre Auger observatory, *J. Cosmol. Astropart. Phys.* **02** (2013) 026.
- [59] H. P. Dembinski, Computing mean logarithmic mass from muon counts in air shower experiments, *Astropart. Phys.* **102**, 89 (2018).
- [60] L. Cazon (EAS-MSU, IceCube, KASCADE Grande, NEVOD-DECOR, Pierre Auger, SUGAR, Telescope Array, and Yakutsk EAS Array Collaborations), Working group report on the combined analysis of muon density measurements from eight air shower experiments, *Proc. Sci. ICRC2019* (2020) 214.
- [61] D. Soldin, Update on the combined analysis of muon measurements from nine air shower experiments, *Proc. Sci. ICRC2021* (2021) 349.
- [62] A. Andronic, P. Braun-Munzinger, K. Redlich, and J. Stachel, Hadron yields, the chemical freeze-out and the QCD phase diagram, *J. Phys. Conf. Ser.* **779**, 012012 (2017).
- [63] K. Werner, A. G. Knospe, C. Markert, B. Guiot, Iu. Karpenko, T. Pierog, G. Sophys, M. Stefaniak, M. Bleicher, and J. Steinheimer, Resonance production in high energy collisions from small to big systems, *EPJ Web Conf.* **171**, 09002 (2018).
- [64] J. Manninen and F. Becattini, Chemical freeze-out in ultrarelativistic heavy ion collisions at  $s(\text{NN})^{1/2} = 130$  and 200-GeV, *Phys. Rev. C* **78**, 054901 (2008).
- [65] F. Becattini and J. Manninen, Strangeness production from SPS to LHC, *J. Phys. G* **35**, 104013 (2008).
- [66] J. Aichelin and K. Werner, Is the centrality dependence of the elliptic flow  $v_2$  and of the average  $\langle p_T \rangle$  more than a core-corona effect?, *Phys. Rev. C* **82**, 034906 (2010).
- [67] J. Aichelin and K. Werner, Centrality dependence of strangeness enhancement in ultrarelativistic heavy ion collisions: A core-corona effect, *Phys. Rev. C* **79**, 064907 (2009); **81**, 029902(E) (2010).
- [68] T. Sjöstrand, S. Mrenna, and P. Z. Skands, PYTHIA 6.4 physics and manual, *J. High Energy Phys.* **05** (2006) 026.
- [69] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, An introduction to PYTHIA 8.2, *Comput. Phys. Commun.* **191**, 159 (2015).
- [70] H. J. Drescher, M. Hladik, S. Ostapchenko, T. Pierog, and K. Werner, Parton based Gribov-Regge theory, *Phys. Rep.* **350**, 93 (2001).
- [71] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, Parton fragmentation and string dynamics, *Phys. Rep.* **97**, 31 (1983).
- [72] A. Ortiz Velasquez, P. Christiansen, E. Cuautle Flores, I. A. Maldonado Cervantes, and G. Paić, Color Reconnection and Flowlike Patterns in  $pp$  Collisions, *Phys. Rev. Lett.* **111**, 042001 (2013).
- [73] C. Bierlich and J. R. Christiansen, Effects of color reconnection on hadron flavor observables, *Phys. Rev. D* **92**, 094010 (2015).
- [74] K. Werner, B. Guiot, I. Karpenko, and T. Pierog, Analysing radial flow features in p-Pb and p-p collisions at several TeV by studying identified particle production in EPOS3, *Phys. Rev. C* **89**, 064903 (2014).
- [75] R. Ulrich, T. Pierog, and C. Baus, The Cosmic Ray Monte Carlo package, CRMC, 10.5281/zenodo.4558705 (2021).
- [76] ALICE Collaboration, The ALICE definition of primary particles, CERN Technical Report No. ALICE-PUBLIC-2017-005, 2017.
- [77] K. Werner, F.-M. Liu, and T. Pierog, Parton ladder splitting and the rapidity dependence of transverse momentum spectra in deuteron-gold collisions at RHIC, *Phys. Rev. C* **74**, 044902 (2006).
- [78] A. M. Sirunyan *et al.* (CMS Collaboration), Measurement of the average very forward energy as a function of the track multiplicity at central pseudorapidities in proton-proton collisions at  $\sqrt{s} = 13$  TeV, *Eur. Phys. J. C* **79**, 893 (2019).
- [79] C. Bierlich *et al.*, Robust independent validation of experiment and theory: Rivet version 3, *SciPost Phys.* **8**, 026 (2020).
- [80] CMS Collaboration, Public Rivet routine for Ref. [78], <https://rivet.hepforge.org/analyses/> (2019).
- [81] From PYTHIA8 manual: Option ColourReconnection:mode=1 aka “The new QCD based scheme”.
- [82] B. I. Abelev *et al.* (STAR Collaboration), Strange particle production in p + p collisions at  $s^{*1/2} = 200$ -GeV, *Phys. Rev. C* **75**, 064901 (2007).
- [83] R. A. Lacey (STAR Collaboration), Long-range collectivity in small collision-systems with two- and four-particle correlations @ STAR, *Nucl. Phys.* **A1005**, 122041 (2021).
- [84] J. Adam *et al.* (STAR Collaboration), Azimuthal Harmonics in Small and Large Collision Systems at RHIC Top Energies, *Phys. Rev. Lett.* **122**, 172301 (2019).
- [85] X. Zhu (STAR Collaboration), Strangeness production in d + Au collisions at  $s(\text{NN})^{1/2} = 200$ -GeV in STAR, *Nucl. Phys.* **A830**, 845C (2009).
- [86] B. I. Abelev *et al.* (STAR Collaboration), Systematic measurements of identified particle spectra in  $pp$ ,  $d^+$  Au and Au + Au collisions from STAR, *Phys. Rev. C* **79**, 034909 (2009).
- [87] A. Botti, I. Goos, M. Perlin, and T. Pierog (to be published).
- [88] S. Ostapchenko, On the re-summation of enhanced Pomeron diagrams, *Phys. Lett. B* **636**, 40 (2006).

- [89] S. Ostapchenko, Total and diffractive cross sections in enhanced Pomeron scheme, *Phys. Rev. D* **81**, 114028 (2010).
- [90] F. Riehn, R. Engel, A. Fedynitch, T. K. Gaisser, and T. Stanev, Hadronic interaction model Sibyll 2.3d and extensive air showers, *Phys. Rev. D* **102**, 063002 (2020).
- [91] F. Gesualdi, H. Dembinski, K. Shinozaki, D. A. Supanitsky, T. Pierog, L. Cazon, D. Soldin, and R. Conceição, On the muon scale of air showers and its application to the AGASA data, *Proc. Sci. ICRC2021* (2021) 473.
- [92] R. Abbasi *et al.* (IceCube Collaboration), Density of GeV muons measured with IceTop, *Proc. Sci. ICRC2021* (2021) 342.
- [93] W. Apel *et al.* (KASCADE-Grande Collaboration), Probing the evolution of the EAS muon content in the atmosphere with KASCADE-Grande, *Astropart. Phys.* **95**, 25 (2017).
- [94] A. Bogdanov, D. Gromushkin, R. Kokoulin, G. Mannocchi, A. Petrukhin, O. Saavedra, G. Trincherro, D. Chernov, V. Shutenko, and I. Yashin, Investigation of the properties of the flux and interaction of ultrahigh-energy cosmic rays by the method of local-muon-density spectra, *Phys. At. Nucl.* **73**, 1852 (2010).
- [95] A. Bogdanov, R. Kokoulin, G. Mannocchi, A. Petrukhin, O. Saavedra, V. Shutenko, G. Trincherro, and I. Yashin, Investigation of very high energy cosmic rays by means of inclined muon bundles, *Astropart. Phys.* **98**, 13 (2018).
- [96] A. Aab *et al.* (Pierre Auger Collaboration), Measurement of the Fluctuations in the Number of Muons in Extensive Air Showers with the Pierre Auger Observatory, *Phys. Rev. Lett.* **126**, 152002 (2021).
- [97] A. Aab *et al.* (Pierre Auger Collaboration), Direct measurement of the muonic content of extensive air showers between  $2 \times 10^{17}$  and  $2 \times 10^{18}$  eV at the Pierre Auger Observatory, *Eur. Phys. J. C* **80**, 751 (2020).
- [98] J. A. Bellido, R. W. Clay, N. N. Kalmykov, I. S. Karpikov, G. I. Rubtsov, S. V. Troitsky, and J. Ulrichs, Muon content of extensive air showers: Comparison of the energy spectra obtained by the Sydney University Giant Air-shower Recorder and by the Pierre Auger observatory, *Phys. Rev. D* **98**, 023014 (2018).
- [99] A. Glushkov, M. Pravdin, and A. V. Saburov (Yakutsk Collaboration) (private communication).
- [100] J. F. Soriano, L. A. Anchordoqui, T. C. Paul, and T. J. Weiler, Probing QCD approach to thermal equilibrium with ultrahigh energy cosmic rays, *Proc. Sci. ICRC2017* (2018) 342.
- [101] J. Alvarez-Muniz, L. Cazon, R. Conceição, J. D. de Deus, C. Pajares, and M. Pimenta, Muon production and string percolation effects in cosmic rays at the highest energies, [arXiv:1209.6474](https://arxiv.org/abs/1209.6474).
- [102] G. R. Farrar and J. D. Allen, A new physical phenomenon in ultra-high energy collisions, *EPJ Web Conf.* **53**, 07007 (2013).
- [103] L. A. Anchordoqui, H. Goldberg, and T. J. Weiler, Strange fireball as an explanation of the muon excess in Auger data, *Phys. Rev. D* **95**, 063005 (2017).
- [104] D. LaHurd and C. E. Covault, Exploring potential signatures of QGP in UHECR ground profiles, *J. Cosmol. Astropart. Phys.* **11** (2018) 007.
- [105] O. Adriani *et al.* (LHCf Collaboration), Measurement of forward photon production cross-section in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the LHCf detector, *Phys. Lett. B* **780**, 233 (2018).
- [106] O. Adriani *et al.* (LHCf Collaboration), Measurement of inclusive forward neutron production cross section in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the LHCf Arm2 detector, *J. High Energy Phys.* **11** (2018) 073.
- [107] O. Adriani *et al.* (LHCf Collaboration), LHCf plan for proton-oxygen collisions at LHC, *Proc. Sci. ICRC2021* (2022) 348.