Constraints from the first LHC data on hadronic event generators for ultra-high energy cosmic-ray physics

David d’Enterria
ICREA & ICC-UB, Univ. de Barcelona, 08028 Barcelona, Catalonia
CERN, PH Department, 1211 Geneva, Switzerland

Ralph Engel, Tanguy Pierog
Karlsruhe Institut of Technology, Postfach 3640, 76021 Karlsruhe, Germany

Sergey Ostapchenko
NTNU, Inst. for Fysik, 7491 Trondheim, Norway
D.V. Skobeltsyn Inst. Nuc. Phys, Moscow State Univ., 119992 Moscow, Russia

Klaus Werner
SUBATECH, 4 rue Alfred Kastler, BP 20722, 44307 Nantes Cedex 3, France

Abstract
The determination of the primary energy and mass of ultra-high-energy cosmic-rays (UHECR) generating extensive air-showers in the Earth’s atmosphere, relies on the detailed modeling of hadronic multiparticle production at center-of-mass (c.m.) collision energies up to two orders of magnitude higher than those studied at particle colliders. The first Large Hadron Collider (LHC) data have extended by more than a factor of three the c.m. energies in which we have direct proton-proton measurements available to compare to hadronic models. In this work we compare LHC results on inclusive particle production at energies $\sqrt{s} = 0.9, 2.36, \text{and} 7 \text{ TeV}$ to predictions of various hadronic Monte Carlo (MC) models used commonly in cosmic-ray (CR) physics ($\text{qgsjet}$, $\text{epos}$ and $\text{sibyll}$). As a benchmark with a standard collider physics model we also show $\text{pythia}$ (and $\text{phojet}$) predictions with various parameter settings. While reasonable overall agreement is found for some of the MC, none of them reproduces consistently the $\sqrt{s}$ evolution of all the observables. We discuss implications of the new LHC data for the description of cosmic-ray interactions at the highest energies.

1. Introduction
In astroparticle physics, the identification and understanding of the sources of high-energy cosmic rays is one of the most important open problems. There is an increasing observational support for the hypothesis of particles being accelerated at shock fronts in supernova remnants. These cosmic rays are expected to populate the observed cosmic-ray spectrum up to $10^{15} - 10^{17} \text{ eV}$ [1]. The sources of cosmic rays beyond that energy are not known and, due to the extreme requirements needed to reach such energies in acceleration processes, subject to various speculations including extensions of the Standard Model of particle physics [2, 3, 4].

A compilation of recent flux measurements of cosmic rays is shown in Fig. 1. The power law of the energy spectrum of cosmic rays changes at about $10^{15.5} \text{ eV}$. This feature is known as the knee, the origin of which is still under debate. Theoretical explanations have been put forward based on a change of slope in the source spectra, effects of leakage from the Galaxy, the assumption of changes in hadronic interactions or the production of exotic new particles [4]. The ankle in the energy spectrum at $\sim 10^{18.5} \text{ eV}$ is often assumed to be the imprint of the change from galactic to extragalactic sources [7, 11] or, alternatively, a signature of the propagation of extragalactic cosmic rays through the microwave background radiation [8, 9, 10]. Finally the suppression of the flux at ultra-high energies could be due

Preprint submitted to Astroparticle Physics
| Energy (eV/particle) |
|---------------------|
| 10^13               |
| 10^14               |
| 10^15               |
| 10^16               |
| 10^17               |
| 10^18               |
| 10^19               |
| 10^20               |

Figure 1: Data on the all-particle flux of cosmic rays. The flux has been scaled by $E^{-2.5}$ to make the features clearly visible. The axis at the top indicates the equivalent c.m. energy if the cosmic ray particles were protons. The nominal collider c.m. energy for different accelerators is also shown. References to the data can be found in [5].

Knowing the elemental composition of cosmic ray particles arriving at Earth is of crucial importance to understand the production and propagation of cosmic rays. Unfortunately, cosmic rays can be measured only indirectly above an energy of $10^{14}$ eV through the cascades of secondary particles, called extensive air-showers (EAS), that they produce in the atmosphere (for a recent review, see [5]). Only by simulating the generation of EAS and comparing the predictions with measurements one can draw conclusions on the primary mass composition of the arriving particles [18]. With the operation of modern large-scale experiments the reliability of air-shower simulations has become the source of the largest systematic uncertainty in the interpretation of cosmic-ray data [19, 20, 21, 22, 23, 24]. While the electroweak interaction processes are reasonably well understood, modeling of hadronic multiparticle production is subject to large theoretical uncertainties that are difficult to estimate [25, 26, 27].

In this context it is not surprising that in some speculative scenarios the knee in the cosmic-ray spectrum has been related to a change in the characteristics of hadronic interactions. In an air-shower only a part of the energy of the primary particle is transferred to electromagnetic particles and low-energy muons that can be detected easily. Assuming that, for example, a rapidly increasing fraction of very high energy muons is produced in interactions just above the knee energy, the non-detection of the energy of these muons would result in a systematically incorrect reconstruction of the primary energy and lead to a knee in the observed spectrum even if there is not such a break in the primary CR flux (see, for example, [28, 29, 30, 31]). These models could not be constrained by Tevatron data as the $2 \times 10^{15}$ eV
equivalent (fixed-target) energy of the Tevatron collider is below the knee.

The Large Hadron Collider (LHC) at the CERN laboratory allows us to access for the first time the energy region above the knee in the laboratory. Therefore an analysis of inclusive particle data taken at the LHC is particularly interesting for constraining existing hadronic interaction models and for testing possible new mechanisms of hadron production \[32\]. Data from LHC experiments published so far have mostly been taken with detectors covering the central phase space region in pseudorapidity ($|\eta| \lesssim 2.5$). This region is most easily accessible in collider experiments and is also the region of the highest rapidity-density of produced particles. On the other hand, since the number of particles in an air-shower is roughly proportional to the energy of the primary particle, the most energetic outgoing particles of an interaction, emitted in the very forward region of a collider experiment – such as in diffractive interactions – are the most important ones for understanding air-showers. While waiting for the results from forward detectors at the LHC \[33, 34, 35\], important constraints can be derived from the already measured central particle production.

In this paper we compare the predictions of several representative hadronic interaction models with single-particle inclusive observables measured at midrapidity at the LHC. We focus on the following observables: (i) the pseudorapidity density of charged particles at midrapidity $dN_{ch}/d\eta|_{\eta=0}$, (ii) the average transverse momentum of hadrons $\langle p_{\perp} \rangle$; and (iii) the event-by-event distribution of the charged multiplicity, $P(N_{ch})$. Implications of the data-theory comparisons are discussed and further observables of direct relevance to EAS simulations, measurable with central and forward detectors at the LHC, are pointed out.

The paper is organized as follows. In Section 2 we recall the main theoretical ingredients of the hadronic interaction models used in our study and discuss their differences. In Section 3 we summarize the recently collected inclusive charged hadron data at the LHC from the ALICE \[36, 37\] and CMS \[38, 39\] experiments as well as results from previous colliders (ISR \[40\], UA1 \[41\], UA5 \[42\] and CDF \[43, 44\]). [Unfortunately, we could not include in this study the latest results from the ATLAS experiment \[45\], which appeared public only when this analysis was being finished.] We discuss the different event trigger selection criteria applied in each experiment as those affect somewhat the data-model comparisons. In Section 4 we compare the measurements to the various event generators, and discuss the results in the context of high energy cosmic-ray interactions with the atmosphere in Section 5. We summarize our main findings in Section 6.

2. High-energy hadronic interaction models

Calculating predictions for the bulk of produced particles in hadronic interactions is not possible yet within first-principles Quantum-Chromodynamics (QCD). One has to resort to phenomenological models that combine fundamental principles of quantum field theory – such as unitarity, analyticity and crossing – together with perturbative QCD (pQCD) predictions including phenomenological fits (e.g. accounting for the parton-to-hadron fragmentation) to experimental hadron spectra.

General-purpose hadronic interaction models used in high-energy physics (HEP), such as PYTHIA \[46\], HERWIG \[47\] and SHERPA \[48\], are developed to learn and interpret the data measured in accelerator experiments with an emphasis on hard-scattering measurements (signals and backgrounds) rather than on the bulk of hadron production at lower transverse momenta. The predictions of models can be adjusted by a large number of parameters that encode fundamental physics quantities on the one hand, and phenomenological assumptions or simplifications on the other. These models are typically optimized for processes that can be calculated in perturbation theory and other assumptions needed for generating complete final states for hadronic interactions are often kept at a more simple level. In addition, these models are mainly designed for proton, pion or photon interactions and not suited for nuclear interactions relevant for CRs collisions with air nuclei in the upper atmosphere.

In contrast, interaction models commonly used in cosmic-ray physics such as QGSJET01 \[49, 50\], QGSJETII \[51, 52, 53\] and SIBYLL \[54, 55, 56\], are supposed to predict hadronic interactions as realistically as possible with the emphasis
on reproducing existing accelerator measurements and providing a reasonable extrapolation to higher energy and to phase-space regions where no data are available. On the one hand, there is typically only one “optimal” parameter set for each of the models and simplifications are made in the implementation of known QCD predictions, in particular, if they are not directly relevant to the production of high energy secondaries. On the other hand, more sophisticated models of soft particle production are implemented and great care is given to the relation of the total, elastic and inelastic cross sections to particle production.

In between these two generic categories there are models such as PHOJET \cite{57, 58, 59}, DPMJET \cite{60, 61} and EPS \cite{62}, which are designed to be more universal and approach the sophistication of HEP models regarding some aspects of hard processes. At the same time the parameters of these models are tuned to describe a large set of accelerator data at various energies and are not supposed to be changed for tuning for different experiments.

In the following we will consider the models DPMJET, QGJET01 and II, SIBYLL 2.1, and EPS v1.99 which are often used in cosmic-ray simulations \cite{18, 5} and the PYTHIA Monte Carlo versions 6.4 and 8 as a benchmark HEP model. The PYTHIA predictions are shown mainly for reference as this model is used in all publications on minimum-bias measurements at the LHC. Since the proton-proton interaction generator in DPMJETIII is identical to PHOJET we will use the latter in our comparisons.

Providing a description of the physics implemented in these models would be beyond the scope of this work. Here we will rather recall only the basic concepts behind the modeling particle production.

2.1. Modeling of central multiparticle production

The inclusive production of particles in high-energy hadronic collisions receives contributions from “soft” and “hard” interactions between the partonic constituents of the colliding hadrons. Soft (resp. hard) processes involve mainly $t$-channel partons of virtualities $q^2$ typically below (resp. above) a scale $Q_s^2$ of a few GeV$^2$. Soft scatterings give rise to production of hadrons with low transverse momenta $p_T$ and dominate hadronic collisions at low energies ($\sqrt{s} \lesssim 20$ GeV).

Although soft processes have a virtuality scale not far from $\Lambda_{QCD} \approx 0.2$ GeV and thus cannot be treated within pQCD, predictions based on basic quantum field-theory principles such as unitarity and analyticity of scattering amplitudes, as implemented in the Gribov’s Reggeon Field Theory (RFT) \cite{63}, give a decent account of their cross sections in terms of the exchange of virtual quasi-particle states (Pomerons and Reggeons). At increasing energies, the Pomeron contribution, identified perturbatively with a multi-gluon exchange, dominates over those from secondary Reggeons (virtual mesons). In particular, the soft Pomeron exchange, of colourless nature, is responsible for particle production in peripheral hadronic collisions and, in particular, for diffractive dissociation which accounts for a noticeable fraction of the total inelastic cross section all the way up to asymptotic energies.

(Semi)hard parton-parton scatterings dominate the inelastic hadron production cross-sections for c.m. energies above a few hundreds of GeV. Hard processes with large $|q^2| \gg \Lambda_{QCD}^2$ can be treated within pQCD in a collinear-factorized approach where one sums up Feynman diagrams of the underlying parton-parton (quarks and gluons) scatterings, which give the leading logarithmic contribution, with each power of the strong coupling constant $\alpha_s(q^2)$ being multiplied by a collinear logarithm $\log q^2$. The scattered quarks and gluons produce then collimated bunches of final-state hadrons (jets) in a branching process dominated by perturbative parton splittings described by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations \cite{64, 65, 66}, and by non-perturbative hadronization when the parton virtuality is below $O(1$ GeV). The RFT approach can also be generalized to include hard processes via “cut (hard) Pomerons” diagrams. The physical picture is that of a colour flux tube, which is in the simplest case treated as a classical string that subsequently fragments into hadrons.\footnote{Strictly speaking one cannot refer to individual partons in case of soft interactions because of confinement. Still it is a practical approach taken in all event generators to extend some of the concepts of hard interactions to the soft domain.}
Historically, Monte Carlo (MC) event generators of high-energy hadronic collisions have evolved either starting up from the RFT approach, exemplified e.g. in the original Dual Parton Model (DPM) [67], extended with a leading-logarithmic pQCD description for high-\( p_T \) production (based on cut-Pomerons) – such as in the PHOJET [57, 58], QGSJET01 and II [49, 51, 52, 53], SIBYLL [54, 55, 56], NYXUS [68, 69], EPOS [62] and DPMJET [60, 61] cases – or they started from a purely collinear-factorized framework – such as in e.g. general-purpose MCs like PHOJET [46] – complemented with an add-on model for truly soft [70] and diffractive [71] scatterings. Thus, on the one hand, the RFT approaches try to extend a consistent framework based on Pomeron degrees of freedom to the hard regime. On the other, the collider MCs contain a description based on partonic degrees of freedom (with scattering cross sections dumped in the infrared, below a “tunable” semihard scale) with soft and diffractive scatterings incorporated in a more or less ad hoc way. In both approaches the final non-perturbative transition of partons to hadrons is modeled based on the ideas of the Lund string fragmentation model [72]. At increasingly higher \( \sqrt{s} \), in both frameworks one has to account for multiple scattering processes between the colliding hadrons, namely one has to include multi-Pomeron exchanges and/or multiple hard scattering processes.

In the RFT framework, the single Pomeron (P) exchange amplitude is characterized by a power-like energy dependence, \( f^P(x, t) \propto x^{-\alpha_0(t)} \), with the Pomeron intercept \( \alpha_0(0) \sim 1 \) leading to a corresponding energy rise of the total cross section \( \sigma_{\text{tot}} = \frac{1}{t} \text{Im} f^P(x, 0) \), which asymptotically violates the so-called Froissart bound (\( \alpha_{\text{tot}} < c \log^2 s \)) [73]. Accounting for eikonal multi-Pomeron exchanges, the cross sections are unitarized, i.e. \( \sigma_{\text{tot}, \text{inel}} \propto \log^2 s \), although due to the Abramovskii-Gribov-Kancheli (AGK) cancellations [74] such multi-Pomeron configurations give zero contribution to inclusive particle spectra. Thus, the total soft charged particle density produced at midrapidity follows the energy-dependence defined by a single Pomeron exchange contribution:

\[
\frac{dN_{\text{ch}}(s, \eta)}{d\eta} \bigg|_{\eta=0} \propto \frac{\text{Im} f^P(x, 0)}{s} \frac{s^{\Delta_\text{P}}}{\log^2 s}, \quad \text{with } \Delta_\text{P} \equiv \alpha_0(0) - 1 \sim 0.1.
\]  

(1)

In pure DGLAP-based models, the central pseudo-rapidity particle density is proportional to the inclusive jet cross section which is given by the parton distributions functions (PDFs) and parton-parton scattering cross sections:

\[
\sigma_{pp}^{\text{jet}}(s, Q_0^2) = \int dx_1 \int dx_2 \sum_{i,j=q,g} f_{i/q}(x_1, p_{T\perp}^2) f_{j/g}(x_2, p_{T\perp}^2) \times \frac{d\sigma_{ij}(x_1, x_2, p_{T\perp}^2)}{dp_{T\perp}^2} \Theta(s - 4p_{T\perp}^2).
\]  

(2)

The hard cross section is divergent in the limit \( p_{T\perp} \rightarrow 0 \) and one needs to introduce a \( p_{T\perp} \)-cutoff \( Q_0 \) to indicate the regime of validity of the perturbative approximation. At increasingly larger \( \text{c.m.} \) energies, one needs to account for multi-parton scatterings and saturation effects. On the one hand, the cross section predicted by the regularized processes exceeds the total inelastic cross section, indicating that several (or multiple) hard scatterings occur per collision. On the other, for decreasing but still perturbative \( p_{T\perp} \) values, the integrals receive major contributions from the region of low parton fractional momenta \( (x = p_{T\perp}/p_{\text{hadron}}) \), where the dominant gluon distribution rises roughly as \( f_{g/p}(x, p_{T\perp}^2) \sim x^{-\Delta_\text{glu}} \) with \( \Delta_\text{glu} \approx 0.3 \). After integrating above the \( p_{T\perp} \)-cutoff \( Q_0 \), one obtains an energy-dependence of the corresponding hard central charged hadron densities of the type:

\[
\frac{dN_{\text{ch}}(s, \eta)}{d\eta} \bigg|_{\eta=0} \sim \frac{\sigma_{pp}^{\text{jet}}(s, Q_0^2)}{\sigma_{\text{inel}}(s)} \frac{Q_0^{\Delta_\text{hard}}}{\log^2 s}, \quad \text{with } \Delta_\text{hard} \approx 0.3.
\]  

(3)

Clearly, the fast growth of the gluon densities at low \( x \) results in the hard part of the particle density (\( \propto s^{\Delta_{\text{glu}}}, \Delta_{\text{hard}} \approx 0.3 \)) to rise with energy much faster than for soft processes (\( \propto x^{\Delta_\text{P}}, \Delta_\text{P} \sim 0.1 \)). However, at sufficiently small \( x \), the number of gluons is so large that new parton multiscattering phenomena have to be accounted for. First, non-linear (gg fusion) effects become important in the PDFs themselves, saturating their growth as \( x \rightarrow 0 \) [75]. The strength of these effects is controlled by the “saturation scale” \( Q_0 \) at which parton branching and fusion processes start to compensate each other. Second, the probability to have simultaneous scatterings of the constituents of the colliding hadrons also increases leading to multiple parton interactions (MDP) in a single collision. In many MC generators one effectively mimics saturation effects by introducing some energy dependence to the infrared \( p_{T\perp} \)-cutoff: \( Q_0^2 = Q_0^2(s) \).
In particular, choosing \( Q_0^2(s) \sim s^\epsilon \) as done e.g. in \texttt{PYTHIA}, the power-law changes as

\[
\frac{dN_{ch}(s, \eta)}{d\eta} \bigg|_{\eta=0} \sim s^{\epsilon_{sat} - \epsilon},
\]

which reduces the relative role of the hard contributions compared to the “linear” predictions for fixed \( Q_0 \) cutoff, Eq. (3). [We note that stricto sensu in \texttt{PYTHIA} there is extra particle production from “purely” soft scatterings [70] which is not accounted for by this simple expression.]

Not only the particle multiplicities but also their transverse momenta are sensitive to the underlying parton dynamics. In soft processes, the produced hadrons have typical transverse momenta in the sub-GeV range (\( p_T^{soft} \sim 0.3 - 0.4 \) GeV), without a pronounced energy-dependence for the average \( p_T \). For the hard contribution, due to the fast drop (\( \propto 1/p_T^2 \)) of the differential jet production cross-section, the peak of the perturbative production sits at gluons whose transverse momentum is close to the cutoff, \( p_T \sim Q_0 \), producing jets of a few GeV which are often called “minijets” since they are perturbatively calculable but often not experimentally observable over the soft hadronic background. Due to the faster energy rise of the hard contribution compared to the soft one – cf. Eqs. (1) and (3) – average hadron transverse momenta rise with \( \sqrt{s} \) from \( p_T \sim p_T^{soft} \) to \( p_T \sim Q_0 \). As a matter of fact, the energy-dependence of the average \( p_T \) is expected to be even steeper when parton saturation effects come into play – due to the suppression of semi-hard parton production in the “dense” saturated region, \( q_T^2 \sim p_T^2 < Q_{sat}^2(s) \) and due to the rise of the saturation scale \( Q_{sat}^2 \propto \log(1/x) \propto \log(\sqrt{s}) \) itself. The effective value of \( \langle Q_{sat}^2 \rangle \) controls thus the mean transverse momentum \( \langle p_T \rangle \) of a significant part of the finally produced hadrons after parton fragmentation. It is noteworthy that the MC procedure of mimicking the saturation and MPI effects via an energy-dependent \( p_T \)-cutoff changes somehow this picture. Having the soft part of the interaction unchanged and reducing the minijet contribution – cf. Eqs. (4) and (5) – the \( p_T \)-rise from \( p_T^{soft} \) to \( Q_0(s) \) proceeds slower than in the linear case, up to very high energies.

All in all, it is clear that both the total particle pseudorapidity density \( dN_{ch}/d\eta|_{\eta=0} \) and the average \( p_T \) of produced hadrons are sensitive to the soft and semi-hard dynamics, to the non-linear parton (saturation) effects, and to their implementations in the models of high-energy hadronic scattering.

### 2.2. The \texttt{PYTHIA} event generator

The basic ingredients of the \texttt{PYTHIA} event generator are leading-order pQCD \( 2 \to 2 \) matrix elements, augmented by initial- and final-state parton showers (ISR and FSR, respectively), folded with parton distribution functions (interfaced here via the \texttt{HERAPDF} v5.8.2 package [76]) on the initial-state and the Lund string model [72] to describe the final parton-to-hadron fragmentation. The infrared \( 1/p_T^2 \) divergence of the hard (multi)parton cross sections is regularized by a cutoff \( Q_0 \), such that \( 1/p_T^2 \to 1/(p_T^2 + Q_0^2)^2 \). The infrared cutoff depends on the hadron-hadron c.m. energy: \( Q_0^2(s) = Q_0^2(s_0) \cdot (s/s_0)^\epsilon \), where \( Q_0(s_0) \) is a reference value at a given c.m. energy \( \sqrt{s_0} \), e.g. \( Q_0 \approx 2 \) GeV at \( \sqrt{s_0} = 1.8 \) TeV is preferred by the Tevatron data [77, 78]. Other non-perturbative ingredients of \texttt{PYTHIA} include a Regge-based modeling of diffractive processes [71], plus a model for the underlying-event (UE) issuing from multiparton interactions (MPI), soft scatterings, and beam-remnants [70]. Multiple parton collisions are treated “perturbatively” – i.e. based on the eikonalization of hard scattering contributions – but require a non-perturbative ansatz for the impact-parameter profile of the colliding hadrons.

In this work we use the \texttt{PYTHIA} event generator in two flavours: the Fortran version 6.422 [46] (Nov. 2009), as well as the C++ version \texttt{PYTHIA} 8.130 [79]. Both codes include the newest description of multiple parton interactions based on the concept of “interleaved evolution” given by \( p_T \)-ordered showers [80, 81] which accounts well for correlations between the hard-scattering and the underlying-event components seen in the Tevatron data. Compared to \texttt{PYTHIA} 6.4, the MPI, ISR and FSR in \texttt{PYTHIA} 8 are all interleaved, and the treatment of diffraction has improved: a diffractive system is viewed as a Pomeron-proton collision, including hard scatterings subject to all the same ISR/FSR and MPI dynamics as for a “normal” parton-parton process.
We consider three different “tunes” of the non-perturbative ingredients (IS and FS showering, UE, beam-remnants, FS colour-reconnection, and hadronization) of the two versions of \textsc{pythia}. For \textsc{pythia} 6.4 we selected (via the \textsc{PYTUNES} switch) the so-called Perugia-0 (320) \cite{82, 83} and Atlas-CSC (306) \cite{84, 85} tunes. Comparisons of LHC data to other existing \textsc{pythia} 6 tunes can be found in the literature (see e.g. \cite{86}). For \textsc{pythia} 8 we use the default “tune 1”. The chosen settings describe one or more of the minimum-bias (MB), underlying-event (UE), and/or Drell-Yan (DY) data sets at the Tevatron \cite{87, 82, 83}, and are reasonably complementary on a number of important points, as illustrated in Table 1.

| Version  | Tuning (PYTUNES) | Diffraction | $Q_0$ cutoff at $\sqrt{s} = 1.8$ TeV | $Q_0$ scaling power $\epsilon$ | PDF | Proton profile | FS colour reconnection | Exp. constraints (Tevatron) |
|----------|------------------|-------------|------------------------------------|-------------------------------|-----|----------------|------------------------|---------------------------|
| 6.422    | Perugia 0 (320)  | simple      | 2.0 GeV                            | 0.13                          | CTEQ5L | ExpOfPower | weak                  | UE, MB, DY               |
| 6.422    | Atlas-CSC (306)  | simple      | 1.9 GeV                            | 0.11                          | CTEQ6L | double-Gauss | weak                  | MB, DY                   |
| 8.130    | Tune 1           | improved    | 2.15 GeV                           | 0.08                          | CTEQ5L | double-Gauss | weak                  | MB                        |

Table 1: Comparison of the various ingredients controlling the non-perturbative and semi-hard dynamics in the three \textsc{pythia} models used in this work. See text for details.

Since the three tunes share the same MPI dynamics, the single most important parameter to extrapolate the hadron multiplicity from Tevatron to LHC energies is the exponent $\epsilon$ which controls the dependence of the MPI infrared cutoff scale on the collision energy, see Eq. (4). Given that a single value of $Q_0$ is used to regularize both the hard scattering and the MPI in the event, and that the generation of additional parton-parton interactions in the underlying event is suppressed below $Q_0$, a higher scaling power of the infrared cut-off implies a slower increase of the overall hadronic activity. The default (“untuned”) \textsc{pythia} 6.4 energy scaling $\sqrt{s}$, $\epsilon = 0.08$, was chosen to follow the scaling of the total cross section ($\Delta P \sim 0.08$). However, most of the subsequent tunes needed to significantly increase the scaling power value to $\epsilon \approx 0.13$, in order to find a better agreement with the $\sqrt{s}$-dependence of the UE and MB measured from $\bar{p}p$S and Tevatron \cite{82, 83}. One of the first conclusions from the MB LHC data, as we will see next, is that this exponent needs to be lowered again.

For the initial-state Atlas-CSC uses the CTEQ6L parton densities \cite{88} plus a double-Gaussian transverse distribution for the partons in the proton, whereas Perugia uses the default CTEQ5L PDFs \cite{89} and a smoother proton form-factor (exponential-of-power, $\exp(-r^n)$, with exponent $n = 1.7$) which decreases the multiplicity fluctuations. The Tune-1 of \textsc{pythia} 8 is closer to Atlas-CSC in terms of the energy-scaling of the pQCD cut-off (it features the smallest exponent, $\epsilon = 0.08$, of the three models) and of the proton profile but uses the default CTEQ5L PDFs. For the final-state, the three tunes have weak final-state colour reconnections, which increase the final particle multiplicities, compared to models with stronger colour correlations.

2.3. Models based on Reggeon Field Theory (RFT)

The RFT-based MCs considered here start off from a construction of the hadron-hadron elastic scattering amplitude to determine the total and elastic cross sections. Experimental data allow one to fix the basic model parameters, such as the intercept and slope of the Pomeron Regge trajectory, the Pomeron-hadron couplings, etc. The inelastic events are then understood in terms of cut Pomerons which correspond to colour flux tubes (treated as strings) extended between constituent partons of the colliding hadrons. Final hadrons are produced in the fragmentation of the strings. The different models used in this work (Table 2) differ in various approximations for the collision configurations (e.g. the distributions for the number of cut Pomerons and for the energy-momentum partition among them), the treatment of diffractive and perturbative contributions as well as of high parton density effects, and the details of particle production from string fragmentation.

\textsuperscript{2}Parameter $\text{PARP}(90) = 2 \cdot \epsilon = 0.16$. 

7
Table 2: Comparison of the main ingredients controlling the non-perturbative and semi-hard dynamics present in the RFT-based event generators used in this work.

| Model (version) | Diffraction                    | $Q_0$ cutoff (GeV) | Saturation effects                      | Other                                      |
|-----------------|--------------------------------|-------------------|----------------------------------------|-------------------------------------------|
| PHOJET 1.12     | 2-channel eikonal              | 2.5               | via running $Q_0(s)$                    | FS+hadronization via PYTHIA 6.115         |
| QGSJET01        | quasi-eikonal                  | 2.0               | none (flat low-x PDFs)                 | –                                         |
| SYBILL 2.1      | 2-channel eikonal              | 1.0               | via running $Q_0(s)$                    | –                                         |
| QGSJETII        | cut enhanced graphs            | 1.6               | enhanced $P$-graphs                    | –                                         |
| EPS v1.99       | diffractive Pomeron            | 2.0               | parametrized                           | final-state collective flow               |

Whereas the Regge-Gribov approximation is applied to hadrons as interacting objects in the case of SYBILL, QGSJET and PHOJET, it is extended to include partonic constituents in EPS. In the latter case, this allows the implementation of energy sharing between the different constituents of a hadron at amplitude level and suppresses final states with very large particle multiplicity.

Except in EPS, the basic implementation of diffraction dissociation follows the Good-Walker formalism\cite{90}, where the colliding hadrons are represented by superpositions of elastic scattering eigenstates which undergo different absorption during the collision. In all these models one uses two eigenstates per hadron (2-channel eikonal approach); in QGSJET01 the second eigenstate is “passive” (has zero coupling to Pomerons) which leads to the so-called quasi-eikonal approach\cite{91}. While in SYBILL v2.1 this mechanism is associated with high-mass diffraction, in the others it is restricted to the production of low-mass diffractive states. High-mass diffraction is treated empirically in QGSJET01, accounting for single diffraction only. It is based on diffractive cuts of triple-Pomeron and Pomeron-loop graphs in PHOJET, and it is based on all-order resummation of cut enhanced $P$-graphs in QGSJETII. In EPS v1.99, a special kind of Pomeron is used to define a diffractive event. Depending on each event configuration it can be non-diffractive, low mass diffraction without central particle production, or high mass diffraction. In all the cases, each initial hadron can be in an excited state.

All the models account for multiple soft and (semi)hard production processes via multi-Pomeron interactions. The effects of high parton density at small $x$ and the treatment of the hadronic remnants are implemented differently in the different generators. SYBILL uses an energy-dependent transverse momentum cutoff for minijet production, $Q_0(s) \sim Q_0 + \exp(\log s)$, based on the geometric criterion that there cannot be more gluons in a hadron than would fit in a given transverse area. This cutoff is independent of the impact parameter of the collision. PHOJET uses also a $\sqrt{s}$-dependent cutoff, of the form $Q_0(s) \sim Q_0 + C \cdot \log(\sqrt{s})$. QGSJET01 is based on old parton densities of the pre-HERA era and the (artificial) flatness of low-$x$ PDFs, $x g(x, Q_0^2) \propto$ const, allows one to partly avoid the problem of high parton densities. This issue is addressed more realistically in the successor model, QGSJETII, which is based on a re-summation of enhanced diagrams corresponding to Pomeron-Pomeron interactions\cite{51, 93, 94}, which lead to impact-parameter and density-dependent parton saturation for soft processes\cite{95}. In EPS, non-linear effects are introduced phenomenologically via the elastic and inelastic fusion of parton-ladders. The elastic process provides screening, therefore a reduction of total and inelastic cross sections. The inelastic process affects particle production, in particular transverse momentum spectra. Finally, in contrast to other MCs, the EPS version 1.99 used in this paper includes also a parametrized\cite{96} collective expansion in the final state for all colliding systems including $p$-$p$ at LHC energies.

\footnote{The latest EPS developments (EPS2)\cite{96, 97} include a full hydrodynamical evolution.}
Proton-(anti)proton interactions at colliders can be roughly divided into three categories according to the overall event topology: elastic, diffractive and non-diffractive processes. Elastic events are usually missed by the experiments unless specific detectors (Roman Pots) are installed inside the tunnel. Inelastic interactions are often collected with a generic minimum bias (MB) trigger that accepts a large fraction of the particle production cross section by requiring a minimum activity in one or various detectors. Often one requires a two-arm trigger with particles in opposite hemispheres to eliminate backgrounds mostly from beam-gas collisions and from cosmic-rays. Such a trigger selection, however, reduces significantly the detection rate of (single) diffractive collisions characterized by the survival of one of the colliding (anti)protons and particle production in just one hemisphere. The associated measurements are often dubbed “non-single-diffractive” (NSD), although the corresponding cross sections are obtained with some level of model-dependence as different event-generators have different descriptions of diffractive production.

The UA5 experiment triggered only on one arm to measure the “single-diffractive” (SD) events and by combining the SD and the NSD results published MB results very close to a true inelastic trigger (INEL) (using MC corrections in both cases). LHC experiments define their inelastic trigger in different ways.

The list of minimum-bias (non-single diffractive and inelastic) event-selection triggers that have been employed by the different experiments presented in this work are listed in Table 3. Since in most cases, and in particular at the LHC, the experimental MB-trigger corrections have been obtained using Pythia and PHOJET, we can compare the MC-truth results from these two generators directly to the data. For the RFT models, we will use the predictions obtained both by using a theoretical definition of an NSD event as well as by applying directly the experimental cuts at the hadron-level to the final states produced in the MC simulation. This will allow us to estimate the effect of the trigger definitions on the measured hadron distributions and to investigate the model dependence of the experimental event-selection corrections.

| Experiment | Collision √s | Non-single diffractive (NSD) trigger | Inelastic (INEL) trigger |
|------------|--------------|------------------------------------|-------------------------|
| ISR        | p+p          | 30 - 65 GeV                        | Cross section based on luminosity |
| UA1        | p+̅p         | 0.2 - 0.9 TeV                      | 1 or more charged particles at each opposite rapidities (1.5 < |η| < 5.5) |
| UA5        | p+̅p         | 0.2 - 0.9 TeV                      | 1 or more charged particles at each opposite rapidities (2.0 < |η| < 5.6) within 2.0 < |η| < 5.6 |
| CDF        | p+̅p         | 1.96 TeV                           | 1 or more charged particles at each opposite rapidities (3.2 < |η| < 5.9) |
| ALICE      | p+p          | 0.9 - 7 TeV                        | 1 or more tracks at each opposite rapidities (2.8 < η < 5.1, -1.7 < η < -3.7) OR 2.8 < η < 5.1 OR -1.7 < η < -3.7 |
| CMS        | p+p          | 0.9 - 7 TeV                        | 1 or more charged particle(s) at each opposite rapidities (2.9 < |η| < 5.2) in the central region |η| < 2.0 |

Table 3: List of hadron-level cuts implemented in the “non-single diffractive” (NSD) and “inelastic” (INEL) triggers used by various proton-(anti)proton collider experiments.

4 True “zero bias” triggers, which only require the passing by of the two beams by the interaction point, are more sensitive to instrumental backgrounds.
Primary charged hadrons are defined as all charged hadrons produced in the collision, including the products of strong and electromagnetic decays, but excluding products of weak decays. Feed-down corrections from weakly decaying strange resonances (mainly $K^0_S$, $\Lambda$, $\bar{\Lambda}$, and $\Sigma^+, \Sigma^0$) have to be accounted for in order to obtain the final hadron spectrum. Such corrections, which depend on the strange particle composition in the MC, reduce by about $\sim 8\%$ the total charged yield at midrapidity. In all the simulations, one takes this into account by decaying all unstable particles for which $\tau < 10$ mm. The sole contribution from charged leptons to the reconstructed tracks in the low-$p_\perp$ range, comes from the Dalitz $\pi^0$ decay amounting to about 1.5% of the charged yield. ALICE does not correct for this contribution, whereas CMS does. We have removed this small contribution from all our model predictions by counting only the produced charged hadrons.

4. Data versus models

4.1. Particle pseudorapidity densities

The pseudorapidity densities, $dN_{ch}/d\eta$, of charged hadrons measured in NSD collisions at the LHC (0.9, 2.36 and 7.0 TeV) by ALICE and CMS (as well as by UA5 at 900 GeV) are shown in Fig. 2 compared to two PHOJET 6.4 tunes, PYTHIA 8 and to PHOJET. In the PYTHIA case, the NSD predictions are obtained switching off the single-diffractive contribution$^4$ without any hadron-level trigger. Since the effects of the LHC MB-selections have been corrected for by the experiments themselves using PYTHIA and PHOJET as a cross-check, this is a consistent comparison.

Figure 2: Pseudorapidity distributions of charged hadrons, $h^+ \equiv (h^+ + h^-)$, measured in NSD $p-p$ events at the LHC ($\sqrt{s} = 0.9, 2.36$ and 7 TeV) by ALICE$^5$ and CMS$^6$ (and by UA5 in $p-p$ at 900 GeV) compared to three different versions of PYTHIA and to the PHOJET MC. The dashed band is the systematic uncertainty of the CMS experiment which is similar to those of the two other measurements.

The Perugia-0 tune underpredicts all the measured midrapidity densities – by about 20% at $\sqrt{s} = 0.9$ and 2.36 TeV and by about 35% at 7 TeV – whereas the Atlas-CSC tune overpredicts by 10% the central multiplicities at 0.9 TeV but reproduces well the data at higher c.m. energies. PYTHIA 8 is 15% (10%) below the experimental central densities at $\sqrt{s} = 0.9$ and 2.36 TeV but agrees well with the experimental shape and normalization at 7 TeV. PHOJET is compatible within uncertainties with the measured hadron $dN_{ch}/d\eta$ distributions at 0.9 and 2.36 TeV, whereas at 7 TeV it is some 15% lower. The Perugia tunes were obtained mostly with Tevatron non-diffractive processes using hadrons with $p_\perp > 0.4$ GeV/c, which affects their prediction accuracy for the lowest multiplicities and lower transverse momenta considered here (the ALICE and CMS experimental hadron distributions are measured from $p_\perp \approx 100$ MeV/c and extrapolated down to zero $p_\perp$). The better agreement of the Atlas-CSC tune and PYTHIA 8 with the observations is linked to the faster rise of particle production with $\sqrt{s}$, as given by their smaller exponents, $\epsilon = 0.11$ and 0.08 respectively.

$^5$MSUB(92)=0 in PYTHIA 6.4, and ParticleDecays:limitTau0 = on, ParticleDecays:tau0Max = 10 in PYTHIA 8.

$^6$MSUB(92)=MSUB(93)=0 in PYTHIA 6.4, SoftQCD:singleDiffraction=off in PYTHIA 8.
Figure 3: Pseudorapidity distributions of charged hadrons, $h^\pm \equiv (h^+ + h^-)$, measured in NSD $p-p$ events at the LHC (0.9, 2.36 and 7 TeV) by ALICE [36, 37] and CMS [38, 39] (and by UA5 [42] in $p-\bar{p}$ at 900 GeV) compared to the predictions of qgsjet 01 and II, sibyll, and epos. The dashed band is the systematic uncertainty of the CMS experiment which is similar to those of the two other measurements.

see Eq. (4) and Table 1. The differences among the tunes cannot be related to the treatment of final-state effects since all of them have hadronization parameters tuned to LEP data and a “weak” option for the FS colour reconnection mechanism, which thus does not further reduce the particle multiplicity.

The comparison of the various cosmic-ray interaction generators to the NSD charged particle densities at the LHC is shown in Figure 3. In general the data-theory consistency is better than for PYTHIA. At 900 GeV, all the RFT models are in a reasonable agreement with the measurements, because inclusive UA5 data were used for the model calibrations. With increasing energy, the differences among the models predictions increase, bracketing the data with a very good agreement for Sibyll and QGSIJET01, and about +10% for QGSIJETII and -20% for EPOS. Such a RFT-model-data comparison is not fully self-consistent because the NSD event-selection corrections applied by the LHC experiments have been obtained using PYTHIA and PHOJET. We have carried out the same comparison but now applying the CMS NSD trigger (Table 2) directly at the hadron-level in the simulations. In Table 4 we list the predicted fractions of single diffraction events, the efficiencies of the SD and NSD event selection, and the resulting corrections

$$C = \frac{dN_{ch}^{NSD-theor}}{d\eta} \cdot \frac{dN_{ch}^{NSD-CMS}}{d\eta},$$

for all the models considered and for different LHC energies. The results obtained with the RFT models using the theoretical and experimental triggers are practically identical, whereas PYTHIA- and PHOJET-based corrections for the NSD event-selection resulted in about a 8% shift down of the measured $dN_{ch}/d\eta_{|\eta|=0}$ [38, 59].

| Model          | PYTHIA01 | QGSIJET01 | QGSIJETII | EPOS1.99 | SIBYLL2.1 |
|----------------|----------|-----------|-----------|-----------|-----------|
| $\sqrt{s}$ (TeV) | 0.9      | 2.36      | 7         | 0.9       | 2.36      | 7         | 0.9       | 2.36      | 7         |
| $\sigma_{pp}/\sigma_{inel}$ | 0.18     | 0.18      | 0.18      | 0.18      | 0.18      | 0.18      | 0.18      | 0.18      | 0.18      |
| $\epsilon_{SD}$ | 6.4%     | 6.0%      | 5.0%      | 16%       | 19%       | 21%       | 29%       | 32%       | 32%       |
| $\epsilon_{NSD}$ | 93%      | 95%       | 96%       | 90%       | 93%       | 95%       | 90%       | 92%       | 94%       |
| $C$             | 0.99     | 0.99      | 1.0       | 0.99      | 1.0       | 1.01      | 0.97      | 0.98      | 0.97      |

Table 4: Fractions of single diffractive (SD) events, efficiencies of the SD ($\epsilon_{SD}$) and NSD ($\epsilon_{NSD}$) event-selections, and correction factors C, Eq. (5), obtained applying the CMS NSD trigger at the hadron-level in the RFT models.

The energy dependence of the charged hadron density at $\eta = 0$ predicted by the different PYTHIA models and by
PHOJET at c.m. energies from $\sqrt{s} = 10$ GeV to $\sqrt{s} = 20$ TeV is presented in Fig. 4 compared to the existing NSD (left panel) and inelastic (right panel) data measured at Sp$\bar{p}$S, Tevatron and LHC. We note that the NSD selection results in measured central densities which are about 15% larger than those obtained with the less-biased INEL trigger, which has less particles produced on average as it includes (most of) diffractive production. We notice also that the theoretical spread is smaller for the inelastic event-selection than for the NSD case, which indicates the different modelling of the non-diffractive contributions in the various MCs. Up to $\sqrt{s} = 2.36$ TeV, PHOJET gives the best overall agreement with the data. Perugia-0 fails to reproduce the pace of increase of the central particle densities at all the energies. The predictions of Atlas-CSC and PYTHIA 8 globally bracket the experimental energy evolution of the hadron multiplicities at $\eta = 0$.

In Fig. 5, the same data are compared to the RFT models used for air-shower simulations. The spread of model predictions at all c.m. energies is smaller than that of the different PYTHIA tunes. Although all RFT models reproduce globally well the pre-LHC data up to $\sqrt{s} \approx 2$ TeV, one sees different extrapolations at the current top LHC energy and beyond. Up to 7 TeV, the older models SIBYLL and QGSHI01 have better predictions for the average multiplicity than the newer QGSHI2 and EPOS: The rate of the multiplicity rise between 900 GeV and 7 TeV compared to the measured one is 20% higher for QGSHI2 and 15% smaller for EPOS.

4.2. Average transverse momentum

The end result of a MC prediction for a given MB observable is not controlled by one single physical effect, but by a combination of various ingredients implemented in the model, such as e.g. those listed in Tables 1 and 2. For example, in PYTHIA the multiplicity density can be increased by allowing more underlying-event activity, and decreased by allowing e.g. more colour reconnections. Hence the same final multiplicity can be obtained through different combinations of MC tunings. For example, since the underlying event pumps energy into the event, in order to maintain the same multiplicity distribution, the hardness of the hadron spectra must then be a function of the underlying activity. Thus, by combining constraints from the experimental particle and transverse-momentum flows, some extra discriminating power can be gained. As discussed in the Introduction, the mean $p_\perp$ in a given $p$-$p$ event is a sensitive probe of the semi-hard dynamics (multiparton interactions and saturation effects). In this section, we discuss the average $\langle p_\perp \rangle$ predicted by the models in comparison to the collider data.
In the case of the Pythia and PHOJET simulations we have computed $\langle p_\perp \rangle$ as done by CMS, i.e. by fitting the midrapidity $p_\perp$-differential charged hadron spectra with the Tsallis function and averaging the $p_\perp$ over that function. For the RFT models we simply average the $p_\perp$ of all the charged particles in the central $\eta$ range. Applying the NSD or full-inelastic selections does not change drastically the values of $\langle p_\perp \rangle$ which differ only by ~5%. Also, the exact pseudorapidity coverage of the measurement around midrapidity (e.g. $|\Delta \eta| < 1$ or $|\Delta \eta| < 2.5$) does not change much the associated mean $p_\perp$ values (~4%) although an extension to full rapidities would decrease its value by about 12%.

The energy dependence of the average transverse momentum of charged hadrons measured from the ISR collider up to LHC energies is compared to the predictions of Pythia and PHOJET (left panel) and of cosmic ray models (right panel) in Fig. 6. The PHOJET and EPOS results are globally in good agreement with the $\sqrt{s}$-dependence of the average $p_\perp$ seen in the data. The Atlas-CSC Pythia tune and Sibyll predict a slower rate of increase at LHC energies. On the contrary, the rate of the increase predicted by Pythia Perugia-0 and by QGSJetII is compatible with the data but their absolute scale is higher by roughly 10% and 20% respectively. The Pythia 8 and QGSJet01 predictions miss the shape and absolute magnitude of $\langle p_\perp \rangle (\sqrt{s})$. It is interesting to notice that the Atlas-CSC Pythia tune which reproduced well the pseudorapidity distribution (Fig. 2) predicts a too low value for the average $p_\perp$, while the Perugia-0 tune which has a too low multiplicity shows a too large $\langle p_\perp \rangle$.

4.3. Multiplicity probability distributions

The multiplicity distribution $P(N_{ch})$, i.e. the probability to produce $N_{ch}$ charged hadrons in an event, is of special interest because it provides extra differential constraints on the internal details of the hadronic interaction models. The low multiplicity part is mostly dominated by the contributions from diffraction (and from single-cut Pomeron exchanges in the RFT approaches), whereas the tail of the distribution gives information on the relative contribution of multiparton scatterings (multi-Pomeron exchanges). The ALICE experiment has measured multiplicity distributions within $|\eta| < 1$ using different triggers (inelastic, ‘Inel>0’ with at least one particle measured in the considered $\eta$ range, and NSD) at 900 GeV, 2.36 TeV and 7 TeV [36, 37]. Such different triggers affect significantly the first few bins of the distributions, where their maxima lie. The CMS collaboration has provided a higher statistics set of results but applying a NSD trigger and, thus, with large uncertainties (up to 40%) in the low multiplicity part of the distributions. In Figs. 7 and 8 we show the $P(N_{ch})$ probabilities for the ALICE ‘Inel>0’ selection at the three
Figure 6: Average $p_\perp$ of charged particles at midrapidity in $p$-p and $p$-$\bar{p}$ collisions as a function of $\sqrt{s}$ compared to the PYTHIA 6 and 8 and to PHOJET models (left panel) and to the predictions of QGSJET 01 and II, SIBYLL, and EPOS (right panel).

The c.m. energies compared to the corresponding results of PYTHIA, PHOJET and RFT models. Figure 7 uses a log-scale to better visualize the behaviour at very large multiplicities. Figure 8 shows a zoom of $P(N_{ch})$ in the low $N_{ch}$ < 20 region for PYTHIA and PHOJET (top panels) and for the RFT models (bottom panels).

PYTHIA 8 and the Atlas-CSC tune of PYTHIA 6 reproduce globally well the high multiplicity tail at 0.9, 2.36 and 7 TeV whereas the Perugia-0 (as well as the D6T one [37, 100]) tune predicts too few hadrons at all the energies, PHOJET is somehow in between: it reproduces the 0.9 and 2.36 TeV results but misses the tail at $\sqrt{s} = 7$ TeV. In the low multiplicity region, Perugia-0 and PYTHIA 8 predict too many hadrons, whereas Atlas-CSC tune and PHOJET are closer to the ALICE results. Since the modeling of diffraction is the same in Perugia-0 and Atlas-CSC, and it is similar in PYTHIA 8 and PHOJET, these results constrain well the concurrent role of diffractive scatterings and of multi-parton interactions implemented in all these MCs.

In the case of RFT models, the high-$N_{ch}$ tail is underestimated by EPOS and QGSJET 01, whereas SIBYLL and QGSJET II get a bit closer, sometimes overestimating the data. In the low-$N_{ch}$ region, $P(N_{ch}) \sim 4$, only EPOS globally reproduces the experimental results whereas the rest of the models overestimate the measurements up to $+30\%$ for SIBYLL. The peak is even shifted towards lower multiplicity in the case of both QGSJET models. Thus, even if the average MB multiplicities at 0.9 and 2.36 TeV are well reproduced by most RFT models (Fig. 3), the details of their probability distributions are missed and indicate possible paths for the improvement of the different model ingredients. For example, EPOS has a pretty good description of the low multiplicity part but does not produce enough high multiplicity events (with more than 50 particles) compared to the data. This explains why the average $dN_{ch}/d\eta$ is too low for this model (see Fig. 3) which seems to give a correct average multiplicity per cut-Pomeron but with too few Pomerons. The QGSJET II model underestimates the data at the low multiplicity end but has a much longer tail. One possible reason for this mismatch is that the model neglects hard Pomeron-Pomeron couplings, hence, can not describe a dynamical evolution of the saturation scale beyond the fixed $Q_0$ cutoff value. For SIBYLL, the simulations oscillate around the data giving a correct average value but the shape of the multiplicity distribution is not described correctly. So even if the average values of inclusive observables – such as the multiplicity and average transverse momentum – are predicted correctly within 10% by the RFT models, there is still room for improvement of the microscopic dynamics of hadronic interactions in order to safely extrapolate the predictions up to the highest cosmic-ray energies.
Figure 7: Multiplicity distributions of charged hadrons, $P(N_{ch})$, measured by ALICE in Inel > 0 p-p events at 0.9, 2.36 and 7 TeV [37] compared to the predictions of 

\text{PYTHIA 6.422 (Atlas-CSC)} \quad \text{PYTHIA 6.422 (Perugia-0)} \quad \text{PYTHIA 8.130 (Tune-1)} \quad \text{PHOJET 1.12 (+PYTHIA 6.11)} 

Figure 8: Low part of the multiplicity distributions of charged hadrons, $P(N_{ch})$, measured by ALICE in Inel > 0 events at 0.9, 2.36 and 7 TeV [37] compared to the predictions of 

\text{PYTHIA 6.422 (Atlas-CSC)} \quad \text{PYTHIA 6.422 (Perugia-0)} \quad \text{PYTHIA 8.130 (Tune-1)} \quad \text{PHOJET 1.12 (+PYTHIA 6.11)} 

\text{SIBYLL 2.1} \quad \text{QGSJET 01} \quad \text{QGSJET II} \quad \text{EPOS 1.99} 

\text{ALICE (Inel=0)} \quad \text{SIBYLL 2.1} \quad \text{QGSJET 01} \quad \text{QGSJET II} \quad \text{EPOS 1.99}
4.4. Discussion

The level of (dis)agreement between the three LHC inclusive hadron observables discussed in the previous sections and each one of the eight hadronic MCs considered in this work, is summarized in Tables 5 and 6. As a general conclusion, none of the models can reproduce completely all the sets of measurements. However, some models need more tuning than others. For example, the tune Perugia-0 of \textsc{pythia} 6 does a poor job reproducing the minimum bias results measured at the LHC. As discussed previously, this is mainly due to the fact that the Perugia tunes were obtained mostly with Tevatron non-diffractive processes using hadrons with larger transverse momenta ($p_\perp > 0.4 \text{ GeV}/c$) than those considered at the LHC. \textsc{pythia} 8 has a better description of the LHC data, thanks mostly to an improved implementation of diffractive scattering, but the default settings under Tune-1 are just a \textit{bona fide} first guess. The Atlas-CSC tune does a better job in general but clearly cannot reproduce diffractive-enhanced data samples measured at the LHC (not shown here, see [101]), and thus it can be mostly considered as providing a useful reference for possible settings – in particular the power-law exponent $\epsilon$ that regulates the c.m. energy dependence of the infrared cutoff for (multi)parton scatterings – that should be tried with the more advanced \textsc{pythia} 8 code. As a matter of fact, recent developments in \textsc{pythia} 8.145 indicate that it is possible to reproduce LHC and previous collider data with an updated set of parameters (Tune 4C) [102].

| Model                  | PY6 Perugia0 | PY6 Atlas | PYTHIA 8 | PHOJET 1.12 |
|------------------------|--------------|-----------|----------|-------------|
| $\sqrt{s}$ (TeV)       | 0.9          | 2.36      | 7        | 0.9         | 2.36        | 7          | 0.9   | 2.36 |
| $dN_{ch}/d\eta_{\eta=0}$ | under        | under     | under    | over        | ✓            | ✓          | ✓     | ✓    |
| $\langle p_\perp \rangle$ | ✓            | ✓         | over     | ✓           | under        | ✓          | ✓     | ✓    |
| $P(N_{ch} < 5)$         | over         | over      | over     | ✓           | ✓            | over       | ✓     | ✓    |
| $P(N_{ch} > 30)$        | ✓            | ✓         | ✓        | ✓           | ✓            | ✓          | ✓     | ✓    |

Table 5: Level of overall agreement between different \textsc{pythia} tunes and PHOJET 1.12 with inclusive charged hadron results measured in $p$-$p$ collisions at 0.9, 2.36 and 7 TeV: pseudorapidity densities $dN_{ch}/d\eta_{\eta=0}$, mean transverse momentum $\langle p_\perp \rangle$, and multiplicity probabilities $P(N_{ch})$ (for low and high values of $N_{ch}$). A tick (✓) indicates a reasonable data-model agreement within experimental uncertainties, and ‘over’ (‘under’) that the MC tends to over (under) estimate the data.

The RFT models (Table 6) give an overall good description of the central pseudorapidity densities at LHC energies, but fail to describe consistently other characteristics of proton-proton collisions measured by the LHC experiments. Thus, further model improvements, particularly, those related to the treatment of inelastic diffraction and of the parton saturation mechanism, are highly desirable.

| Model                  | QGSJET01 | QGSJETII | SIBYLL 2.1 | EPOS 1.99 |
|------------------------|----------|----------|------------|-----------|
| $\sqrt{s}$ (TeV)       | 0.9      | 2.36     | 7          | 0.9       | 2.36     | 7          | 0.9   | 2.36 |
| $dN_{ch}/d\eta_{\eta=0}$ | ✓         | ✓         | ✓          | ✓         | ✓         | ✓          | ✓     | under |
| $\langle p_\perp \rangle$ | over     | over     | ✓          | over      | ✓         | ✓          | ✓     | ✓    |
| $P(N_{ch} < 5)$         | over     | over     | over      | over      | ✓         | ✓          | ✓     | ✓    |
| $P(N_{ch} > 30)$        | ✓        | ✓        | ✓         | ✓         | ✓         | ✓          | ✓     | ✓    |

Table 6: Level of overall agreement between QGSJET01, QGSJETII, SIBYLL 2.1 and EPOS 1.99 with inclusive charged hadron results measured in $p$-$p$ collisions at 0.9, 2.36 and 7 TeV: pseudorapidity densities $dN_{ch}/d\eta_{\eta=0}$, mean transverse momentum $\langle p_\perp \rangle$, and multiplicity probabilities $P(N_{ch})$ (for low and high values of $N_{ch}$). A tick (✓) indicates a reasonable data-model agreement within experimental uncertainties, and ‘over’ (‘under’) that the MC tends to over (under) estimate the data.
5. Implications for ultra-high energy cosmic-ray physics

5.1. Extrapolations to ultra-high cosmic-ray energies

In this section we show the $\sqrt{s}$-evolution of the model predictions for the central rapidity density $dN_{ch}/d\eta|_{\eta=0}$ and the average transverse momentum ($p_{t}$) for charged hadrons produced in $p$-$p$ collisions. We present the predicted c.m. energy dependence of these observables up to the highest cosmic-ray energies. Both the High Resolution Fly’s Eye (HiRes) experiment \cite{103} and the Pierre Auger Observatory \cite{104} have measured UHECR up to energies of the order of $10^{20}$ eV. Beyond this energy, the CR flux from cosmologically distant sources ($D > 100$ Mpc) is significantly reduced due to the interactions of the primary cosmic-ray particles with the cosmic microwave background radiation on their way from the extragalactic sources to the Earth. This so-called Greisen-Zatsepin-Kuzmin (GZK) cutoff \cite{11,12} corresponds to $p$-$p$ c.m. energies of order $\sqrt{s}_{\text{GZK}} \approx 300$ TeV, more than twenty times higher than those reachable at the CERN LHC.

In Figs. 9 and 10 we present the energy evolution predicted by all the models for $dN_{ch}/d\eta|_{\eta=0}$ up to GZK energies for NSD and inelastic $p$-$p$ interactions respectively. The left panels show the predictions of the various \textsc{pythia} versions considered and of \textsc{phojet}, and the right panels those of the rest of the RFT generators. \textsc{phojet} predictions stop at around $\sqrt{s} \approx 100$ TeV, which is the maximum c.m. energy where the model can be safely extrapolated to within its current implementation.

The first thing to notice is that all \textsc{pythia} tunes as well as \textsc{phojet}, \textsc{qgsjet01} and \textsc{epos} feature, with different tunes of \textsc{pythia} and by \textsc{phojet} (left panel) and by \textsc{qgsjet01} and \textsc{epos} (right panel) MCs up to the GZK cutoff energies. The data points are the same as in Fig. 11(left).

The data points are the same as in Fig. 4 (left).

Figure 9: Collision-energy dependence of the midrapidity charged hadron invariant yields in non single-diffractive (NSD) $p$-$p$ collisions predicted by different tunes of \textsc{pythia} and by \textsc{phojet} (left panel) and by \textsc{qgsjet01} and \textsc{epos} (right panel). The midrapidity $dN_{ch}/d\eta|_{\eta=0}$ scales – of the midrapidity particle-density. On the other hand, \textsc{qgsjetII} show faster and slower dependencies on energy, respectively. Disregarding for the moment the \textsc{pythia} Perugia-0 and \textsc{epos} 1.99 predictions, which are already lower than the particle multiplicity measured at 7 TeV, we see that the rest of the models can produce more or less reproduce the collider data up to LHC energies. Still they have very different extrapolations at the GZK energy range. The highest central rapidity densities at $\sqrt{s}_{\text{GZK}}$ are predicted by \textsc{qgsjetII} which reaches $dN_{ch}/d\eta|_{\eta=0} \approx 50$ (off-scale in the plot), \textsc{qgsjet01} and \textsc{pythia} 8 predict values of the order $dN_{ch}/d\eta|_{\eta=0} \approx 20$, the Atlas-CSC \textsc{pythia} 6.4 tune indicates $dN_{ch}/d\eta|_{\eta=0} \approx 14$, and finally \textsc{epos} (as \textsc{epos}) a low $dN_{ch}/d\eta|_{\eta=0} \approx 10$ value at the GZK cutoff.
Figure 10: Collision-energy dependence of the midrapidity charged hadron invariant yields in inelastic \( p-p \) collisions predicted by different tunes of \textsc{pythia} and by \textsc{phojet} (left panel) and by \textsc{qgsjet}01 and II, \textsc{sibyll}, and \textsc{epos} (right panel) MCs up to the GZK cutoff energies. The data points are the same as in Fig. 4 (right).

In Fig. 11 we show the energy evolution of the average \( p_{\perp} \) predicted by \textsc{pythia} 6 and 8 and by \textsc{phojet} (left panel) and by the cosmic-ray event generators (right panel). Above \( \sqrt{s} \approx 100 \text{ GeV} \), all the models show a power-law behaviour (with varying exponents) of \( \langle p_{\perp} \rangle \) with collision energy. Interestingly, all the RFT MCs with the exception of \textsc{epos} predict a very moderate increase of \( \langle p_{\perp} \rangle \) with energy: the \( \langle p_{\perp} \rangle \) amounts to \( \sim 0.6 \text{ GeV}/c \) at GZK energies which is only 0.05 GeV/c above the current CMS result at 7 TeV. On the other hand, \textsc{pythia} 8, \textsc{pythia} 6.4 (Atlas-CSC) and \textsc{epos} indicate \( \langle p_{\perp} \rangle_{\text{GZK}} \approx 1 \text{ GeV}/c \) (\textsc{phojet}, with an extrapolated \( \langle p_{\perp} \rangle_{\text{GZK}} \approx 0.8 \text{ GeV}/c \), is somewhat in between).

The different energy behaviour of the \( \langle p_{\perp} \rangle \) reflects directly the assumptions made in the models related to the low-\( x \) behaviour of the PDFs. For example, in models with saturation of parton densities, the mean transverse momentum of the produced hadrons is of the order of the saturation scale \( Q_0 \) in the high-energy limit. In the case of the two \textsc{qgsjet} models, the \( \langle p_{\perp} \rangle \) does initially increase with energy and later approaches an asymptotic value. This behaviour is related to the fact that parton saturation is neglected in \textsc{qgsjet}01 while in \textsc{qgsjet}II there is no dynamical evolution of the saturation scale above the fixed \( Q_0 \) value. In contrast, assuming saturation also for highly virtual partons, saturation effects lead to a continuous increase of the average \( p_{\perp} \) faster than \( \log s \), as found in \textsc{phojet} or \textsc{pythia}. \textsc{epos} would have had a behaviour similar to \textsc{qgsjets}, but because of the final-state collective parton expansion is taken into account in this model, the generated flow increases the average \( p_{\perp} \) rapidly above \( \sqrt{s} \approx 20 \text{ TeV} \).

5.2. Implications for the interpretation of extensive air showers

Extensive air showers initiated by interactions of primary cosmic ray particles (protons and nuclei) with air nuclei in the upper atmosphere constitute multi-step cascade processes. The backbone of an air shower is the hadronic cascade of interactions of both primary and secondary hadrons – (anti)nucleons, pions, and kaons – which quickly dissipates the initial energy between many hadronic particles and which pumps the energy into secondary electromagnetic (e.m.) cascades, mostly via decays of the produced neutral pions. In turn, high energy e.m. cascades proceed mainly via pair production of electrons and positrons from photons and by \( e^\pm \) bremsstrahlung, with a relatively weak feedback to the hadronic cascade via photonuclear interactions.

Experimental methods of EAS detection include measurement of charged particles arriving at ground level and studies of longitudinal shower development – mostly via observations of fluorescence and Čerenkov light produced
by charged particles ($e^\pm$ mainly) – see [105] for a review of detection techniques. The most important shower observables are the position of the shower maximum, $X_{\text{max}}$, i.e. the depth in the atmosphere (in g/cm$^2$) where the number of charged particles reaches its maximum, the number of particles at maximum $N_{\text{max}}$, and the number of electromagnetic particles ($e^\pm$, $\gamma$) and muons ($\mu^\pm$) at ground. Most of EAS characteristics measured at ground are closely related to $X_{\text{max}}$ and to $N_{\text{max}}$ [106]. For example, the number of charged particles (mainly $e^\pm$) at ground $N_e$ strongly increases for deeper showers (where the maximal number of charged particles is reached close to the ground) and falls down for decreasing $X_{\text{max}}$ – due to the quick shower attenuation in the atmosphere. On the other hand, the number of muons $N_\mu$ at ground, which emerge mainly from charged pion and kaon decays, has a weak dependence on $X_{\text{max}}$, being instead sensitive to charged particle multiplicity of hadron-air and nucleus-air collisions. This explains the importance of the latter observable for CR composition studies with ground-based detectors. Finally it should be noted that the number of muons is also dependent on the physics of low-energy interactions in air showers. With pions typically decaying only after their energy is reduced to $\sim 20$ GeV, low-energy interactions of c.m. energy in the range of $10^{-5} - 50$ GeV become important [107, 108].

The relation between hadronic interactions at high energies and EAS observables has been studied numerically in detail in [27]. Here we recall only the most important interaction parameters that determine the longitudinal development of air showers. The depth of shower maximum depends mainly on the characteristics of multiparticle production of the first few generations of hadronic interactions in a shower. It is mainly related to the inelastic cross section of the primary incoming particle for interaction with air nuclei and on the corresponding energy fraction transferred to secondary particles but the most energetic (“leading”) secondary particle, relative to the primary particle, which is called inelasticity. Additionally, although less strongly, $X_{\text{max}}$ depends on the inelastic cross sections and inelasticities for interactions of secondary hadrons with air nuclei. The third feature of direct relevance to the position of the shower maximum is the multiplicity of the primary and subsequent very high-energy interactions, which defines how the energy is distributed to secondary particles and corresponding sub-showers, i.e. whether many particles with small energy or few particles with large energy are produced.

The LHC measurements of charged hadron pseudorapidity density and multiplicities presented here, have considerable importance for CR physics. First, the observations by the ATLAS, CMS and ALICE collaborations indicate that $dN_{ch}/d\eta_{p\nu0}$ changes smoothly in the lab energy range from $10^{15}$ to $3 \times 10^{16}$ eV, and are well bracketed, at the
±10% level at the highest energies, by the predictions of current interaction models used for EAS simulations. The bulk of hadron production data did not reveal serious deficiencies in the overall description of high energy hadronic collisions implemented in the event generators currently used in cosmic ray physics. This gives a strong support to the overall interpretation of the experimental results in terms of primary CR spectrum and nuclear mass composition in the knee energy range \(10^{15.5}\) eV obtained by various collaborations that applied these models for the corresponding data analysis \([109, 20, 110, 111]\). At the same time, none of the models is in perfect agreement with all the hadronic observables measured at the LHC (see Table 6), underlining the need for re-tuning model parameters and reconsidering model assumptions. New versions of epos \([96, 97]\) and qgsjet \([112]\) are under development, partially influenced by the input from the LHC data.

Second, the new experimental results strongly disfavour speculative ideas that the knee in the observed cosmic ray spectrum may be due to a sudden change in the hadronic interaction mechanism above 2 TeV c.m. energy (see, for example, \([28, 29, 30, 31]\)). Indeed, any generic model that has about 20% of primary energy being transferred to particles that are not observed in EAS experiments just above \(3 \times 10^{15}\) eV lab, would predict that proton showers of primary energy \(3 \times 10^{16}\) eV, corresponding to about 7 TeV c.m. energy, would have to be reconstructed with a 25% lower energy in order to attribute the knee in the spectrum solely to the production of such new exotic particles. Of course, so far only data from central LHC detectors have been published, covering just a small fraction of the energy released in final state particles. However, it is difficult to imagine how such a significant change of the interaction properties from 2 to 7 TeV c.m. energy should only affect the particles close to the beam axis. More realistically, one has to include also nuclei as cosmic rays in the knee energy range, leading to even larger energy fractions that need to be channeled to undetected particles in hadronic interactions.

Third, the first measurements at \(\sqrt{s} = 7\) TeV support a conventional extrapolation of the known features of multi-particle production to higher energies. Still it is not clear, whether the new data of the Pierre Auger Collaboration \([23]\), indicating a change to rather small shower-to-shower fluctuations that would be typical for a primary mass composition dominated by medium or heavy elements, could not also be attributed to a change of the characteristics of hadronic interactions at energies above \(3 \times 10^{18}\) eV \([113, 114]\). But the current wide range of predictions for the particle densities, \(dN_{\pi,\eta}/d\eta_{p=0} \approx 10\) (epos, sibyll) – 50 (qgsjetII), as well as for the mean hadron transverse momentum, \(\langle p_{\perp} \rangle \approx 0.6\) (sibyll, qgsjet01) – 1 (epos) GeV/c, justifies the concurrent use of various MCs to gauge the uncertainties connected to hadronic interaction models in the interpretation of the cosmic ray data at GZK-cutoff energies.

Last but not least, the measurements of \(p_{\perp}\)-differential spectra and mean \(p_{\perp}\) of hadrons at LHC midrapidity do not have a direct impact on the interpretation of air shower data because the shape of lateral distributions of electrons (positrons) and muons at ground level is rather defined by multiple Coulomb scattering in the atmosphere and by the transverse momentum spectra of secondaries at much lower (fixed-target) energies \([107, 108]\). Nevertheless, the corresponding results are of importance for testing the overall physics consistency of soft and hard interaction mechanisms implemented in the models as well as for the understanding of the parton saturation and multi-scattering phenomena discussed above.

5.3. Importance of forthcoming LHC measurements

The first multiplicity measurements at \(\sqrt{s} = 7\) TeV (\(2.5 \times 10^{16}\) eV in lab system) have put already serious constraints on the interpretation of cosmic ray data. Coming \(p-p\) data at the nominal c.m. energy of \(\sqrt{s} = 14\) TeV, corresponding to cosmic-ray protons of \(10^{17}\) eV in the lab frame, will further reduce the uncertainty linked to the extrapolation of the interaction models to ultra-high energies, as they will the expected proton-nucleus runs \((pPb\text{ collisions at } \sqrt{s_{NN}} = 8.8\) TeV) \([35, 113]\). First results on the \(dN_{\pi,\eta}/d\eta_{p=0}\), and its centrality dependence, measured in Pb-Pb collisions at \(\sqrt{s_{NN}} = 2.76\) TeV \([116]\) will provide extra important cross-checks on the role of initial-state gluon saturation effects in the hadronic wave functions.

There are many more measurements that can be done already at the current LHC energy to reduce the uncertainties of air shower predictions. Key information is expected, for example, from the measurement of the distribution of forward neutral hadrons (\(\pi^0\) and neutrons) by LHCf \([117]\) and possibly other experiments using zero degree and other forward calorimeters \([118, 119, 120]\) in both \(p-p\) and Pb-Pb interactions. Of similar importance would be the
measurement of the energy flow and particle spectra in the very forward direction, in the pseudorapidity range from 5 to 10. This is an angular range that is very difficult to access in collider experiments but it is partially covered by various detectors in the LHC experiments [121], for example, TOTEM [98] and CASTOR (CMS) [122]. Another fundamental input for tuning and extrapolating models are the total, elastic, and diffractive cross sections accessible to measurement thanks to various forward proton detectors (existing such as TOTEM, ATLAS ALFA [123], or proposed [124]) in the LHC tunnel area. Examples of the impact of the current uncertainty of the cross section on air shower observables are discussed in [125].

Figure 12: Examples of midrapidity distributions whose measurement in p-p collisions at $\sqrt{s} = 7$ TeV will help to refine the modeling of (anti)baryon production in current event generators. Left: Collision-energy dependence of the ratio of $\bar{p}$ to $\pi^-$ multiplicities. Right: Ratio of $\bar{p}$ to charged particle multiplicities as a function of the central pseudorapidity density.

It is worth mentioning that other measurements that can be performed with central detectors at the LHC, such as e.g. of the rapidity spectra of identified hadrons, are of significant importance for cosmic ray physics. Indeed, due to different lifetimes of charged pions and kaons, knowledge of their relative contributions to the multiplicity of secondaries is required for precise calculations of the total number and of the energy spectra of muons produced during EAS development. Also the production of antibaryons is a theoretically poorly understood aspect of hadronic interactions and data on baryon pair production are sparse. As demonstrated in [126], the muon content of air showers may be significantly enhanced if the rate of production of (anti)baryons is up to 30% higher than assumed in models like QGSJET and SIBYLL. Baryon-induced subshowers lead to a higher number of muons at ground than meson-induced ones. The energy dependence of the overall baryon production rate, its relation to the centrality of the collision, and the momentum distribution of the baryons are important quantities to be measured. This is illustrated in Fig. 12 where CR model predictions for antiproton production in p-p at $\sqrt{s} = 7$ TeV are shown. The number of fragmenting strings leads to a so-called delayed energy threshold for baryon pair production. This threshold is lowest in SIBYLL, being well below the Tevatron energy, and largest in EPOS where it is in the energy range of the LHC (Fig. 12 left). Also their predicted asymptotic limits are very different. The dependence of antiproton production on the midrapidity particle density in proton-proton collisions at the LHC is shown in Fig. 12 (right). The peak at low central multiplicities is related to the fact that antibaryons are pair-produced with baryons in a single string, consuming a large fraction of the energy stored in the corresponding string. Peripheral interactions are typically simulated with only two color-strings maximizing the energy per string. EPOS differs not only quantitatively but also qualitatively in the antibaryon-baryon pair production model from the other models [127].
6. Summary

We have compared first CERN LHC data on inclusive charged hadron production – central pseudorapidity densities \( dN_{ch}/d\eta \), multiplicity probabilities \( P(N_{ch}) \), and mean transverse momentum \( \langle p_T \rangle \) – with hadronic interaction models (\textsc{qgsjet}01 and \textsc{siyl}), \textsc{epos} 1.99 used for the simulation of extensive air showers generated by ultrahigh-energy cosmic rays in the Earth atmosphere. We have also included in this comparison commonly used collider physics event generators, such as \textsc{pythia} (various tunes of version 6.4, and the new \textsc{pythia} 8) and \textsc{phojet}, in order to identify the model ingredients most sensitive to bulk hadron production. In all the cases we considered only those versions of the interaction models that were developed and tuned before the LHC data became available.

In contrast to most Monte Carlo event generators for collider physics, models developed for air shower simulation are based on Reggeon Field Theory (RFT) and are optimized to predict the overall characteristics of the final state particles including those of predominantly soft interactions. Moreover, these models are designed to deal with proton-nucleus and nucleus-nucleus interactions and allow for an extrapolation to energies as high as \( \sqrt{s} \approx 400\text{TeV} \). Therefore the comparison to LHC data is an important benchmark for the quality of the models and hence the reliability of air shower simulations currently used to interpret cosmic ray (CR) data.

The quality of the LHC data description varies from model to model and differs for different observables. A first general observation is that none of the models considered provides a very good description of all the LHC data considered here. Yet, the CR models bracket the LHC central rapidity densities and the low region of the multiplicity probabilities more naturally than most of the \textsc{pythia} tunes. The main difference appears to be on the more advanced treatment of diffractive scattering in the RFT models compared to the latter. The measured midrapidity charged-particle density is found to follow a simple power-law in energy, \( s^{\epsilon} \), from 10 GeV up to 7 TeV c.m. energy with exponent \( \epsilon \approx 0.10 \). This observation can be used to retune some basic model ingredients and improve their extrapolations to the highest energies. Such a result constrains in particular the way in which multiparton interactions and gluon saturation are implemented in various MCs (e.g. in \textsc{pythia}, \textsc{phojet} and \textsc{siyl}), via an energy-dependent infrared transverse momentum cutoff, \( Q_0(\sqrt{s}) \), for (multi)parton scatterings.

The studies presented here have also shown that it is very difficult, if not impossible, to compare model predictions with data at the sub-5% level if the trigger used for data-taking cannot be implemented without a full detector simulation. It is highly desirable to have the LHC inelastic, minimum bias and diffractive data published with a simple hadron-level prescription for emulating the data trigger, to which the measurements have been corrected to. Only then we will be able to make best use of LHC results to lower the uncertainty of various non-perturbative model ingredients and to improve the description of the transition from soft to hard parton dynamics (multiparton interactions, gluon saturation, ...).

With \( \sqrt{s} = 7\text{TeV} \), LHC is the first collider reaching an energy higher than the knee in the energy spectrum of cosmic rays. The LHC hadron multiplicity measurements give strong support to the conventional interpretation that the break in the power-law index of the observed CR spectrum at \( 10^{\text{25.2}}\text{eV} \) is indeed due to a feature of the primary cosmic ray flux. Alternative interpretations of the knee being a side-effect of rapidly changing properties of hadronic interactions above \( \sqrt{s} \approx 2\text{TeV} \) are strongly disfavoured. Similarly the LHC measurements support the interpretation of air shower data in the knee energy range as reflecting a change from a light to a more heavy mass composition. No new or exotic physics assumptions or extrapolations are needed for describing the overall event features measured in the central pseudorapidity region at the LHC. While re-tuning of model parameters to match LHC data will improve the reliability of air shower simulations, there is no indication from the LHC results that the extrapolations have to be changed significantly. At the highest CR energies of \( \mathcal{O}(10^{20}\text{eV}) \) – i.e. more than twenty times higher than those c.m. energies reachable in \( p-p \) at the LHC – the current wide range of predictions for the particle densities, \( dN_{ch}/d\eta_{lab} \approx 10 \) (\textsc{epos}, \textsc{siyl}) – 50 (\textsc{qgsjet}II), as well as for the mean hadron transverse momentum, \( \langle p_T \rangle \approx 0.6 \) (\textsc{siyl}, \textsc{qgsjet}01) – 1 (\textsc{epos}) GeV/c, justify today the concurrent use of all RFT MCs to gauge the uncertainties linked to the underlying hadronic interactions in the interpretation of the cosmic ray data at GZK-cutoff energies.

22
Future measurements at the LHC will further cross-check interaction models and help to understand better the underlying hadron production processes. Of direct relevance to the interpretation of air shower data are the multiplicity and spectrum of produced high-energy particles (leading particles, neutral and charged) emitted at forward rapidities, the contribution of different particle types to the overall final state multiplicity, the inelastic cross section and its breakdown in various diffractive contributions, as well as the production of baryon-antibaryon pairs at mid-rapidities in proton-proton, nucleus-nucleus and proton-nucleus collisions.

Acknowledgments

We thank Ferenc Sikler and Gabor Veres (CMS) and Jan-Fiete Grosse-Oetringhaus (ALICE) for valuable discussions on the LHC experimental data. Enlightening discussions with Torbjörn Sjöstrand on PYTHA 6 and 8 are also acknowledged. This work has been supported in part by Bundesministerium für Bildung und Forschung (BMBF) grant No. 05A08VK1. D.d’E. acknowledges support by the 7th EU Framework Programme (contract FP7-ERG-2008-235071). S.O. acknowledges the support of the FP7 Marie Curie program (contract FP7-IEF-2007-220251) and by the program Romforsking of Norsk Forskningsradet.

References

[1] A. M. Hillas, J. Phys. G31 (2005) R95–R131.
[2] P. Bhattacharjee and G. Sigl, Phys. Rept. 327 (2000) 109–247.[astro-ph/9811011]
[3] D. F. Torres and L. A. Anchordoqui, Rept. Prog. Phys. 67 (2004) 1663–1730 and [astro-ph/0402373].
[4] M. Kachelrieß, Comptes Rendus Physique 5 (2004) 441–452 and [hep-ph/0406174].
[5] J. Blümer, R. Engel, and J. R. Hördanel, Prog. Part. Nucl. Phys. 63 (2009) 293–338 and [arXiv:0904.0725] [astro-ph.HE].
[6] J. R. Hördanel, Astropart. Phys. 21 (2004) 241–265 and [astro-ph/0402356].
[7] T. Wibig and A. W. Wolfendale, J. Phys. G31 (2005) 255–264 and [astro-ph/0410624].
[8] V. S. Berezhinsky and S. I. Grigor’eva, Astron. Astrophys. 199 (1988) 1–12.
[9] V. S. Berezhinsky, A. Z. Gazizov and S. I. Grigor’eva, Phys. Rev. D74 (2006) 043005 and [astro-ph/0210095].
[10] R. Aloisio, V. Berezinsky, P. Blasi, and S. Ostapchenko, Phys. Rev. D77 (2008) 025007 and [arXiv:0706.2834] [astro-ph].
[11] K. Greisen, Phys. Rev. Lett. 16 (1966) 748–750.
[12] G. T. Zatsepin and V. A. Kuzmin, J. Exp. Theor. Phys. Lett. 4 (1966) 78.
[13] A. M. Hillas, Ann. Rev. Astron. Astrophys. 22 (1984) 425–444.
[14] P. Blasi, R. L. Epstein, and A. V. Olinto, Astrophys. J. 533 (2000) L123 and [astro-ph/9912240].
[15] J. Arons, Astrophys. J. 589 (2003) 871–892 and [astro-ph/0208444].
[16] P. L. Biermann, G. A. Medina-Tanco, R. Engel, and G. Pugliese, Astrophys. J. 604 (2004) L29–L32 and [astro-ph/0401150].
[17] A. Calvée, A. Kusenko, and S. Nagataki, Phys. Rev. Lett. 105 (2010) 091101 and [arXiv:1004.2535] [astro-ph.HE].
[18] J. Knapp, D. Heck, S. J. Sciutto, M. T. Dova, and M. Risse, Astropart. Phys. 19 (2003) 77–99 and [astro-ph/0206414].
[19] T. Antoni et al. (KASCADE Collab.), Astropart. Phys. 16 (2002) 245–263 and [astro-ph/0102443].
[20] T. Antoni et al. (KASCADE Collab.), Astropart. Phys. 24 (2005) 1–25 and [astro-ph/0505413].
[21] M. Amenomori et al. (Tibet ASy Collab.), Phys. Lett. B636 (2006) 58–64 and [astro-ph/0511469].
[22] T. Abu-Zayyad et al. (HiRes-RIA Collab.), Phys. Rev. Lett. 84 (2000) 4276 and [astro-ph/9911144].
[23] J. Abraham (Pierre Auger Collab.), Phys. Rev. Lett. 104 (2010) 091101 and [arXiv:1002.0699] [astro-ph.HE].
[24] R. U. Abbasi et al. (HiRes Collab.), Phys. Rev. Lett. 104 (2010) 161101 and [arXiv:0910.4184] [astro-ph.HE].
[25] J. Knapp, D. Heck, and G. Schatz, in Wiss. Berichte FZKA 5828, Forschungszentrum Karlsruhe, 1996.
[26] M. Zha, J. Knapp, and S. Ostapchenko, Proc. of 28th Int. Cosmic Ray Conf., Tsukuba (2003) 515.
[27] R. Ulrich, R. Engel, and M. Unger, Phys. Rev. D 83 (2011) 054026 and [arXiv:1010.4310] [hep-ph].
[28] V. B. Petkov, Int. J. Mod. Phys. A20 (2005) 6846–6848.
[29] A. A. Petrukhin, Nucl. Phys. Proc. Suppl. 151 (2006) 57–60.
[30] R. Belgioio, M. Masip, and I. Mastromatteo, JCAP 0906 (2009) 027 and [arXiv:0903.5237] [hep-ph].
[31] A. Dixit, P. Jain, D. W. McKay, and P. Mukherjee, Eur. Phys. J. C 68 (2010) 573–580 and [arXiv:0912.0980] [astro-ph.CO].
[32] B. Alessandro et al. [arXiv:1101.1852] [hep-ex].
[33] R. Engel and H. Rebel, Acta Phys. Polon. B35 (2004) 321–330.
[34] T. Sako et al., Nucl. Instrum. Meth. A578 (2007) 146–159.
[35] D. d’Enterria, R. Engel, T. McCauley, and T. Piorog, Indian J. Phys. 84 (2010) 1837 and [arXiv:0806.0944] [astro-ph].
[36] K. Aamodt et al. (ALICE Collab.), Eur. Phys. J. C68 (2010) 89 [arXiv:1004.3034] [hep-ex].
[37] K. Aamodt et al. (ALICE Collab.), Eur. Phys. J. C 68 (2010) 345 and [arXiv:1004.3514] [hep-ex].
[38] V. Khachatryan et al. (CMS Collab.), JHEP 02 (2010) 041 and [arXiv:1002.0621] [hep-ex].
[39] V. Khachatryan et al. (CMS Collab.), Phys. Rev. Lett. 105 (2010) 022002 and [arXiv:1005.3299] [hep-ex].
[40] A. M. Rossi et al., Nucl. Phys. B84 (1975) 269.
[41] C. Albajar et al. (UA1 Collab.), Nucl. Phys. B335 (1990) 261.
[107] H.-J. Drescher, M. Bleicher, S. Soff, and H. Stoecker, Astropart. Phys. 21 (2004) 87–94 and astro-ph/0307453.
[108] C. Meurer, J. Blümer, R. Engel, A. Haungs, and M. Roth, Czech. J. Phys. 56 (2006) A211 and astro-ph/0512536.
[109] M. Aglietta et al. (EAS-TOP Collab.), Astropart. Phys. 21 (2004) 583–596.
[110] M. Amenomori et al. (Tibet ASγ Collab.), Astrophys. J. 678 (2008) 1165–1179 and arXiv:0801.1803 [hep-ex].
[111] N. M. Budnev et al., Nucl. Phys. Proc. Suppl. 190 (2009) 247–252 and arXiv:0902.3156 [astro-ph.HE].
[112] S. Ostapchenko, Phys. Rev. D83 (2011) 014018 and arXiv:1010.1869 [hep-ph].
[113] T. Wibig, Phys. Lett. B678 (2009) 60–64.
[114] T. Wibig, Phys. Rev. D79 (2009) 094008.
[115] C. A. Salgado et al., “Proton-nucleus at the LHC: Scientific opportunities and requirements”, J. Phys. G to be submitted; J. Jowett and C. Carli, “The LHC as a proton nucleus collider”, in Proceeds. EPAC 2006, Edinburgh, pp. 550-552. http://accelconf.web.cern.ch/AccelConf/e06/.
[116] K. Aamodt et al. (ALICE Collab.), Phys. Rev. Lett. 106 (2011) 032301 and arXiv:1012.1657 [nucl-ex].
[117] Technical Design Report of the LHCf experiment, CERN-LHCC-2006-004 (2006).
[118] O. A. Grachov et al., AIP Conf. Proc. 867 (2006) 258.
[119] S. White, Nucl. Instrum. Meth. A 617 (2010) 126.
[120] C. Oppedisano et al., Nucl. Phys. Proc. Suppl. 197 (2009) 206.
[121] D. d’Enterria, arXiv:0708.0551 [hep-ex].
[122] V. Andreev et al., Eur. Phys. J. C 67 (2010) 601.
[123] S. Ask, arXiv:0706.0644 [hep-ex].
[124] M. G. Albrow et al. (FP420 R&D Collab.), JINST 4 (2009) T10001 and arXiv:0806.0302 [hep-ex].
[125] R. Ulrich, R. Engel, S. Muller, F. Schussler, and M. Unger, Nucl. Phys. Proc. Suppl. 196 (2009) 335–340 and arXiv:0906.3075 [astro-ph.HE].
[126] T. Pierog and K. Werner, Phys. Rev. Lett. 101 (2008) 171101 and astro-ph/0611311.
[127] F. M. Liu et al., Phys. Rev. D67 (2003) 034011.