LHC results and High Energy Cosmic Ray Interaction Models

Tanguy Pierog
Karlsruhe Institute of Technology (KIT), IKP, D-76021 Karlsruhe, Germany
E-mail: Tanguy.Pierog@kit.edu

Abstract. Currently the uncertainty in the prediction of shower observables for different primary particles and energies is dominated by differences between hadronic interaction models. The LHC data on minimum bias measurements can be used to test Monte Carlo generators and these new constrains will help to reduce the uncertainties in air shower predictions. In this article, after a short introduction on air showers we will show the results of the comparison between the updated version of high energy hadronic interaction models with LHC data. Results for air shower simulations and their consequence on the comparison with air shower data will be discussed.

1. Introduction
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 $10^{14}$ \( \text{eV} \) through the cascades of secondary particles, called extensive air-showers (EAS), that they produce in the atmosphere (for a recent review, see [1]). 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 [2]. 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 [3, 4, 5, 6, 7, 8]. While the electroweak interaction processes are reasonably well understood, modeling of hadronic multi-particle production is subject to large theoretical uncertainties that are, moreover, difficult to estimate [9, 10, 11].

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 (Fig. 1). 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 [13]. Data from LHC experiments published so far have mostly been taken with detectors covering the central phase space region in pseudorapidity ($|\eta| < 2.5$). This region is most easily accessible in collider experiments and is also the region of the highest rapidity-density of produced particles. The first data have been compared to cosmic ray models in [14]. 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.

In this paper, we will discuss changes in the hadronic model predictions after LHC data and their consequences on air shower observables. In the first section, we will explain the so-called Heitler model to extract from a simple toy model the main hadronic observables which drive the development of air showers. We will then compare the results of the hadronic interaction models with LHC data for such observables. Finally using detailed Monte Carlo simulations done with CONEX [15], the new predictions for $X_{\text{max}}$, for the number of muons and for the missing energy will be presented.

2. Heitler’s Model

To qualitatively describe the dependence of shower development on some basic parameters of particle interaction, decay and production, a very simple toy model can be used. Although initially developed for electromagnetic (EM) showers [16] it can also be applied to hadronic showers [17].

2.1. Electromagnetic showers

For simplicity, instead of having three particle types ($\gamma$, $e^+$ and $e^-$) like in electromagnetic showers, we will consider only one particle with energy $E$ with only one EM interaction producing two new particles with energy $E/2$ after a fixed interaction length of $\lambda_e$, see Fig. 2.

Denoting with $n$ the number of generations (consecutive interactions), the number of particles

Figure 1. Flux of cosmic ray arriving at Earth rescaled by the energy to the power 2.5. Data references are given in [12].
at a given depth $X = n \cdot \lambda_e$ follows from

$$N(X) = 2^n = 2^{X/\lambda_e},$$

with the energy $E$ per particle for a given primary energy $E_0$ being

$$E(X) = \frac{E_0}{2^{X/\lambda_e}}.$$

Defining the critical energy $E_c$ ($\sim 85$ MeV in air) as the energy below which energy loss processes dominate over particle production, one can make the assumption that the shower maximum is reached at a depth at which the energy of the secondary particles reaches $E_c$. Then two main shower observables are given by

$$N_{\text{max}} = \frac{E_0}{E_c} \quad \text{and} \quad X_{\text{max}}^e(E_0) \sim \lambda_e \cdot \ln \left( \frac{E_0}{E_c} \right).$$

This simplified picture does not reproduce the detailed behavior of an EM shower, but two important features are well described: the number of particles at shower maximum is proportional to $E_0$ and the depth of shower maximum depends logarithmically on the primary energy $E_0$.

### 2.2. Hadronic showers

Generalizing this idea, a hadronic interaction of a particle with energy $E$ is assumed to produce $n_{\text{tot}}$ new particles with energy $E/n_{\text{tot}}$, two thirds of which are charged particles $n_{\text{ch}}$ (charged pions) and one third are neutral particles $n_{\text{neut}}$ (neutral pions), as shown Fig. 3. Neutral particles decay immediately into EM particles ($\pi^0 \rightarrow 2\gamma$). After having travelled a distance corresponding
to the mean interaction length $\lambda_{\text{line}}$, charged particles re-interact with air nuclei if their energy is greater than some typical decay energy $E_{\text{dec}}$.

2.2.1. Energy transfer In each hadronic interaction, one third of the energy is transferred to the EM shower component. After $n$ generations the energy in the hadronic and EM components is given by

$$E_{\text{had}} = \left(\frac{2}{3}\right)^n E_0,$$  \hspace{1cm} (4)

$$E_{\text{EM}} = \left[1 - \left(\frac{2}{3}\right)^n\right] E_0,$$ \hspace{1cm} (5)

where $n$ will be calculated later. Simulations show that the number of generations is typically about 5 to 6 [18].

Even in an air shower initiated by a hadron, most of the energy is carried by EM particles ($\sim 90\%$ for $n = 6$). Hence the depth of shower maximum is given by that of the EM shower component, $X_{\text{max}}^e$. As the first hadronic interaction produces EM particles of energy $\sim E_0/n_{\text{tot}}$ one gets

$$X_{\text{max}}(E_0) \sim \lambda_{\text{line}} + X_{\text{max}}^e(E_0/n_{\text{tot}}),$$ \hspace{1cm} (6)

$$\sim \lambda_{\text{line}} + \lambda_e \cdot \ln \left(\frac{E_0}{n_{\text{tot}}E_c}\right),$$ \hspace{1cm} (7)

where $\lambda_{\text{line}}$ is the hadronic interaction length. This simplified expression for the depth of maximum neglects the EM sub-showers initiated by hadrons of later generations. The inclusion of higher hadronic generations does not change the structure of Eq. (7), only the coefficients change (see, for example, [19]).

2.2.2. Muon component To keep the picture simple, we assume that all charged hadrons decay into muons when their energy reaches $E_{\text{dec}}$. In a real shower, this limit can be seen as the characteristic energy where interaction length and decay length of charged pions are similar (about 150 GeV for pions). By construction, charged particles will reach the energy $E_{\text{dec}}$ after $n$ interactions

$$E_{\text{dec}} = \frac{E_0}{(n_{\text{tot}})^n}.$$ \hspace{1cm} (8)

Since one muon is produced in the decay of each charged particle, we get

$$N_{\mu} = n_{\text{ch}}^n = \left(\frac{E_0}{E_{\text{dec}}}\right)^\alpha,$$ \hspace{1cm} (9)

with $\alpha = \ln n_{\text{ch}}/\ln n_{\text{tot}} = 1 + \ln R/\ln n_{\text{tot}} \approx 0.82 \ldots 0.95$ [19] where $R = n_{\text{ch}}/n_{\text{tot}} < 1$. The number of muons produced in an air shower depends not only on the primary energy and air density, but also on the total particle multiplicities and in a much more sensitive way [20] of the charged over all particle ratio of hadronic interactions.

It should be kept in mind that the parameters of the model are only effective quantities and are not identical to the respective quantities measured at accelerators. In particular, the approximation of all secondary particles carrying the same energy is only motivated by the fact that it allows us to obtain simple, closed expressions. The well-known leading particle effect, typically quantified by the (in)elasticity of an interaction, can be implemented in the model [17] but will not be considered here.
3. Hadronic Interaction Models and LHC data

It is clear that such a model is only giving a very much over-simplified account of air shower physics. However, the model allows us to qualitatively understand the dependence of many air shower observables on the characteristics of hadronic particle production. Accordingly the parameters of hadron production being most important for air shower development are the cross section (or mean free path), the multiplicity of secondary particles of high energy, and the production ratio of neutral to charged particles. Until the start of LHC, these parameters were not well constrained by particle production measurements at accelerators. As a consequence, depending on the assumptions of how to extrapolate existing accelerator data, the predictions of hadronic interaction models differ considerably.

There are several hadronic interaction models commonly used to simulate air showers. Here we will focus on the two high energy models which were updated to take into account LHC data at 7 TeV: QGSJETII-03 [23, 24] changed into QGSJETII-04 [27] and EPOS 1.99 [21, 22] replaced by EPOS LHC. There is no major change in these models but in addition to some technical improvements, some parameters were changed to reproduce TOTEM [28] cross sections. Both are based on Gribov-Regge multiple scattering, perturbative QCD and string fragmentation. The former versions reproduce accelerator data and even first LHC data reasonably well [14] and Figs. 5 and 6 but predict different extrapolations above \( E_{\text{cms}} \sim 1.8 \text{ TeV} \) (\( E_{\text{lab}} \sim 10^{15} \text{ eV} \)) that lead to very different results at high energy [2, 26] which can be improved using LHC data.

3.1. Cross section

As shown in 7, the cross section is very important for the development of air showers and in particular for the depth of shower maximum. As a consequence, the number of electromagnetic particles at ground is strongly correlated to this observable (if the shower maximum is closer to ground, the number of particles is higher).

The proton-proton scattering total cross section is usually used as an input to fix basic parameters in all hadronic interaction models. Therefore it is very well described by all the models at low energy, where data exist [29]. And then it diverges above 2 TeV center-of-mass (cms) energy because of different model assumptions. As shown on Fig. 4 the new point measured by the TOTEM experiment at 7 TeV reduces the difference between the models by a factor of 5 (50 to 10 mb). In all the figures EPOS LHC is represented by a full (blue) line, QGSJETII-04 by a dotted (red) line, EPOS 1.99 by a dashed (black) line and QGSJETII-03 by a dashed-dotted (green) line.

3.2. Multiplicity

According to eq. 7, the multiplicity plays a similar kind of role as the cross section, but with a weaker dependence (log). On the other hand, the predictions from the models had much larger differences for the multiplicity compared to the cross section. As shown Fig. 5 left-hand side, the particle density at mid-rapidity is well reproduced by all the models up to 2 TeV where Tevatron data [31] constrain the results, but at the highest energies in \( \pi \)-air, the difference can be as high as a factor of 10 (Fig. 5 right-hand side). After re-tuning at 7 TeV to be compatible with CMS data [32] or ALICE data [33] on Fig. 6, the difference even at high energy in \( \pi \)-air is less than a factor of 2. On the right-hand side of Fig. 6, we can see that not only the averaged multiplicity had been changed after re-tuning, but the fluctuations are now very similar for QGSJETII-04 and EPOS LHC. This will be important for the fluctuations of the air shower maximum.

So for both cross section and multiplicity, when the models are constrained by LHC data up to 7 TeV, the extrapolation to the highest energy is not so different any more. This will have a strong impact on \( X_{\text{max}} \) uncertainty in air shower simulations.
Figure 4. Total and elastic p-p cross section calculated with EPOS LHC (full line), QGSJETII-04 (dotted line), EPOS 1.99 (dashed line) and QGSJETII-03 (dashed-dotted line). Points are data from [30] and the stars are the LHC measurements by the TOTEM experiment [28].

3.3. Baryon production

Another important observable for EAS is the number of muons reaching the ground. Using eq. 9 and the definition of $\alpha$ and $R$, it has been shown in [20] that the number of (anti)baryons plays an important role in the value of $R$ especially if we take into account the leading particle effect. As a consequence the number of muons in EAS is sensitive to the number of (anti)baryons produced in the hadronic interactions and it is important to check the production of such particles in LHC data.

Both ALICE [34] and CMS [35] experiments published very nice results on identified spectra used to constrain models used for air shower simulations. As shown in Fig. 7, these data helped a lot to reduce the differences between the models especially because it could resolve an ambiguity on the phase space used to produce some anti-proton over pion ratio with Tevatron data at 1.8 TeV. LHC data are much better defined and can be used to constrain the production of baryon pairs at mid-rapidity (largely dominated by string fragmentation).

It is important to notice that not only (and not all) (anti)baryons are entering in the definition of the ratio $R$. All particles which do not decay into an electromagnetic particle can play a similar role and keep the energy of the shower into the hadronic channel to produce muons. For instance in QGSJETII-04 the newly introduced $\rho^0$ resonance as excited state of the pion remnant in pion
Figure 5. Particle density at $\eta = 0$ for non single diffractive events (NSD) (left-hand side) and multiplicity for $|\eta| < 2.5$ of $\pi$-air collisions (right-hand side) as a function of center of mass energy. Simulations are done with EPOS LHC (full line), QGSJETII-04 (dotted line), EPOS 1.99 (dashed line) and QGSJETII-03 (dashed-dotted line). Points are data from old experiments and red stars are from CMS experiment [32].

Figure 6. Pseudorapidity distribution $dN/d\eta$ for events with at least one charged particle with $|\eta| < 1$ (left-hand side) and corresponding multiplicity distribution (right-hand side) for p-p interactions at 7 TeV. Simulations with EPOS LHC (full line), QGSJETII-04 (dotted line), EPOS 1.99 (dashed line) and QGSJETII-03 (dashed-dotted line) are compared to data points from ALICE experiment [33].

Interactions has a very strong influence on the muon production. Since forward $\pi^0$, transferring a lot of energy in the electromagnetic channel, are replaced by particles which decay in charged pions, the energy is kept in the hadronic channel. This is clearly illustrated by the Fig. 8 where we can see that QGSJETII-04 reproduce nicely both $\pi^0$ and $\rho^0$ forward spectra while QGSJETII-03 producing no $\rho^0$ had to hard $\pi^0$ spectra. This effect was already in EPOS 1.99, being one source of difference between the 2 models. EPOS models seems to underestimate $\rho^0$
Figure 7. Transverse momentum distribution for $|y| < 1$ for non single diffractive (NSD) p-p scattering at 7 TeV (left-hand side) and pion over anti-proton ratio at $y = 0$ for p-p collisions (right-hand side) as a function of center of mass energy. Simulations are done with EPOS LHC (full line), QGSJETII-04 (dotted line), EPOS 1.99 (dashed line) and QGSJETII-03 (dashed-dotted line). Points are data from ALICE and PHENIX [34] experiments and triangles are from CMS experiment [35].

Figure 8. $\pi^0$ longitudinal momentum fraction distribution (left-hand side) and $\rho^0$ longitudinal momentum fraction distribution (right-hand side) for $\pi$-p interactions at 250 GeV. Simulations with EPOS LHC (full line), QGSJETII-04 (dotted line), EPOS 1.99 (dashed line) and QGSJETII-03 (dashed-dotted line) are compared to data points from NA22 experiment [36, 37].

production but this is due to the mixing with $\omega$ resonance. The sum of the 2 is correct. On the other hand in EPOS 1.99 another process producing forward (anti)baryons was missing at high energy. As a consequence the reduced rate of (anti)baryons production at mid-rapidity is compensated by more forward (anti)baryons production which is even more important for muon production. Unfortunately there is very little data to constrain this production channel especially in collider experiments. It leads to a large uncertainty for the extrapolation at the
Figure 9. Mean $X_{max}$ for proton and iron induced showers as a function of the primary energy. Predictions of different high-energy hadronic interaction models, full lines for proton and dashed lines for iron with full triangles for EPOS 1.99, open squares for QGSJETII-04, open circles for QGSJETII-03, and full stars for EPOS LHC, are compared to data. Refs. to the data can be found in [1] and [7].

highest energy as shown in the next section. NA61 and LHCf data may help to constrain this process in the future.

4. EAS Simulations
Using the air shower simulation package CONEX and the new versions of the high-energy hadronic interaction models, we can get an estimate of the resulting uncertainties.

In the following EAS simulation results using EPOS LHC and QGSJETII-04 are presented and compared to former results using QGSJETII-03 and EPOS 1.99.

As shown in Fig. 9, the mean depth of shower maximum, $X_{max}$, for proton and iron induced showers simulated with CONEX is still different for EPOS LHC and QGSJETII-04. But now the elongation rate (the slope of the mean $X_{max}$ as function of the primary energy) is the same in both cases while EPOS 1.99 had an elongation rate larger than QGSJETII-03. The difference between the 2 models is a constant shift of about 20 g/cm$^2$ (close to the experimental systematic error in Pierre Auger Observatory (PAO) [7]) while before the difference were increasing up to 50 g/cm$^2$ at the highest energies.

This is very important to study the primary cosmic ray composition. If the models converge to a similar elongation rate, it will allow us to have a more precise idea on possible changes in composition at the “ankle” for instance where the PAO measured a break in the elongation rate of the data.

Concerning the number of muons at ground (for 40 inclined shower at the height of 1500 m),
Figure 10. Mean number of muons at ground divided by the primary energy for proton and iron induced showers as a function of the primary energy. Predictions of different high-energy hadronic interaction models: full lines for proton and dashed lines for iron with full stars for EPOS 1.99, open squares for QGSJETII-04, open circles for QGSJETII-03, and shaded area for the preliminary results of EPOS LHC (see text).

The difference between the new QGSJETII-04 and the old QGSJETII-03 is even more impressive. We can see on Fig. 10 that QGSJETII-04 predicts now about the same number of muons than EPOS 1.99 which is about 20% more than QGSJETII-03. It is due to the change in baryon, strangeness and mostly resonance production as described in section 3.3. Concerning the preliminary predictions of EPOS LHC a shaded area is presented. It reflects the uncertainty (about 5% to 15% around the EPOS 1.99 and QGSJETII-04 predictions) due to the forward (anti)baryon production which is not fixed by accelerator data yet. So even if the number of muons is very similar now for the 2 most recent hadronic models, there is still an uncertainty of about 10% and furthermore the energy spectrum of the muons at ground is different between the models. This can be an important factor for the attenuation length of the muons in the atmosphere. The average energy of the muons is for the moment larger in QGSJETII-04 than in EPOS.

A consequence of the larger mean energy of the muons in EAS simulated with QGSJETII-04 is that the missing energy is larger for showers simulated with this model. The factor to correct for the missing energy in fluorescence based telescope is presented on Fig. 11. The difference between QGSJETII-04 and QGSJETII-03 is less than 3% but it has a direct influence on the energy scale considered to compare simulations and data. It shouldn’t be neglected.
Figure 11. Inverse of missing energy factor for proton and iron induced showers as a function of the primary energy. Predictions of different high-energy hadronic interaction models are shown with triangles for EPOS 1.99, circles for QGSJETII-04 and squares for QGSJETII-03. Lines are averages between models.

5. Summary
Using a simple cascade model, it is possible to find the main parameters of hadronic interactions that influence air shower predictions. For the mean depth of shower maximum, $\langle X_{\text{max}} \rangle$, these parameters are the inelastic cross sections, the secondary particle multiplicity, and the inelasticity (not studied here). Using recent LHC data at 7 TeV it is possible to reduce the uncertainty in the extrapolation of the hadronic interaction models used for EAS simulations. Using pre- and post-LHC versions of the QGSJETII and EPOS models, it has been showed that the difference in multiplicity between these models has been reduced by a factor of 5 at the highest energy, resulting in a very similar elongation rate. There is still a systematic shift in $X_{\text{max}}$ of about 20 $g/cm^2$ due to remaining differences in the multiplicity (and elasticity) of the models. This uncertainty is comparable to the experimental uncertainty in the measurement of $X_{\text{max}}$. As a consequence the interpretation of the data using post-LHC data will be more reliable especially concerning the possible change in mass composition with energy as summarized in [38].

For the number of muons, the ratio between particles producing hadronic sub-showers and the total number of particles is very important. LHC data are important to constrain (anti)baryon and strangeness production at mid-rapidity. Lower energy data of fixed target experiment are also important to measure forward production of $\pi^0$ for instance. Taking into account both aspect, the new version of the QGSJETII and EPOS models predict very similar results close to EPOS 1.99 model but with harder spectrum.

The difference between EPOS 1.99 and the preliminary results of EPOS LHC is not very large because most of the changes are taking place at mid-rapidity. This phase space is good
to test the physics of the model but is not very important for air shower development. A contrary, large differences between QGSJETII-03 and QGSJETII-04 are observed. With a larger $\langle X_{\text{max}} \rangle$ the average mass is heavier than before at Auger energies. In addition the number of muons increased by about 20% to a value very similar to EPOS 1.99 but with a larger mean energy for the muons at ground.

To conclude we can say that LHC data contribute a lot to reduce the uncertainties in air shower simulations providing better tools to analyse cosmic ray data. The differences between the hadronic models has been reduced but one should keep in mind that there is still uncertainties in the models them-self which have to be better quantified and transferred to the calculation of the systematic errors in EAS analysis.

Acknowledgments
The author would like to thank R. Engel and D. Heck for their help and fruitful discussion and S. Ostapchenko for providing the code of QGSJETII-04.

References
[1] Blümer J, et al 2009 Prog. Part. Nucl. Phys. 63 293
[2] Knapp J, Heck D, Sciutto S J, Dova M T and Risse M 2003 Astropart. Phys. 19 77
[3] Antoni T, et al (for the KASCADE Collaboration) 2005 Astropart. Phys. 24 1
[4] Amenomori M, et al (for the Tibet ASγ Collaboration) 2006 Phys. Lett. B632 58
[5] Abu-Zayyad T, et al (for the HiRes-MIA Collaboration) 2000 Phys. Rev. Lett. 84 4276
[6] Knapp J, Heck D and Schatz G 1996 Wiss. Berichte, Forschungszentrum Karlsruhe FZKA 5828
[7] Engler, Engel R and Unger M 2011 Phys. Rev. D 83 054026
[8] Engel R, Heck D and Pierog T 2011 Ann. Rev. Nucl. Part. Sci. 61 467
[9] Bergmann T, et al 2007 Astropart. Phys. 26 420
[10] Matthews J 2005 Astropart. Phys. 22 387
[11] Meurer C, et al 2006 Czech. J. Phys. 56 A211
[12] Abramowicz A, et al 2008 Phys. Rev. C74 044902
[13] Caso C, et al 1998 Eur. Phys. J. C3 1
[14] Chatrchyan S, et al (for the CMS Collaboration) 2012 Phys. Lett. B716 102
[15] Czyz T, et al (for the TOTEM Collaboration) 2012 Prog. Theor. Phys. Suppl. 193 180
[16] Chojnacki M (for the ALICE Collaboration) 2011 J. Phys. G 38 124074
[17] Chatrchyan S, et al (for the CMS Collaboration) 2012 [arXiv:1207.4724 [hep-ex]]