

Air Shower Simulations¹

Marco Alania*, Ignacio J. Araya[†], Adolfo V. Chamorro Gómez*, Humberto Martínez Huerta**, Alejandra Parra Flores[‡] and Johannes Knapp[§]

**Centro de Tecnologías de Información y Comunicaciones,
Universidad Nacional de Ingeniería, Lima, Perú*

[†]*Departamento de Astronomía y Astrofísica, Pontificia Universidad Católica de Chile,
Av. Vicuña Mackenna 4860, Santiago, Chile*

***Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla,
Puebla, México*

[‡]*Facultad de Ciencias Físico Matemáticas,
Benemérita Universidad Autónoma de Puebla, Puebla, México*

[§]*School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, UK*

Abstract. Air shower simulations are a vital part of the design of air shower experiments and the analysis of their data. We describe the basic features of air showers and explain why numerical simulations are the appropriate approach to model the shower simulation. The CORSIKA program, the standard simulation program in this field, is introduced and its features, performance and limitations are discussed. The basic principles of hadronic interaction models and some general simulation techniques are explained. Also a brief introduction to the installation and use of CORSIKA is given.

Keywords: cosmic rays, air showers, simulations, hadronic interactions

PACS: 02.70.Uu, 13.85.Tp, 96.50.sd, 96.50.sf

Introduction

Cosmic rays (CR) are charged particles, mostly protons and ionised nuclei, from outer space. CR energies range from 10^8 to $> 10^{20}$ eV, following an approximate power law. The flux of the highest energy cosmic rays is very low. Above 10^{20} eV it is less than 1 particle per km^2 per century. Therefore high-energy CRs are only detectable by their interaction with the Earth atmosphere.

The main questions about cosmic rays are: Where do Cosmic Rays come from? What are they? How do they get their energies? To answer these questions requires to measure the energy and arrival directions and identify the primary CRs first.

Numerical simulation of the shower development in the atmosphere is a powerful approach to link the properties of the measured shower to those of the primary particle. In this work we discuss how simulations are used in CR physics, we describe the main ingredients needed in a numerical model, present the CORSIKA EAS simulation program and discuss its strengths and limitations. Finally we present a short tutorial on how to install and use CORSIKA.

¹ This article is based on a series of lectures given at the “Third School on Cosmic Rays and Astrophysics”, August 25 to September 5, 2008, Arequipa - Perú. (<http://www.physics.adelaide.edu.au/>)

Air Showers and Simulations

The spectrum of cosmic rays at Earth is shown in figure 1. The flux falls steeply with rising energy. Therefore, high energy particles are very rare and consequently require huge detectors. These can only be realised with natural media, such as the atmosphere. The primary CRs strike the atmosphere and produce extensive air showers (EAS) of secondary particles. At energies $> 10^{14}$ eV the only practical detection method is to detect these air showers and reconstruct the properties of the primary CR from the shape and particle content of the shower. The shower particles travel in a thin disk with an area up to many km^2 for the highest energies. The secondaries can be detected in arrays of particle detectors at ground level (e.g. scintillator counters or water tanks). The trace of the shower in the atmosphere can be seen by fluorescence telescopes which detect the near UV-light which is emitted by the nitrogen molecules when excited by the passage of the charged secondaries.

Air showers are also used for TeV gamma ray astronomy. Imaging Cherenkov Telescope systems (e.g. HESS [7] or CTA [8]) are observatories that investigate TeV gamma ray sources by recording the Cherenkov light emitted by air shower particles in the upper atmosphere.

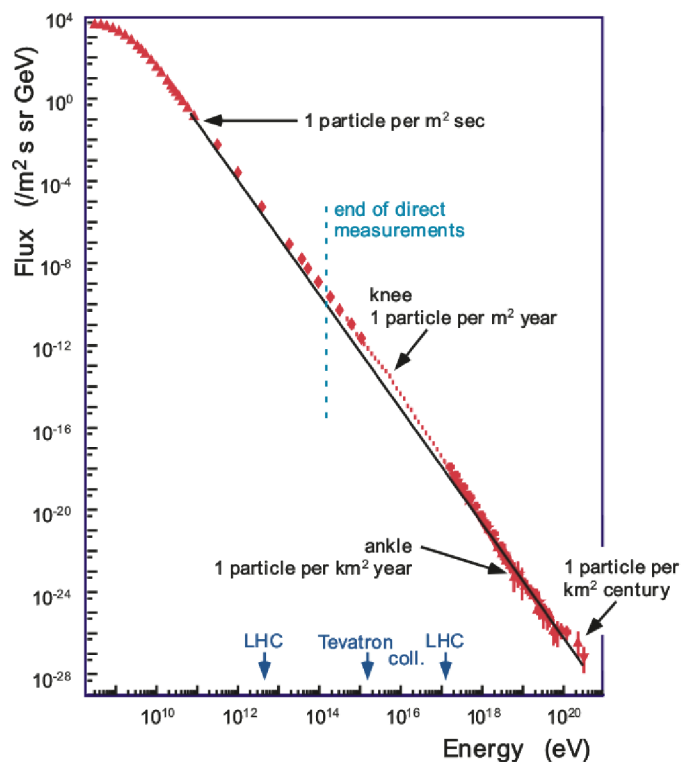


FIGURE 1. The energy spectrum of cosmic rays extends over >11 orders of magnitude in energy, and over 32 orders of magnitude in flux.

The Pierre Auger Observatory measures the highest energy cosmic rays, by combining a large surface detector (SD) array of $3000km^2$ area and a system of 24 fluorescence telescopes (FD) at the same site. While the SD provides the well-defined aperture and the 100% duty cycle needed to achieve high statistics, the FD provides, for a subset of about 10% of showers, calorimetric energy information and the possibility to calibrate

at 10^{19} eV. The charged particles ionise the air and produce Cherenkov and fluorescence light. The former is emitted in the forward direction, collimated in an opening angle of about 1° around the shower axis, the latter is emitted isotropically.

The air shower analysis aim primarily at the energy spectrum, the arrival direction distribution and the mass composition of the primary cosmic rays. The direction of a cosmic ray is easily determined from the arrival times of shower particles at distributed detectors in an array. The energy of the primary is deduced from the “size” of the shower, i.e. from the number of secondaries produced. The mass of the primary is reflected in the height of the shower development, the muon to electron ratio in the shower and the amount of fluctuations. On average, Fe showers have a higher 1st interaction and a higher shower maximum than proton showers. This is due to $\sigma_{\text{Fe-Air}}$ being larger than $\sigma_{\text{p-Air}}$ and the energy per nucleon is smaller. Fe interactions produce more secondaries (since $N_{\text{sec}} \sim \ln(E)$), and therefore more muons and less electrons-positrons and photons at ground level. They exhibit smaller fluctuations, since each shower is a composition of 56 subshowers which average out larger fluctuations. The height of the development and the increased muon number leads also to a faster signal rise time. To quantify how exactly primary energy and mass affect the observable quantities in an air shower is the subject of air shower analysis. However, the analysis is complicated by the fact that at high energies also the hadronic interaction are rather uncertain. There are no testbeams at CR energies, where the relevant hadronic reactions could be measured, nor is there a fundamental theory (such as QED for the electromagnetic interaction) on which extrapolation to the highest energies could be based. A variety of phenomenological models, tuned to the existing collider data, have been proposed. These models describe the overall features of air showers quite well, but fall short, if a precision in the prediction of less than about 30% is needed, (e.g. for telling proton and iron showers apart). Then the fundamental uncertainties in the details of the hadronic and nuclear interactions are the limiting factor. Thus, *understanding* the astro-particles means to find a combination of primary spectrum & composition and hadronic interaction model such that simulations and data are consistent with each other in all respects.

With recent air shower experiments progress has been made in various fields:

- TeV gamma ray experiments [7, 9, 11, 12] have detected more than 80 sources with a wealth of results on all many astrophysical objects and acceleration mechanisms. Thus, TeV astronomy has been established as a field with tremendous potential.
- The spectrum and the composition in the knee region has been measured by the KASCADE Collaboration [10].
- At highest energies the existence of the Greisen-Zatsepin-Kuz'min cut-off in the CR spectrum has been found [6, 13], limits on primary neutrinos [5] and photons [4] have been placed, and, for the first time, an anisotropy in the CR arrival directions has been detected, hinting at nearby AGN as sources of the CRs [2, 3].

EAS Models and the CORSIKA Program

To build a realistic air shower model requires a number of ingredients. The composition and layering of the interaction medium (atmosphere) needs to be specified. The *projectile* particles (primary and secondaries) can be virtually any known particles being produced in collisions (nuclei, p, n, e^\pm , γ , μ^\pm , π^\pm , π^0 , K^\pm , ρ , Ω , Δ , Λ , Σ , ...). Their interactions (cross section and particle production) with the nuclei in air and their decay properties need to be known. Tracking, deflection in the Earth magnetic field, multiple scattering, energy loss, absorption, ionization, Cherenkov light production are to be taken into account.

Most of these ingredients are well known. Thus, the Monte Carlo technique provides the ideal method to provide a convolution of all the elementary processes to form the complex model of the whole shower. If the individual parts are simulated properly, then also the overall shower is modelled realistically. The main uncertainties of the shower simulation come from the hadronic and nuclear interaction, where there are still large uncertainties.

An example of such an Air Shower Monte Carlo is CORSIKA (Cosmic Ray Simulations for Cascade) [1]. CORSIKA is a program for very detailed simulation of extensive air showers initiated by high energy cosmic ray particles. It used proven components wherever possible. For the uncertain hadronic interactions at high energies a number of interaction models have been implemented: QGSJET, DPMJET, EPOS, (models are based on the successful Gribov-Regge theory of multiple Pomeron exchange), SIBYLL (a minijet model with Gribov-Regge type features), and FLUKA and UrQMD are used for hadronic interactions at low energies.

For electromagnetic interactions a modified version of the shower program EGS4 (or the analytical NKG formulae) can be used. Options for the simulations of Cherenkov radiation and neutrinos exist. CORSIKA can be used from 10^6 up to 10^{21} eV and beyond. For references to the models, see the references in ref. [1]. CORSIKA allows production of very educational shower images, and movies of the shower development (see e.g. [14, 15]). Examples of simulated showers are shown in figs. 7 and 8. Note the richness of details and the complexity of individual trajectories.

Hadronic Interactions

The features of hadronic interaction models that are most relevant for the air shower development are the interaction cross sections and the secondary particle productions. The latter is usually described by the inelasticity, the fraction of the projectile's energy spent to produce secondary particles. Larger cross sections and bigger inelasticities produce short showers. Conversely, lower cross sections and smaller inelasticities lead to longer showers.

Hadronic interaction can be of very different types:

Elastic scattering leaves projectile and target unchanged. Very little energy and momentum and no charge or colour are transferred. The scattering of the projectile is very small. As these interactions are elastic, i.e. do not produce secondaries, they do not

contribute to the air shower development. They can be ignored in the simulations.

Inelastic, diffractive scattering has very similar characteristics to elastic scattering (small scattering, small energy and momentum transfer, no charge or colour transfer), but one or both of the collision partners are put into an excited state, which decays in a small number of secondaries that are emitted in the forward direction (i.e. the direction of one of the incoming collision partners). About 10% of all inelastic p-p reactions are found to be diffractive, These interactions have a very small inelasticity and therefore make the showers longer. The actual fraction (and its change with energy) is therefore one of the crucial inputs to an hadronic interaction model.

Inelastic, non-diffractive, soft scattering is the lion's share of the interactions happening in an air shower (about 90% for p-p). The energy and momentum transfer is still small, but as colour is transferred between quarks of the collision partners, 2 colour strings are stretched between a quark from one collision partner to a di-quark from the other one, and vice versa. Upon fragmentation of the colour strings many secondaries are produced which carry away much of the available energy. The secondaries are still largely emitted along the initial direction of the collision partners. As these reactions have only a small momentum transfer (i.e. they are "soft"), they cannot be calculated within QCD. This is the domain of the Gribov-Regge theory, that is also used to describe elastic scattering.

Inelastic, non-diffractive, hard scattering has also colour transfer, creating strings and many secondaries. In addition, hard reactions have a large momentum transfer, and thus these reactions can be computed in QCD. This is the domain of particle physics and collider experiments. Due to the large momentum transfer the secondary particles are emitted under large angles to the incoming collision partners, and can easily be recorded in detectors placed around the beam pipe. These reactions are best suited for producing new particles (such as W, Z, t-quarks or the HIGGS), but the cross sections for hard interactions are 10^2 to $10^8 \times$ smaller than those for soft interactions. Therefore, hard interactions play only a minor role in air shower physics.

The change of cross sections with energy has been measured for p-p collisions and shows the resonance region ($p < 2$ GeV/c), a smooth decrease ($2 < p < 100$ GeV/c), and then a steady rise ($p > 100$ GeV/c), and are usually explained by the exchange of complex, hypothetical exchange particles, called the "Reggeon" (responsible for the falling part of the cross section) and the "Pomeron" (for the rising part). The highest available measurement is still at Tevatron energies (i.e. 2×900 GeV in the centre of mass system, corresponding to $p = 1.7 \times 10^6$ GeV/c), but for the highest-energy cosmic rays cross sections up to $> 10^{20}$ eV = 10^{11} GeV are needed. The large extrapolation (over 5 orders of magnitude) depends crucially on the assumptions made on the energy dependence of the p-p cross section. The more relevant p-nucleus, π -nucleus (see fig. 3) and nucleus-nucleus cross sections as function of energy are then computed via the Glauber formalism (see fig. 4, where nucleus-nucleus collisions are effectively reduced to nucleon-nucleon collisions of a subset of participating nucleons, while the other nucleons remain spectators as illustrated in fig. 4). The uncertainties in σ_{p-p} translate directly into similar uncertainties in σ_{p-Air} , $\sigma_{\pi-Air}$ and $\sigma_{nucleus-Air}$. Despite a sizeable reduction over the last 10 years, typical systematic uncertainties are still of the order of $\pm 25\%$ at high energies. An average value at $E=10^{18}$ eV, for instance, is $\sigma_{p-Air} \approx 500$ mb. At much lower energies, where measurements from colliders constrain the models,

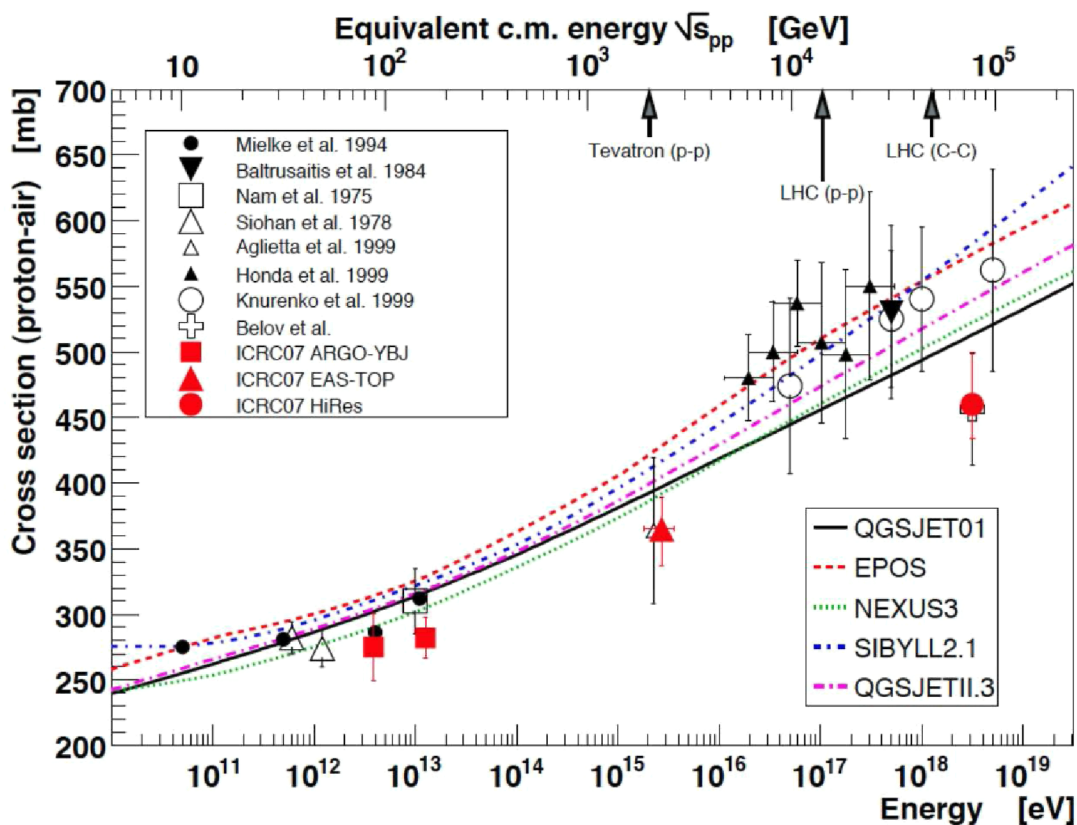


FIGURE 3. Inelastic p-Air cross-section as a function of energy. The scatter of the experimental points and the models indicates the size of the experimental and theoretical uncertainties.

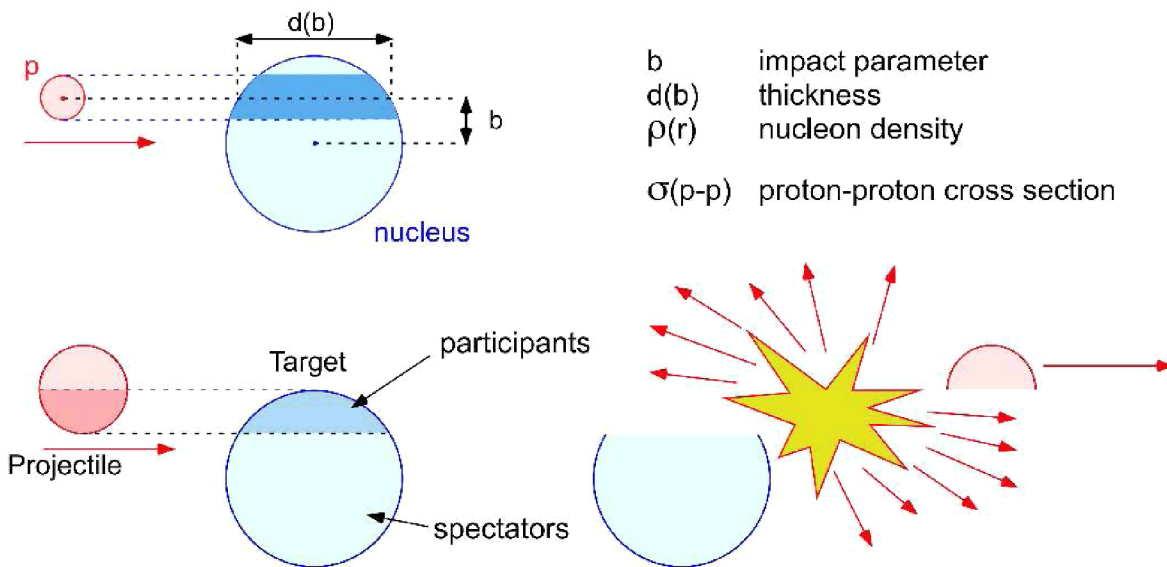


FIGURE 4. Schematic of proton-nucleus and nucleus-nucleus collision. Only those nucleons can participate in the collision which are in the actual collision region. The other nucleons remain spectators that are largely unaffected. The interacting nucleons are governed by the proton-proton cross section $\sigma_{p-p}(E)$.

they usually agree within a few percent.

The particle production depends on the details of the production of colour strings and their subsequent fragmentation into observable particles. Again, at low energies the models are constrained by collider results, but at higher energies rather large differences occur. Typical distributions used to tune models are the distribution of the number of secondary particles produced, the distribution of their emission angles (often in form of a plot of “rapidity” or “pseudorapidity” distributions), and the number and energy of particles produced in the very forward direction. Not surprisingly, the number of secondaries is much less important than the number of secondaries in the very forward direction, which is crucial for the air shower development. Distributions of the number of participating nucleons in projectile and target of nuclear collisions provide rather basic tests of the quality of the simulations, already disqualifying the more simple implementations of the Glauber formalism.

In the end, the influence of details of hadronic interaction models on measurable quantities in an air shower experiment have to be checked. The longitudinal development of a shower (i.e. the number of secondaries as function of atmospheric depth) can be directly measured by fluorescence detectors. The most characteristic feature is the position of the shower maximum X_{\max} . Unfortunately, the fluctuations from shower to shower are so large ($\approx \pm 100 \text{ g/cm}^2$ for proton showers), that only average values can be reasonably compared. The remaining uncertainties in $\langle X_{\max} \rangle$ due to the models are of the order of $\pm 20 \text{ g/cm}^2$, whereas the difference between proton and iron showers is of the order of 75 g/cm^2 . The lateral distribution of particles in a shower, as it would be measured in an air shower array, is usually well described, but the fraction of muons is very dependent on the models. The rather recent model EPOS, for instance, predicts for 10^{19} eV showers more than 30% larger muon numbers at 1 km core distance than the other models, with almost the same density of electrons and photons. Also for other array observables, the residual uncertainty due to models is of the order 10-30%. Thus, confronting simulations with experiments to analyse subtle differences below that level, e.g. to tell protons from iron showers, is still seriously limited by model uncertainties.

Over the last 10 years the agreement of simulations with experimental data has greatly improved. Today the simulations describe *consistently*, within about 20-30%, air shower experiments from TeV (i.e. the γ -ray regime), to PeV (i.e. the region of the knee in the energy spectrum), to EeV (where Haverah Park, AGASA and other arrays have taken data) and up the 10^{20} eV range of the Pierre Auger experiment. Given the limited data base on which the models have to be built, and the large extrapolation needed to cosmic ray energies, this is a remarkable achievement.

In the near future accelerator experiments at the LHC and RHIC will provide new experimental inputs to cross-sections, diffraction and hadronic particle production under small angles, further constraining the hadronic interaction models and thereby improving the predictive power of air shower simulations at the highest energy. New astroparticle experiments will provide new constraints at higher energies and data with improved quality (e.g., from KASCADE-GRANDE, Auger, ICE cube, AMS, and others). Only High Energy Physicists and Astroparticle Physicists together can solve the problem of origin of the high energy cosmic rays (the highest-energy particles in the Universe) and its hadronic interactions with the atmosphere.

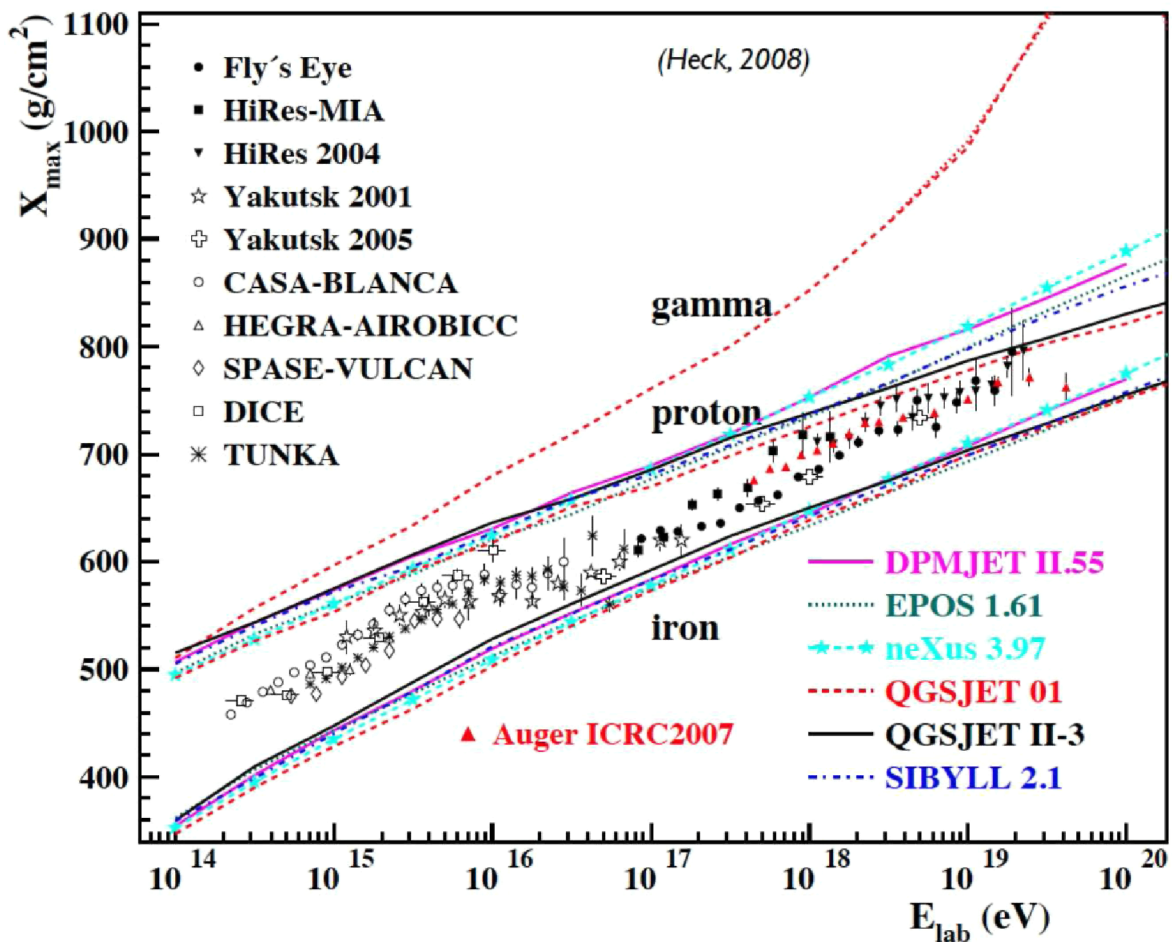


FIGURE 5. X_{\max} as a function of energy. The experimental data in the energy range 10^{14} to 10^{20} eV lie between the simulations of the generally accepted extremes (pure protons and pure Fe composition), thus being consistent with the assumption of a mixed composition. The scatter of the lines indicates the differences between hadronic interaction models used in CORSIKA.

Selected Aspects of EAS simulation & Simulation Techniques

Random Numbers: At the heart of a Monte Carlo program is the generation of random numbers, to pick random values from probability distributions, needed in the simulation of interactions, decay and transport of particles. True random numbers however, are not suitable in simulation programs, as a simulation needs to be fully reproducible, e.g. to find programming errors in large programs. Therefore, pseudo-random numbers are used, which are produced by a random number generator following a fully predictable algorithm, but the numbers produced are showing all properties of truly random numbers.

The most basic random number generator is one that produces a uniform distribution of numbers between 0 and 1. All more complicated distributions are deduced from uniform random numbers. It is clear that a computer can only represent a finite set of numbers between 0 and 1, and as the algorithm to create random numbers is deterministic, any random number sequence repeats itself at some stage. Thus, good random number generators are those that have a very long sequence length. Modern generators (such

as RANMAR or RANLUX from the CERN library) can produce more than 9×10^8 independent sequences of a period length of 10^{43} .

There are several methods to produce random numbers following a certain distribution from uniformly distributed random numbers.

Gaussian distributed random numbers, for instance, can be produced by the following method. If x_1 and x_2 are two independent, uniformly distributed random numbers, then $y_1 = \sqrt{-2 \ln(x_1)} \cdot \cos(2\pi x_2)$ and $y_2 = \sqrt{-2 \ln(x_1)} \cdot \sin(2\pi x_2)$ are independent, Gaussian distributed random numbers.

More general, if random numbers are needed that follow the probability density function $f(x)$, then one can form the integral function $F(x) = \int_{-\infty}^x f(x) dx$. By construction, the values of $y = F(x)$ are between 0 and 1. In the case we can form the inverse function $G(y) = F^{-1}(y)$, we can draw a random number y uniformly from the interval 0 to 1, then compute $z = G(y)$, and get z which is distributed as $f(x)$. This is a very fast and elegant method, but it requires that we are able to invert $F(x)$, which is not always the case.

If $f(x)$ is a function that cannot be integrated and inverted, the brute-force rejection method can be used (see fig. 6). One draws first a uniformly distributed random number

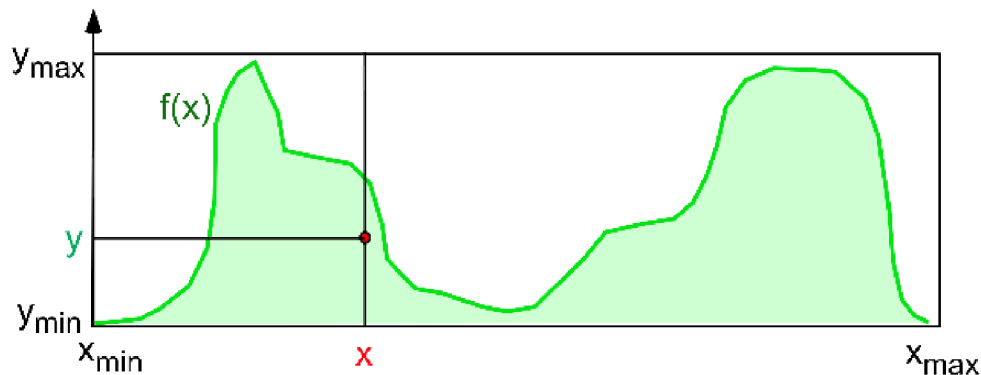


FIGURE 6. Production of random numbers following the function $f(x)$ using the rejection method.

x from the range x_{\min} to x_{\max} . Then one draws a second random number y that is uniformly distributed between y_{\min} to y_{\max} . If then $y < f(x)$, then x is accepted as random number, otherwise it is rejected and the process starts again. Like this, eventually the accepted numbers x will be distributed like $f(x)$, but for each accepted random number x at least 2 uniformly distributed random numbers have to be drawn, and the rejection makes this a rather ineffective method. To produce good random numbers is computing time intensive. CORSIKA spends about 30% of the computing time just to produce random numbers.

Atmosphere: The atmosphere is composed of 78% Nitrogen, 21.5 % Oxygen and 0.5% Argon. All other elements in the air are neglected because they make only 1% of the total. The atmospheric density is roughly given by $\rho(h) \approx \rho(0) \cdot e^{-h/h_0}$ with $\rho(0) = 0.00123 \text{ g/cm}^3$ at sea level and $h_0 \approx 7 \text{ km}$. From this, the total atmospheric overburden at any height h , $T(h) = \int_h^{\infty} \rho(h) dh$ can be computed. $T(0)$, the overburden at sea level, is about 1000 g/cm^2 , but the atmospheric layering varies with pressure and temperature at diurnal and seasonal time scales, requiring a more detailed atmospheric description. The amount of matter to be traversed by a shower to reach a detector at

height h is then $T(h)/\cos(\theta)$, with θ being the zenith angle of the shower. A typical development of the number of shower particles with atmospheric depth t (in g/cm^2) is often described by $N(t) \approx c \cdot t^a \cdot e^{-bt}$, with t^a giving the rising part and e^{-bt} giving the falling part of the shower. A 10^{20} eV proton shower coming in vertically maximises around at sea level, a shower at a zenith angle of 70° will have to penetrate >3000 g/cm^2 such that its electromagnetic component is practically absorbed and only muons arrive at ground level. The atmospheric model in CORSIKA uses various atmospheres that are parameterised in 5 layers with piece-wise fitted exponential. Special versions for different locations and seasons can be introduced.

Decay versus interaction: For secondary particles which are unstable, it is complicated to determine the location of the next action point, as collision and decay compete. The location of the next collision depends on the matter traversed (x) according to the following simple probability distribution: $dN/dx \sim e^{-x/\Lambda_0}$, where Λ_0 is the collision mean free path. The location of the next decay depends on the time, according to $dN/dt \sim e^{-t/\tau}$ with $\tau = \gamma\tau_0$ being the proper life time τ_0 multiplied with the Lorentz factor γ . Since $t = s/c = x/\rho c$ it follows that $dN/dx \sim e^{-x/\rho c\tau}$, where ρ is again an exponential depending on the height. Thus an analytic evaluation of the combined probability for *collision or decay* becomes quite difficult. The Monte Carlo method solves this problem elegantly, by computing independently a next action point due to a collision and the one due to decay, and selecting the one that happens first:

- (i) determine x_c from $dN/dx \sim e^{-x/\Lambda_0}$,
- (ii) determine t_d from $dN/dt \sim e^{-t/\tau}$, and compute $x_d = t_d c \rho$
- (iii) if $x_c < x_d$ then do a collision, otherwise a decay.

This is a very fast procedure and it is exactly how nature determines the fate of an unstable particle.

Electromagnetic showers – from a toy model to EGS4: The main reaction of high-energy photons is e^+e^- pair production and of energetic electrons it is bremsstrahlung. Both reactions have about the same characteristic length, the radiation length ($X_0 \approx 37$ g/cm^2 in air), and lead to 2 particles in the final state. The Bethe-Heitler model describes qualitatively the development of an electromagnetic shower of energy E_0 , by using a very simple description of these reactions: After each X_0 , the particle number doubles, and the energy per particle is halved, until the energy of the secondaries falls below the critical energy E_c , where particle multiplication stops. Here is the number of secondaries maximal. Thus, we find at depth $t = k \cdot X_0$ (with $k = 1, 2, \dots$) that $N(t) = 2^k$ and $E(t) = E_0/2^k$. At the maximum of the shower development we find that

$$N_{\max} = E_0/E_c \quad \text{and} \quad t_{\max} = X_0 \cdot \ln(E_0/E_c)/\ln(2) ,$$

i.e. the number of secondaries at the maximum is proportional to the primary energy E_0 , and the position of the maximum increases with $\ln(E)$. The latter follows directly from $E_c = E_0/2^{k_{\max}}$ and $t_{\max} = X_0 \cdot k_{\max}$. These findings hold also for the much more complicated full simulation.

A better, but still analytic description of electromagnetic showers is provided by the Nishimura-Kamata-Greisen (NKG) formulae. The electron number as a function

of atmospheric depth is

$$N_e(t) = \frac{0.31 \cdot e^{t(1-1.5\ln s)}}{\sqrt{\ln(E_0/E_c)}} \quad \text{with} \quad s = \frac{3t}{t + 2\ln(E_0/E_c)} .$$

s is the so-called *age* of the shower. In the shower maximum, i.e. $t = t_{\max}$, the value of $s \equiv 1$. The lateral distribution of electrons at ground level is given by:

$$\rho_e(r) = \frac{N_e}{2\pi r_m^2} \cdot \frac{\Gamma(4.5 - s)}{\Gamma(s)\Gamma(4.5 - s)} \left(\frac{r}{r_m}\right)^{s-2} \left(1 + \frac{r}{r_m}\right)^{s-4.5} ,$$

where ρ_e is the electron density as function of the core distance r , and $r_m = (0.78 - 0.21s) \cdot r_{\text{mol}}$, with the Moliere radius $r_{\text{mol}} \approx 9.6 \text{ g/cm}^2$. The NKG formulae do describe the shower shape of purely electromagnetic showers quite well, but they fail to deliver information on the fluctuations and are not strictly valid for hadronic showers. Nevertheless, the same functional forms are used to describe also hadronic cosmic ray showers. The variables r_m and the powers $s - 2$ and $s - 4.5$ lose then their physical meaning and become free parameters in the fit.

The well-proven Electron Gamma Shower code (EGS) simulates all reactions of e^+ , e^- and photons in much greater detail and, therefore, provides a much better description of an electromagnetic shower, including fluctuations, lateral extension, energy distribution of shower particles and so on. EGS has been checked and verified by many researches in a big variety of different experiments. As EGS is based on QED, the extrapolation to the highest energies have fairly small uncertainties. Therefore, it is recommended to rely on EGS whenever possible. However, due to the large number of electrons and photons in an air shower and due to its complexity, EGS is relatively slow and dominates by far the computing time.

Simulation Speed-up: Computing time and disk space needed for a shower simulation grow approximately linear with primary energy. As a 10^{15} eV shower needs about one hour to be simulated and produces close to 300 MB particle output on the disk, it is straight forward to estimate that a single shower at 10^{20} eV, with about 10^{11} secondary particles, needs about 10^5 hours (= 11.4 years) and a disk space of 30 Tera bytes. Obviously this is not practical.

Therefore, statistical subsampling (or thinning) [17] is used to discard most of the shower particles and follow only a representative subset of them, which carry a weight to account for the discarded ones. Thinning sets in once the particle energies drop below a certain threshold (typically 10^{-6} of the primary energy) allowing the initial part of the shower to develop without thinning. While the energy and the average properties of the shower are preserved, the statistical fluctuations are enhanced due to thinning and may bias the simulations in regions where only few particles are followed, e.g. at very large core distances. The fluctuations are especially large when particle acquire very high weights. Therefore, a limitation of weights is introduced which costs some computing time, but keeps the artefacts due to thinning relatively small. For a good compromise of thinning and weight limitation, typical seed-up factors of 10^4 - 10^6 are reached. Thus, the simulation of one 10^{20} eV shower is reduced to about 20 hours.

With CORSIKA also unthinned showers at high energies can be simulated, by distributing subshowers onto many processors to reduce the computing time. The outputs of the subshowers are then stored in thousands of output files, which need to be combined for the analysis.

Install and run CORSIKA

The first point of contact and the place to find user manuals and general informations is the CORSIKA webpage: <http://www-ik.fzk.de/corsika/>. CORSIKA is a set of routines written in Fortran and C. The source and all files belonging to CORSIKA can be downloaded from the anonymous ftp server <ftp-ik.fzk.de>. However, to get access, You have to be registered. Send a request to Tanguy Pierog (e-mail: tanguy.pierog@ik.fzk.de) or Dieter Heck (e-mail: dieter.heck@ik.fzk.de), with the hostname of the computer (or static IP addresses) from which you intend to access the CORSIKA ftp site. Once the hostname of your computer has been registered you will receive instructions on how to download of CORSIKA code from this server. To install the software on LINUX systems one needs

- gcc - the GNU C/C++ compiler
- make - a Compilation tool
- g77 - the GNU Fortran 77 compiler.

For the version to make shower plots one needs in addition

- libpng12 - the PNG library (used by PLOTSH2 option)
- libpng12-dev - the development header files for libpng12 (used for PLOTSH2 option).

The CORSIKA set is distributed as a compressed tape archive (.tar.gz) file for each version (e.g. `corsika-6735.tar.gz`). Once you have got the CORSIKA file, unpack and compile CORSIKA by typing:

```
$ tar zxvf corsika-6735.tar.gz
$ cd corsika-6735/
$ ./corsika-install
```

The `corsika-install` script lets you configure the version of CORSIKA you want to use. For information about the program variants consult the CORSIKA Users Guide. By always typing <Enter> you will install the default version.

The PLOTSH and PLOTSH2 are necessary to produce plots of the air shower. Finally, if all goes well, you end up with an message similar to the following :

```
--> "corsika6735Linux_QGSJET_gheisha" successfully installed in:
    /home/marko/corsika-6735/run/
--> You can run CORSIKA in /home/marko/corsika-6735/run/ using:
    ./corsika6735Linux_QGSJET_gheisha < all-inputs > output.txt
```

To run a simulation, change to the “run/” subdirectory of the installation, where you find the executable generated by the compilation process. Execute it by typing:

```
$ ./corsika6735Linux_QGSJET_gheisha < all-inputs > output.txt
```

CORSIKA uses the “all-inputs” file to set the options for the simulation (such as primary particle type, energy, zenith angle, etc.). You must edit this file to modify the options according to what you want to simulate. All options are explained in the User’s Guide. Here we want to explain in a bit more detail how shower plots are created. When installing and compiling the program, you must use the CORSIKA option “PLOTSH” to be able to generate .map and .track files which are later used by the plotting programs. Then add “PLOTSH T” in the “all-inputs” configuration file. It will generate the following extra files

```

DAT000001.em_xy.map
DAT000001.em_xz.map
DAT000001.em_yz.map
DAT000001.hd_xy.map
DAT000001.hd_xz.map
DAT000001.hd_yz.map
DAT000001.mu_xy.map
DAT000001.mu_xz.map
DAT000001.mu_yz.map
DAT000001.track_em
DAT000001.track_hd
DAT000001.track_mu

```

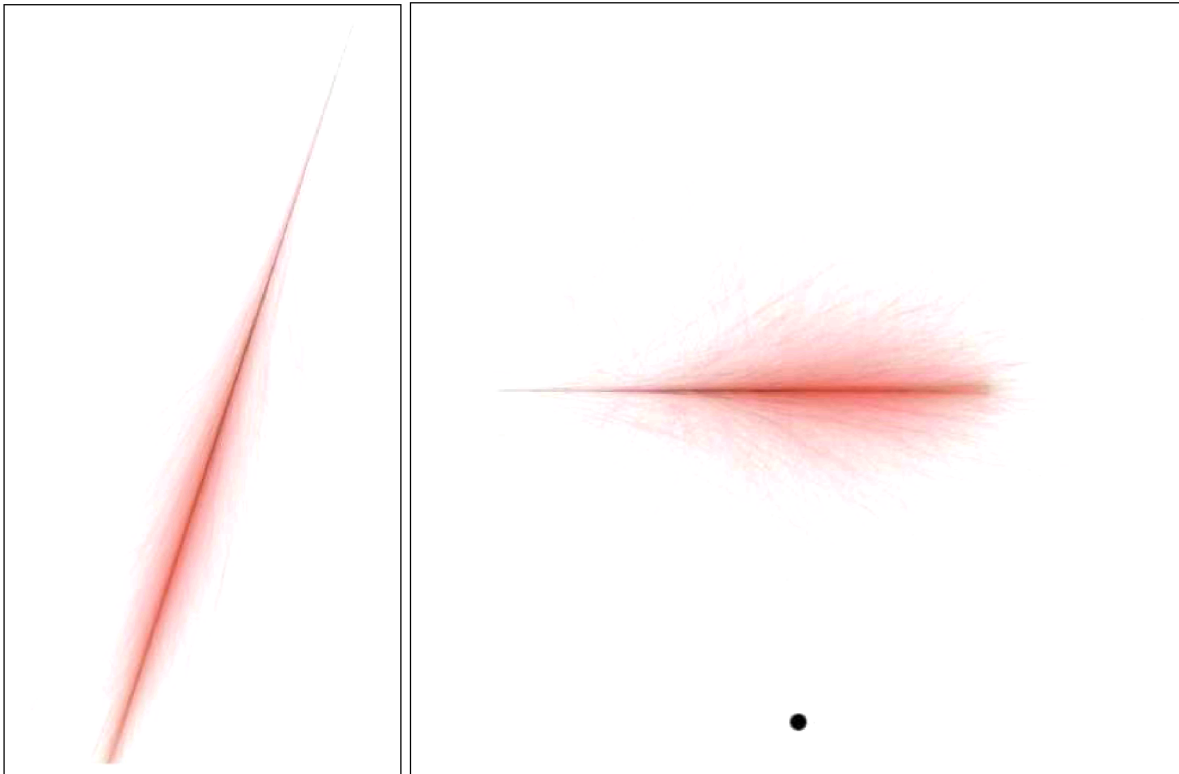


FIGURE 7. Proton, 10 TeV, $\theta = 20^\circ$; left: x-z projection, right: x-y projection ($-1 < x < 9$ km, $-5 < y < 5$ km, $0 < z < 22.5$ km).

The program *map2png* uses the .map files to generate a single PNG image. Execute the program by typing:

```
$ ./map2png <CORSIKA DATnnnnnn file> <output PNG file>
```

The colours used by default in the plot are red for electromagnetic, green for muon, blue for hadronic, and black for the background. Also by default, we get a projection xz of the shower. However, it is possible to specify other colours or projections to be used. To see a list of options run *map2png* without arguments. To control the axis range displayed, the keyword “PLAXES” followed by 6 parameters (xMin xMax yMin yMax zMin zMax) must be used in the “all-inputs” file. The default values (in cm) are

```
PLAXES -5.E5 5.E5 -5.E5 5.E5 0. 3.E6
```

In Figs. 7 examples of high-resolution shower plots made with *map2png* are shown.

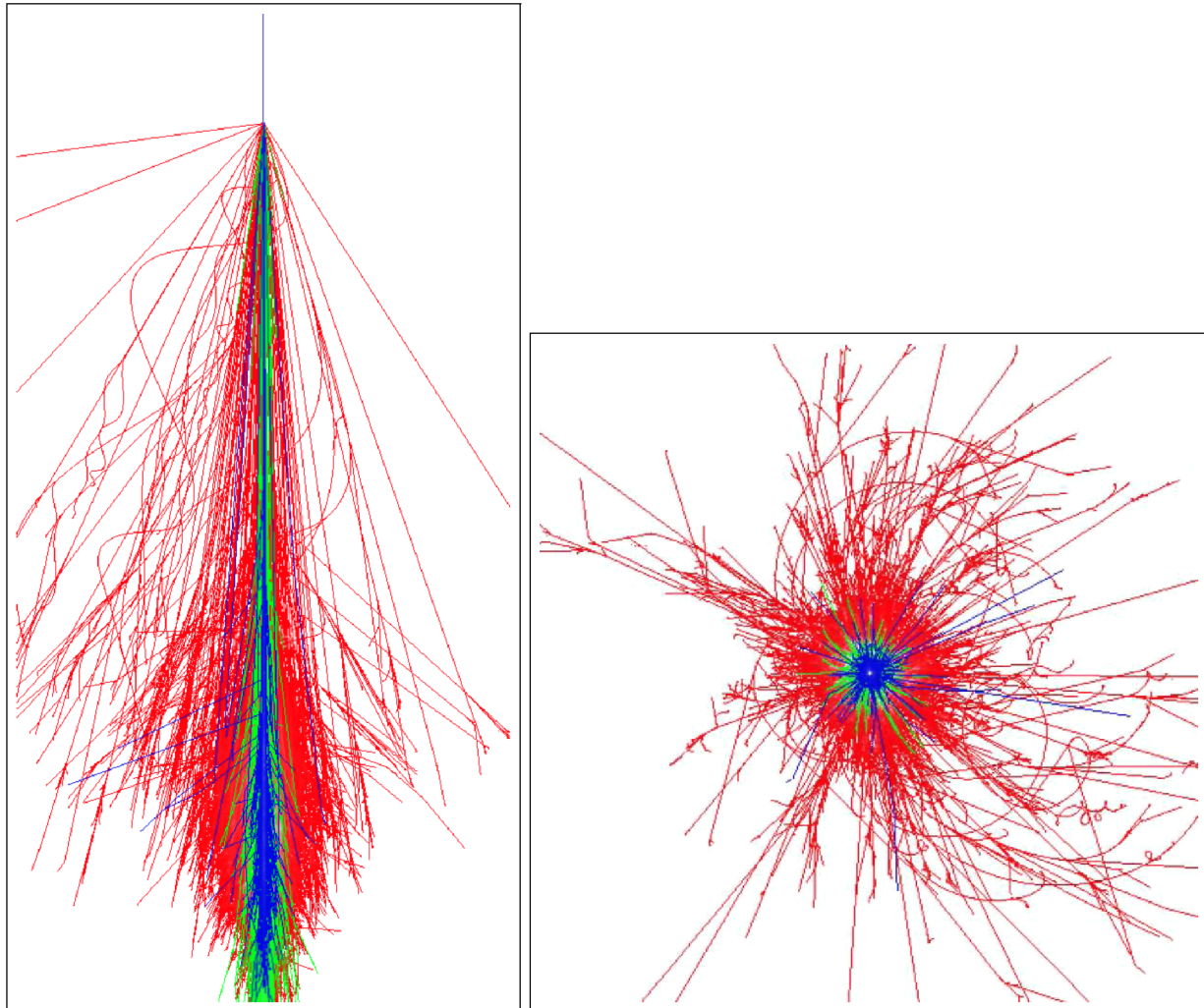


FIGURE 8. Iron 10 TeV, $\theta = 0^\circ$, x-z and x-y projections; $-5 \text{ km} < x, y < 5 \text{ km}$, $0 < z < 28 \text{ km}$

The program *plottracks* uses the *.track* files to produce separately one plot of the tracks of electromagnetic particles, one of muons, one of hadronic particles, and one combined plot. The colors used in the plots are red for electromagnetic, green for muonic and blue for hadronic particles. Start the interactive program by typing “./plottracks”. Define interactively the projection to be used, the radius range in the plane x-y, the background colour in the plot and the energy cuts to be used for electromagnetic, muon and hadronic particles. Additionally, if one wants to control the height range of tracks

plotted, one needs to modify the Fortran code *plottracks3c.f* (available in the directory *src/*) in the following lines:

```
c height : 28 km
      zmin = 0.110
      zmax = 28.
```

and recompile it. Fig. 8 shows an Iron-induced shower at 10 TeV, produced with *plottracks*.

Acknowledgements: The authors are very grateful for the invitation to the 3rd School on Cosmic Rays and Astrophysics. JK acknowledges the financial support from the organisers. AVCG acknowledges support from the Centro de Tecnologías de Información y Comunicaciones, Universidad Nacional de Ingeniería, Lima, Perú.

REFERENCES

1. CORSIKA web site: <http://www-ik.fzk.de/corsika/>
2. J Abraham et al., Auger Collaboration, *Science* 938 (2007) 318
3. J Abraham et al., Auger Collaboration, *Astropart. Phys.* 27 (2007) 244
4. J Abraham et al., Auger Collaboration, *Astropart. Phys.* 27 (2007) 155
J Abraham et al., Auger Collaboration, *Astropart. Phys.* 29 (2008) 243
5. J Abraham et al., Auger Collaboration, *PRL* 100 (2008) 211101
6. J Abraham et al., Auger Collaboration, *PRL* 101 (2008) 061101
7. HESS website: <http://www.mpi-hd.mpg.de/hfm/HESS/>
8. CTA web site: <http://www.mpi-hd.mpg.de/hfm/CTA/>
9. Magic web site: <http://wwwmagic.mppmu.mpg.de/>
10. KASCADE web site: http://www-ik.fzk.de/KASCADE_home.html
11. Tibet air shower array web site: <http://www.icrr.u-tokyo.ac.jp/em/index.html>
12. MILAGRO web site <http://www.lanl.gov/milagro/>
13. HIRES web site: <http://www.cosmic-ray.org/>
14. CORSIKA images: <http://www.ast.leeds.ac.uk/~fs/>
15. CORSIKA movies: <http://www.ast.leeds.ac.uk/~knapp/movies/> and
<http://www-ik.fzk.de/corsika/movies/>
16. M Nagano, AA Watson, *Rev. Mod. Phys.* 72 (2000) 689
17. M Kobal, *Astropart. Phys.* 15 (2001) 259