Implementing the De-thinning Method for High Energy Cosmic Rays Extensive Air Showers Simulations

Alex Estupiñán[a], Hernan Asorey[a,b], Luis A. Núñez[a,c]

[a]*Escuela de Física, Universidad Industrial de Santander, Carrera 27 and Calle 9, (680002) Bucaramanga, Colombia*
[b]*Laboratorio Detección de Partículas y Radiación, Centro Atómico Bariloche & Instituto Balseiro, Av. Bustillo 9500, (8400) Bariloche, Argentina*
[c]*Departamento de Física, Universidad de Los Andes, Sector La Hechicera S/N, (5101) Mérida, Venezuela*

Abstract

To simulate the interaction of cosmic rays with the Earth atmosphere requires highly complex computational resources and several statistical techniques have been developed to simplify those calculations. It is common to implement the thinning algorithms to reduce the number of secondary particles by assigning weights to representative particles in the evolution of the cascade. However, since this is a compression method with information loss, it is required to recover the original flux of secondary particles without introducing artificial biases. In this work we present the preliminary results of our version of the de-thinning algorithm for the reconstruction of thinned simulations of extensive air showers initiated by cosmic rays and photons in the energy range $10^{15} < E/\text{eV} < 10^{17}$.

Keywords: Thinning, De-thinning, Extensive Air Showers, Cosmic Rays

1. Introduction

The Earth atmosphere interacts continuously with a flux of particles of galactic and extragalatic origin. The interaction of one of these cosmic rays with an atomic element in the atmosphere produces a cascade of particles (the so called Extensive Air Shower, EAS), in which the number of particles could reach billions at the maximum of the development of the cascade.

To understand and simulate these processes requires highly complex computational resources. To reduce the number of particles that have to be followed at the highest energies, it is common the usage of the thinning algorithm (see for example [1]), a statistical method that reduces the number of secondary particles by assigning weights to representative particles in the evolution of the cascade. However, since this is a compression method with loss of information, it is required to recover the original flux of secondary particles without introducing artificial biasses.

The so called de-thinning method [2] is one of the exis-

tent methods designed to deal with this information loss. Since the EAS is an stochastic process, in this work we will compare the distribution of particles at ground level for several non-thinned showers, with the corresponding reconstructed thinned (with two different levels of thinning, $\epsilon_{th} = 10^{-7}$ and $\epsilon_{th} = 10^{-8}$) and de-thinned showers obtained from exactly the same set of initial parameters. All the simulations were made by using CORSIKA air shower simulation program [3] and several analysis routines developed within the LAGO (Latin American Giant Observatory) project simulation chain [4, 5].

2. Thinning Method

The implementation of this algorithm is applied during the shower simulation over the secondary particles when this condition is fulfilled:

$$E_0 \epsilon_{th} > \sum_{j=1}^{n} E_j, \qquad (1)$$

where E_j is the energy of the secondary particle, E_0 is the energy of the primary particle and $\epsilon_{th} = E_j/E_0$ is defined as the level of thinning.

In this case, only one secondary particle i survives probability:

$$P_i = E_i / \sum_{j=1}^{n} E_j. \qquad (2)$$

Otherwise, if the total sum of the energy of the n secondary particles is greater than the thinning energy threshold, i.e.:

$$E_0 \epsilon_{th} < \sum_{j=1}^{n} E_j, \qquad (3)$$

then the secondary particle with energy below the thinning threshold will survive with a probability:

$$P_i = E_i / E_0 \epsilon_{th}. \qquad (4)$$

In both cases, the particles that survive have their weight multiplied by a factor of $w_i = P_i^{-1}$.

3. De-thinning Method

The main questions to answer for any reconstruction method of thinned showers are:

- How do you determine the accuracy of the survival sample of secondary particles?

- How do you use the thinned sample to completely rebuild the original shower avoiding the introduction of artificial biases?

- What is the maximum value of ϵ_{th}, for which the sample is a good representative of the original shower for a particular type of reconstruction method?

Several techniques has been developed to asses these questions. Our implementation of the de-thinning algorithm is based in the original development of Stokes et al. [2], and consist in the successive application of the following steps over the thinned sample of secondary particles:

1. choose a vertex point on the trajectory of the weighted particle in the way given in the next paragraph;

2. centred in the weighted particle path, chose a random direction of propagation for the new particles (also called "daughter" particles) by using a 2D gaussian distribution with zero mean and a fixed standard deviation;
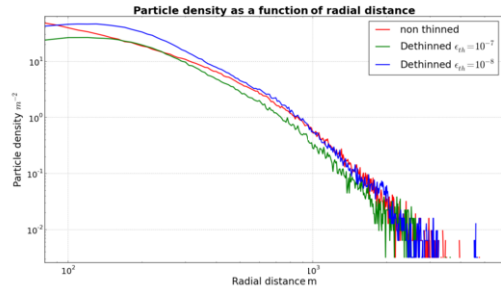

Figure 1: Particle density as a function of the core distance for a vertical proton of energy 5×10^6 GeV. We use the non-thinned shower (red line) as the reference, and compare it with the reconstructed density profile obtained after the application of our de-thinning method over the same showers but with thinning levels of $\epsilon_{th} = 10^{-7}$ (dark green line) and 10^{-8} (blue line). We optimize our method to get good results on intermediate distances to the core, where typical detectors used in many astroparticles observatories are triggered but not saturated.

3. project the daughter particle to ground level, and calculate the travel distance in units of atmospheric depth;

4. determine the energy of the inserted particle using a gaussian random distribution with mean E_i and $\sigma = 0.1E_i$, where E_i corresponds to the energy of the weighted particle;

5. determine the probability of atmospheric absorption of the inserted particle by using a fixed atmospheric interactions length, and decide if it will reach the ground;

6. depending on the daughter particle trajectory and its starting point, calculate the time of flight of the inserted particle;

7. repeat steps two to six, $(w - 1)$ times to build a complete set of secondary unweighted particles.

To assure temporal consistency of the particles reaching ground level, the distance between the vertex where the sub-shower of the de-thinned particles begins and the ground should be less than D_{max}, which is given by:

$$D_{max} = \frac{c^2(t_i - t_0)^2 - | \vec{x_i} - \vec{x_0} |^2}{2(c(t_i - t_0) - (\vec{x_i} - \vec{x_0}) \cdot \hat{p_i})} \qquad (5)$$

where c is the speed of light. Any shorter separation will generate de-thinned particles temporally consistent with the development of the EAS.

We have implemented the above de-thinning algorithms, and introduce some improvements in the overall algorithm, such as:
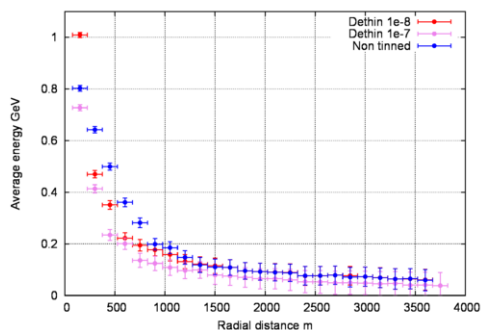
Figure 2: Average energy of secondaries as a function of the radial distance for a shower initiated by a vertical proton with energy $E_0 = 5 \times 10^6$ GeV without thinning (blue circles), and de-thinned with original thinned levels of $\epsilon_{th} = 10^{-7}$ (magenta thin circles) and $\epsilon_{th} = 10^{-8}$ (blue circles).

- the standard deviation used to determine the daughter trajectory depended on the type (electromagnetic, muon or hadron) and the energy E_i of the weighted particle;

- location dependent atmospheric model to determine the atmospheric depth as a function of the altitude and trajectory of the weighted particle;

- the atmospheric interaction length used to determine the probability of reaching ground level also depends on the type of secondary particle;

Our de-thinning code was implemented in python 2.7 and will be licensed under GPLv3.

4. Results

Our main results for this first approach to this problem are displayed in Figures 1 and 2, where it is possible to compare the secondary particle density and the average energy as a function of the distance to the shower core. As explained above, we simulate showers at fixed energies and arrival directions for different primaries and different values of thinning levels: $\epsilon_{th} = 0$ (non-thinned, used as reference) and $\epsilon_{th} = 10^{-7}$ and $\epsilon_{th} = 10^{-8}$. The set of parameters used for each simulation, including the CORSIKA random generators seeds, were exactly the same to allow proper comparisons between the de-thinned and the non-thinned showers.

With this implementation, we obtained very good results with a significant improvement in the usage of computing resources, which is specially useful for simulate large numbers of high energy showers in remote

ϵ_{th}	None	10^{-8}	10^{-7}	10^{-6}
Comp. time (h)	48.92	7.42	5.69	1.12
Storage (MiB)	895	750	645	431
	Reconstructed showers			
Comp. time (h)	–	7.74	5.81	1.19
Storage (MiB)	–	753	725	731

Table 1: Simulation time and final secondary particles file sizes for different thinning levels (ϵ_{th}) for a shower initiated by proton of $E_0 = 10^8$ GeV.

simulation clusters or GRID-CORSIKA implementations, as the cloud storage and network transfers of the large outputs obtained without thinning could be extremely difficult. The outcomes of the computing times and the storage for the different configurations tested are presented in the Table 1. A significant reduction in the computing times and final file sizes was found.

5. Conclusions and Acknowledgements

In this work we show the first results of a python-based implementation of the so called de-thinning algorithm, which is applied for the reconstruction of thinned CORSIKA simulations of the interaction of primary cosmic rays with the atmosphere. After comparing the results of the application of this algorithm over thinned showers with the corresponding non-thinned equivalent ones, we were able to tune different parameters of the reconstruction method, such as the energy dependent angular aperture of the cone for the injected new particles and specific atmospheric absorption coefficients for each type of secondary particle. When totally implemented, this code will be publicly released under the license GPLv3.

References

[1] P. Billoir, A sampling procedure to regenerate particles in a ground detector from a ?thinned? air shower simulation output, Astroparticle Physics 30 (5) (2008) 270–285.

[2] B. Stokes, R. Cady, D. Ivanov, J. Matthews, G. Thomson, De-thinning extensive air shower simulations, Astroparticle Physics 35 (11) (2012) 759–766.

[3] D. Heck, J. Knapp, J. Capdevielle, G. Schatz, T. Thouw, CORSIKA : A Monte Carlo Code to Simulate Extensive Air Showers, FZKA 6019 (1998) 1–98.

[4] The LAGO Collaboration [H. Asorey], The LAGO Solar Project, in: Proceedings of the 33th International Cosmic Ray Conference ICRC 2013, Vol. in press, Río de Janeiro, Brazil, 2013, pp. 1–4.

[5] H. Asorey, The Latin American Giant Observatory Project, in: Proceedings of the X SILAFAE, Medellín, Colombia, 2014.