On Cosmic-ray Spectral Studies

R. Sina[1] **and E.S. Seo**[1]
[1]*Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA*

Abstract

Careful measurements of cosmic-ray spectra are crucial for constraining many of the viable acceleration models. For experiments having limited statistics at very high energies, data and statistical analysis must be done sufficiently accurately to extract reliable information. In this paper we discuss the simulations and analysis we have performed to study spectral measurements above TeV range energies. Our analysis includes both Poisson fluctuations in the energy bins and finite energy resolution of the detector. We will also discuss de-convolution techniques for obtaining the true spectra from the observed ones.

1 Introduction

The origin and acceleration of cosmic rays is not fully understood. Many testable scenarios have been proposed on where and how the particles are accelerated, how they are propagated, and why they have power law spectra. Ultimately, observations will decide which models are viable and which are precluded. One set of models currently in favor is that the bulk of cosmic-rays are accelerated by supernova shock waves as they interact with the interstellar medium or with stellar winds (Völk & Biermann, 1988; Gaisser, 1990). The maximum energy attained in these models, $E_{\rm max}$, is proportional to the charge of the particles and is predicted to be appreciably larger for blast waves in stellar winds than in interstellar medium (Biermann, 1993; Biermann, Gaisser, & Stanev, 1995). For the latter, $E_{\rm max} \approx 100\ Z\ {\rm TeV}$. These models therefore predict a cutoff or a kink in the spectrum of individual elements in the TeV energy range, with protons having the lowest cutoff energy. A kink around 10^{15} eV in the all-particle spectrum, which is referred to as the knee, has already been observed (Gaisser, 1990). No cutoff in the proton spectrum has directly been observed yet, due to insufficient data at high energies. Because of Poisson fluctuations in the number of incoming protons, observing a kink requires a detector of sufficiently larger energy-reach than the kink energy. Moreover, the data will be smeared out by finite energy resolution of the detector. Low counting rate in high energy bins will also complicate the data analysis, so observing and pinpointing a kink in the spectrum is not a trivial task. There are now several experiments under study that extend their energy-reach up to 1000 TeV. In a previous paper (Sina & Seo, 1999), we demonstrated that the proposed experiments do show promise for observing and localizing reasonable steepening at 100 TeV. The present paper extends the previous work by treating both the kink location and the index difference as free parameters and measuring them each by the Maximum Likelihood fitting method. We have also included possible energy dependence of the resolution in our simulations.

2 Simulations

In order to understand what kink energies and what steepenings can be unambiguously observed, we have made sets of one thousand simulations of a broken power law spectra starting at 100 GeV with differential index of γ =-2.76. Four energy bins per decade have been assigned, and the total number of events in each bin i (for each simulation) is allowed to fluctuate by a Poisson distribution around a mean number f_i. The values of f_i are set by the absolute normalization of the expected flux. In these simulations the flux is chosen such that, in the absence of any kink, 10 events would be detected above 1000 TeV, which is consistent with the planned experiments under study. With the index and normalization we have chosen (Wiebel-Sooth, Biermann & Meyer,1998), this corresponds to a effective time integrated geometry factor of 367 $\rm m^2\ sr\ days$.

In reality the energies of the incoming particles are uncertain due to finite energy resolution. A particle that belongs to energy bin i might be assigned an energy bin j. The number of events observed in bin i, n_i, is therefore different from the true value N_i (whose expectation is f_i). We have allowed for this important consideration by including a correction factor in the simulations. Assuming that the energy resolution is Gaussian, the probability of an event in bin j ending up in bin i is

$$P(j,i) = \frac{\exp \frac{-(E_j - E_i)^2}{2\,(\delta E_j)^2}}{\sum_i \exp \frac{-(E_j - E_i)^2}{2\,(\delta E_j)^2}}. \tag{1}$$

Here δE_j is the energy resolution which is in general energy dependent.

The actual value of N_i is therefore related to n_i by a random number whose mean value is set by the above probabilities. De-convolution techniques are needed to distinguish the true number of events from the observed ones. One technique involves matrix inversions, which are discussed in detail in books on statistical applications to experimental physics (Cowan 1998; Roe 1992). In practice, we can compensate for the difference by using a correction factor $C_i = N_i/n_i$ which we can obtain using a large number of simulations and averaging over all simulated events. Care must be taken for high energy bins where small events are expected. The correction factor for each bin is, in principle, a function of the input spectrum, which is unknown prior to fitting. However, it varies slowly for steep power law spectra. In our simulations all initial guesses in the range -2.5 to -2.9, for the purpose of measuring the correction factors, have resulted in accurately fitting the index below the kink to the -2.76 input index. In short, the exact spectrum is not needed to calculate the correction factors. We have, however, assumed that the exact form of the resolution will be known in advance. We will relax this assumption in a later paper.

Once the values of N_i are estimated in the simulation, then they are fitted to a broken power law and the inferred kink energy and steepening are obtained. It should also be pointed out that one can also fit the observed (convoluted) data to a broken power law. The measured spectrum in such cases would be steeper than the true value.

3 Data Analysis

The cosmic-ray spectrum is measured by observing the number of events in each energy bin. A simplistic method is to fit the data on a log-log scale. In this case the expected power law spectrum is represented in a linear term, and linear regression can be applied to the data to obtain the best fit. Since the log of the number of events is taken, the errors is each bin must be transformed to a log scale. There are several problems with this procedure. First, in transforming between linear and log scales, numerical errors may be introduced in the problem. Second, there are ambiguities in the weight factor for each bin. Fitting by Least Squares is also inappropriate because it assumes large number of events in each bin, $|f_i - N_i| \ll N_i$. This condition is not satisfied in high energy bins for which $f_i \approx N_i \approx 1$.

We shall use the method of maximum likelihood to obtain the best fit.

4 Method of Maximum Likelihood

We define L to be the likelihood function, i.e. the product of the probabilities over each bin of energy for obtaining N_i when the true value is f_i

$$L \equiv \prod p(N_i, f_i) = \prod \frac{{f_i}^{N_i} \exp -f_i}{N_i!}. \tag{2}$$

Instead of maximizing L, it is customary to minimize the natural log function of L.

$$\ln L = \sum\,[N_i\,\ln f_i\,-\,f_i\,-\ln N_i]. \tag{3}$$

Let us suppose there is a kink in the spectrum at $E = E_*$, so that

$$f_i = a\,{E_i}^b \qquad\qquad E \ll E_* \tag{4}$$

$$f_i = a\,{E_i}^b (E_i/E_*)^c \qquad\qquad E \gg E_*, \tag{5}$$

where b and c are both negative numbers. By minimizing $\ln L$, we set the derivatives with respect to values a, b, c, and E_* to zero. However, as it is, the function is not differentiable at E_*. A smooth approximation for the spectrum f_i is given by the exact representation

$$f_i = a\, E_i{}^b\, g_i \tag{6}$$

$$g_i \equiv \lim_{\tau \to \infty} \left(\exp(-\left[\frac{E_i}{E_*}\right]^\tau) + \left[\frac{E_i}{E_*}\right]^c \left[1 - \exp(-\left[\frac{E_i}{E_*}\right]^\tau)\right] \right). \tag{7}$$

The approximation is made by using a finite value of τ for the above representation in the numerical modeling. In practice care must be given in choosing τ to avoid numerical overflows and underflows. A typical value of 5 to 8 should be optimum. If we write $f_i = a\, E_i{}^b\, g_i$, then the ML function is given by

$$\ln L = \sum_i [N_i \,\ln a \,+\, N_i\, b\, \ln E_i \,+\, N_i \,\ln g_i \,-\, a\, E_i{}^b\, g_i]. \tag{8}$$

Therefore

$$\frac{\partial \ln L}{\partial a} = \sum (N_i/a \,-\, E_i{}^b\, g_i) \tag{9}$$

$$\frac{\partial \ln L}{\partial b} = \sum (N_i \,-\, a\, E_i{}^b\, g_i)\, \ln E_i \tag{10}$$

$$\frac{\partial \ln L}{\partial E_*} = \sum [(N_i/g_i \,-\, a\, E_i{}^b) \frac{\partial g_i}{\partial E_*}] \tag{11}$$

$$\frac{\partial \ln L}{\partial c} = \sum [(N_i/g_i \,-\, a\, E_i{}^b) \frac{\partial g_i}{\partial c}]. \tag{12}$$

These are four coupled non-linear algebraic equations which must be solved to obtain the estimators of the four parameters. If the true value of kink energy, in the simulations, is sufficiently small, or if the steepening is sufficiently large, then the data analysis for a majority of one thousand simulations in each set give values of the four parameters consistent with the simulated values. In such case, we would consider the kink to be observable and localizable. If, on the other hand, the obtained values are inconsistent among themselves or with the simulated values, then one can not expect to reliably extract information about the kink from the data.

5 Simulation Results

In a separate paper (Sina & Seo, 1999) we used the Kologoromov–Smirnov test (Lupton, 1993) to check what kink energies and index differences are observable. The present work, based on different scheme, confirms the findings that a kink at 100 TeV and with index difference of 0.3 can be observed. Figures 1 through 3 show the number of fits for the kink energy for the cases where the resolution is constant, $\delta E/E = 0.4$; resolution becomes worse as energy increases, $\delta E/E \propto E^{1/8}$; and the resolution improves with energy, $\delta E/E \propto E^{-1/8}$ respectively. Simulations were made for an input kink at 100 TeV and index changes of 0.3 and 0.5. We have not made any assumptions that the kink location, or indices, are known in advance. We have found that, depending on the form of resolution, close to 90% of events could be fitted to infer the kink energies.

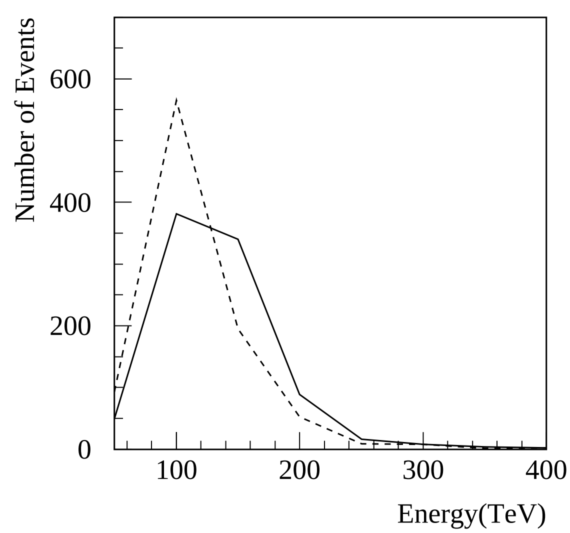

Figure 1: Inferred kink energies for constant resolution of 40%. The solid and dashed lines correspond to index change of 0.3 and 0.5 respectively.

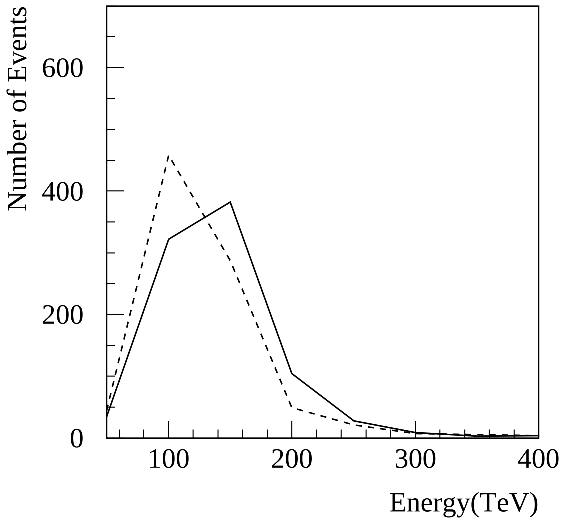

Figure 2: Inferred kink energies for $\delta E/E \propto E^{1/8}$. In this case resolution is 40% at 10 TeV and becomes worse as energy increases. The solid and dashed lines correspond to index change of 0.3 and 0.5 respectively.

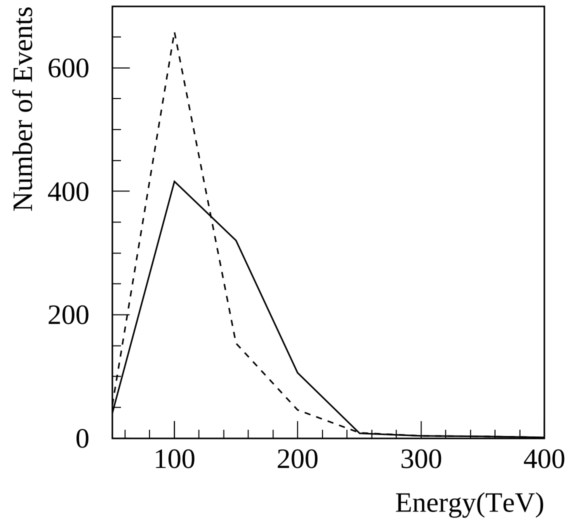

Figure 3: Inferred kink energies when $\delta E/E \propto E^{-1/8}$. In this case the resolution is 40 % at 10 TeV and improves as energy increases. The solid and dashed linescorrespond to index change of 0.3 and 0.5 respectively.

References

Biermann, P.L. 1993, A & A 271, 649

Biermann, P.L., Gaisser T.K., & Stanev T.,1995, Phys. Rev. D. 51, 3450

Cowan, G. 1998, Statistical Data Analysis (Oxford: Oxford University Press)

Gaisser, T.K. 1990, Cosmic Rays and Particle Physics (Cambridge: Cambridge University Press)

Lupton, R. 1993, Statistics in Theory and Practice (Princeton: Princeton University Press)

Roe, B.P. 1992, Probability and Statistics in Experimental Physics (New York: Springer-Verlag)

Völk H.J.& Biermann, P.L., 1988, ApJ Lett. 333, L65

Sina, R. & Seo, E. S. 1999, Adv. Sp. Res. in press

Wiebel-Sooth, B., Biermann, P.L. & Mayer, H., 1998 A &A 330, 389