Chiquita Flux Calculation

Elina Brobeck

November 2, 2006

Original fit to the data (revised derivation)

The most basic fit one can use on the Chiquita data is a straight line in log space. To do this, we assume that the number flux of cosmic rays follows the power law

$$\frac{dN_0}{dE_0} = CE_0^{-\gamma}$$

where $\gamma_0 = 3$ and $\log C \sim 24.5$. Here E_0 is the actual energy of the incoming particle (as opposed to the measured energy E).

For convenience in fitting the energy range measured by Chiquita, this function will be rewritten to be centered on $10^{17.8}$ eV.

$$\frac{dN_0}{dE_0} = C_{17.8}\left(\frac{E_0}{10^{17.8}}\right)^{-\gamma}.$$

The relationship between C and $C_{17.8}$ is $C = C_{17.8} \times 10^{17.8\gamma}$.

We are measuring the number flux per interval in observed energy, $\frac{dN}{dE}$. This is the convolution of the actual number flux with the function $R(E, E_0)$, which describes the distribution of reconstructed energies at each value of E_0.

$$\frac{dN}{dE} = \int dE_0 A(E_0) R(E, E_0) \frac{dN_0}{dE_0}.$$

The function $A(E_0)$ is the acceptance of the array (the fraction of showers which pass all cuts on the data at a given energy).

The histogram of Chiquita data is binned in log space, with bin width $\log(E_{\text{upper}}) - \log(E_{\text{lower}}) = 0.2$. In terms of the central energy E_C of the bin, the bin width is

$$E_{\text{upper}} - E_{\text{lower}} = 10^{0.1}E_C - 10^{-0.1}E_C = 0.46E_C.$$

Each data point in the histogram therefore contains a number of showers given by

$$N = \int_{10^{-0.1}E_C}^{10^{0.1}E_C} \frac{dN}{dE} dE.$$

To obtain the number of Chiquita showers in a given bin, we substitute the expression for $\frac{dN}{dE}$ above:

$$N(E_i) = \int_{\mathrm{E_i\ bin}} dE \int_{\mathrm{all\ E_0}} dE_0 A(E_0) R(E, E_0) \frac{dN_0}{dE_0}.$$

We have the function $R(E, E_0)$ in the form of a data table for a fixed set of values of E_0, so instead of a continuous integral, we need a sum. We make the approximation that the entire flux in the bin centered around E_0 is being transferred to the bin centered around the observed energy E with the efficiency $R(E, E_0)$.[1]

$$N(E_i) = \sum_{E_{0j}} A(E_{0j}) R(E_i, E_{0j}) \int_{\mathrm{E_{0j}\ bin}} dE_0 \frac{dN_0}{dE_0}$$

The integral of the flux over the E_0 bin is given by

$$\begin{aligned}
\int_{\mathrm{E_0\ bin}} C E_0^{-\gamma} dE_0 &= \int_{10^{-0.1E_0}}^{10^{0.1E_0}} C E_0^{-\gamma} dE_0 \\
&= \left. \frac{C E_0^{-\gamma+1}}{-\gamma+1} \right|_{10^{-0.1E_0}}^{10^{0.1E_0}} \\
&= \frac{C}{-\gamma+1} \left[\left(10^{0.1}\right)^{-\gamma+1} - \left(10^{-0.1}\right)^{-\gamma+1} \right] E_0^{-\gamma+1}
\end{aligned}$$

For $\gamma = 3$,

$$\int_{\mathrm{E_0\ bin}} C_0 E_0^{-\gamma} dE_0 = 0.477 C_0 E_0^{-\gamma+1}.$$

The function which should be used to fit the data[2] is thus

$$\begin{aligned}
N(E_i) &= \sum_{E_0} 0.477 A(E_0) R(E_C, E_0) C E_0^{-\gamma+1} \\
&= \sum_{E_0} 0.477 A(E_0) R(E_C, E_0) C'_{17.8} \left(\frac{E_0}{10^{17.8}} \right)^{-\gamma'}
\end{aligned}$$

where $\gamma' = \gamma - 1$.

We want to do the fit to C and γ in log space, so this is rewritten as

$$N_{\mathrm{data}} = \sum_{E_0} 0.477 A(E_0) R(E_C, E_0) 10^{\log C'_{17.8} - \gamma'(\log E_0 - 17.8)}.$$

The factor $\log C_{17.8}$ has been replaced by $\log C'_{17.8}$, where

$$\log C_0 = \log C_{17.8} + 17.8\gamma = \log C'_{17.8} + 17.8\gamma'.$$

What we want to plot is

$$F \times E^3 = \frac{dN}{dE} E^3 = C E^{-\gamma} E^3.$$

This is calculated as

$$F \times E^3 = \frac{dN}{dE} E^3 = C E^{-\gamma} E^3 = C E^{-(\gamma-1)} E^{-1} E^3 = C'_{17.8} \left(\frac{E}{10^{17.8}} \right)^{-\gamma'} E^2.$$

[1] The data table of values for $R(E, E_0)$ was calculated using simulated showers with an E^{-3} distribution over the width of the bin. It is not the efficiency for moving showers from the central energy of one bin to the central energy of another bin, but rather an approximation to the fraction of showers in one bin which get moved to another bin due to error in the reconstructed energies. This is why substituting $R(E_i, E_{0j})$ for $R(E, E_0)$ takes care of the integral over the E_i bin.

[2] In the original derivation, I had $\left[10^{0.1} - 10^{-0.1}\right]$ where I should have had $\frac{1}{-\gamma+1}\left[\left(10^{0.1}\right)^{-\gamma+1} - \left(10^{-0.1}\right)^{-\gamma+1}\right]$. This gave a constant factor of 0.465, which just happens to be very close to the value of the second expression when $\gamma = 3$. Ideally the entire expression should be included when fitting for γ.

Iterative fit method

In this method, we start by defining the observed flux $J(E) = \frac{dN}{dE}$ as before:

$$J(E) = \frac{dN}{dE} = \int dE_0 A(E_0) R(E, E_0) \frac{dN}{dE_0} = \int dE_0 A(E_0) R(E, E_0) J_0(E_0).$$

The number of showers observed in a given bin centered on energy E was approximated by[3]

$$N_{\rm data}(E) = 0.46E\, A(E)\, J_0(E)$$

where $A(E) = \int A(E_0) R(E, E_0) dE_0$, and $0.46E$ is the width of the bin centered on E. $A(E)$ is calculated by taking the number of showers which reconstruct in bin E over the number of simulated showers thrown in bin E. (This will not be the correct ratio if the showers were not thrown with an $E^{-\gamma_0}$ spectrum.)

The flux $J(E)$ is assumed to have the form $J(E) = CE^{-\gamma}$. We can define a function $C_{\rm data}(E) = C_2 E^{3-\gamma_2}$ such that

$$J(E) = C_{\rm data}(E) E^{-3} = C_2 E^{-\gamma_2}.$$

This gives

$$N_{\rm data}(E) = 0.46E\, A(E)\, C_2 E^{-\gamma_2} = 0.46\, A(E)\, C_{\rm data}(E)\, E^{-2}.$$

We can therefore plot

$$C_{\rm data}(E) = \frac{N_{\rm data} E^2}{0.46 A(E)}$$

and use those points to find best fit values for C_2 and γ_2.

This is now repeated assuming an input spectrum of $J_2(E) = C_2 E^{-\gamma_2}$. A data set is generated as

$$N_{\rm data}(E) = 0.46E\, A(E)\, C_2 E^{-\gamma_2}.$$

In the code, this is actually calculated somewhat more carefully as

$$N_{\rm data}(E) = \sum_{E_0} 0.46\, A(E_0) R(E, E_0)\, C_2 E_0^{-\gamma_2+1}.$$

A new function, $C'_{\rm data}(E)$ is defined by the points

$$C'_{\rm data}(E) = \frac{N_2 E^2}{0.46 A(E)}.$$

The flux is now estimated as

$$J(E) = \frac{C'_{\rm data}(E)}{C_{\rm data}(E)} C_2 E^{-\gamma_2}.$$

[3] This expression is not as good an approximation as the one given in the previous section. It would be better to keep the sum over E_0, and include $R(E, E_0)$ explicitly. It also has the wrong numerical factor in front, which should be either 0.477 (for $\gamma = 3$) or left as a function of γ.