Astroparticle Physics 102 (2018) 89–94

Contents lists available at ScienceDirect

Astroparticle Physics

journal homepage: www.elsevier.com/locate/astropartphys

Computing mean logarithmic mass from muon counts in air shower experiments

H.P. Dembinski

Bartol Institute, University of Delaware & Max-Planck-Institute for Nuclear Physics, USA

ARTICLE INFO

Article history:
Received 24 November 2017
Revised 11 April 2018
Accepted 22 May 2018
Available online 23 May 2018

Keywords:
Air showers
Muons
Cosmic rays
Statistics
Hadronic interactions
Fluctuations

ABSTRACT

I discuss the conversion of muon counts in air showers, which are observable by experiments, into mean logarithmic mass, an important variable to express the mass composition of cosmic rays. Stochastic fluctuations in the shower development and statistical fluctuations from muon sampling can subtly bias the conversion. A central theme is that the mean of the logarithm of the muon number is not identical to the logarithm of the mean. It is discussed how that affects the conversion in practice. Simple analytical formulas to quantify and correct such biases are presented, which are applicable to any kind of experiment.

1. Introduction

The mean logarithmic mass $\langle \ln A \rangle$ is a common variable to summarize the mass composition of cosmic rays. Most ground-based experiments infer the mass by counting muons in cosmic-ray induced air showers [1]. This paper discusses the conversion of muon number to mean logarithmic mass from the point of view of the data analyst, with a focus on the effect of stochastic fluctuations in the shower development [2,3] and the detector response on the conversion. The fluctuations can bias estimates of $\langle \ln A \rangle$ in several ways. Biases here are defined in the usual statistical sense; if $\hat{x}$ is an estimate of the true value x that fluctuates according to a probability density $f(\hat{x})$, then the bias is the expectation $\mathrm{E}[\hat{x} - x] = \int (\hat{x} - x) f(\hat{x}) \mathrm{d}\hat{x}$. We generally want $\hat{x}$ to have zero bias, so that the sample average $\langle \hat{x} \rangle$ converges to x for large samples.

The results in this paper are not specific to a particular type of experiment. It is assumed throughout this paper that an experiment provides an unbiased estimate $\hat{N}_\mu$ of the total number of muons N_μ produced in an air shower and an estimate $\hat{E}$ of the shower energy E. This is far from trivial and much of the difficulty in running an experiment deals with this. The total number of muons N_μ produced in an air shower cannot be directly measured, because experiments can only count muons that reach the ground, while some decay on the way. The experimental distinction between muons and other charged particles at the ground is not easy either [4–7]. But in principle, $\hat{N}_\mu$ can be inferred for a given geometry and shower energy from the measurement by applying an average correction obtained from air shower simulations. Highly-inclined air showers recorded by Haverah Park and the Pierre Auger Observatory have been analyzed in this way [8–11]. Similarly, an estimate $\hat{E}$ of the shower energy can be inferred from the number of electrons and photons that reach the ground, or by recording the longitudinal shower profile with telescopes.

The paper deals with the comparably easier part of the conversion of the unbiased estimates $\hat{E}, \hat{N}_\mu$ to $\langle \ln A \rangle$. Fluctuations occur in the shower development and arise from the sampling of an air shower by a detector. It is important to distinguish between these two kinds of fluctuations, because they are approximately independent [12]. Both randomly shift the estimates $\hat{E}, \hat{N}_\mu$ away from their true values E, N_μ, and these random shifts cause some subtle biases in the conversion to $\langle \ln A \rangle$. We quantify these biases. Knowing their sizes allows one to safely neglect them if they are small, and to correct them otherwise.

2. From muon number to mass

It is instructive to introduce fluctuations step-by-step. We start by ignoring fluctuations from detector sampling, only stochastic fluctuations in the shower development are considered. The true muon number N_μ and the shower energy E shall be exactly known and the energy E shall be same for all showers. Stochastic fluctuations in the hadronic interactions are still causing the muon number N_μ to vary randomly.

E-mail address: hdembins@mpi-hd.mpg.de

https://doi.org/10.1016/j.astropartphys.2018.05.008

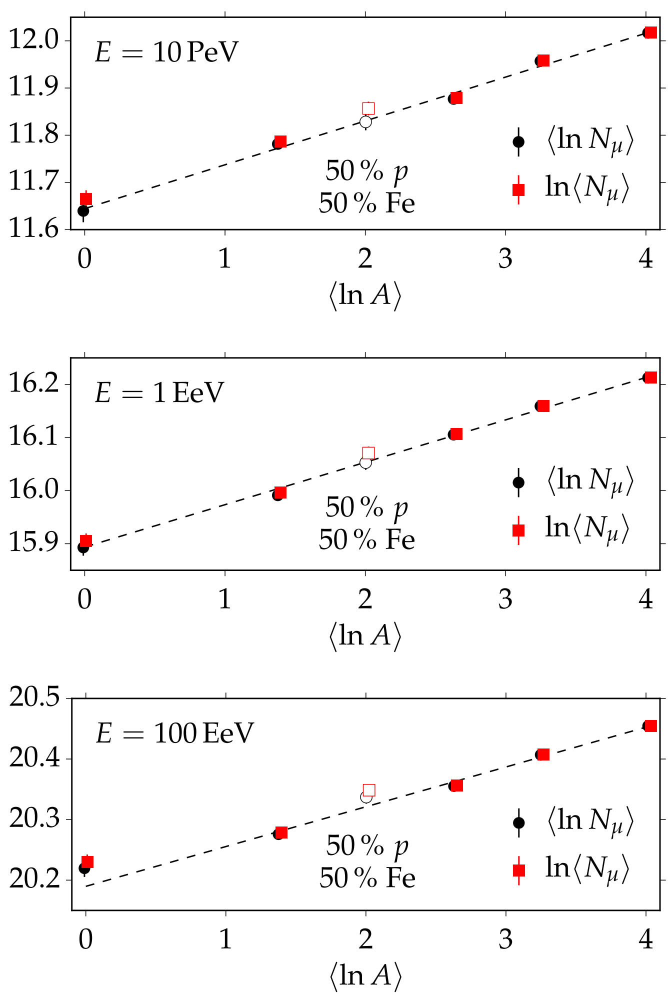

Fig. 1. Average logarithm of the number of muons $\langle \ln N_\mu \rangle$ (circles) and logarithm of the average number of muons $\ln \langle N_\mu \rangle$ (squares) in simulated vertical air showers produced by primary particles with A nucleons. A fitted straight line (dashed) is shown for comparison. Solid markers stand for averages computed over showers from a single primary, open markers stand for an equal mix of proton and iron showers. Error bars indicate the statistical uncertainty of the finite sample (2200 showers per primary and energy).

The first point to make is that $\langle \ln A \rangle$ is best computed from the mean logarithmic muon number $\langle \ln N_\mu \rangle$, and not the mean of the muon number $\langle N_\mu \rangle$. In either case, the average here is computed over many air showers with the same shower energy E.

The following argument is similar to the one developed by the Pierre Auger collaboration for the depth of shower maximum [13]. The relationship between N_μ and A can be understood within the Matthews–Heitler model of a hadronic shower [14]. The analytical model treats air showers in a simplified way, but describes surprisingly many features of air showers correctly. According to the model, the total muon number N_μ for a cosmic ray with A nucleons scales with a power of the number of nucleons

$$N_\mu = A^{1-\beta}\, N_\mu^p, \tag{1}$$

where N_μ^p is the number of muons in a proton-induced air shower, and $\beta \simeq 0.9$ is a constant.

This behavior is well reproduced in full air shower simulations. In the Matthews-Heitler model, stochastic fluctuations in the shower development are neglected. To show that Eq. (1) holds for the real showers, several sets of vertical showers with identical primary particles were simulated with CORSIKA [15] compiled with the CONEX option, using the hadronic interaction models SIBYLL-2.3 [16] and GHEISHA [17]. The showers were simulated in a US standard atmosphere until a slant depth of 1050 g cm^{-2}. The number of muons N_μ in each shower were taken from the maximum of the longitudinal muon profile. Proton, helium, nitrogen, silicon, and iron primaries were simulated. For each primary, the averages $\ln \langle N_\mu \rangle$ and $\langle \ln N_\mu \rangle$ were computed. The two are subtly different, because the expectation is noncommutative with a non-linear mapping $f(x)$, $\mathrm{E}[f(x)] \neq f(\mathrm{E}[x])$. The dependence on A is shown in Fig. 1 for a wide range of primary energies. If A is a constant, both $\ln \langle N_\mu \rangle$ and $\langle \ln N_\mu \rangle$ scale with $\ln A$ as predicted by Eq. (1). This result is independent of the hadronic interaction models and shower inclination.

To use Eq. (1) to get an estimate of $\langle \ln A \rangle$ for real air showers, we consider the realistic case where the mass A is another stochastic variable that changes from shower to shower. For a single primary, the simulations show that $\langle N_\mu \rangle = A^{1-\beta}\, \langle N_\mu^p \rangle$. If f_A is the fraction with which a primary with A nucleons occurs, a superposition of primaries yields

$$\sum_A f_A \langle N_\mu \rangle = \sum_A f_A A^{1-\beta}\, \langle N_\mu^p \rangle = \langle N_\mu^p \rangle \sum_A f_A A^{1-\beta}$$
$$\Leftrightarrow \langle N_\mu \rangle = \langle N_\mu^p \rangle\, \langle A^{1-\beta} \rangle. \tag{2}$$

Unfortunately, we cannot convert $\langle A^{1-\beta} \rangle$ to $\langle A \rangle$ or $\langle \ln A \rangle$, because these are non-linear functions of A. The solution is to start from $\langle \ln N_\mu \rangle = (1-\beta) \ln A + \langle \ln N_\mu^p \rangle$ for a single primary, which is also supported by the simulations. Then the result of the superposition is

$$\langle \ln N_\mu \rangle = (1-\beta) \langle \ln A \rangle + \langle \ln N_\mu^p \rangle, \tag{3}$$

where we used that $\langle ax + by \rangle = a\langle x \rangle + b\langle y \rangle$ for constants a, b and stochastic variables x, y.

Both β and $\langle \ln N_\mu^p \rangle$ can be obtained from air shower simulations. If $\langle \ln N_\mu^{\mathrm{Fe}} \rangle$ is available, it can be used to substitute β. The two related formulas for $\langle \ln A \rangle$ are

$$\langle \ln A \rangle = \frac{\langle \ln N_\mu \rangle - \langle \ln N_\mu^p \rangle}{1-\beta} \tag{4}$$

$$\langle \ln A \rangle = \frac{\langle \ln N_\mu \rangle - \langle \ln N_\mu^p \rangle}{\langle \ln N_\mu^{\mathrm{Fe}} \rangle - \langle \ln N_\mu^p \rangle} \ln 56. \tag{5}$$

This approach is very elegant, because the equations are true whatever the probability distributions are for A, N_μ, N_μ^p, and N_μ^{Fe}.

As previously stated, the mean of the logarithm is not the same as the logarithm of the mean, $\ln \langle N_\mu \rangle$ is always higher than $\langle \ln N_\mu \rangle$. Still, the two are quite close and the bias of substituting one for the other may be negligible in some situations. To judge when this is safe, a simple formula to compute the bias is given in Section 3. Some analyses [18] do not produce an estimate of the muon number event-by-event, only the average $\langle N_\mu \rangle$ over many showers. In these cases, the formula can be used to correct the difference $(\langle \ln N_\mu \rangle - \ln \langle N_\mu \rangle)$.

So far fluctuations introduced by detector sampling were neglected, but N_μ is not known in practice, only an estimate $\hat{N}_\mu$ which fluctuates around N_μ. Some muons decay on the way to the ground, the detector does not count all muons that arrive, and so on. It is assumed that these losses are corrected on average, but they introduces additional fluctuations. Since the mean of the logarithm is not the logarithm of the mean, we find $\langle \ln \hat{N}_\mu \rangle \neq \langle \ln N_\mu \rangle$ even if $\hat{N}_\mu$ is an unbiased estimate of N_μ. How to correct for this effect is discussed in Section 4.

Finally, one has to consider that the average $\langle \ln \hat{N}_\mu \rangle$ is not computed over showers with the same energy E in practice, but for showers that fall into the same energy bin. The energy E is also not known exactly, only an estimate $\hat{E}$ of it. The quantitative impact of that is calculated in Section 5.