Nuclei Identification by Multiple Energy Losses in Detectors of the PAMELA Spectrometer.

Alekseev V.[1,2], Golub O.[2], Epifanov A.[2],
Lukyanov A.[1], Mayorov A.[1,2]

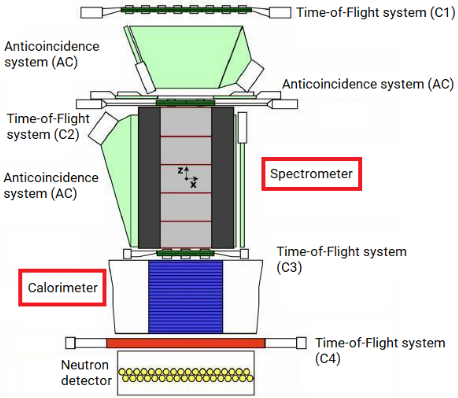
[1]Yaroslavl State University
[2]National Research Nuclear University MEPhI

Moscow, Nucleus – 2022
July 15, 2022

Problem Statement

Goal: to identify kind of a particle by a vector of energy loss values.

- Magnetic spectrometer (tracker): $6 \times 2 = 12$ values;
- Calorimeter: 4 first values (Y projection).



Event selection

Training and test samples were generated with Geant4 software.

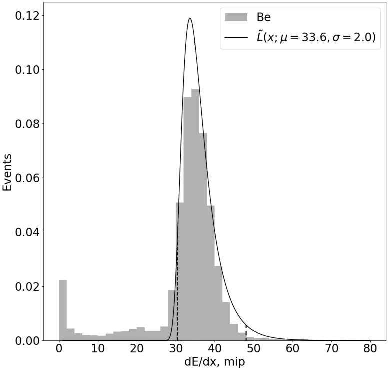
Selection criteria:

- No signals in anticoincidence systems.
- No more than one paddle triggered in each plane of the time-of-flight system.
- Particle track reconstructed by at least 3 points in the deflecting projection X, and by 2 points in projection Y.

Classification algorithm

First step

For each particle, for small rigidity ranges, find 4% and 95% quantiles (q_0 and q_1) of the enegry loss distribution (from empirical distribution).



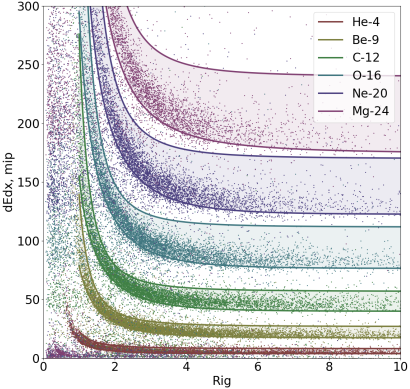
Classification algorithm

The second step

For each particle, fit q_i/Rigidity dependences by smooth curve of special form:

$$q_i(x) = \big(\exp(ax^b + c) + d\big)^2 \text{ (for tracker);}$$

$$q_i(x) = \big(\exp(ax^b + c) + dx^e + f\big)^2 \text{ (for calorimeter)}$$



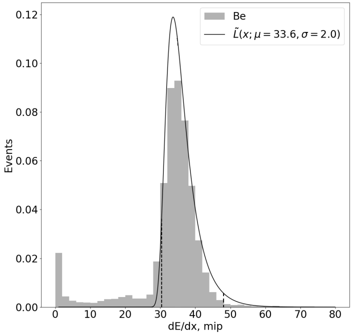
Classification Algorithm

The third step

We assume that dE/dx distribution for fixed particle and rigidity takes a form:

$$L(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma} + e^{-\frac{x-\mu}{\sigma}}\right)\right).$$

For given rigidity x we can compute quantiles $q_0(x)$ and $q_1(x)$, than reconstruct μ and σ parameters by two quantiles.



Likelihood function for a nuclei type i is a function:

$$MLE_i(x; R) = \sum_{j=1}^{N} \ln L(x_j, \mu_{ij}(R), \sigma_{ij}(R)),$$

where $j = 1, \ldots, N$ are detector numbers, $x = (x_1, \ldots x_N)$ are energy losses in correspondent detectors, R is a magnetic rigidity of the particle.

Result of classification algorithm is a nucleus $\hat{n}$, which minimizes $MLE_i(x)$.

$$\hat{n} = \arg\min_i MLE_i(x; R).$$

Error measurement

Confusion matrices

Confusion matrix M consists of a priori classification errors.

$$M_{ij} = P(\text{answer} = j \mid \text{real} = i).$$

Error measurement

Confusion matrices

Confusion matrix M consists of a priori classification errors.

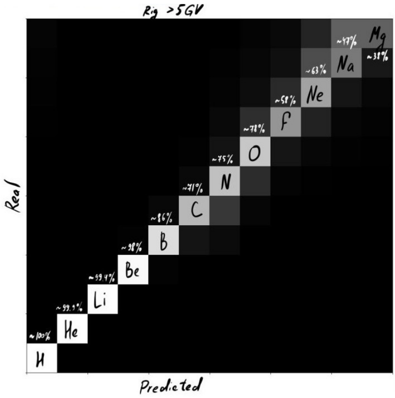
$$M_{ij} = P(\text{answer} = j \mid \text{real} = i).$$

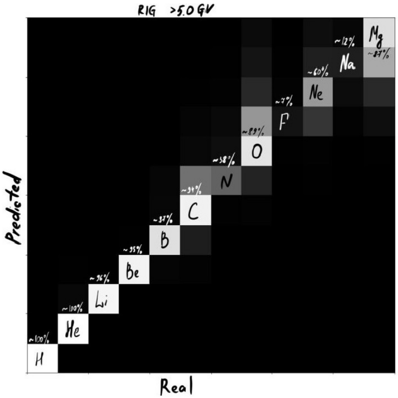
Now assume we know a nuclei distribution in cosmic rays. Confusion matrix Q consists of a posteriori classification errors.

$$Q_{ij} = P(\text{real} = j \mid \text{answer} = i) = \frac{P(\text{real} = j) M_{ij}}{\sum_k P(\text{real} = k) M_{ik}}.$$

Results

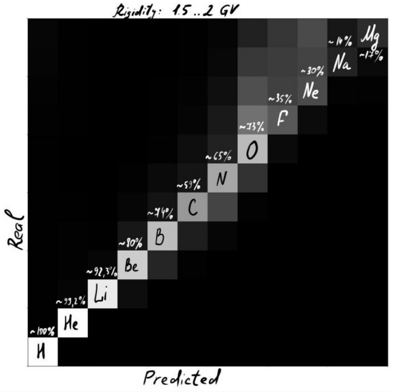
$R > 5GV$

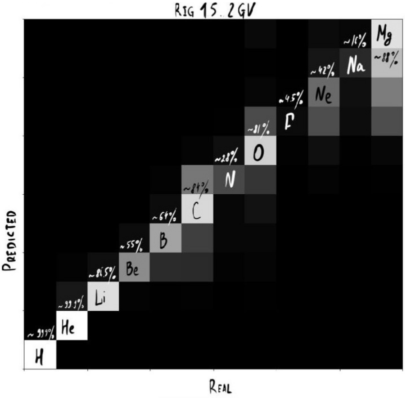




Results

$R = 1.5 \ldots 2.0 GV$





Left: our algorithm.
Right: gradient boosting method, $R > 1$ (power spectrum), with calorimeter data.