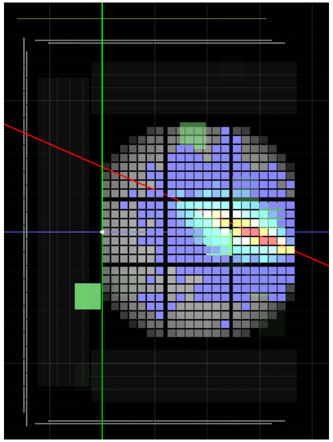
Electron-proton discrimination with the HERD calorimeter

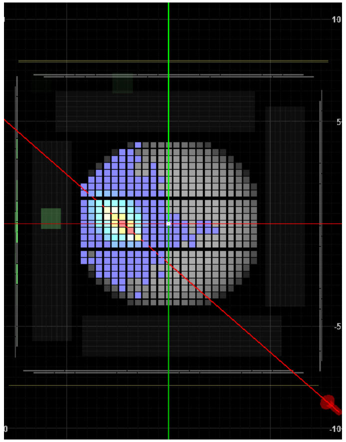
Pietro Betti
CALO meeting
21/02/2024

Electron-Proton discrimination

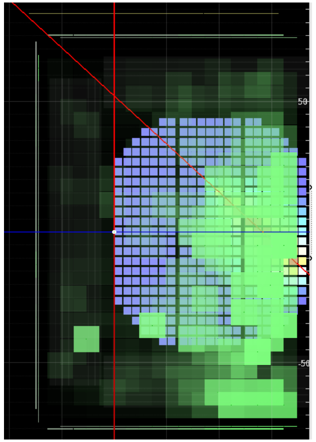
Discriminate protons from electrons is the main task in electron flux analysis
At 1 TeV about 2000 protons for every electron, at higher energy even worse

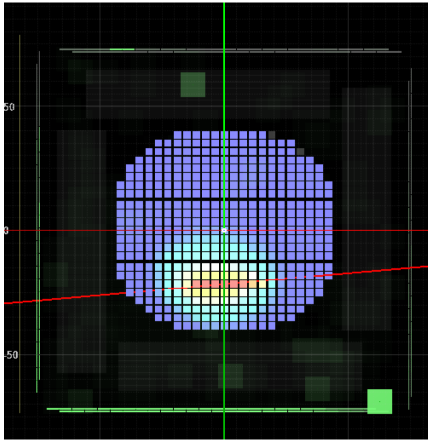
Discriminate the two particles through the different shape of electromagnetic and hadronic showers


Lateral view


Top view

Electron induced shower

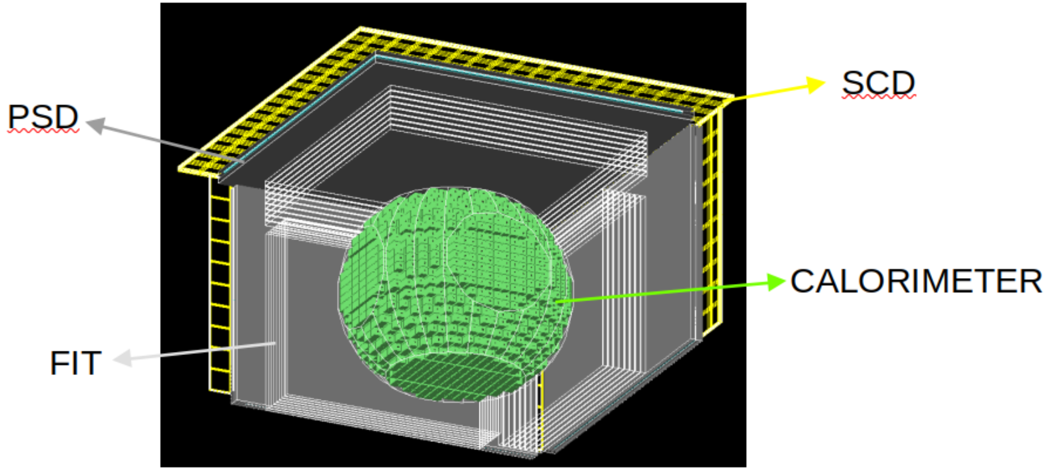

Lateral view


Top view

Proton induced shower

Work environment

- Since HERD will start to acquire data in 2027, we can now work with simulations only.
- The detector geometry for simulations is developed in Geant 4
- The calorimeter is built with LYSO crystals and no mechanical structure (not already designed at the moment of the start of Monte Carlo production). A surface that contains the calorimeter is defined and called envelope. The empty spaces in the envelope are filled with carbon fiber, in order to simulate the mechanical structure that will be mainly constituted by carbon fiber trays.



MonteCarlo production

Energy range simulated electrons: 100 GeV – 20 TeV
Energy range simulated protons: 100 GeV – 1 PeV

Particles statistics used in the following analysis:

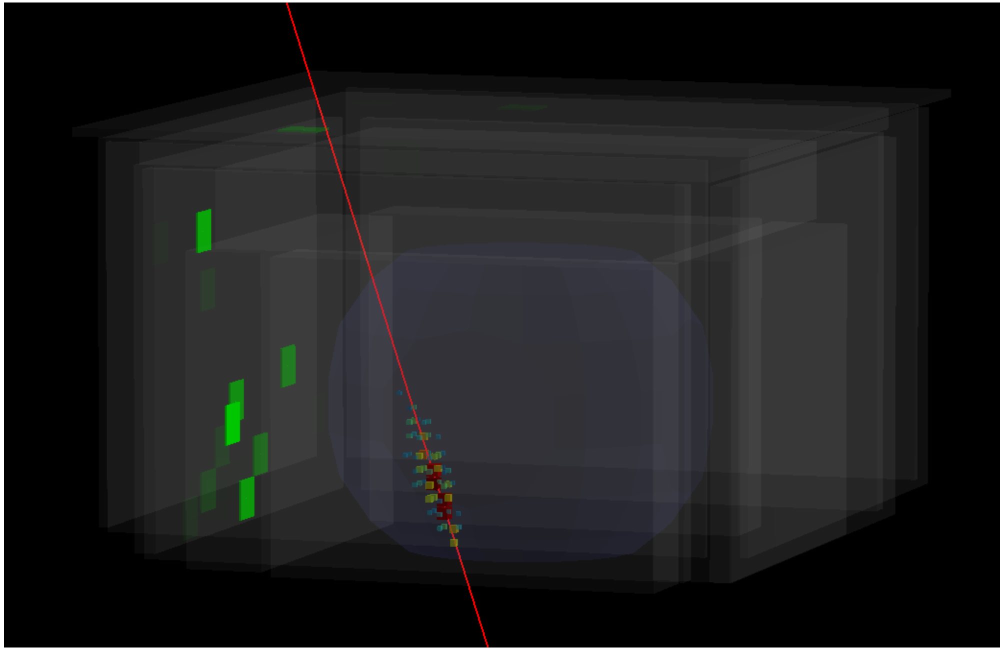
- e- (100 GeV – 1 TeV) ~2.5 M events
- e- (1 TeV – 20 TeV) ~ 3.2 M events
- p (100 GeV – 1 TeV) ~ 4 M events
- p (1 TeV – 10 TeV) ~ 4 M events
- p (10 TeV – 100 TeV) ~ 3.9 M events
- p (100 TeV – 1 PeV) ~ 2.6 M events

Production of more than 1.5 years

Simulations running on Florence INFN farm, CNAF and RECAS
Enormous machine time: 312h to simulate 100 events in the energy bin (100TeV – 1PeV) !
Enormous disk space: over 380 Tb of simulated data!

Energy spectrum for production: ~E^{-1} → equal statistics in equal log-x scale for same statistic in simulated energy bin
(not realistic cosmic rays spectrum → need to reweigh the events)

- Generation surface of particles: spherical surface centered in the calorimeter center, with a radius of 3 m
- Isotropic generation (cosmic rays flux is isotropic with an high degree of approximation)



Preliminary binning definition for analysis

border 0: 100 GeV
border 1: 125.893 GeV
border 2: 158.489 GeV
border 3: 199.526 GeV
border 4: 251.189 GeV
border 5: 316.228 GeV
border 6: 398.107 GeV
border 7: 501.187 GeV
border 8: 630.957 GeV
border 9: 794.328 GeV
border 10: 1000 GeV
border 11: 1258.93 GeV
border 12: 1584.89 GeV
border 13: 1995.26 GeV
border 14: 2511.89 GeV
border 15: 3162.28 GeV
border 16: 3981.07 GeV
border 17: 5011.87 GeV
border 18: 6309.57 GeV
border 19: 7943.28 GeV
border 20: 10000 GeV
border 21: 12589.3 GeV
border 22: 15848.9 GeV
border 23: 19952.6 GeV

Energy decade divided in ten energy bins

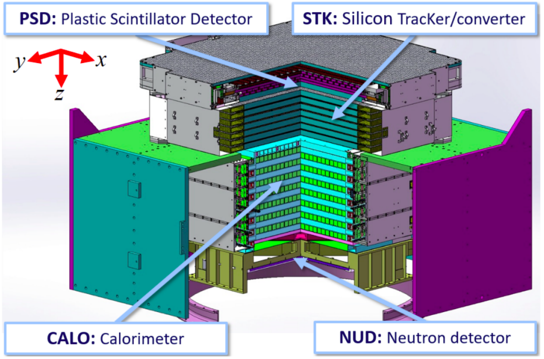
Preliminary binning, to be tuned in the following of the analysis work on the base of resolution and statistics after reweigh of the events

Starting point

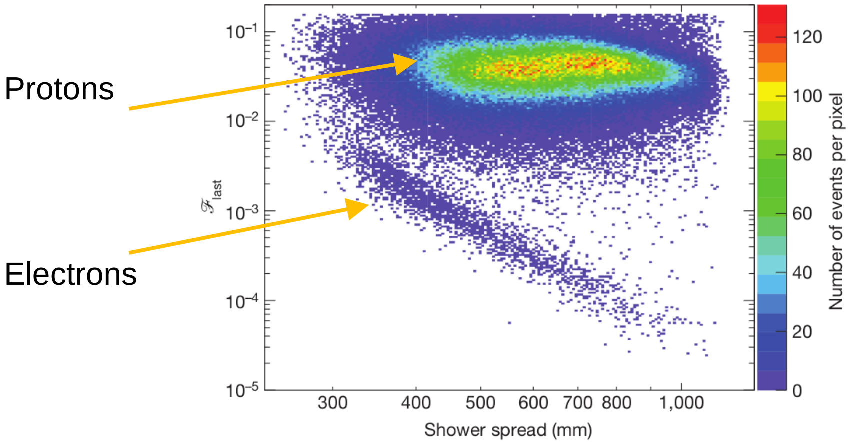
Two powerful rejection variables used by both CALET and DAMPE:

- fraction of energy released in the last layer of the calorimeter respect to the total energy released
- Lateral spread of the shower

DAMPE detector



DAMPE two variables rejection analysis

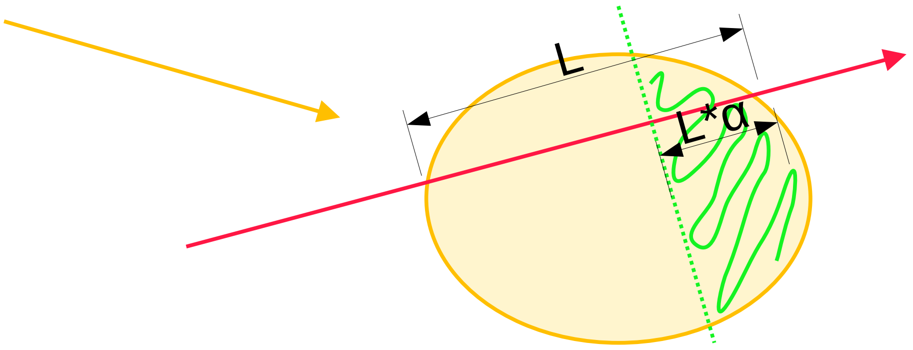


Try to introduce the same type of variables for HERD

Fraction of energy in the last part of the shower

HERD not built by layers
Events entering the detector from five faces are accepted

- L_α : fraction of energy released in the CALO volume that contains the last α% fraction of the track, respect to total energy released in the CALO
- $L_\alpha X_0$: same as L_α but with the fraction of track calculated in X_0
- $\text{depth}_\alpha X_0$: fraction of energy released in a step of 3 X_0 of track at α fraction of the track, respect to total energy released
- over $_\alpha X_0$: same as $\text{depth}_\alpha X_0$, but from that step up to the end of the track

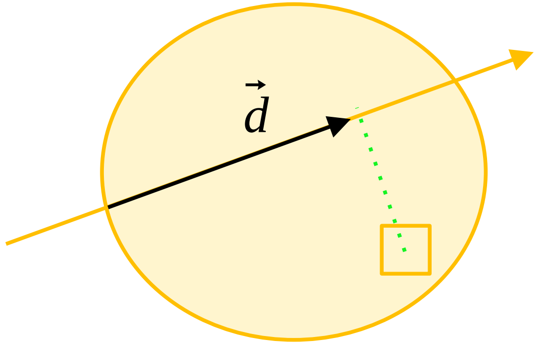


CALO envelope
(not realistic)

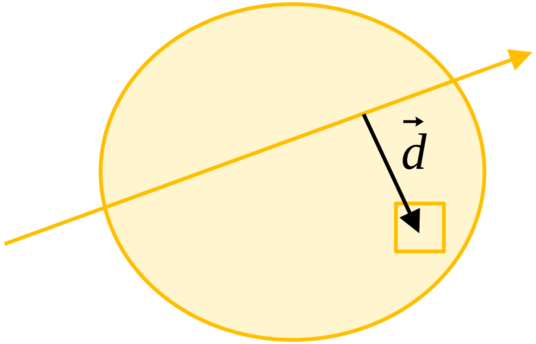
Shower momenta

$$M_n = \left| \sum_{i=crystals} E_i \cdot (\vec{d})^n \right|$$

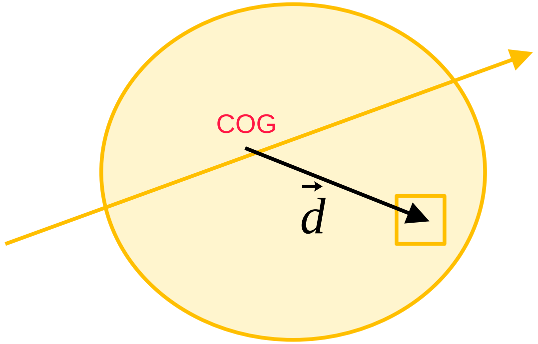
Longitudinal



Lateral



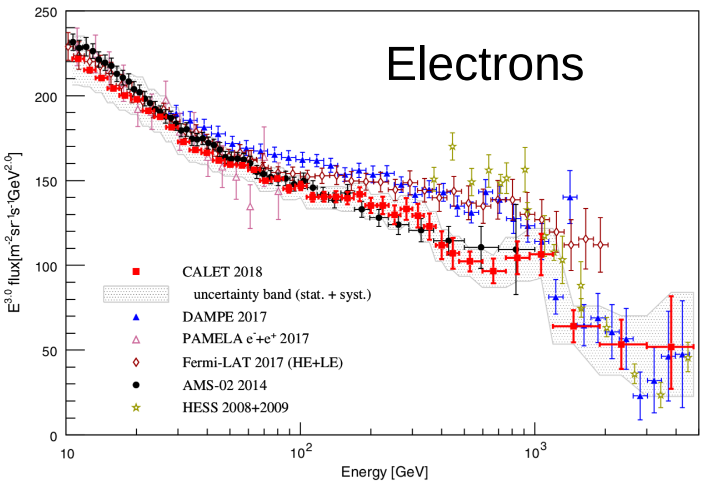
Vector

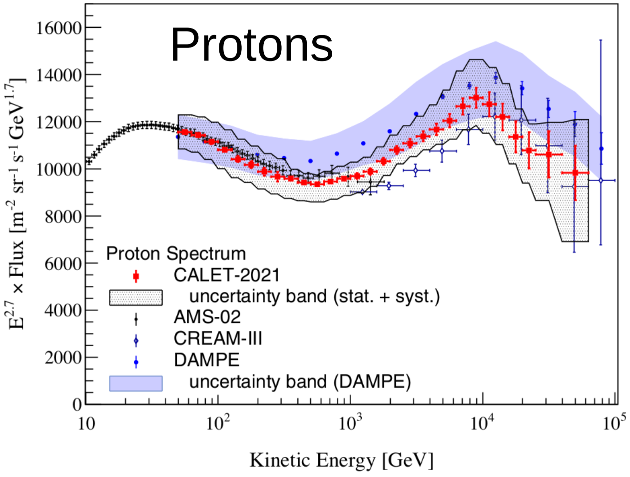


Reweighing of the events

Simulated spectrum: E^{-1}
Electron+positron spectrum: $\sim E^{-3}$
Proton spectrum: $\sim E^{-2.7}$

We need to reweigh the events for the real spectrum: CALET measurements as reference spectrum





Fit of CALET data as reference

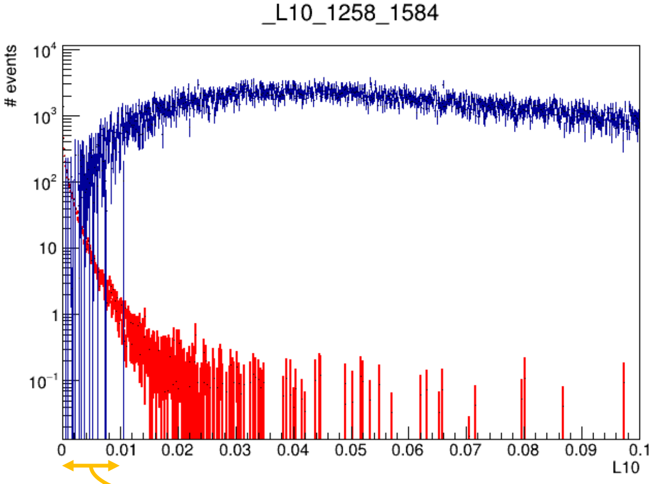
$$\Phi = \Phi_0 (E/100\ \mathrm{GeV})^{-\gamma_1} [1 + (E/E_b)^{-(\gamma_1 - \gamma_2)/\Delta}]^{-\Delta}$$

$$\phi(E) = \left[1 + \left(\frac{E}{E_0}\right)^s\right]^{\frac{\Delta\gamma}{s}} \times \left[1 + \left(\frac{E}{E_1}\right)^{s_1}\right]^{\frac{\Delta\gamma_1}{s_1}}$$

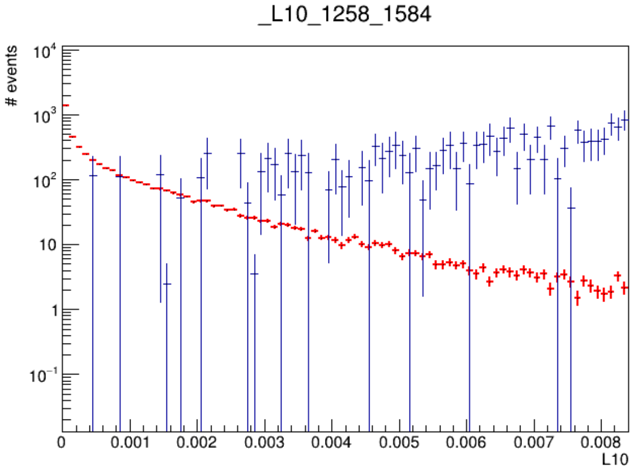
First results on electron proton discrimination

The best two rejection variables found up to now are:

- L_{10} : fraction of energy released in the CALO volume that contains the last 10% fraction of the track, respect to total energy released in the CALO
- Lateral 4th momentum

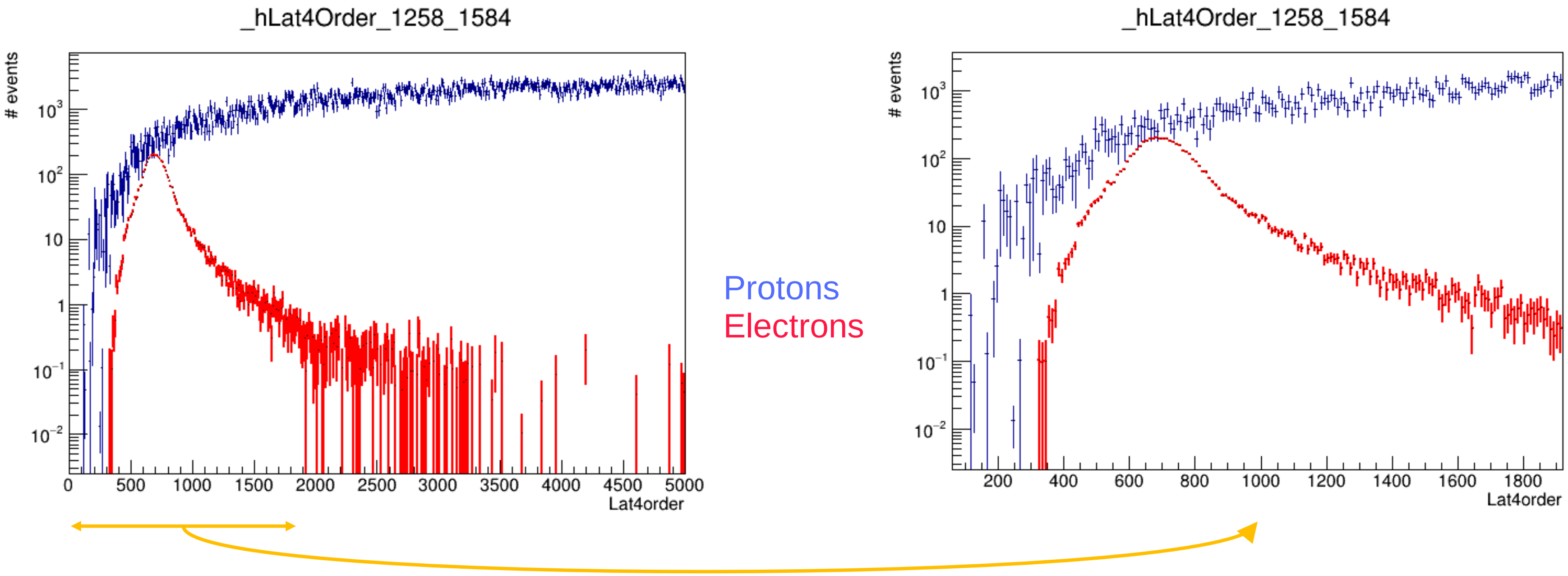


Protons
Electrons

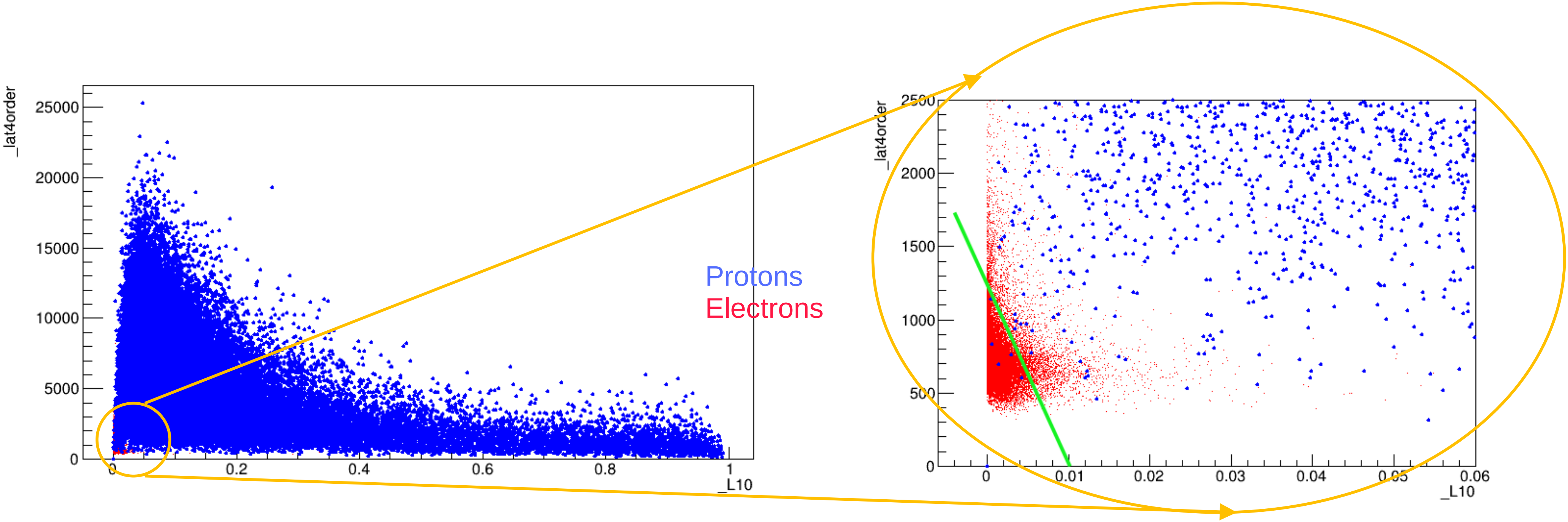


L_{10} has a rejection power by itself

Lateral 4th momentum has no rejection power by itself: in the momentum range of electrons there are more protons than electrons

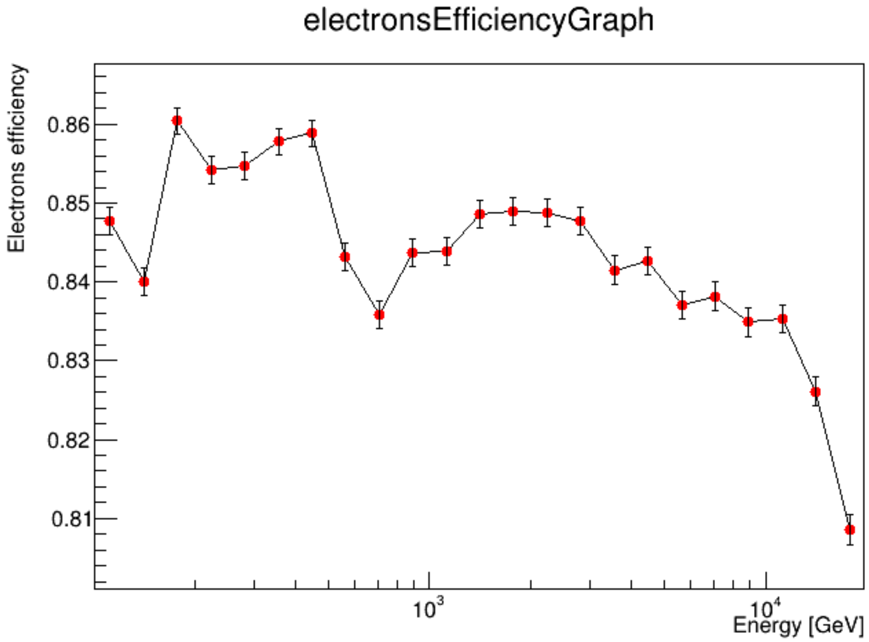


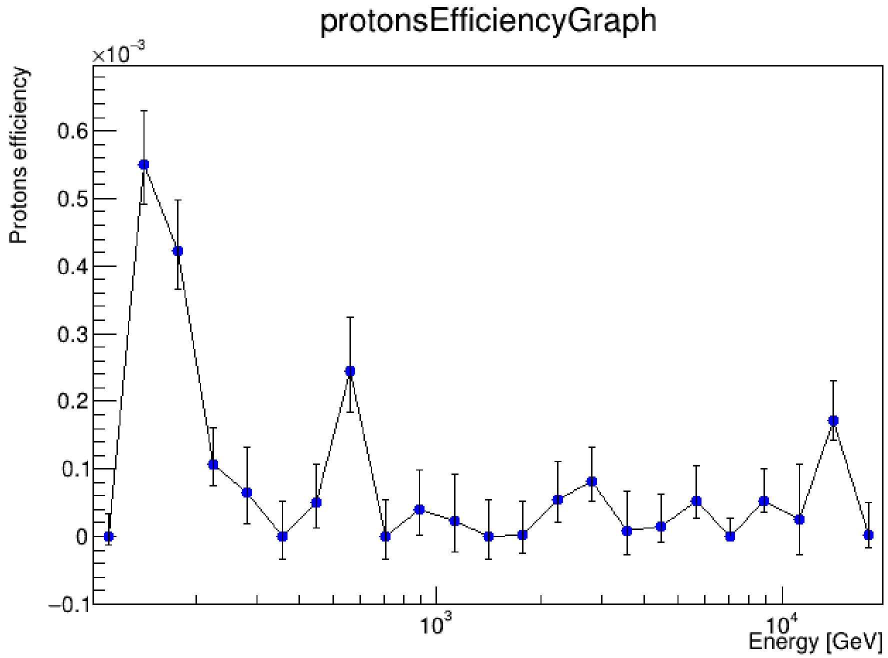
If we correlate L_{10} and 4th lateral momentum:



A small spot of electrons in a sea of protons

Applying some cuts like the green line in figure….

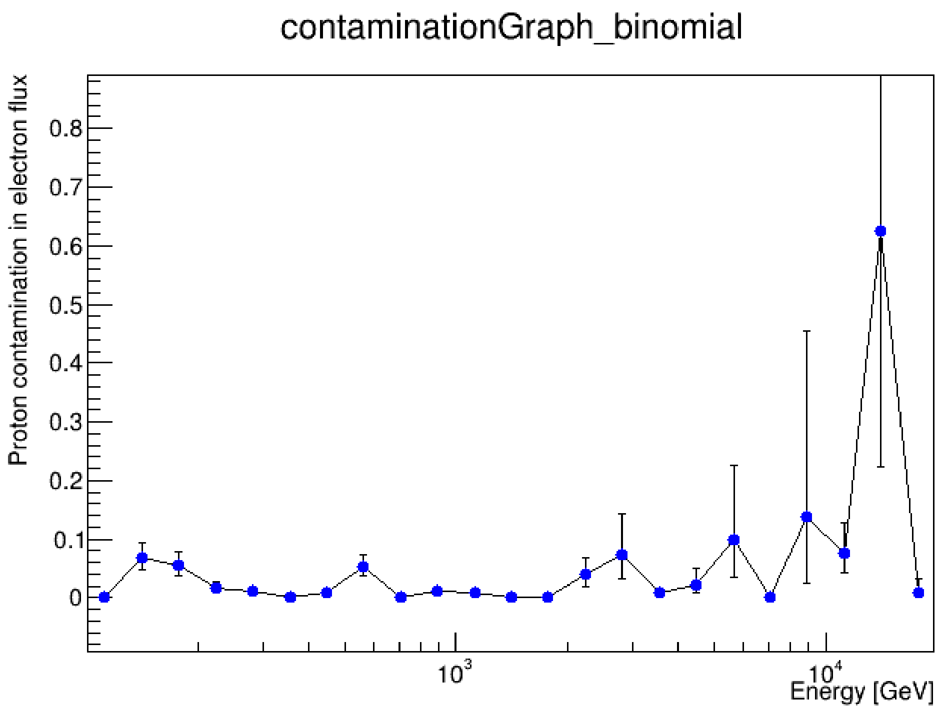




For the error calculation on efficiency Wilson method has been chosen:

- Clopper-Pearson conservative

- Feldmann-Cousins best theoretical performance but practically low computation time and not significant differences



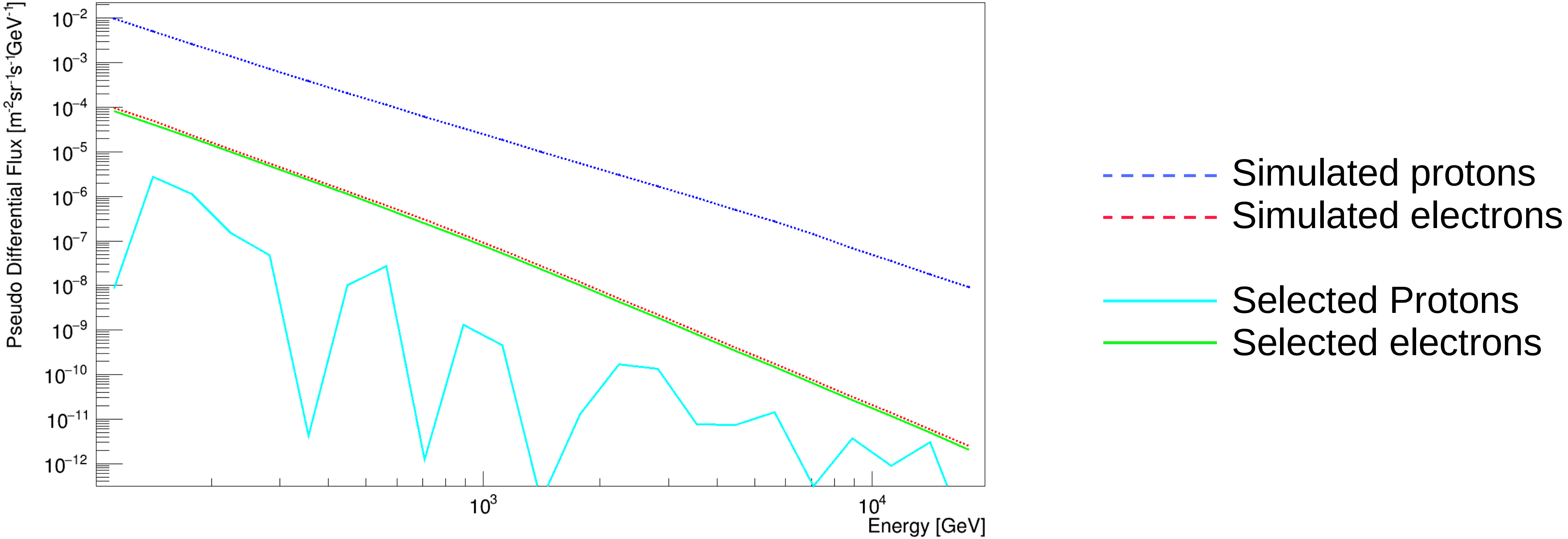
For the error calculation on the contamination:

- assuming selected electrons and protons distributed as Poisson distributions (a small probability to select them in a very big sample of protons+electrons)

- the ratio between two Poisson distributions is a Binomial distribution

- the confidence interval of a Binomial distribution can be estimated exactly with Clopper-Pearson method

Dividing the number of events for the geometric factor, the acquisition time and the width of the energy bin:



What’s next?

- Use digitized data (digitization is ongoing) and the full statistics
- Study rejection performance using more variables
- Use boosted decision tree in rejection studies
- Perform error calculation in the electron flux analysis

- Data analysis of PS-SPS2023 beam test for MC validation (data analysis is ongoing)