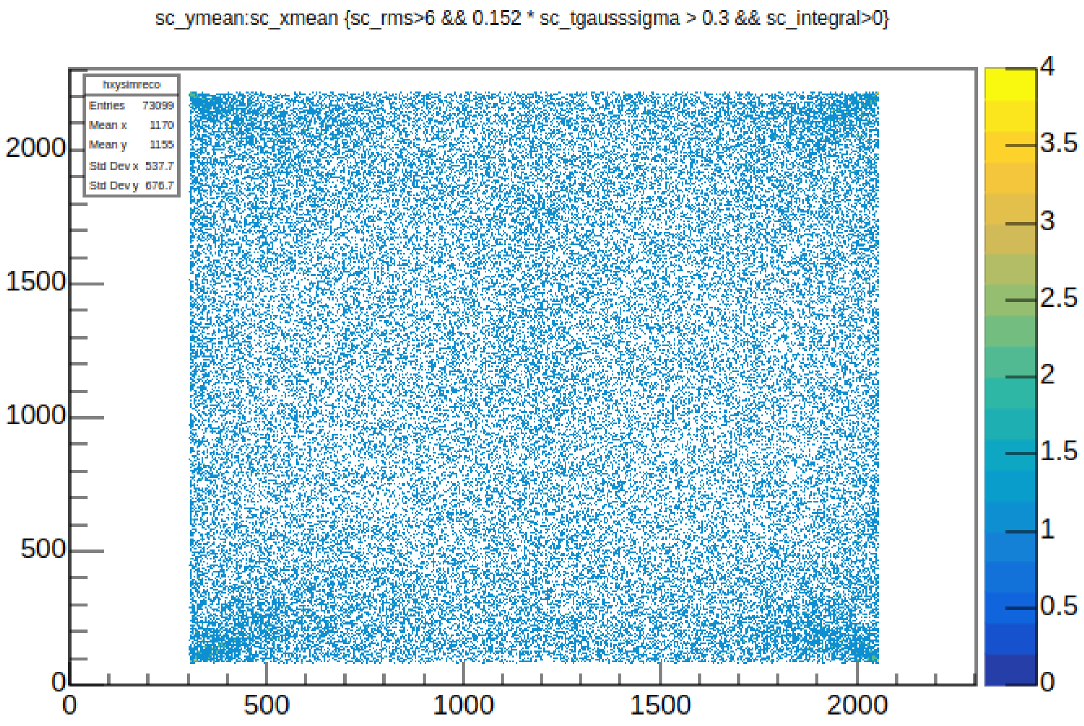
Data/MC comparison Run2

F. Di Giambattista, CYGNO Simulation Meeting 27/06/2023

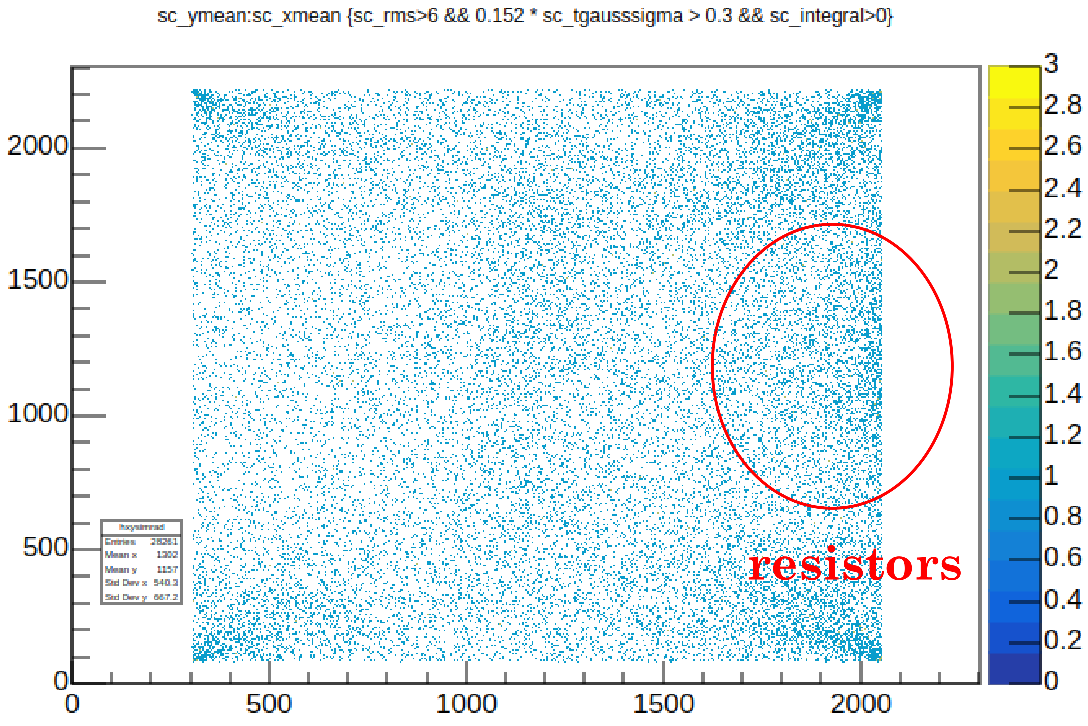
Background simulation chain

- Monte Carlo (MC) simulation of expected background in LIME underground with GEANT4
    - Main contributions:
        - External gammas produced in the lab rock
        - Radioactivity of LIME materials (field rings, cathode, acrylic box, GEMs, field cage resistors, camera)
- Digitization is applied
    - Energy deposits in the sensitive volume of gas are diffused, charge amplification and saturation are applied, vignetting effect is included, and images with the same granularity of real data are produced
    - Energy and spatial distribution of simulated events are retained, and the images are comparable with data
- Events are reconstructed with Winter23 version of the code (same as the one used for run2)

X-Y distribution – simulation



External gammas

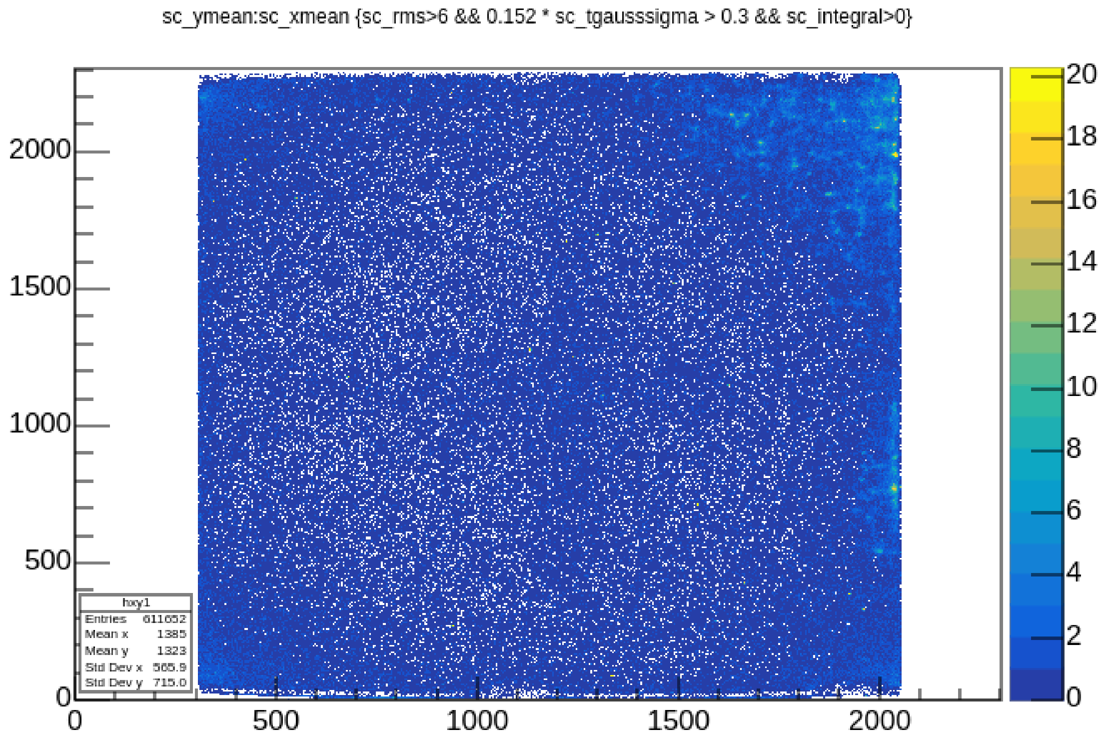


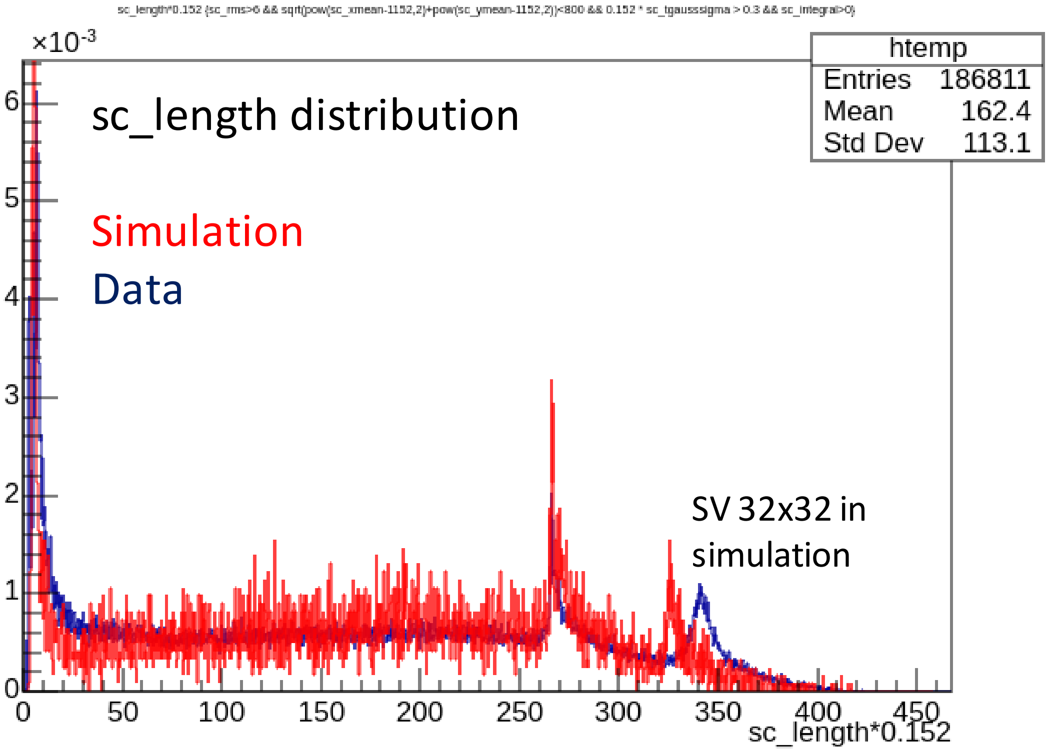
Internal radioactivity

X-Y distribution

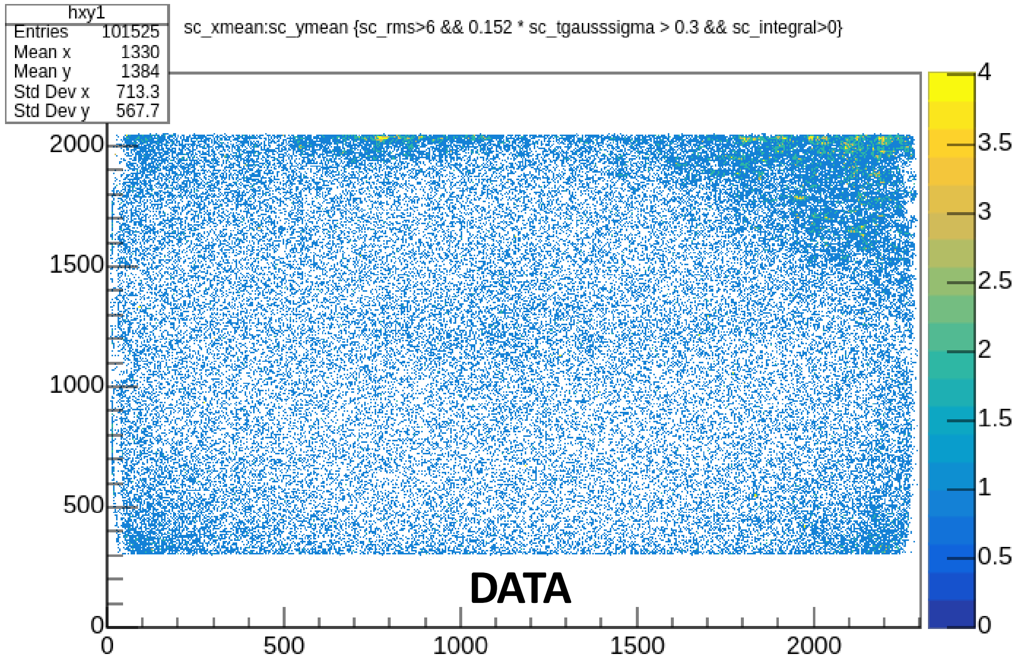
- Central cross attributed to long tracks, whose barycenters tend to create this shape
  - Not really visible in simuation

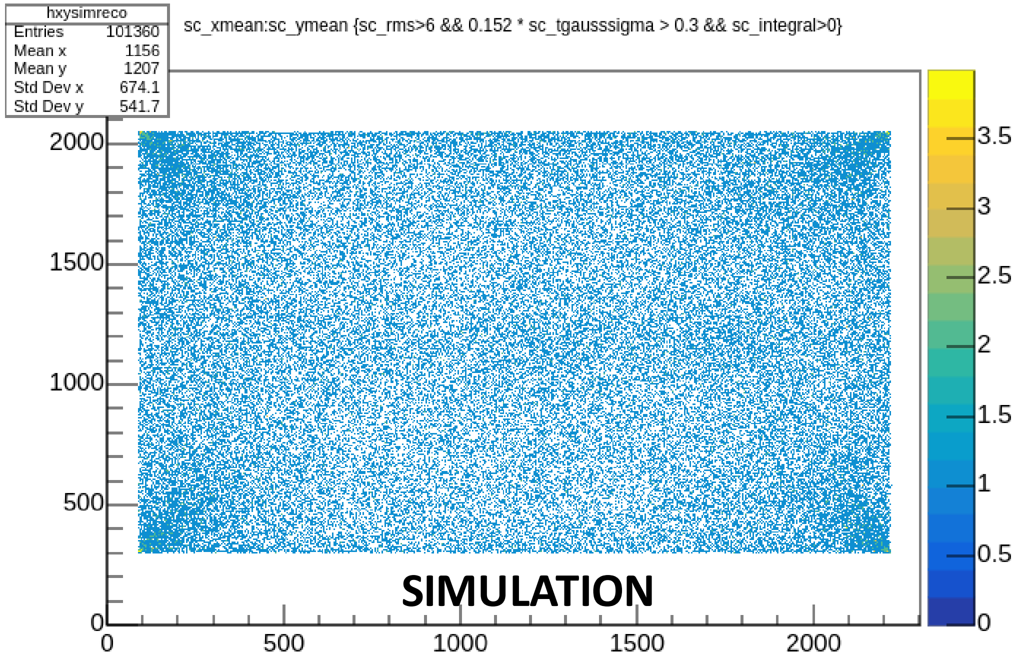
- Maybe the simulation is missing long tracks?
  - *Partially* yes, because I digitized only tracks with E<500keV, **BUT**:





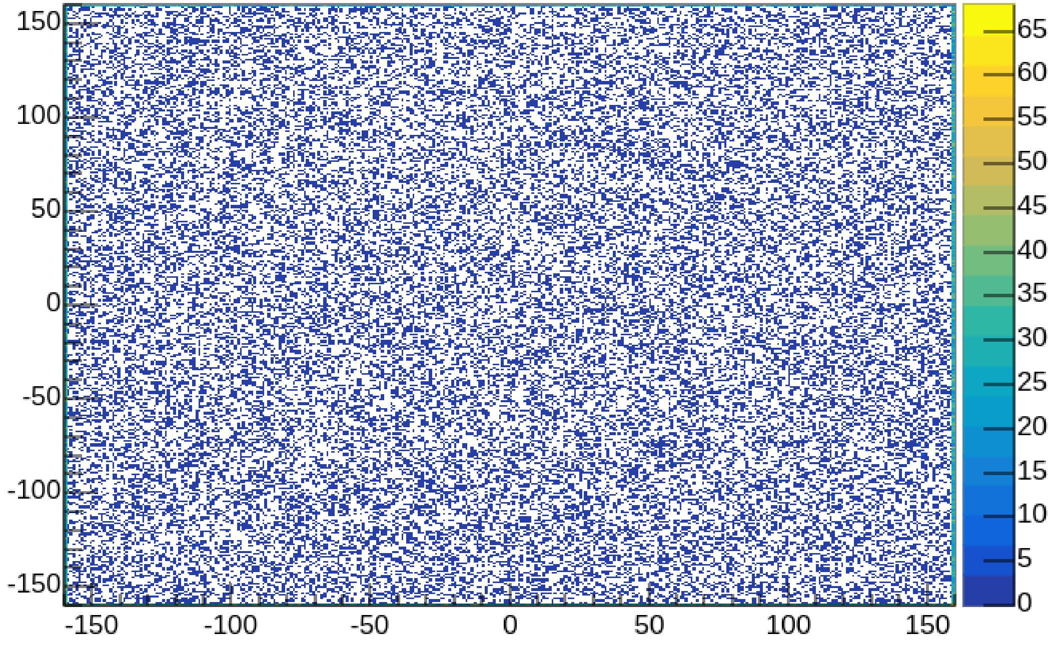
X-Y distribution

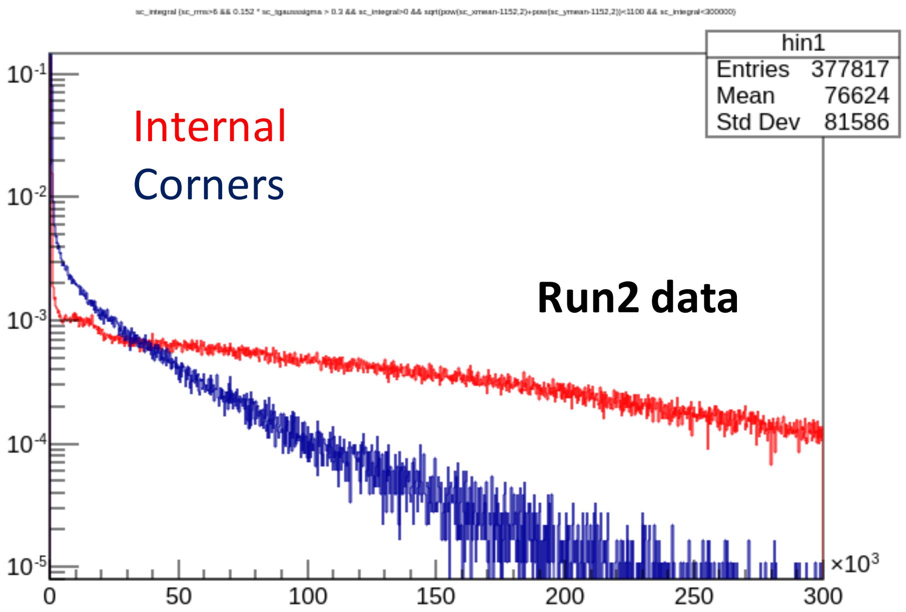




- Plotting a limited number of entries in data shows a similar shape to the simulation distribution, both regarding the central cross and the excess of events in the corner
- Main difference is the excess of clusters in the upper part (fake clusters from the camera, all below 1keV)

X-Y distribution – MC truth

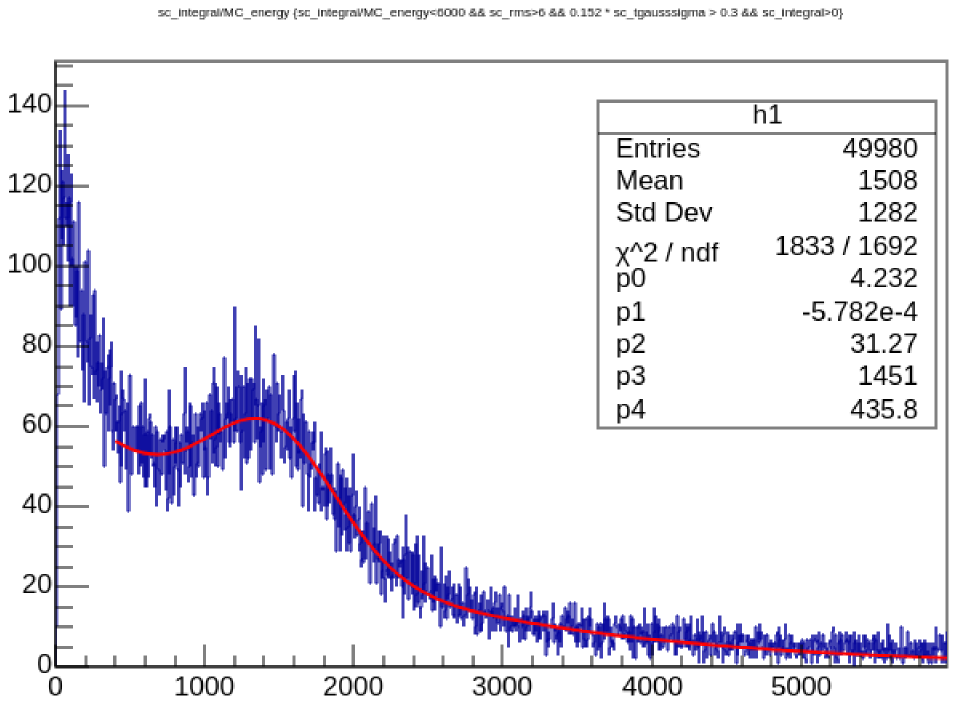


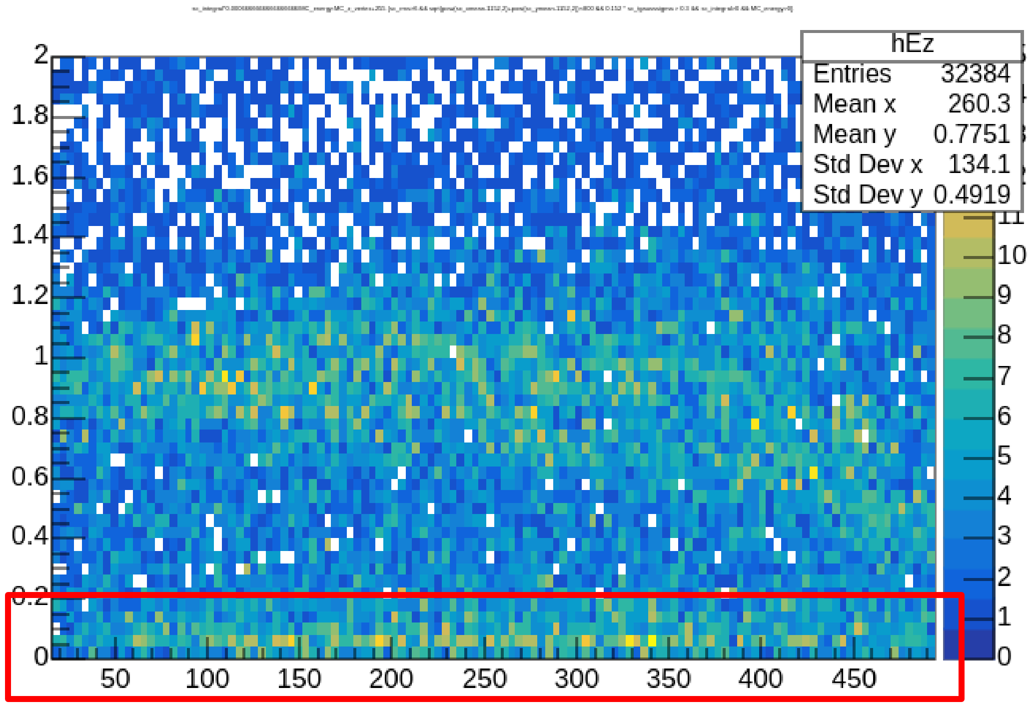


The excess in the corners is absent in the MC truth, so it's an effect of reconstruction (it's present in data and in simulation of external gammas and radioactivity)

Energy (uncalibrated) distribution of events in the central region (R<1152) is different from the corners distribution

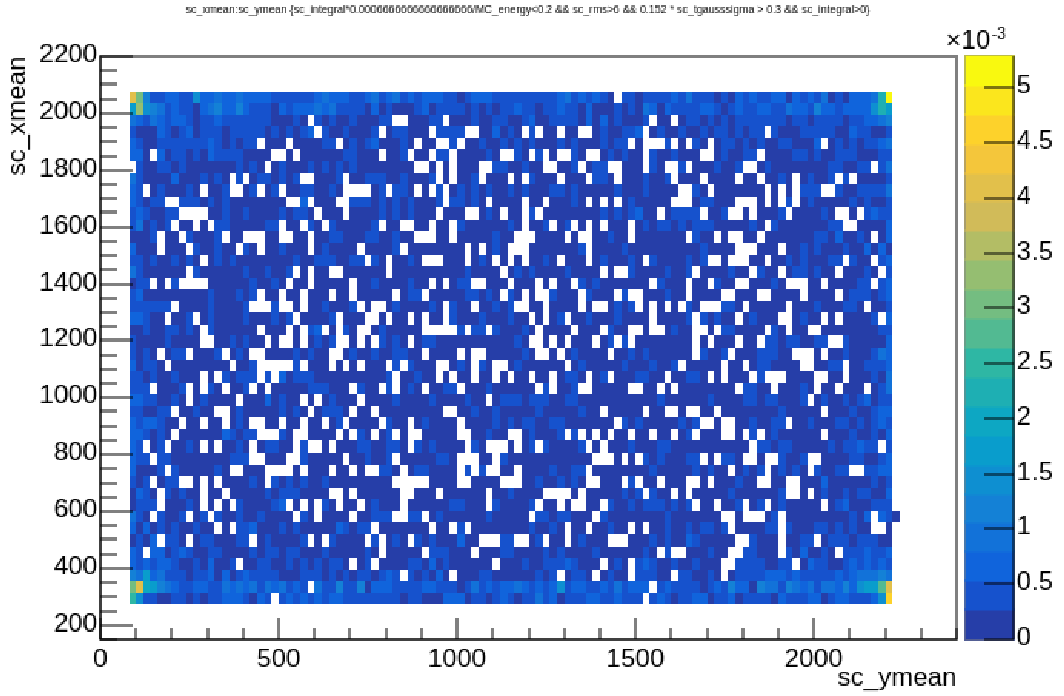
Energy calibration - MC



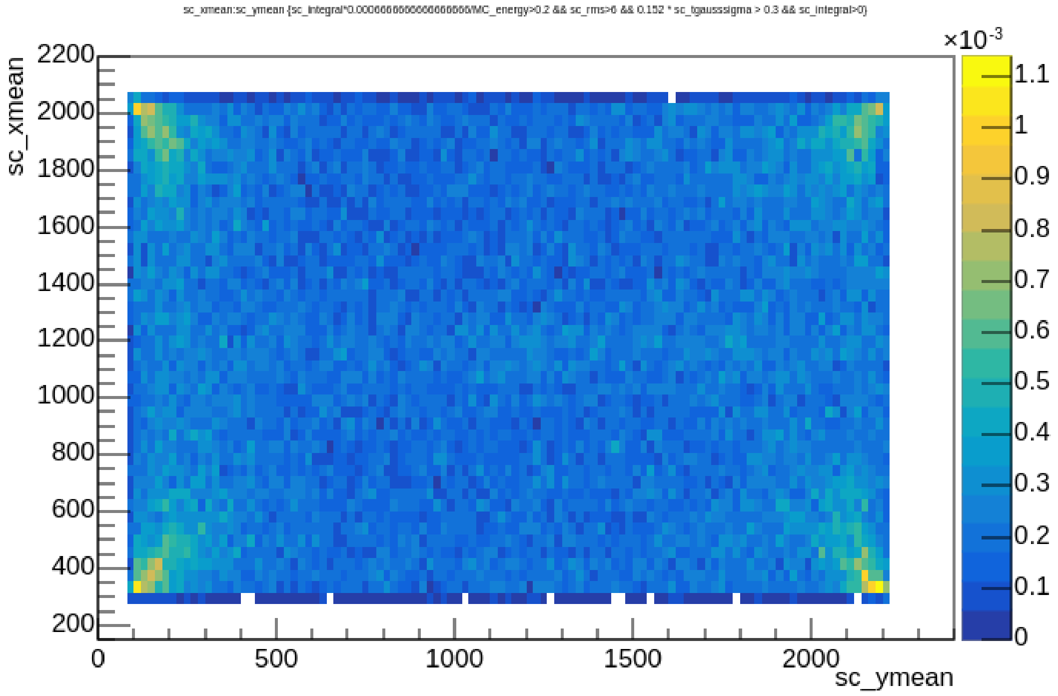


- From distribution of sc_integral/MC_energy I fitted with a gaussian+exponential and took the mean of the gaussian as a constant calibration factor for simulations
- Calibration factor seems constant and then drops at large distances from the GEMs. Problem with saturation?

Energy calibration - MC



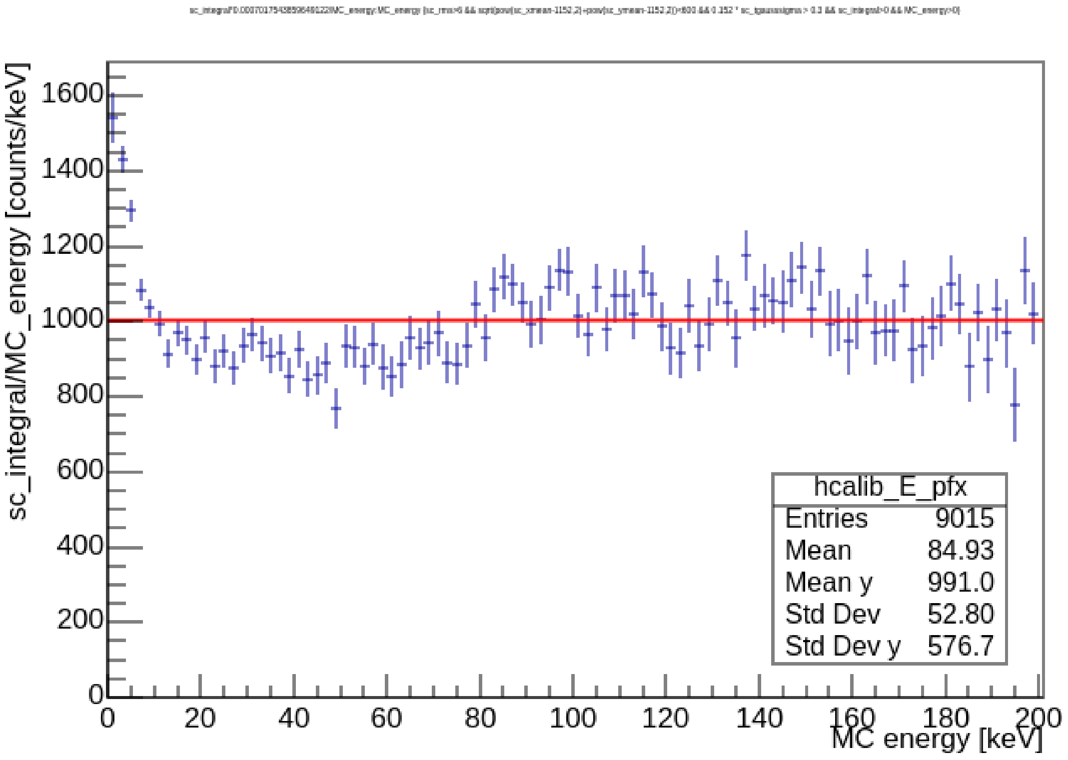
sc_integral/MC_energy<0.2



sc_integral/MC_energy>0.2

The events in the low band of calibration factor are mainly in the corners and in the upper and lower bands (tracks that are cut in the reco when excluding those bands?)

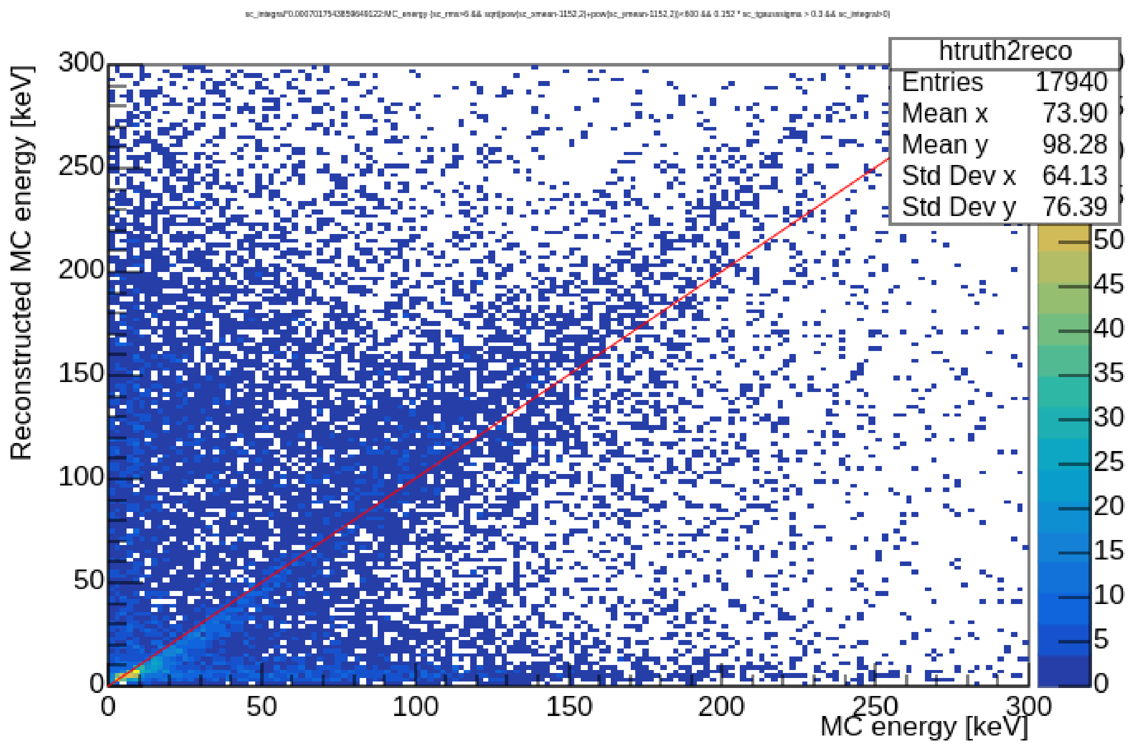
Energy calibration - MC

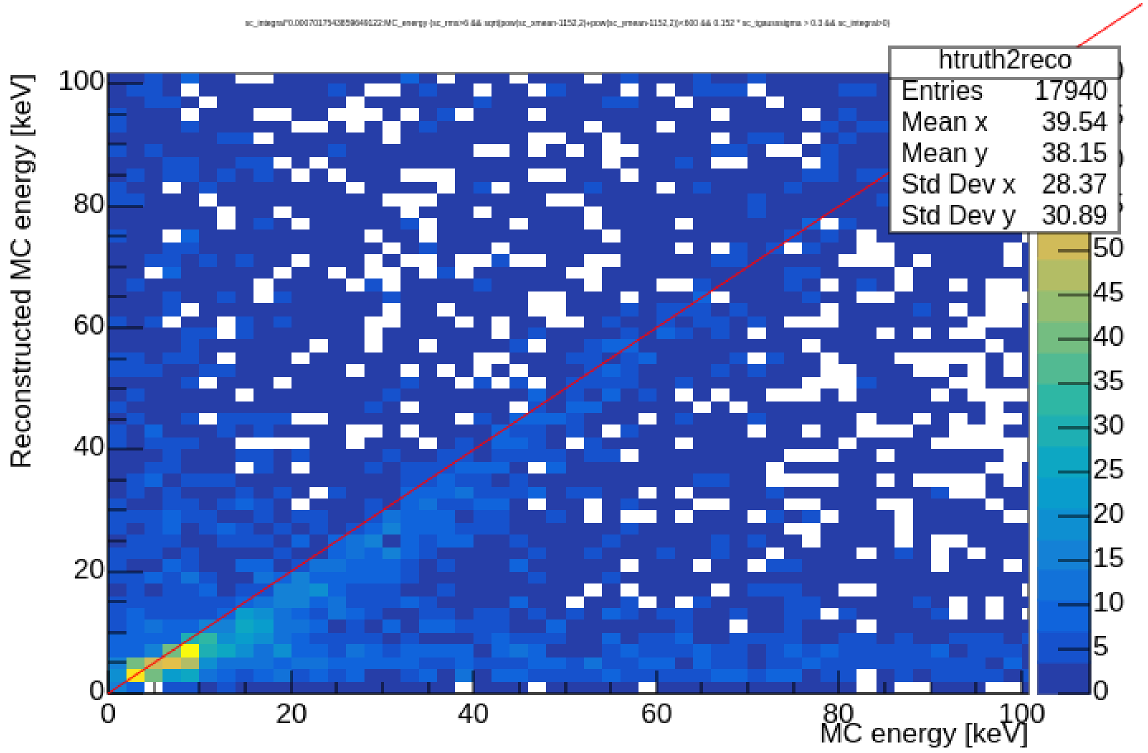


Another attempt at calibrating:

- Constant fit of sc_integral/MC_energy as a function of the MC energy
- Fit between 1 and 200 keV (avoid fake clusters and low statistics/cut tracks at high energy)

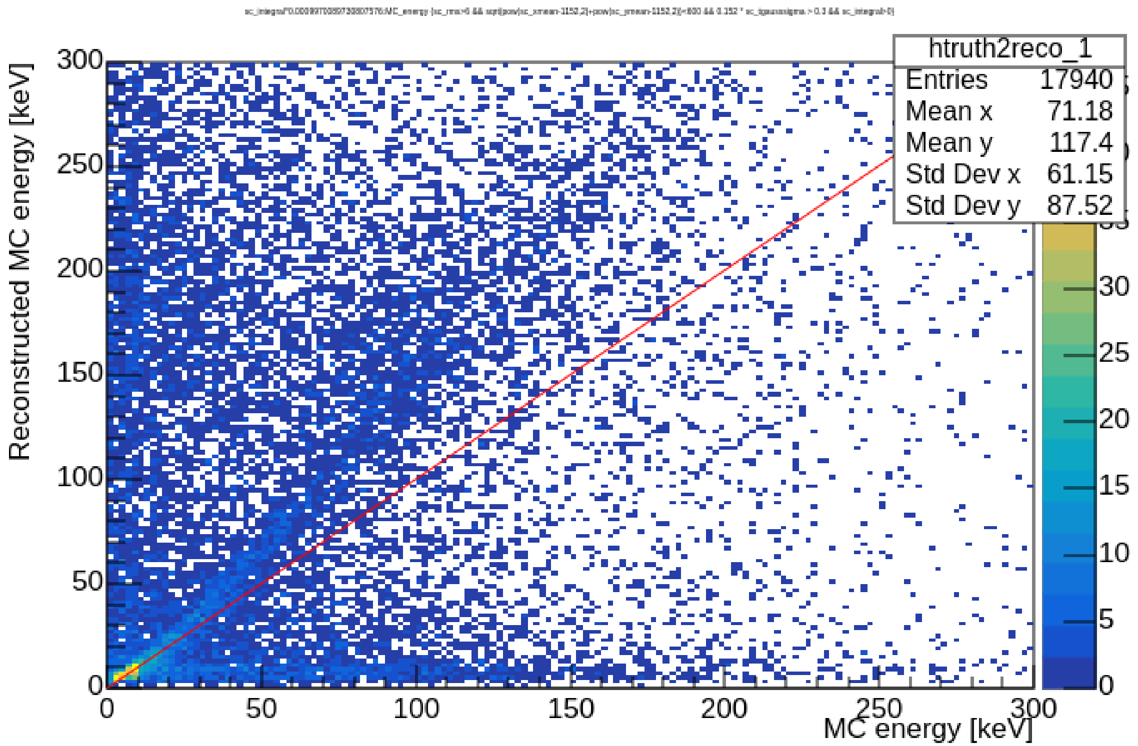
MC truth energy vs post reco energy

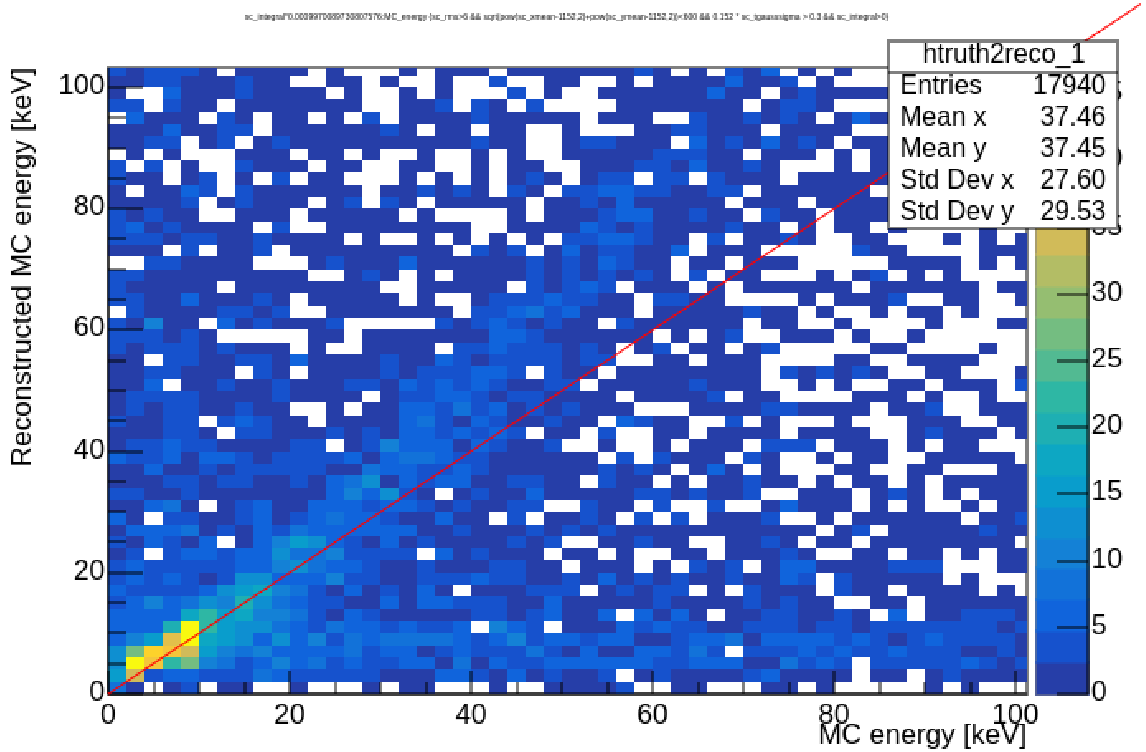




Note: calibrated from gaussian peak

MC truth energy vs post reco energy

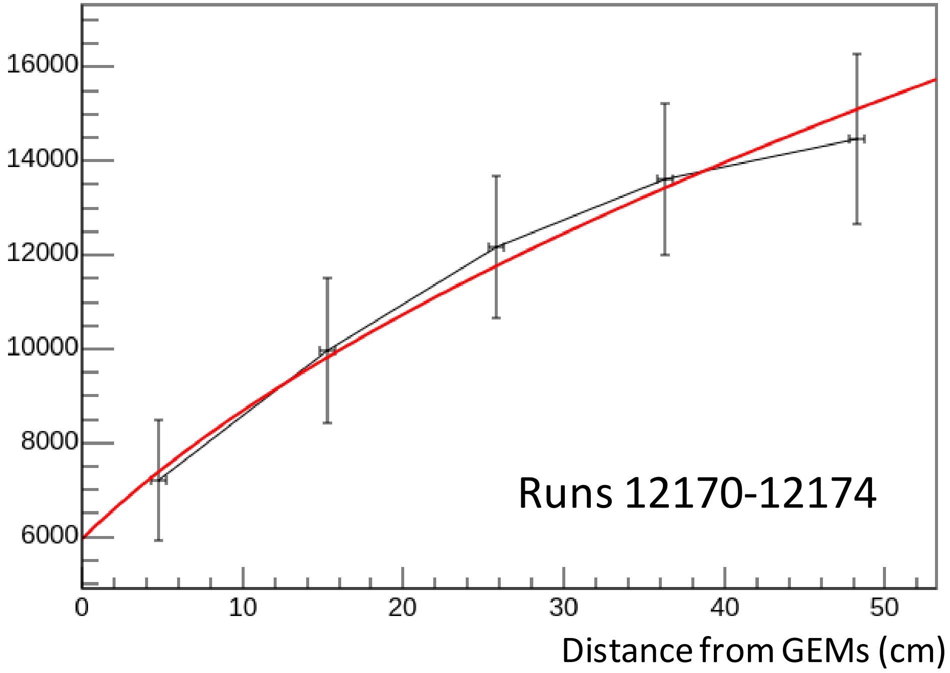




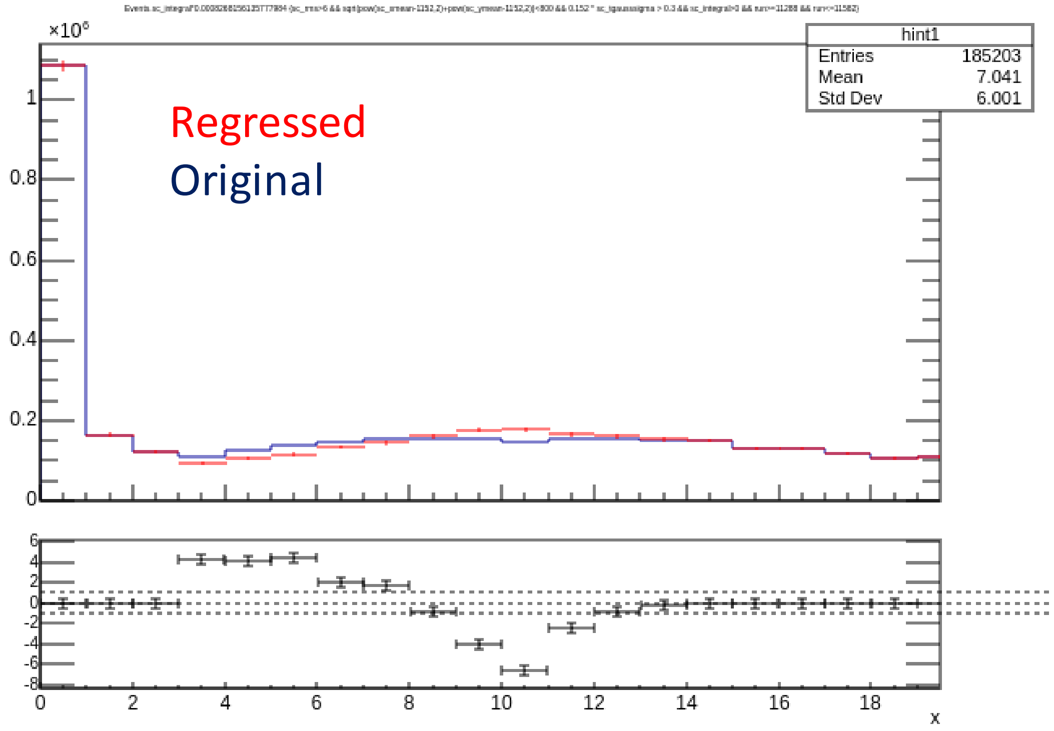
Note: calibrated from pol0 fit

Energy calibration - data

- Fe55 calibration runs, each sc_integral distribution is fitted with a gaussian; mean and sigma are taken to fill a graph as a function of position (for each set of calibration runs)
- I fitted the graphs with p0*sqrt(Z+p1)
- In average I get p0=2050, p1=8
- First approach: use 3rd position (25 cm distance)
- The calibration factor found in the simulation would correspond to a distance of
    - **9cm** from the GEMs (from gaussian peak)
    - **3mm** from the GEMs (from pol0 fit)
        - Another hint that the simulation of saturation doesn't work?
- Second approach: use the extrapolated calibration at 0cm distance from GEMs to compare with MC



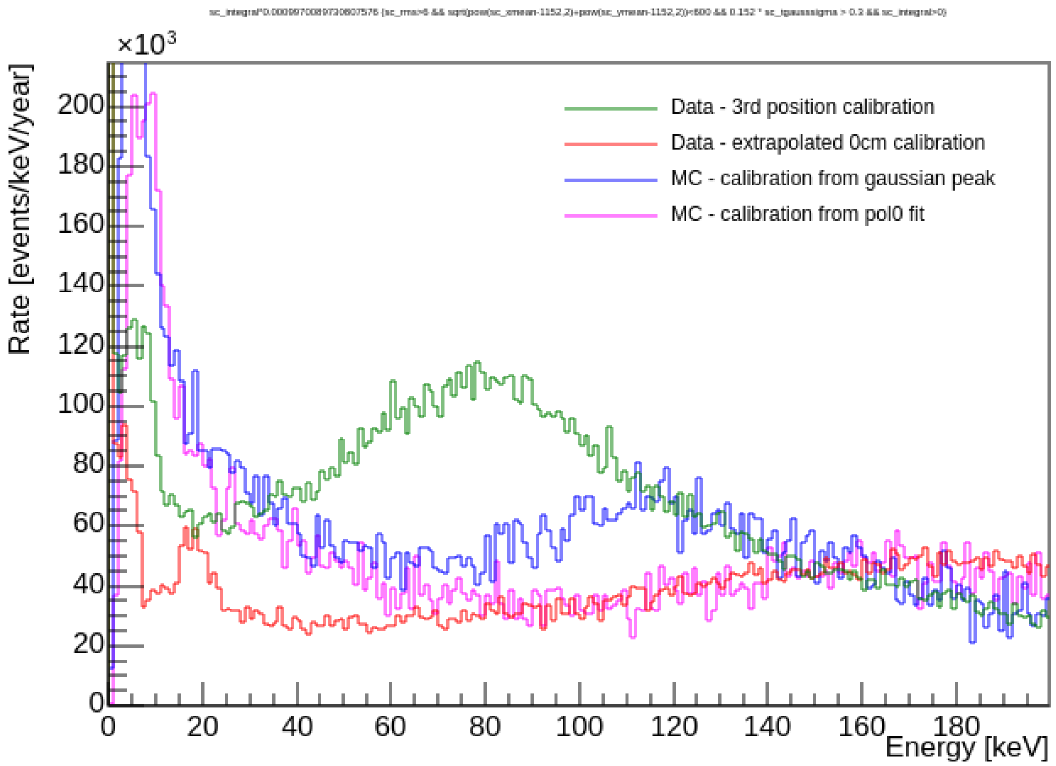
Energy and z regression

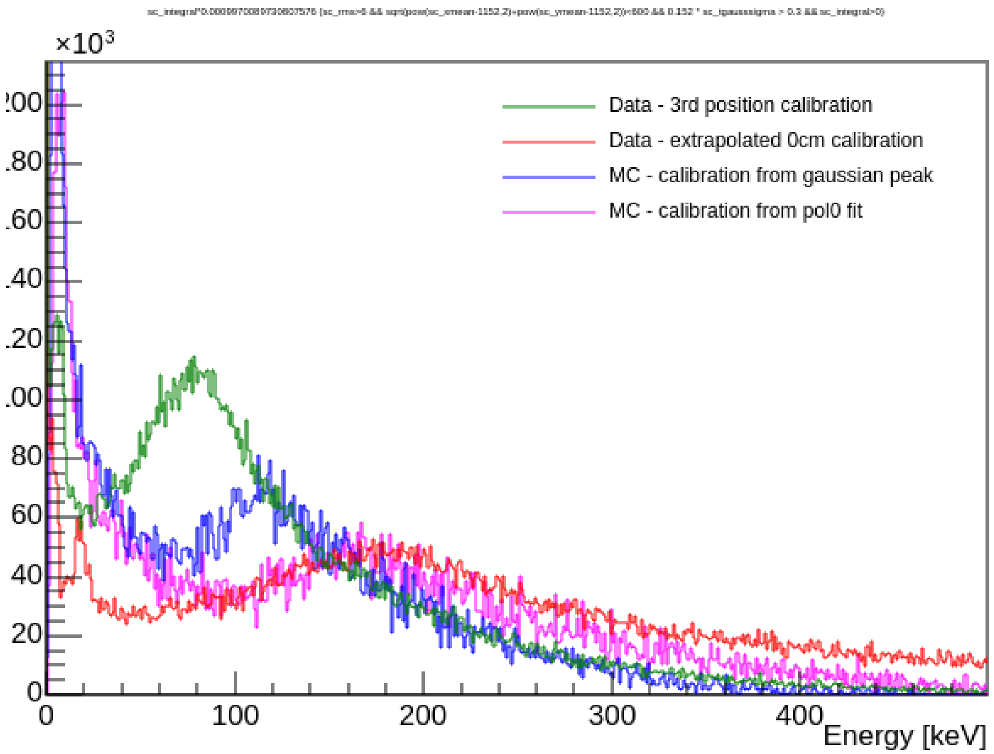


- Regressed and non-regressed energy spectra are the same except for the region around the Fe55 peak (which it was trained on)
- The z position is estimated only in the same energy region

- We should train the regression in E and z on MC simulation to enlarge this region

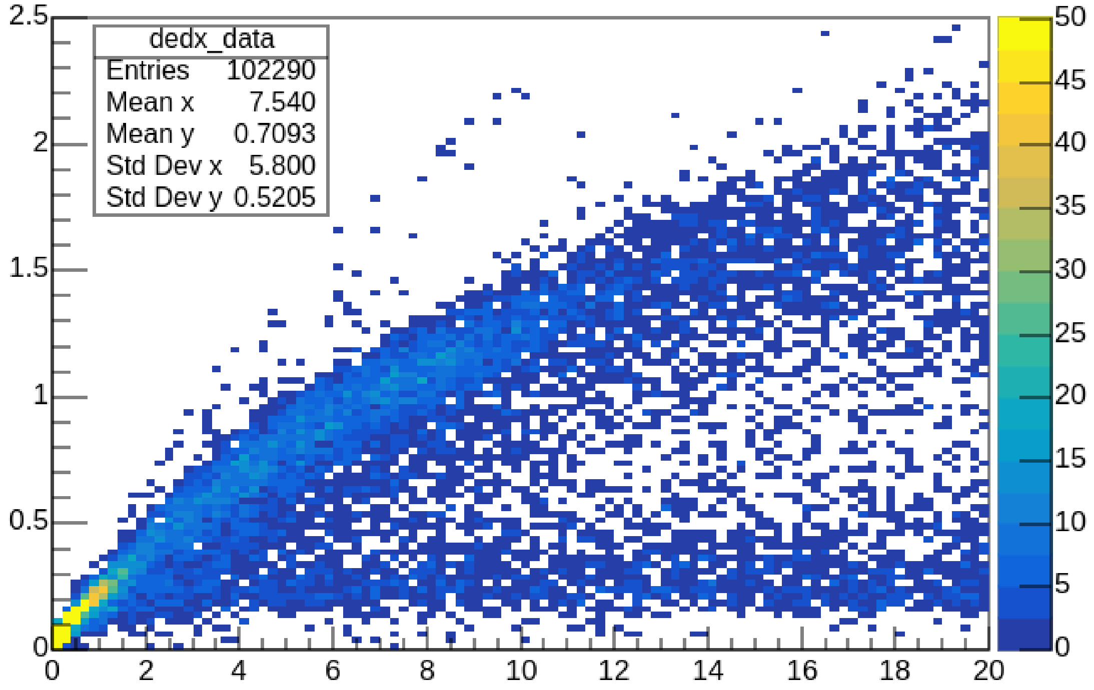
Energy spectrum comparison

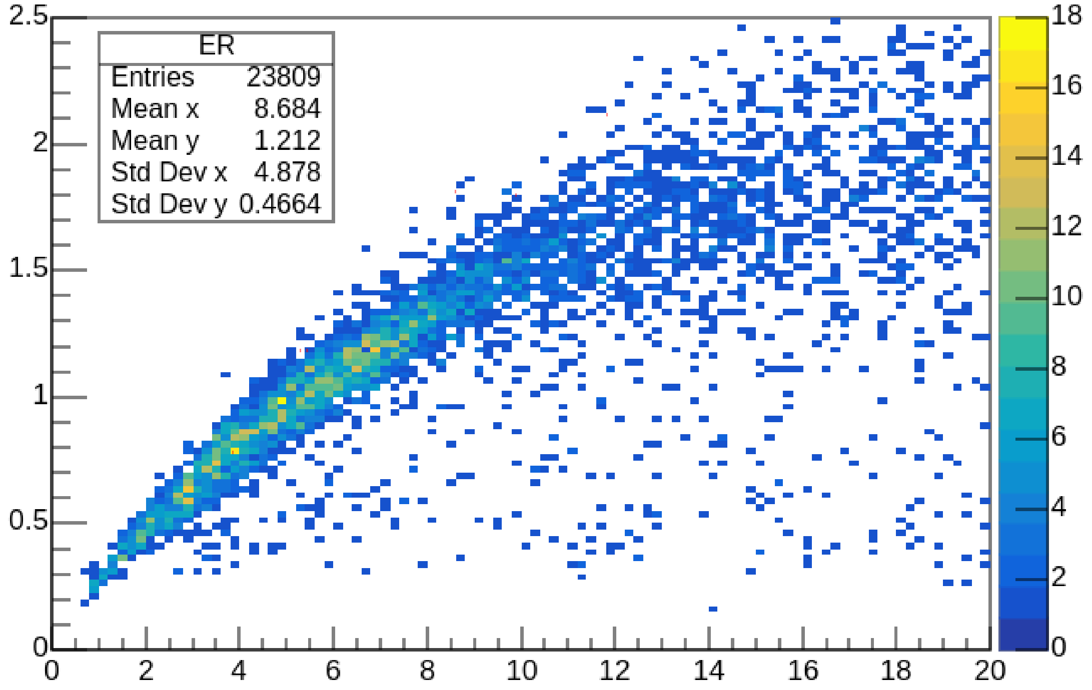




- Up to 60 keV: no match in shape with data (and the calibrated MC are the same)
- Above 60keV pink and red match; above 110keV blue and green match
- Note on normalization: I used the total run time for the data, which corresponds to assuming there were no real events during the dead time

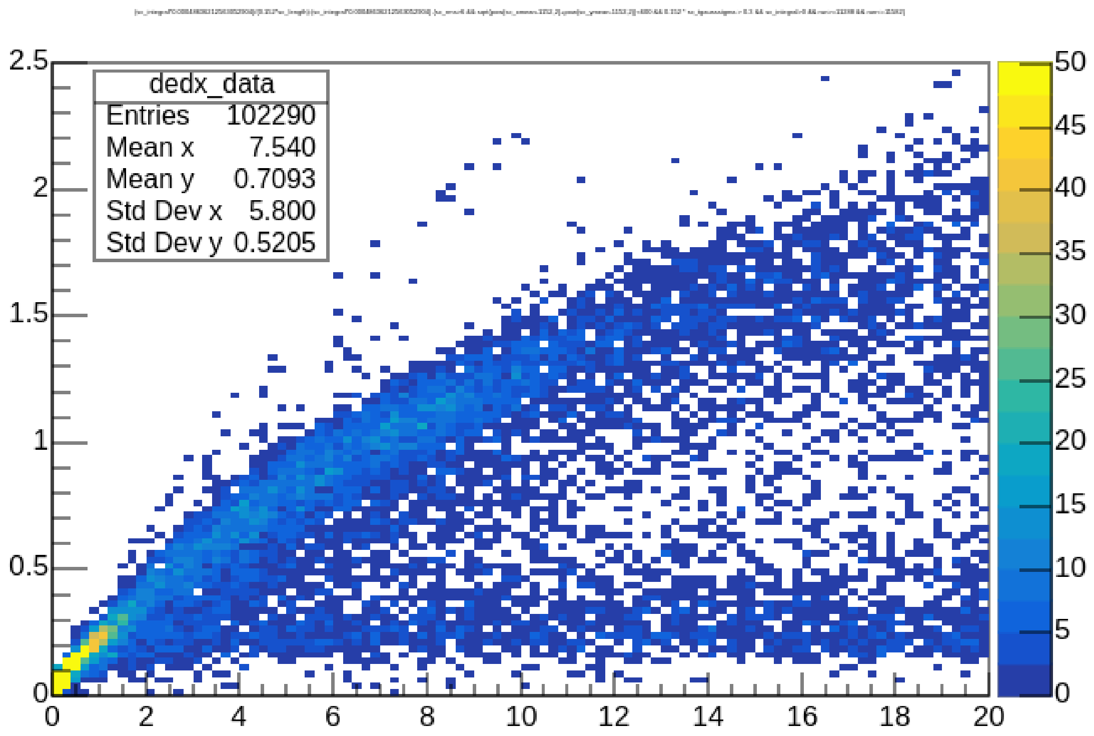
dE/dx vs energy

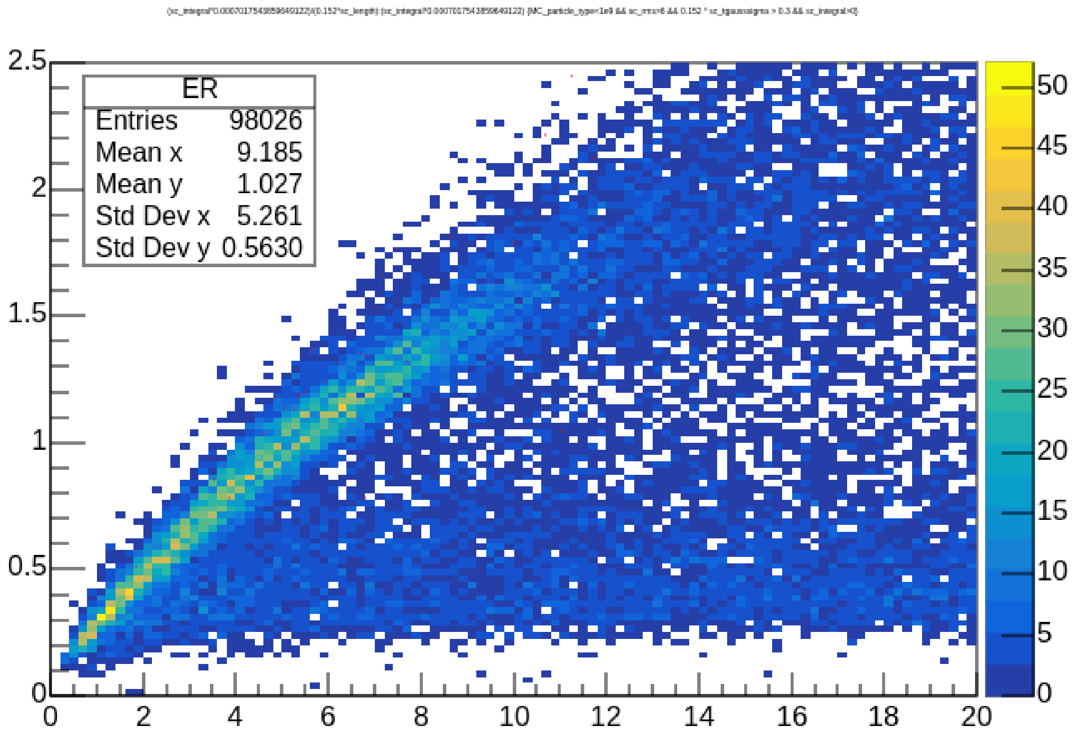




Why there is no "MIP" band in the simulation? We ruled out the absence of long tracks

dE/dx vs energy

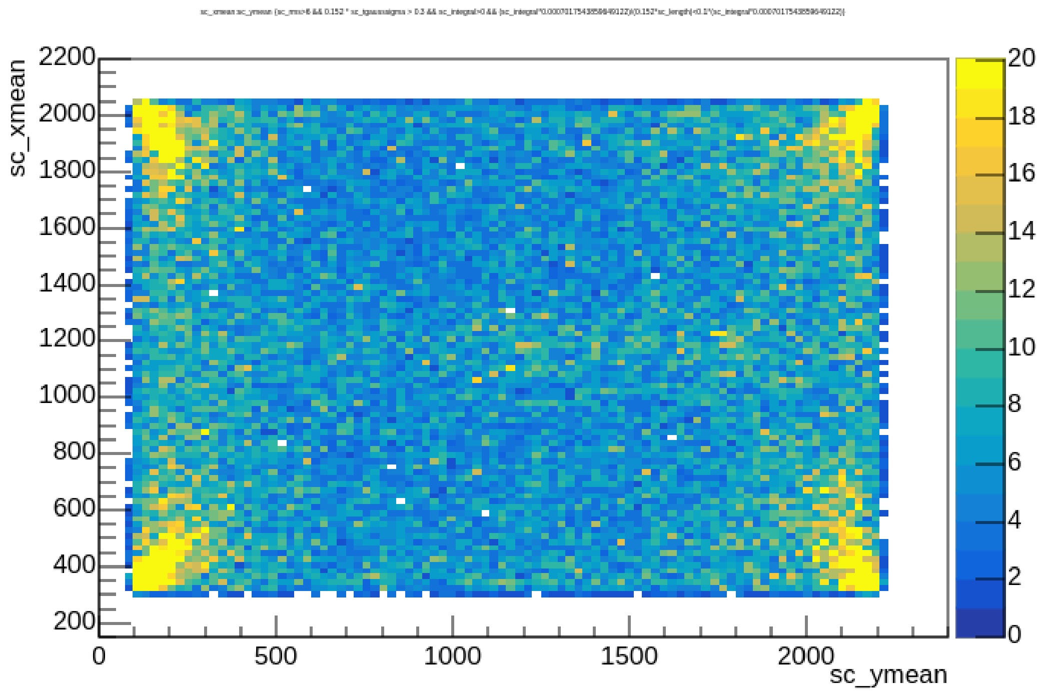




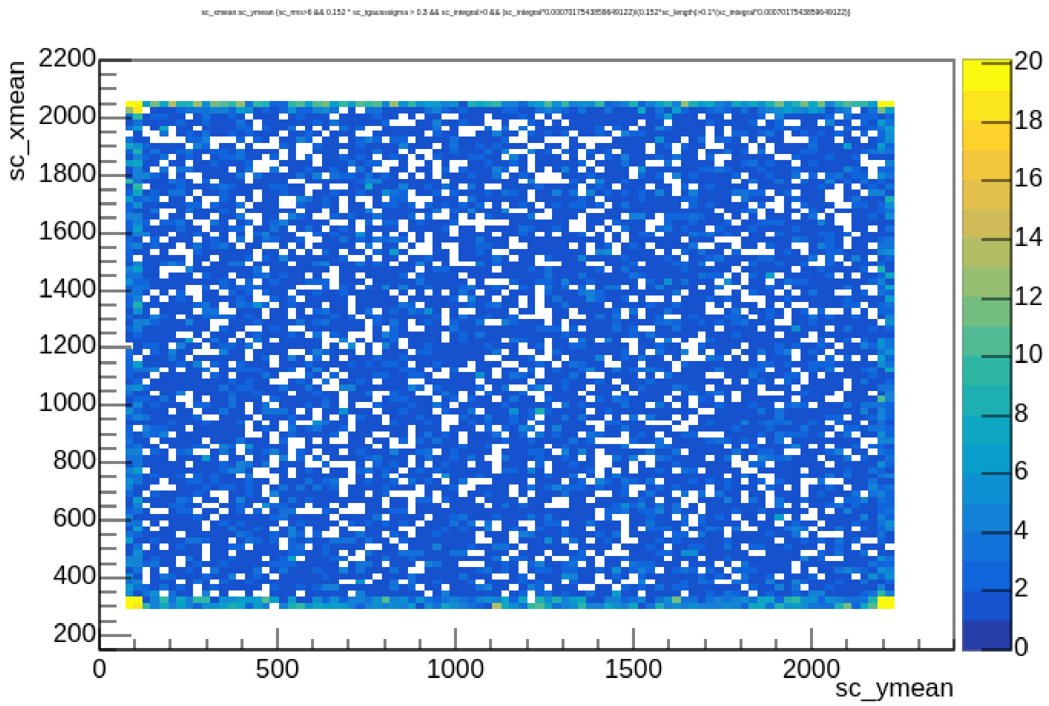
I removed the geometrical cut from the simulation, and there it is

dE/dx vs energy – xy map from MC

I selected the events in the "MIP" band now found in the simulation, and the xy map shows that those events happen mainly in the corners:

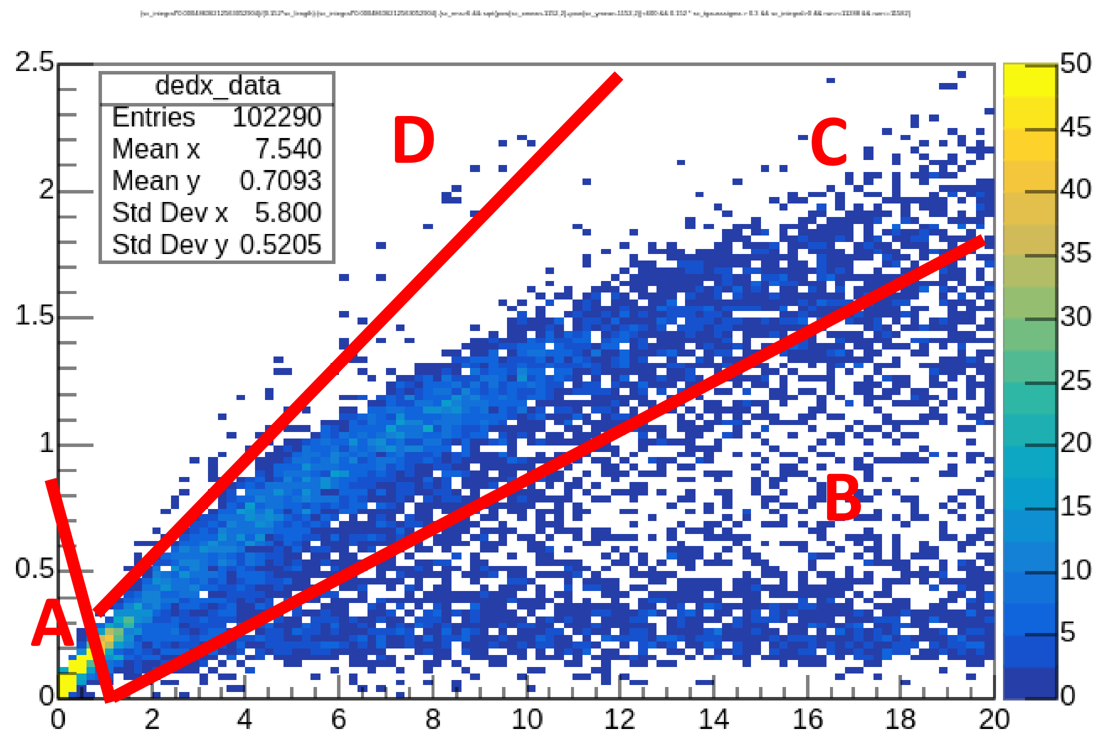


Events in the "MIP" band



Events above the "MIP" band

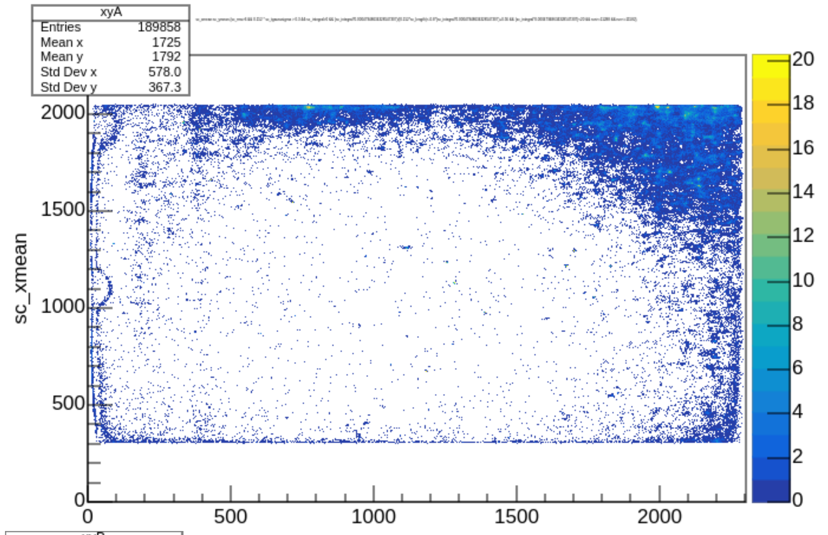
dE/dx vs energy - data



I divided the events in four regions to check where these events are in the xy map (in the region below 20keV)

What we can expect:
A = noise of the camera
B = MIP
C = ER
D = NR/sensor tracks

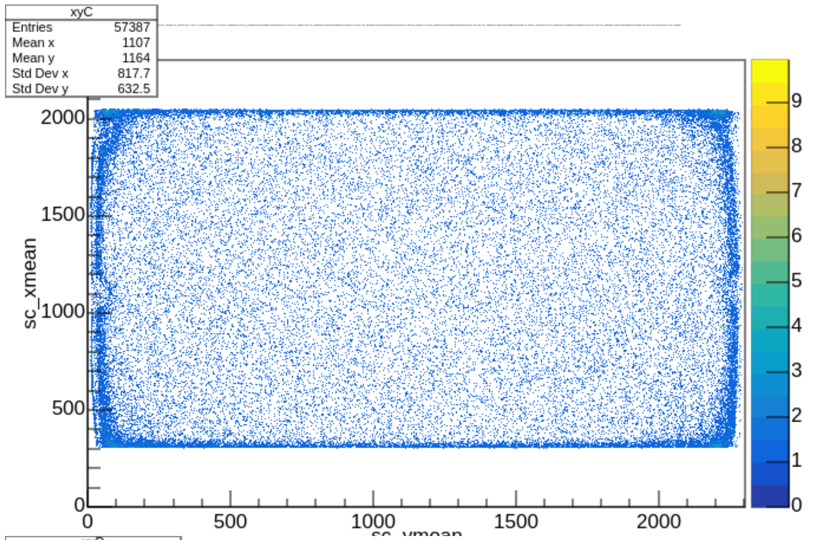
dE/dx vs energy – xy map from data



A

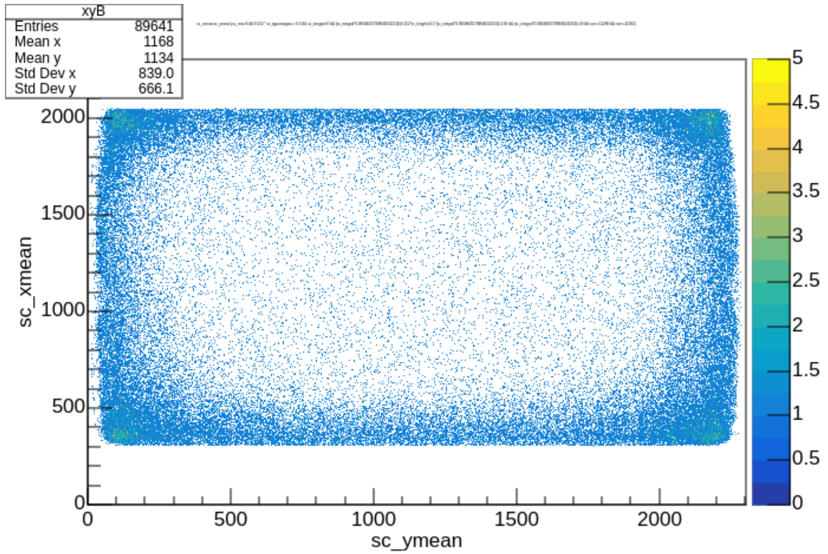
Mainly noise from the camera

What are the weird shapes on the left? (also present in C)



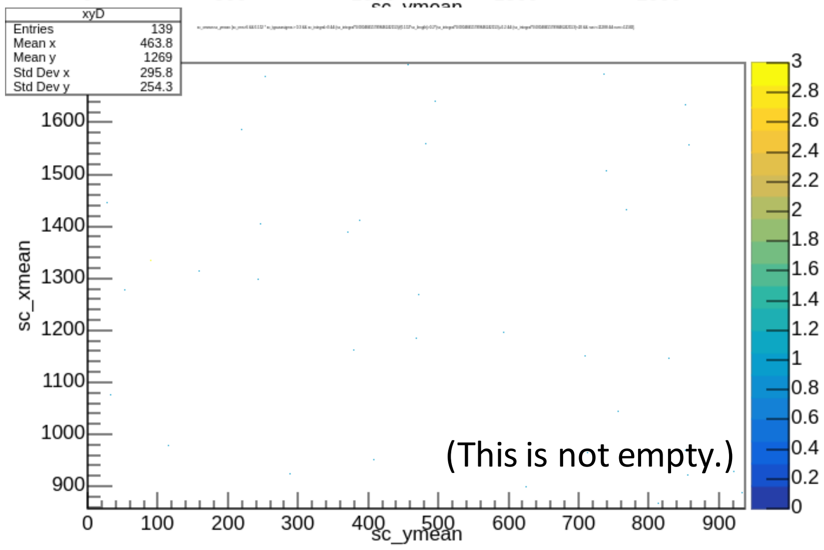
C

This should correspond mainly to real events



B

"MIP" band
Mainly in corners and around the border
Note: if I include tracks >20keV the cross appears in the center



(This is not empty.)

D

Probably tracks in the sensor?

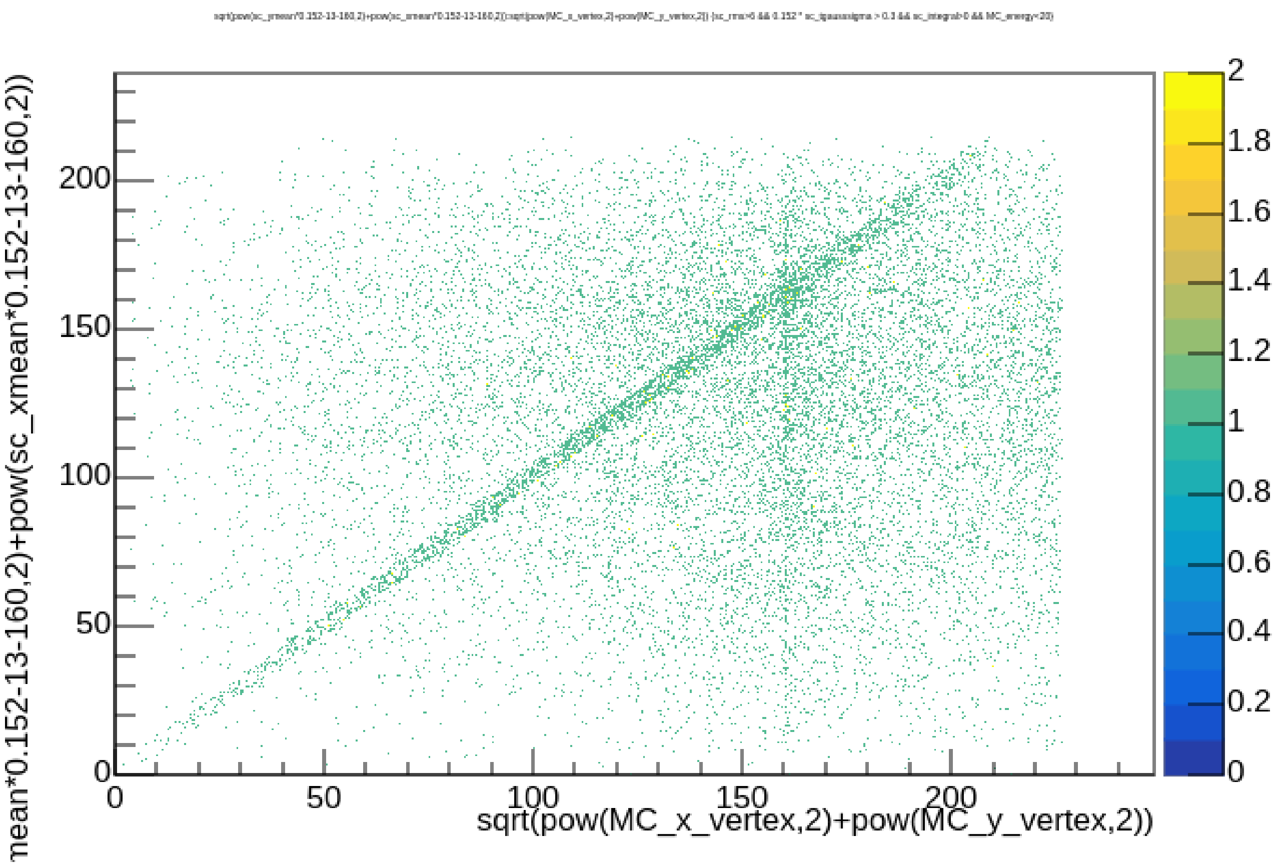
To do

- **Calibrate** data and MC properly – exluding the fake corners events might help
  - Is there a problem with saturation in MC?
- Look at the **re-reconstructed** data
  - Re-run the reconstruction on the MC simulation
    - Should I use a different pedestal?
- Properly account for the dead time to **normalize** correctly
  - We need a rough estimate of the true rate from a continuous acquisition of the PMT/images to understand what is the probability of missing an event during the dead time (i.e. time during which we acquire the pictures and the PMT waveforms)
  - Study the probability of cutting tracks during the time to expose the camera pixels (adding this to digitization might slow it down a lot...)

- Once the fake events in the borders and corners are fixed
  - Study reconstruction efficiency as a function of the energy
  - Study the fiducial cuts
    - We could study the internal background distribution from the border events, where they should be more concentrated

Backup

Track position – MC truth vs reco



Energy calibration