Clustering and Preliminary Results

03-06-2019

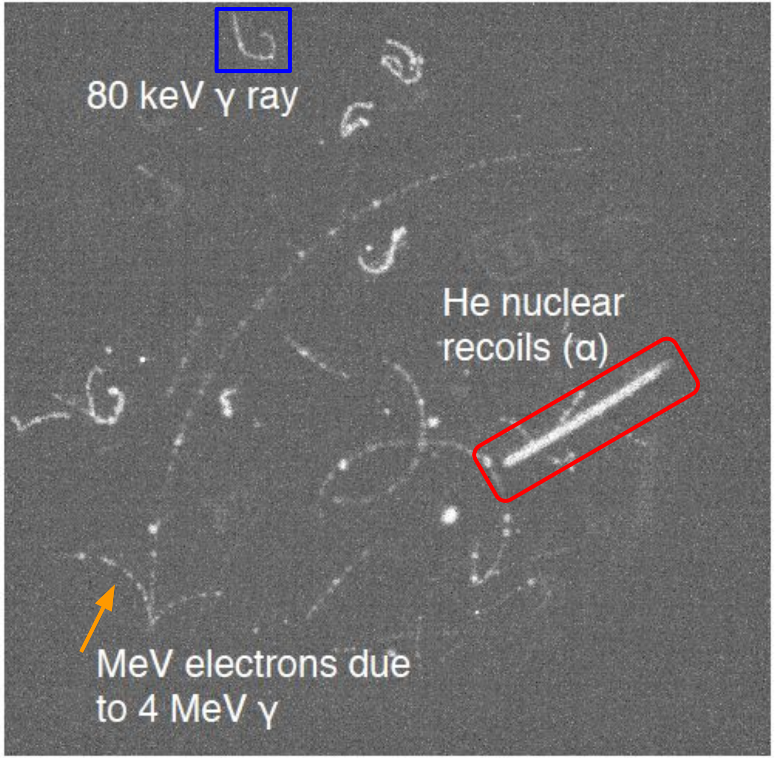
Setup used to took this data

❑ Took using the ORANGE detector;

❑ AmBe Neutron source

Using this configuration we expect to see three types of signals:

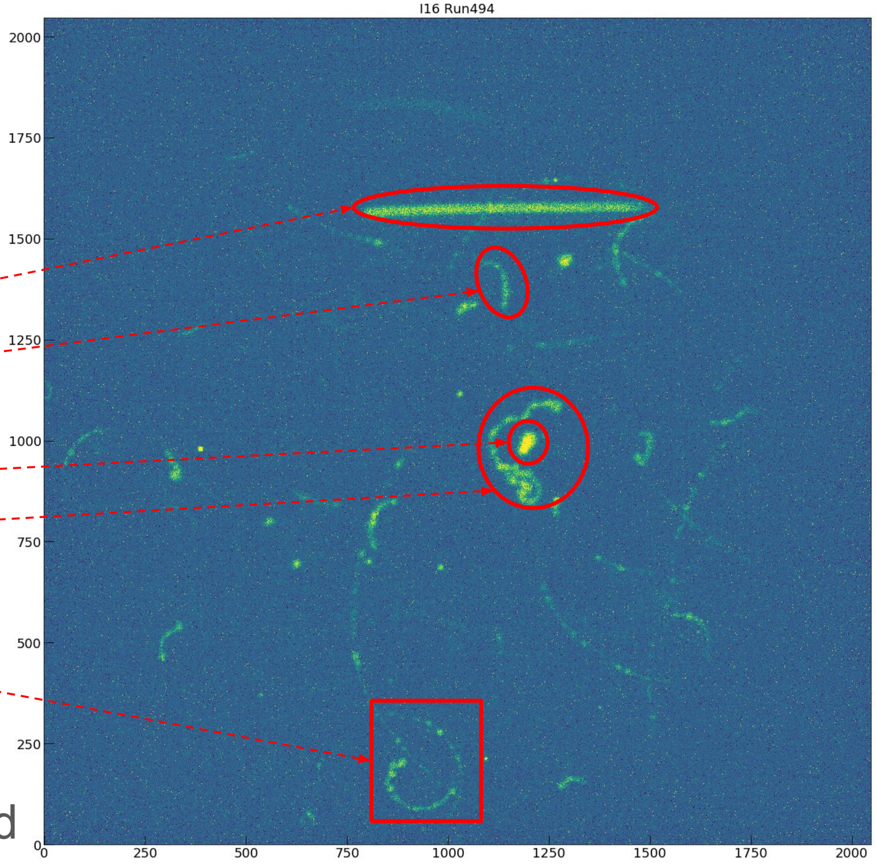
- He nuclear recoils (α); ●
- Low energy electrons due to X rays; ●
- MeV electrons due to 4 MeV γ. ●



Example of data

This image is a example of what we can have using a specific source and what we need to identify:

- ❑ Brighter and long tracks
- ❑ Lighter tracks
- ❑ Brighter and rounded tracks
- ❑ Close tracks
- ❑ Overlapped tracks
- ❑ etc..



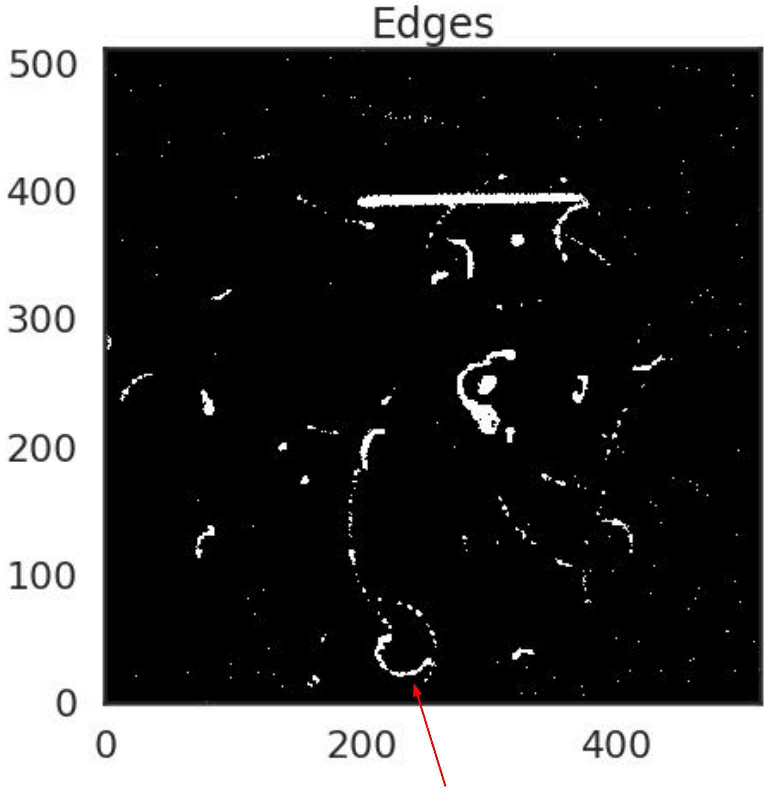
In order to recognize these differents kinds of tracks we tried a few clusterization methods and one works better to one kind like was expected.

Clustering Task

Therefore we thought about using DBSCAN (that can be set to find different tracks) in a iterative approach:

1. Run DBSCAN on a image to look to 'noise' clusters and remove them from the image;
2. Search first for tracks with **high** density of pixels;
    a. Remove them from the image;
3. Search first for tracks with **medium** density of pixels;
    a. Remove them from the image;
4. Search for others tracks;
5. Go to next image

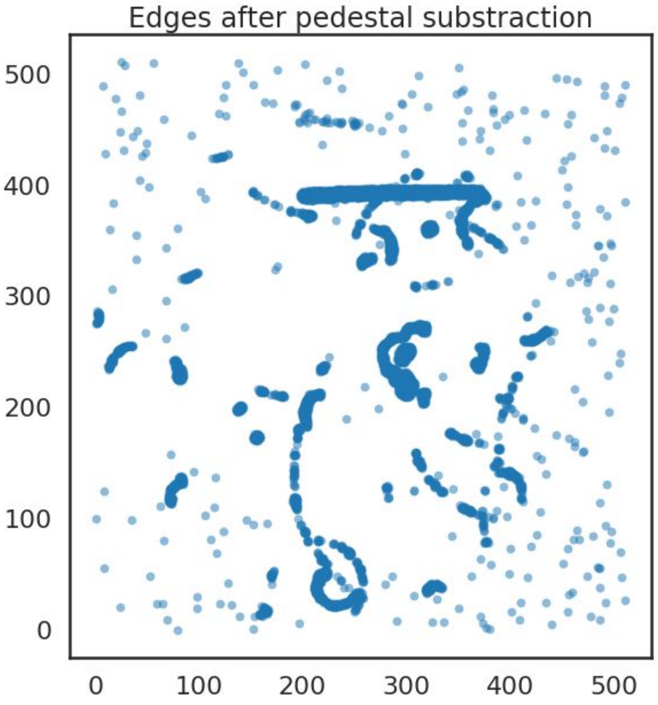
The parameters was decided looking through ~20 different images and many DBSCAN settings.

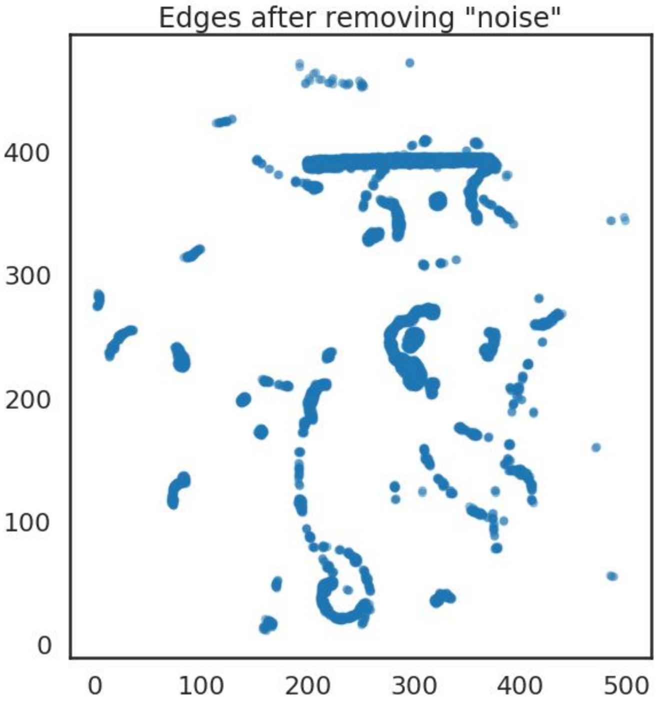


Theses X and Y are the input for the clustering algorithm.

Iterative DBSCAN Method - Step Zero

Remove ‘noise’ clusters (don’t have any near neighbors)





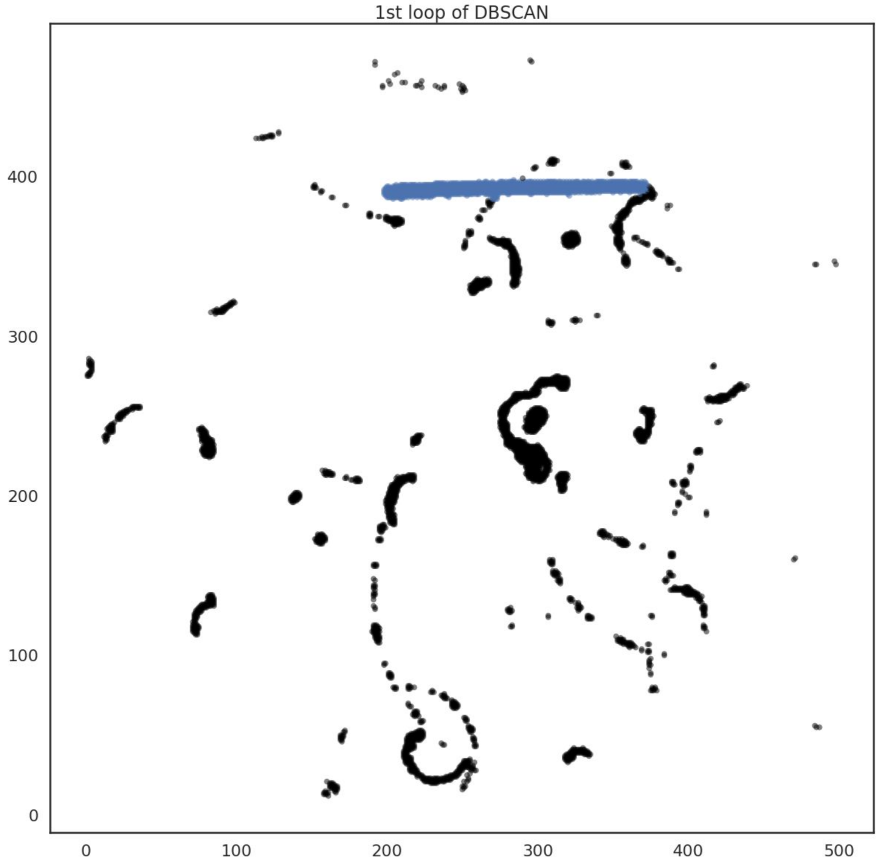
Iterative DBSCAN Method - Step One

In this loop DBSCAN was set to look for groups of pixels that have **high** density.

When the algorithm find a cluster in this step it is labelled as ‘long’.

Then, the found clusters are removed from the image to proceed to the next step.

(In the image at right different colors means different clusters and the black ones represents the not found clusters)

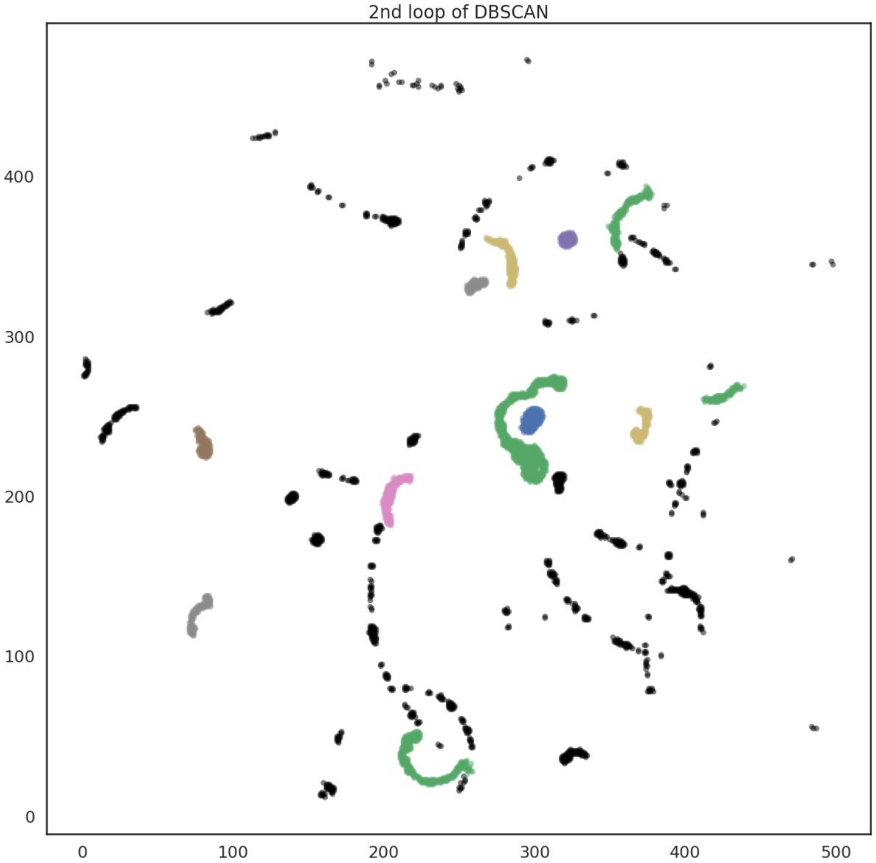


Iterative DBSCAN Method - Step Two

The second loop try to find groups of pixels with not so high density, let’s say **medium** density.

And, as in the first step the found clusters are labelled as ‘medium’ and removed from the image to proceed to the next step.

(In the image at right different colors means different clusters and the black ones represents the not found clusters)

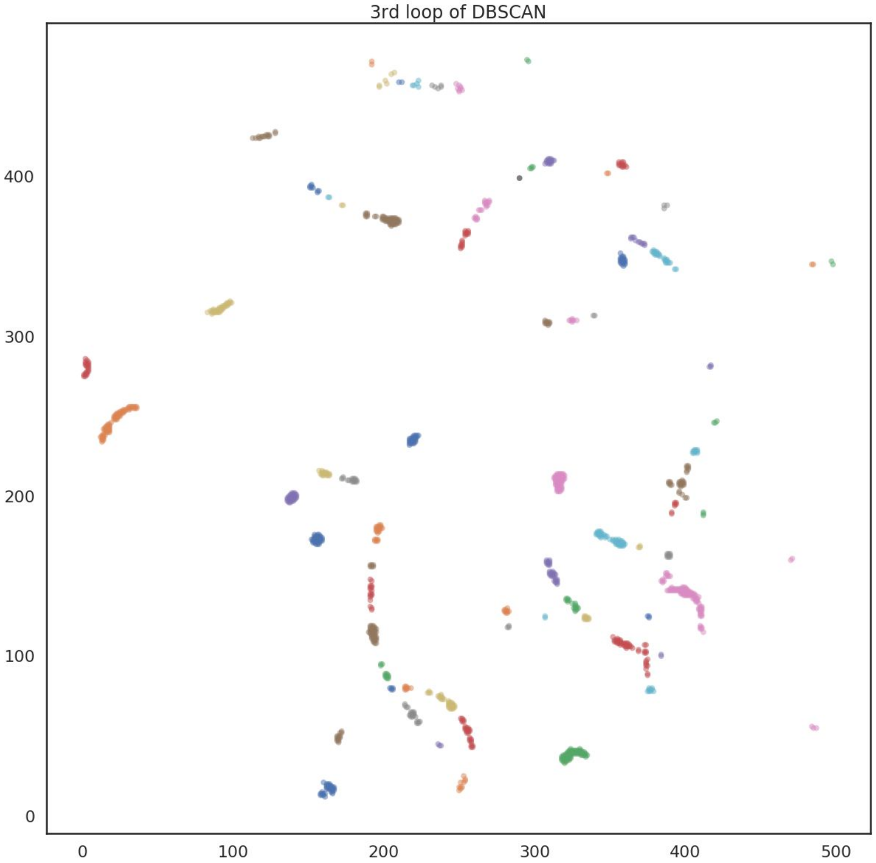


Iterative DBSCAN Method - Step Three

The last one is more flexible and the goal here is find the signals that aren't found yet.

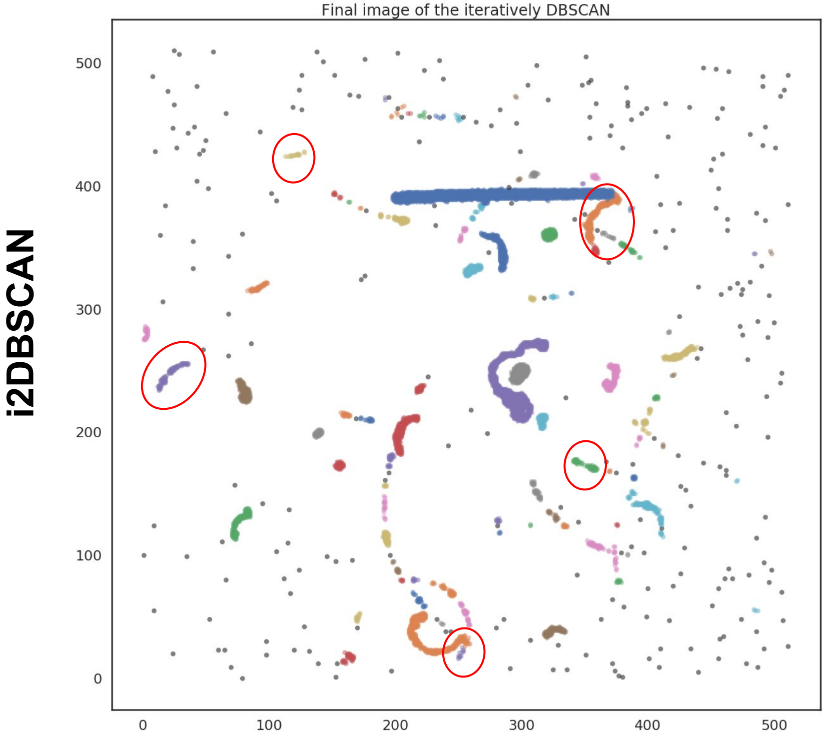
In this case the label is 'small' and the output of all steps is save for further analysis.

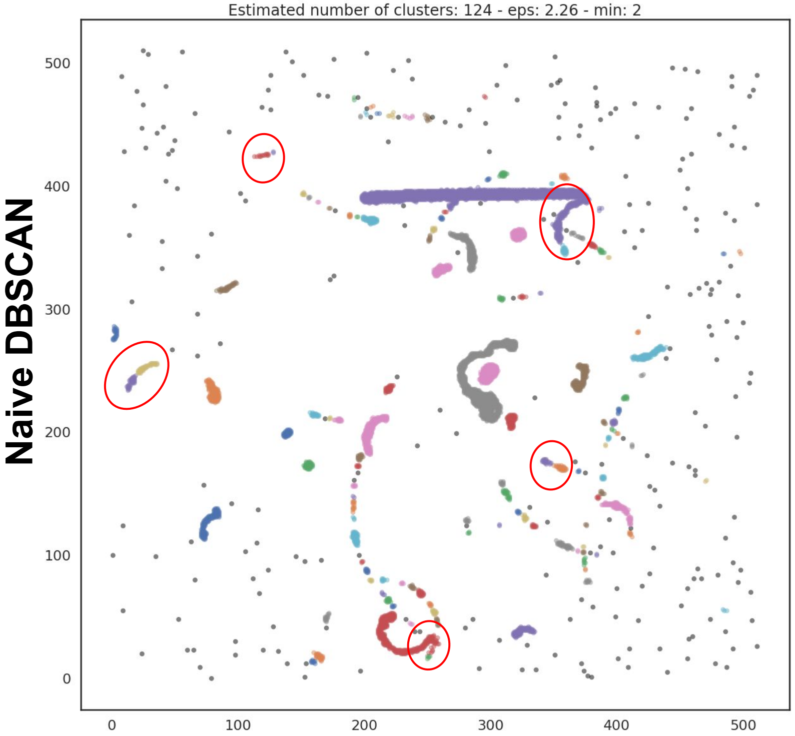
(In the image at right different colors means different clusters and the black ones represents the not found clusters)



i2DBSCAN x DBSCAN

So, the difference between the two methods are not so big by looking at one image. However, it can be a improvement considering a hundred of events.

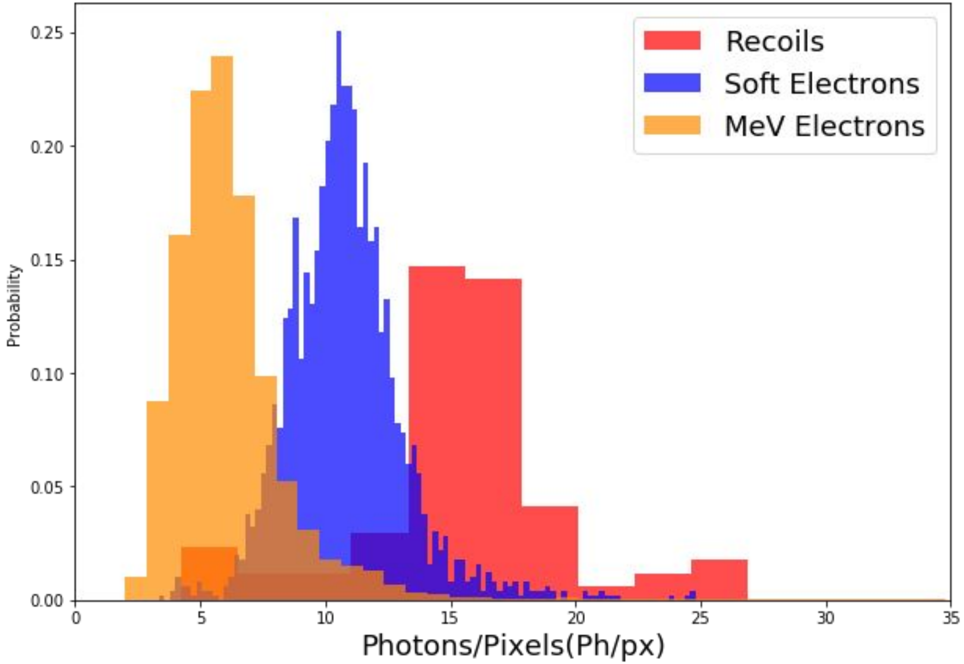




Preliminary Results

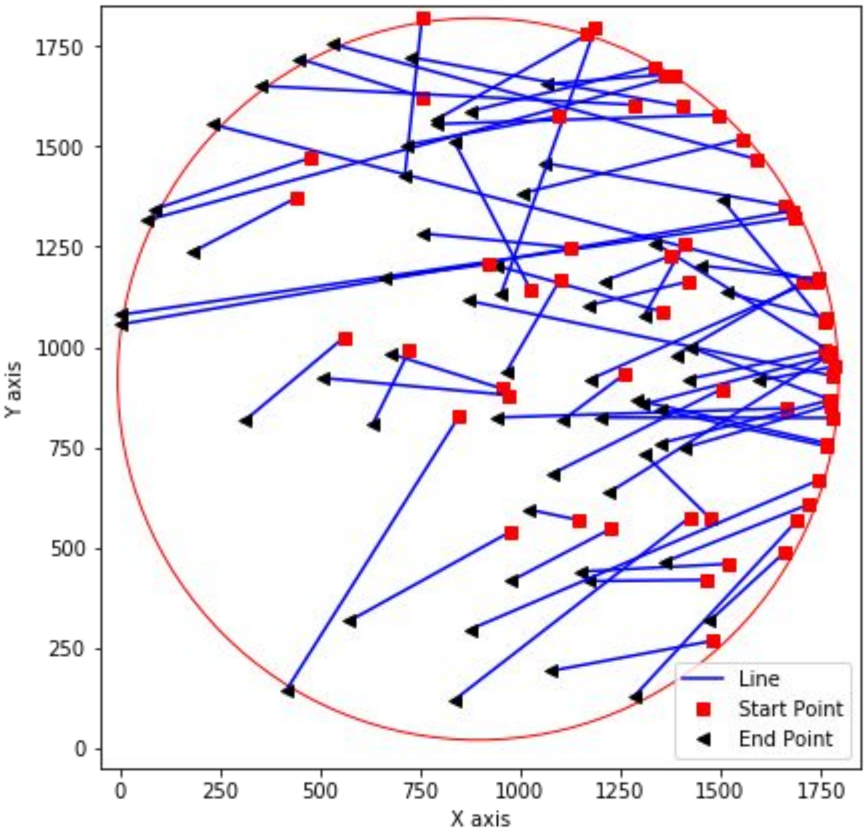
Very preliminary analysis

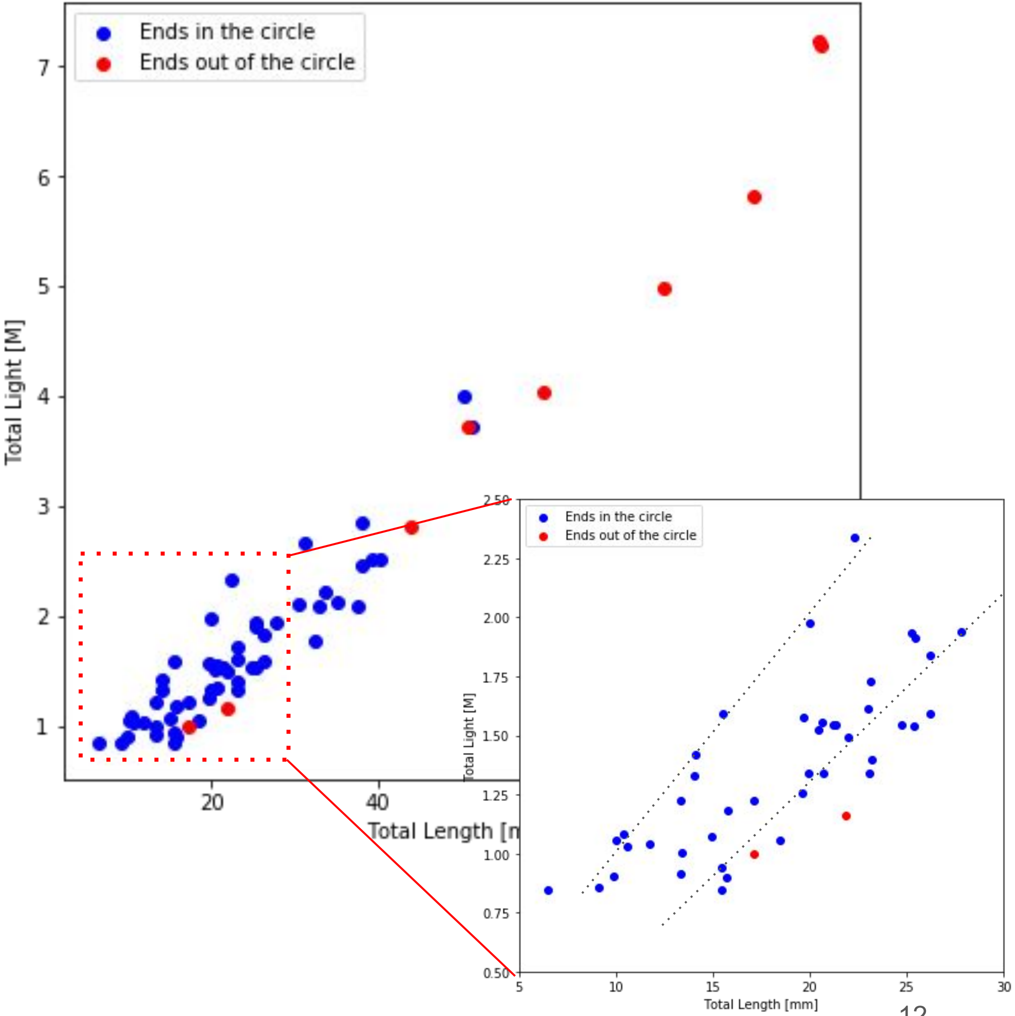
If we consider that the first iteration only will find 'Recoils', the second one 'Soft Electrons' and the third 'MeV Electrons', we can construct the following graphs:



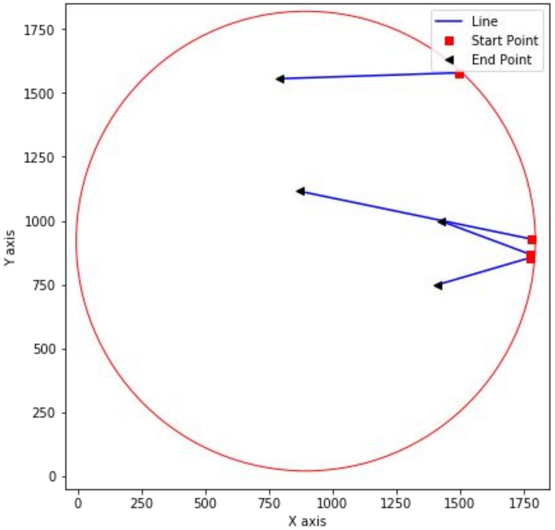
- What is the expected behaviour of protons and alpha particles with an energy about few hundred keV?
- And the length?
- Where are we in the Bragg Curve?

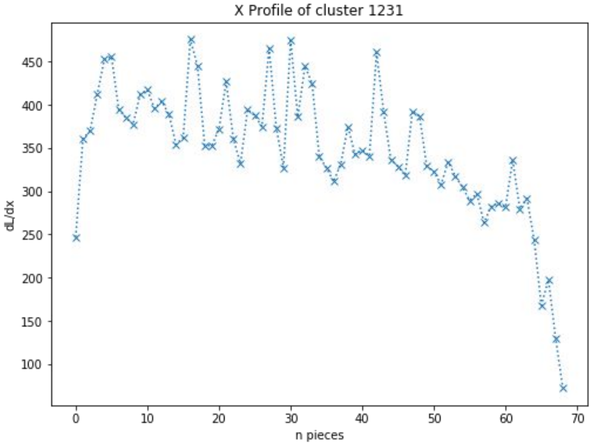
Looking only at the ‘Recoils’

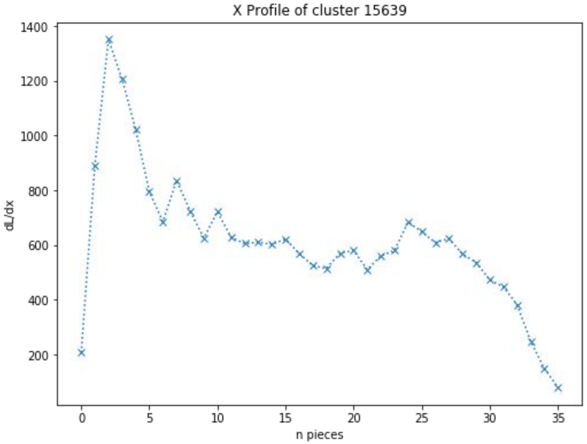


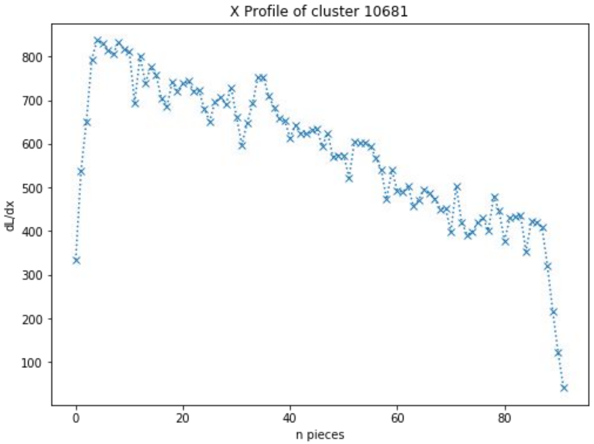


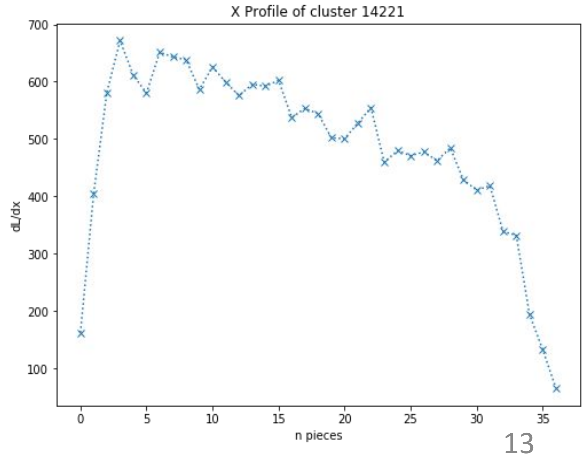
dL/dx Profile - few examples



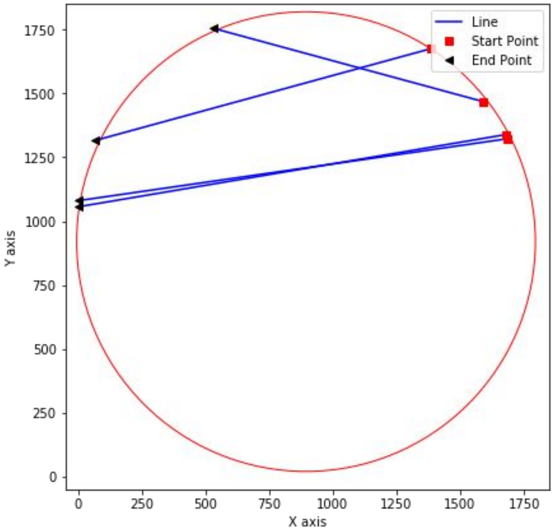

X Profile of cluster 1231


X Profile of cluster 15639

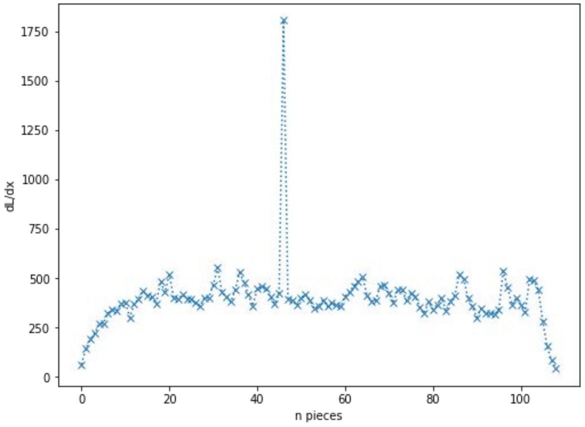

X Profile of cluster 10681


X Profile of cluster 14221

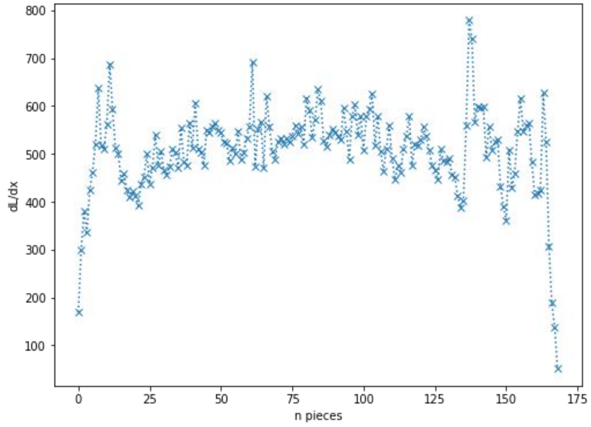
dL/dx Profile - few examples of long tracks



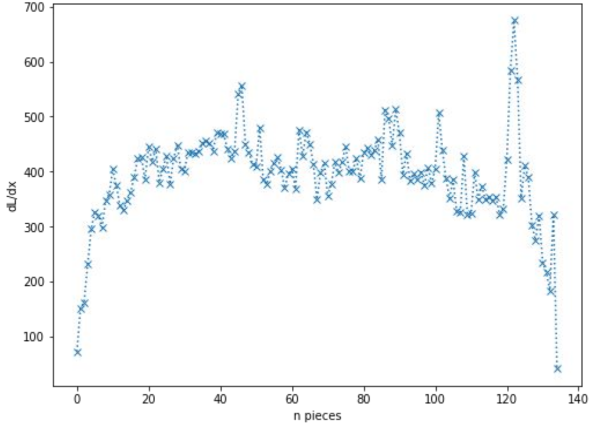
X Profile of cluster 13883



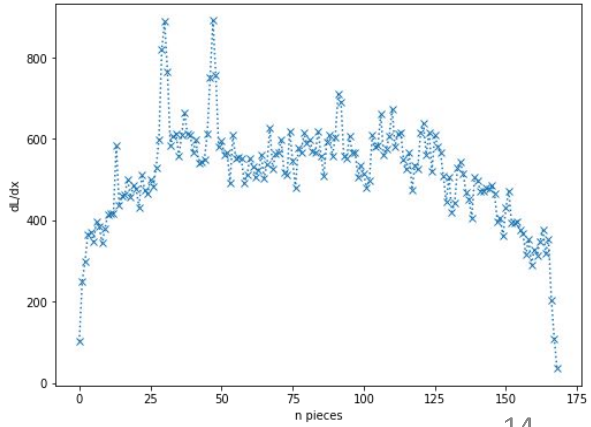
X Profile of cluster 8679



X Profile of cluster 5825



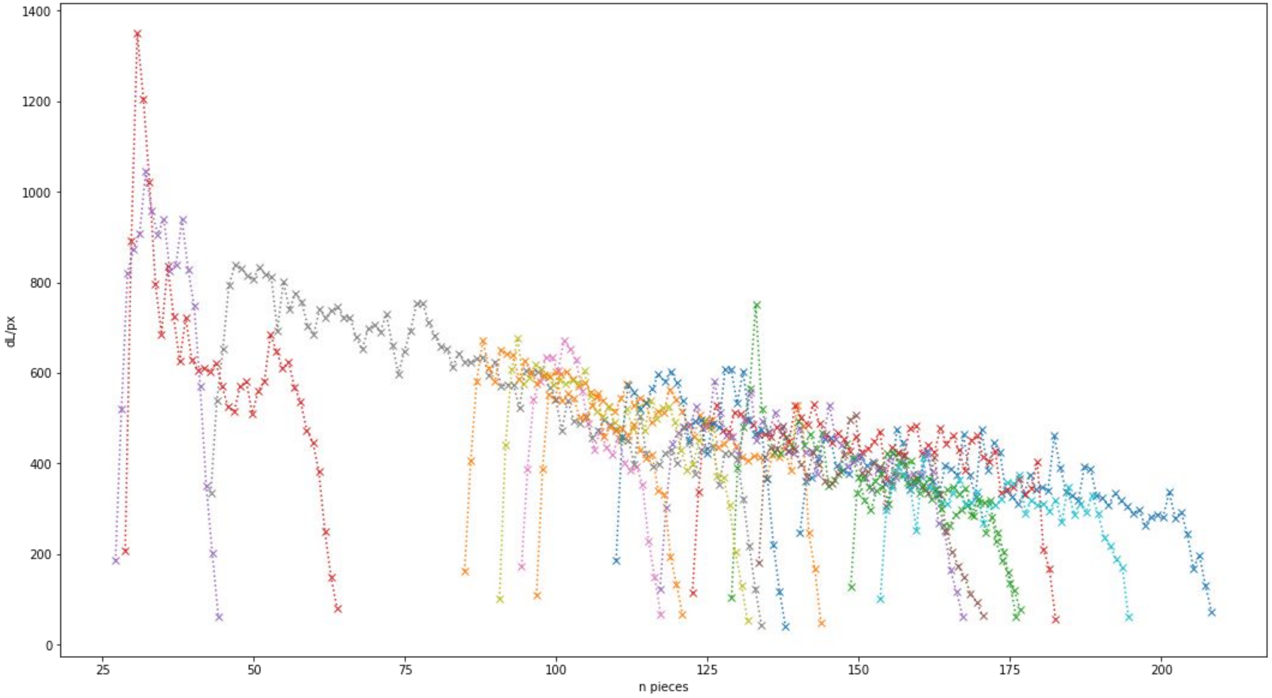
X Profile of cluster 5935



These tracks are crossing the entire detector and we was expecting a lower value of dL/dx to fit then in the beginning of the Bragg Curve.

Aligning the profiles lookin at the dL/dx

Looking at the profiles we couldn’t see a **bragg peak**, so our thought was to try to align the different profiles, by assuming that each profile was showing a different part of the **bragg curve** to see if we could ‘produce’ the entire curve.



Summary

- Developed the *i2DBSCAN* algorithm:
  - Create a table with the necessary information;
  - The algorithm is taking 8-10 minutes to run over 300 images and save the **table**.
- We chose to not generate all the variables in the same time that we are doing the clusterization.
- But the ‘table’ has all the information needed to create the variables.

- **On Going:**
  - We are focus on characterizing first the ‘long’ tracks;
  - Create the needed variables;
  - And then, go to the classification task.