UNIVERSIDADE FEDERAL DE JUIZ DE FORA

PPEE
Programa de Pós-Graduação em Engenharia Elétrica UFJF

Directional-iDBSCAN

a proposal to CYGNO

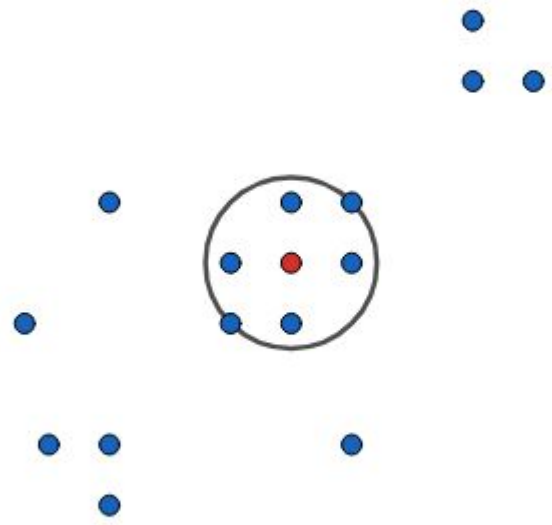
Igor Pains

Igor Abritta and Rafael A Nobrega

Motivations

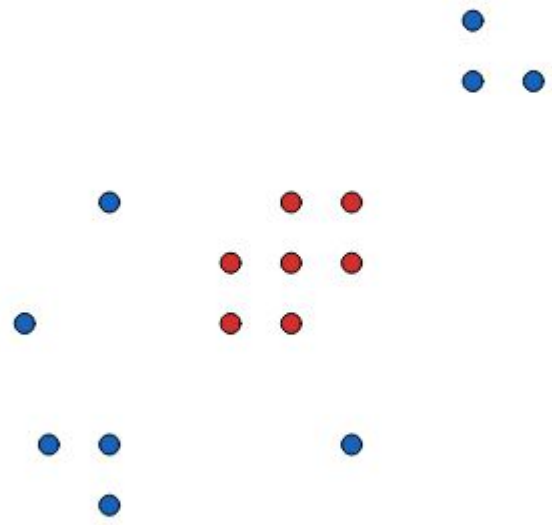
- The current CYGNO reconstruction algorithm is composed of iDBSCAN and the Supercluster. The last step was developed to optimize the clusterization process for long natural radioactive tracks;
- However, during the development of iDBSCAN there was always the idea of implementing some directional clusterization algorithm capable of following the long tracks;
- With this in mind, we studied and implemented the **Directional-iDBSCAN** (iDDBSCAN) algorithm;
- *In this presentation, we just want to present the idea in order to get a feedback on whether it's worth keep going with this work.*
    - *If yes, the next step would be obtain quantitative results (e.g. using simulation data).*

Directional-iDBSCAN algorithm



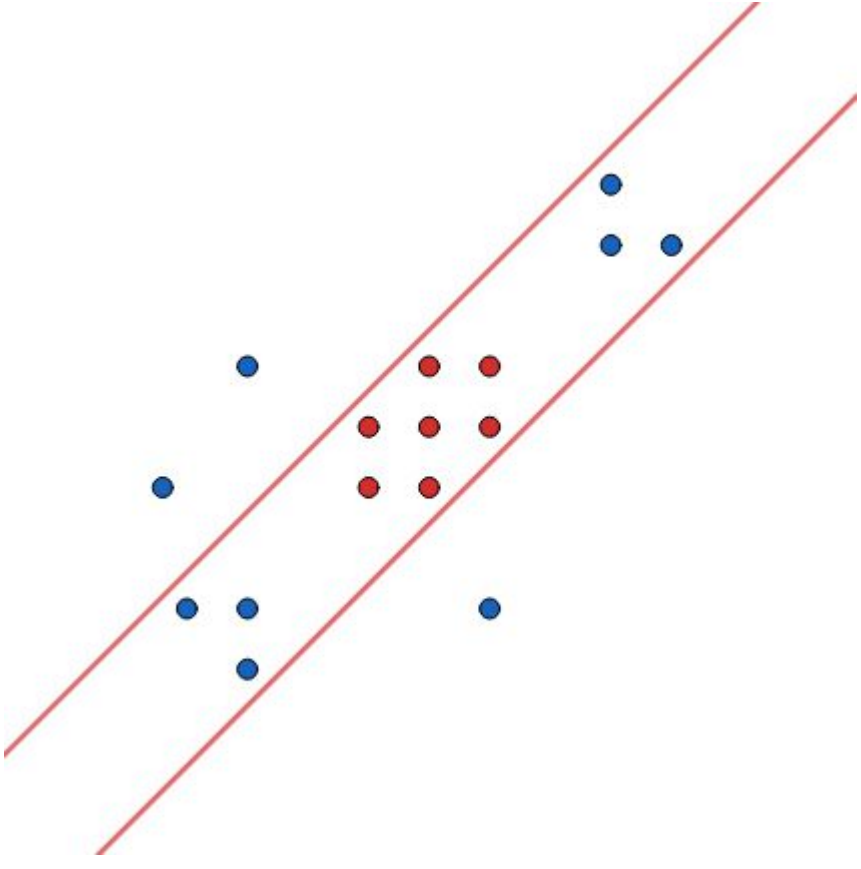
- The algorithm works as the iDBSCAN, except for the extra step of directional search. The flow is as it follows:
  - Cluster search begins at a high density point (core point)

Directional-iDBSCAN algorithm



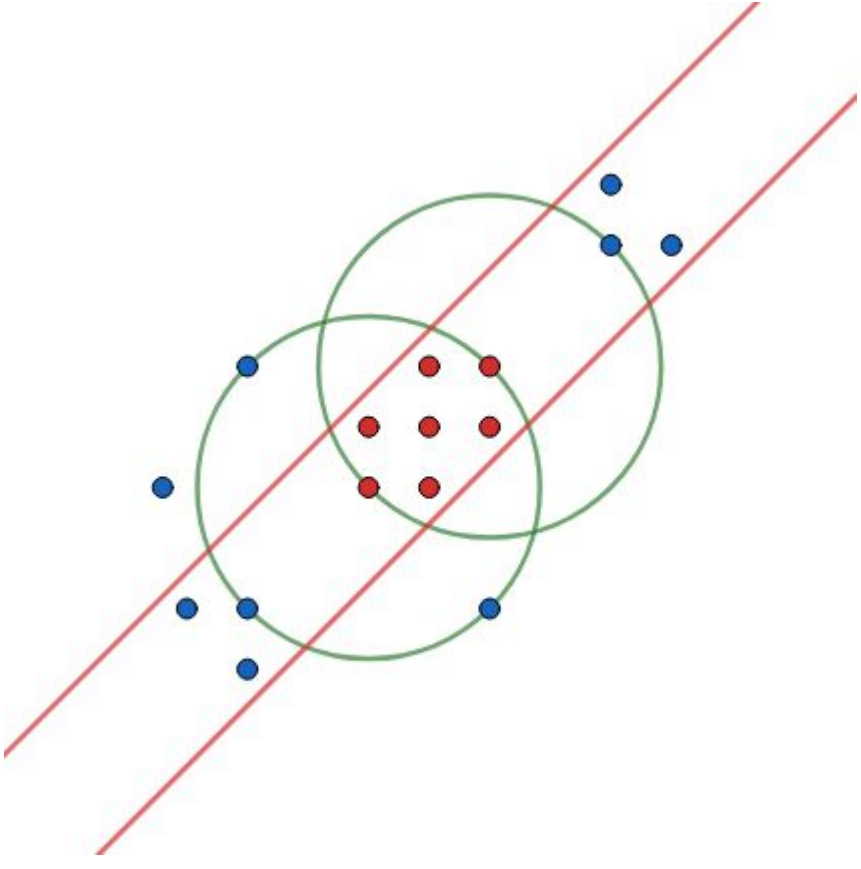
- The algorithm works as the iDBSCAN, except for the extra step of directional search. The flow is as it follows:
  - Cluster search begins at a high density point (core point)
  - Find a cluster

Directional-iDBSCAN algorithm



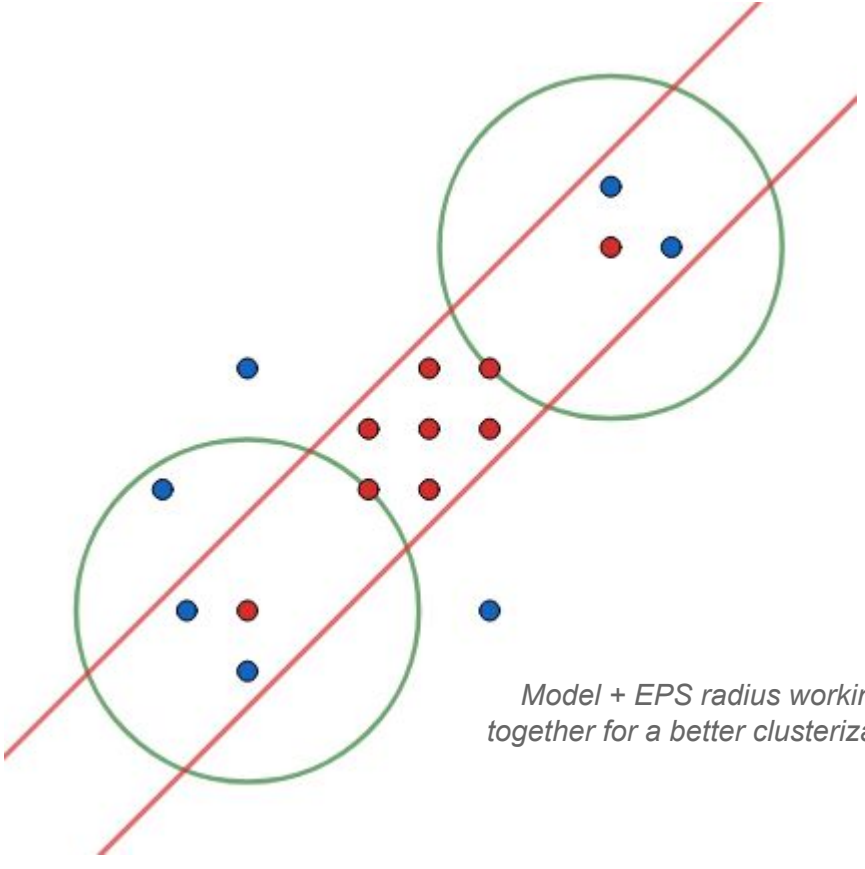
- The algorithm works as the iDBSCAN, except for the extra step of directional search. The flow is as it follows:
  - Cluster search begins at a high density point (core point)
  - Find a cluster
  - Find a polynomial model for that cluster

Directional-iDBSCAN algorithm



- The algorithm works as the iDBSCAN, except for the extra step of directional search. The flow is as it follows:
  - Cluster search begins at a high density point (core point)
  - Find a cluster
  - Find a polynomial model for that cluster
  - Find new points that are fitted in that model and are in the cluster neighborhood not found in the first interaction

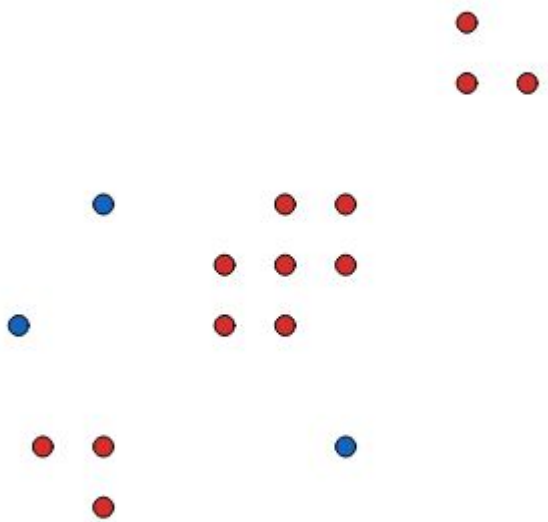
Directional-iDBSCAN algorithm



Model + EPS radius working together for a better clusterization

- The algorithm works as the iDBSCAN, except for the extra step of directional search. The flow is as it follows:
  - Cluster search begins at a high density point (core point)
  - Find a cluster
  - Find a polynomial model for that cluster
  - Find new points that are fitted in that model and are in the cluster neighborhood not found in the first interaction

Directional-iDBSCAN algorithm



- The algorithm works as the iDBSCAN, except for the extra step of directional search. The flow is as it follows:
  - Cluster search begins at a high density point (core point)
  - Find a cluster
  - Find a polynomial model for that cluster
  - Find new points that are fitted in that model and are in the cluster neighborhood not found in the first interaction

Directional-iDBSCAN algorithm

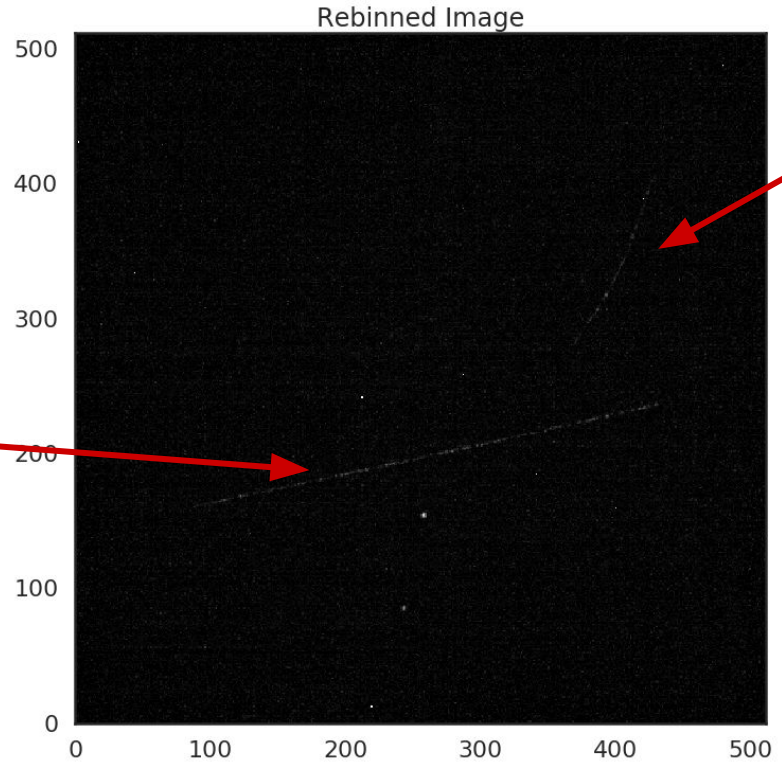
- The directional search is done by a linear regression algorithm named RANSAC[1] with a polynomial equation as its base estimator.

- The model is updated whenever new points are added to the cluster, so it will not be outdated.

- Since the directional search is only enabled when a good RANSAC model is found, the result in datasets without long tracks tends to be the same as for the iDBSCAN's.

Results

- The dataset used to develop and test the algorithm was the Run 2065*, with a lot of long tracks.
- This is the event 15 of the Run 2065, used to present the operation of the D-iDBSCAN.



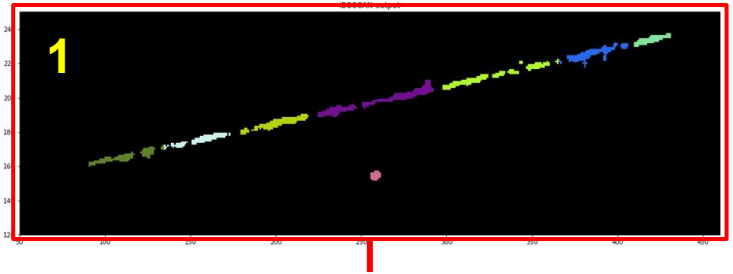
Example of curved track

Example of straight long track

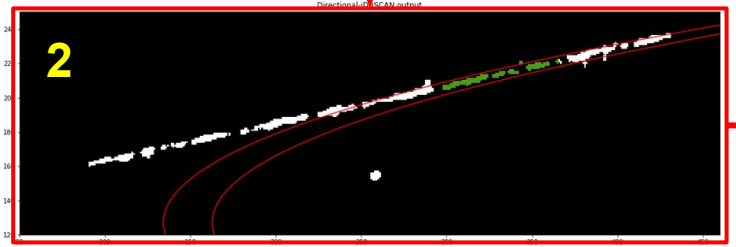
** LEMOn detector with Pb case and 55Fe source.*

Example of the D-iDBSCAN

Output of the D-iDBSCAN first step



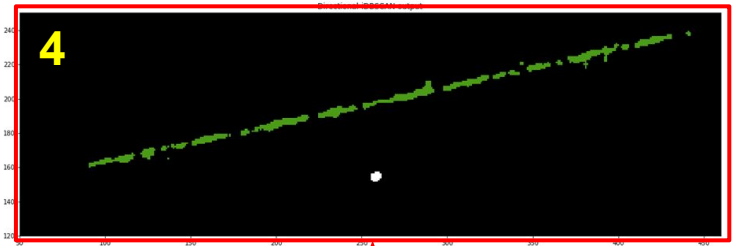
Directionality begins at the highest density cluster and find a first fit model using just that cluster.



*Note that the probability of a misclassification of a Fe55 near long tracks is low due to the thin fit.

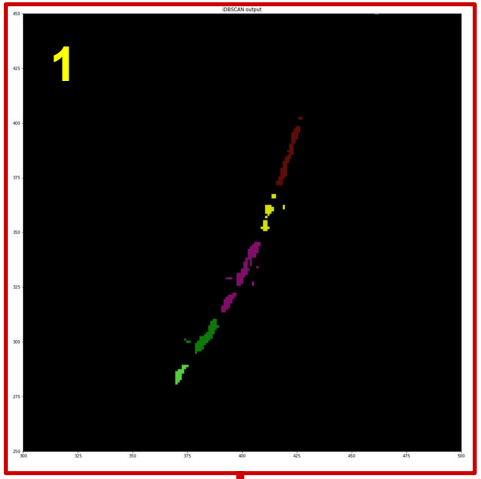
Then, search for new points using the fit, improve the fit with the new information and repeat until nothing else is found.

Final cluster



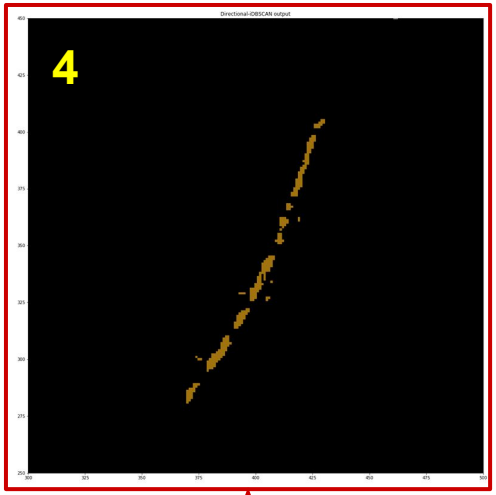
Example of the D-iDBSCAN

Output of the D-iDBSCAN first step



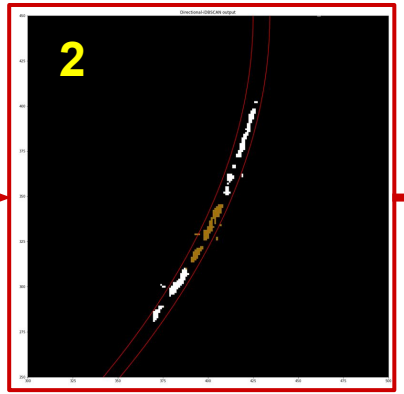
Directionality begins at the highest density cluster and find a first fit model using just that cluster.

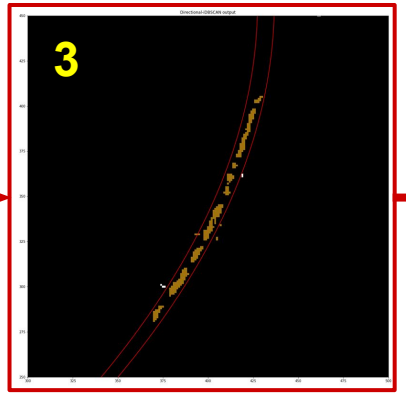
Then, search for new points using the fit, improve the fit with the new information and repeat.



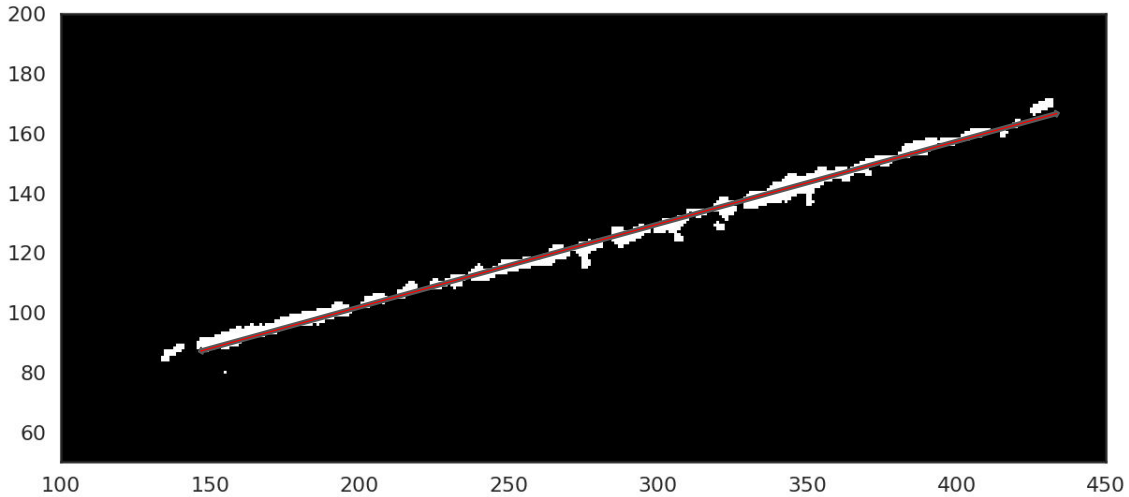
Final cluster with all pieces together

Splitted in 4 pieces





Parameters of the track



$p(x) = 0.00016x^2 + 0.19x + 58$

These are the parameters of the second degree polynomial fit by RANSAC that maybe can be used to extract information about the tracks.

Conclusion

- So far, the results with the Directional-iDBSCAN were qualitative due to the absence of a truth dataset. But we can still conclude the following:
  - This algorithm is highly recommended to images with faint long tracks, because event after the filtering step, those tracks can still be recognized.
  - Even in images without that characteristic, it can also be used because the result will be the same of the current iDBSCAN.

- Next step would be to get quantitative results by analyzing simulated images.

Next steps

- These will be the next steps adopted:
  - Analyze some more images to have a visual feeling about the iDDBSCAN behavior in some particular cases, as overlapped tracks.
    - Analyzed runs: 2097 and 2098 (“Dollar Trillogy” - 60/40 mixture); 2317, 2318 and 2320 (70/30 mixture).
  - See if we can use the directional-DBSCAN as a long track background removing.
  - Analyse the energy distributions using iDDBSCAN, iDBSCAN and Super-iDBSCAN for No-source (NRAD), ER(Fe55) and NR(AmBe).