Status of new pixel clustering

Cecile Lapoire (University of Bonn), Giacinto Piacquadio (CERN), Andreas Salzburger (CERN)
On behalf of the pixel clusterization TF group
Software and computing workshop
CERN
05/04/11

Outline

- Motivation
- Clustering algorithms
- NN-based algorithm :
  - Output
  - Impact on b-tagging performance
- Conclusion

Motivation

- ATLAS current clustering algorithm not optimized for tracking in high p_T jets (dense environment → merged clustered (shared hits) → low efficiency in the core of the jets)

- Pixel Clusterization TF set up to improve clustering algorithm and tracking in these jets

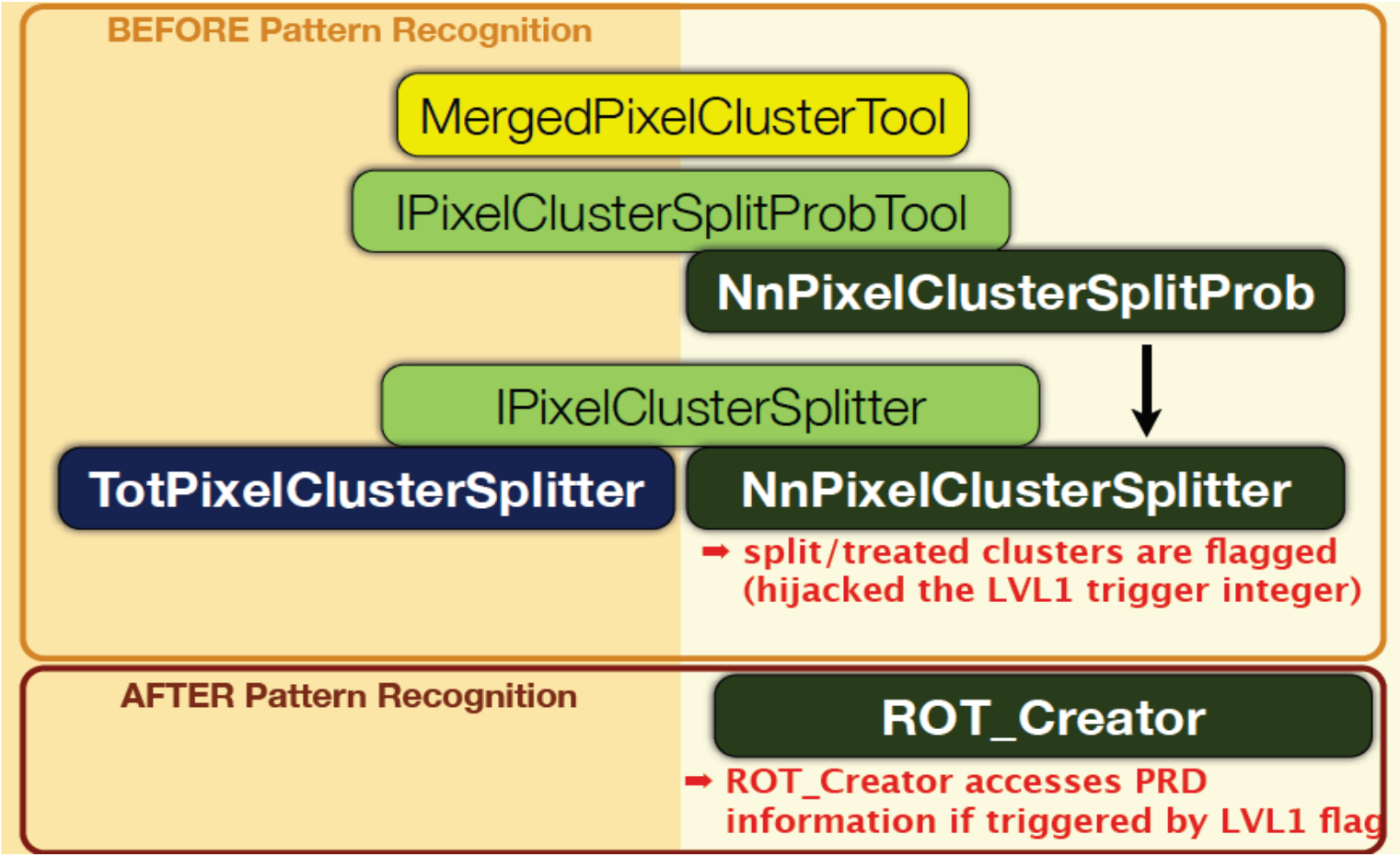
- Huge impact on b-tagging performance at high p_T (useful for H→ bb, SUSY, Z'→bb,..)

Andi

Clustering algorithms

2 new clustering algorithms :

- **NnSplitter** : NN trained with truth input, corrects single clusters & splits
- **TotSplitter**: based on charge information (ToT), searches for split signatures (minimum ToT between 2 maxima) & splits



NN-based algorithm

Giacinto

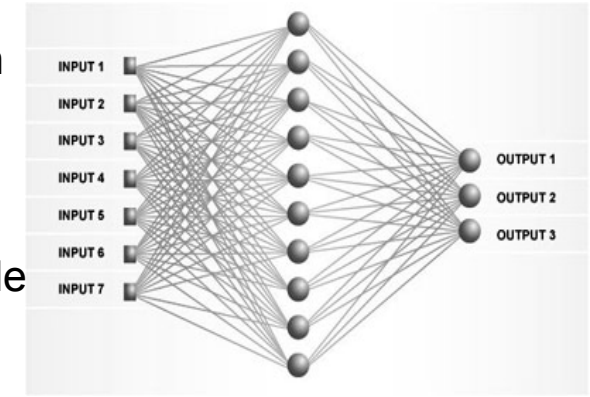
INPUT nodes

7x7 pixel matrix with calibrated charges

vector of z pitches
(normal or long pixel)

track incidence angle
(after pattern recognition)

layer number



OUTPUT nodes

How many subclusters?

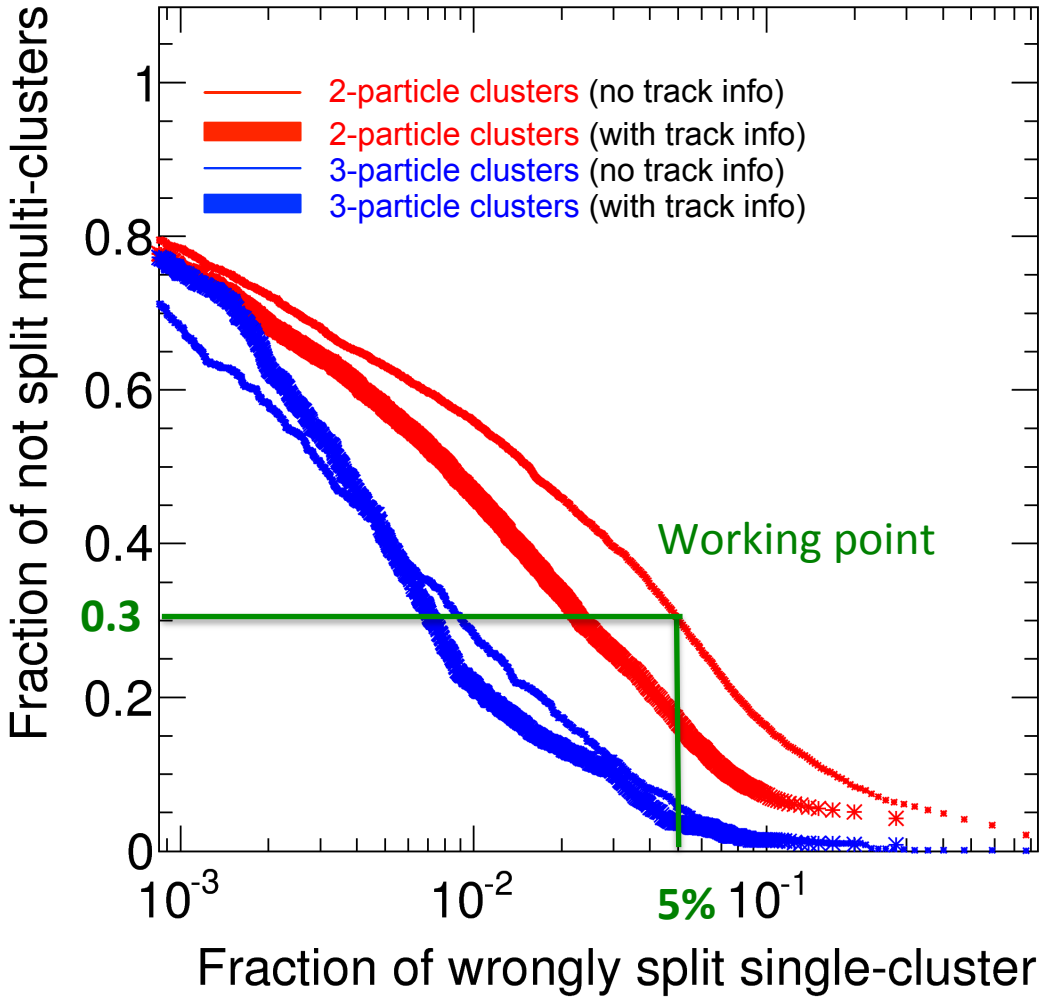
What is their position?

What is their associated distribution? → (error)

- Applied **before** and **after pattern recognition** (different training)
- Output : Number of sub-clusters with associated probability
  → Performance according to cut (working point, next slide)
  → Associated position (slide 7)
- After splitting, 1 given pixel is no more associated to 1 given cluster

NN based alg. : performance

Giacinto



- 1-particle vs. 2-particle case :
- 5% wrongly split 1-particle :
  - 15 % not split 2-particle (with track info = after pattern recognition)
  - 30 % not split 2-particle (without track info = before pattern recognition)

ATLAS W/O IBL

NN based alg. : residuals

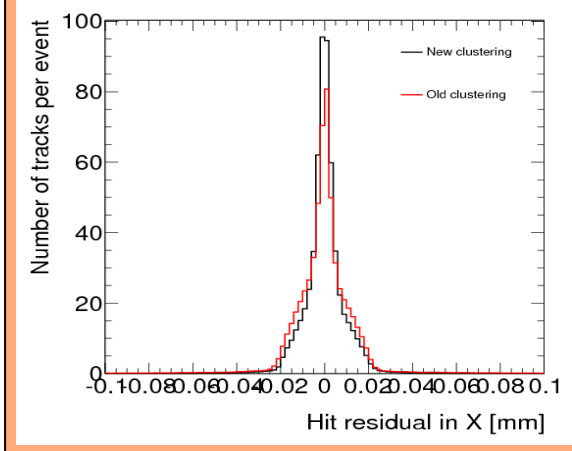
Giacinto

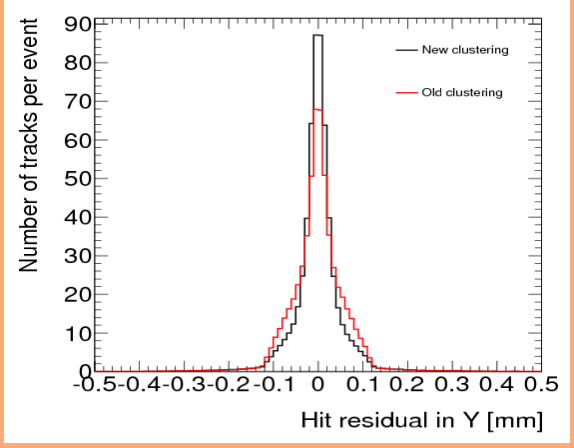
ATLAS + IBL

Distance in X and Y between track and cluster

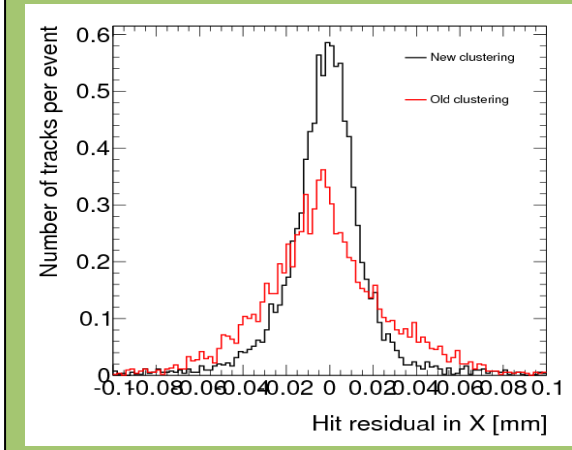
- Better position even for 1-particle clusters (better treatment of **delta-rays**)

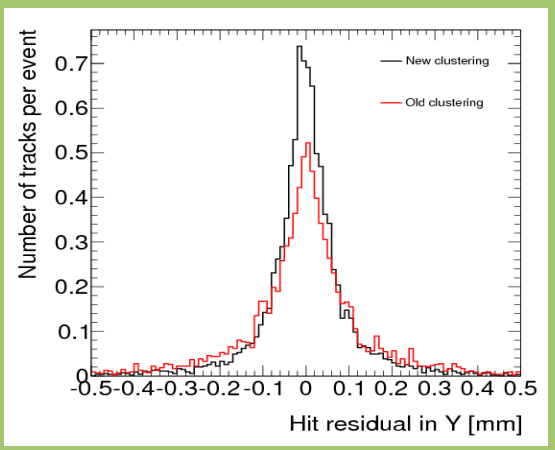
- Residuals of 1-particle clusters:





- Residuals of 2-particle clusters:





- Resolution **2 times better** in **X** and significantly better in Y

Tracking performance : resolution

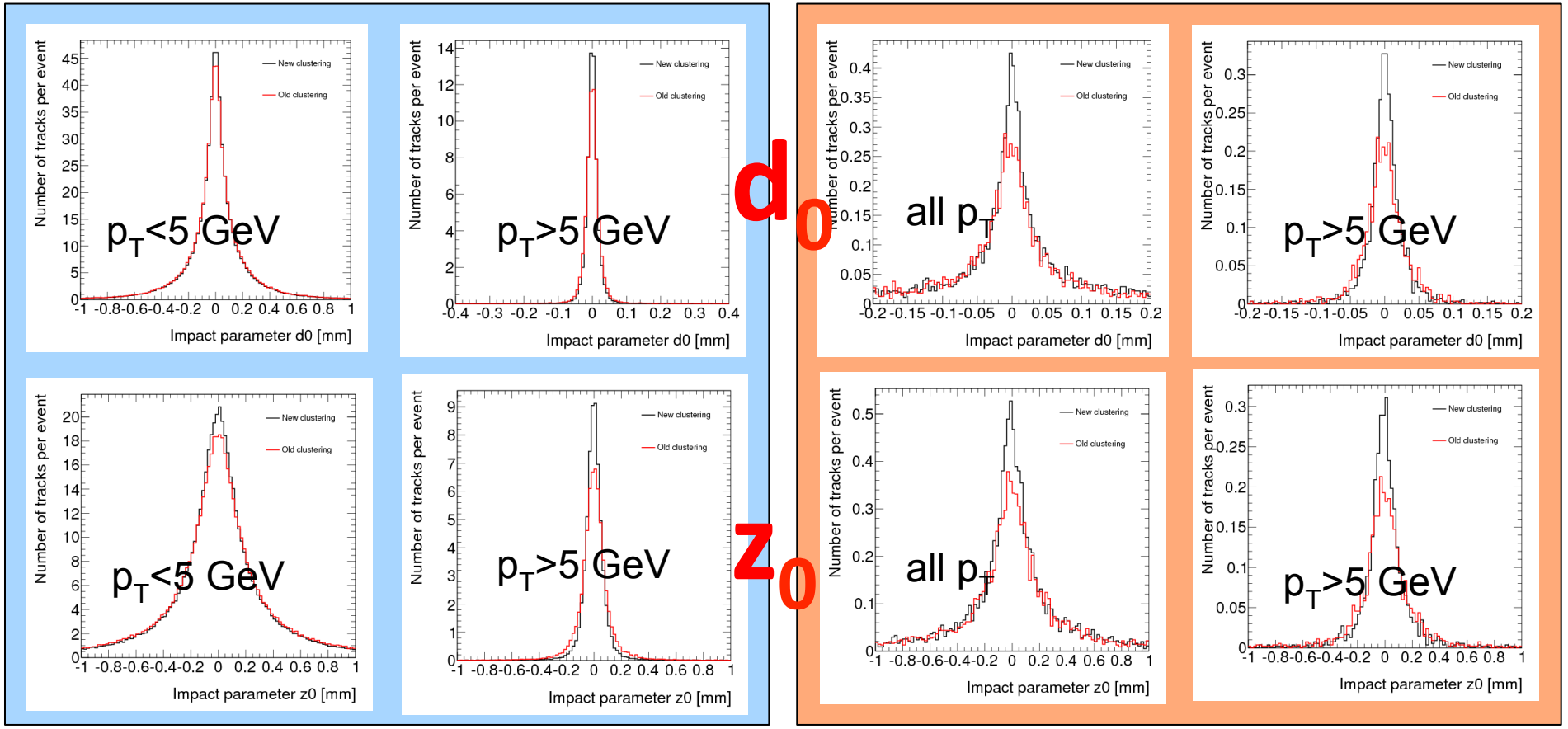
Giacinto

ATLAS + IBL

NN-based alg.

Track with isolated (“1-particle”) IBL cluster

Track with “2-particle” IBL cluster



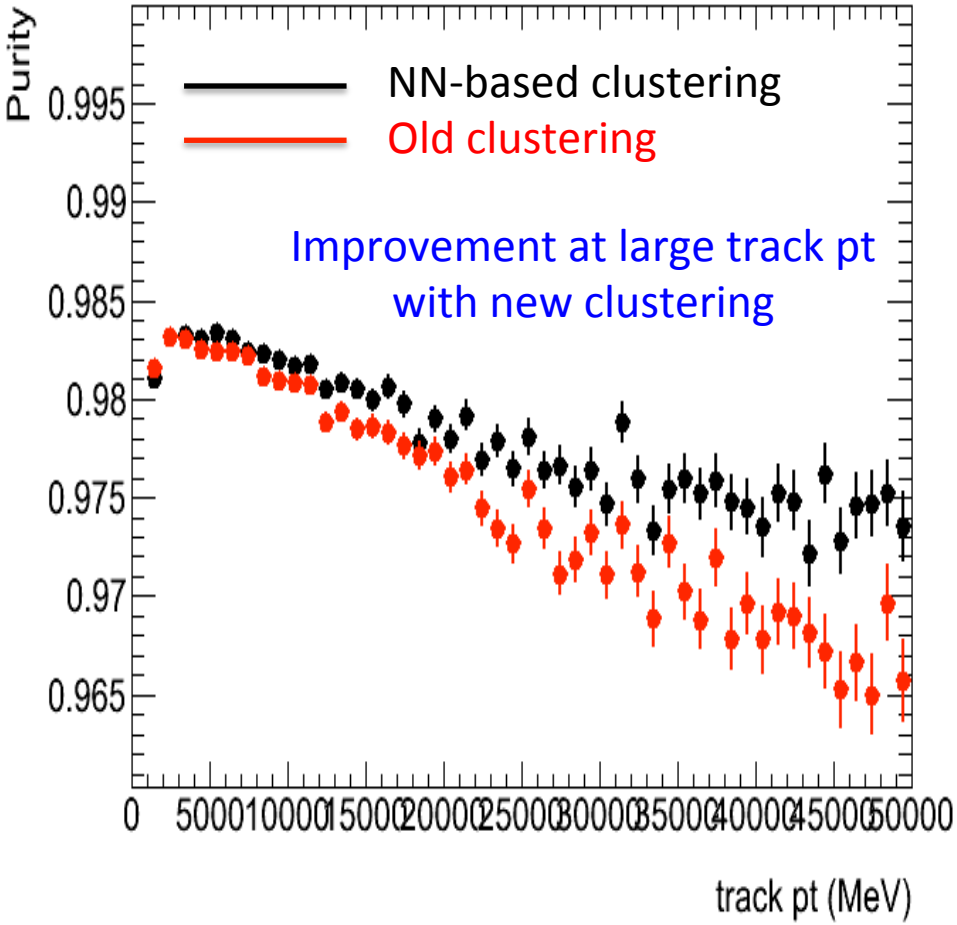
- **Improvement in “isolated” clusters** most likely due to better treatment of **delta rays**
- **Splitting merged clusters** clearly **improves** the **resolution** and the **reconstruction efficiency**

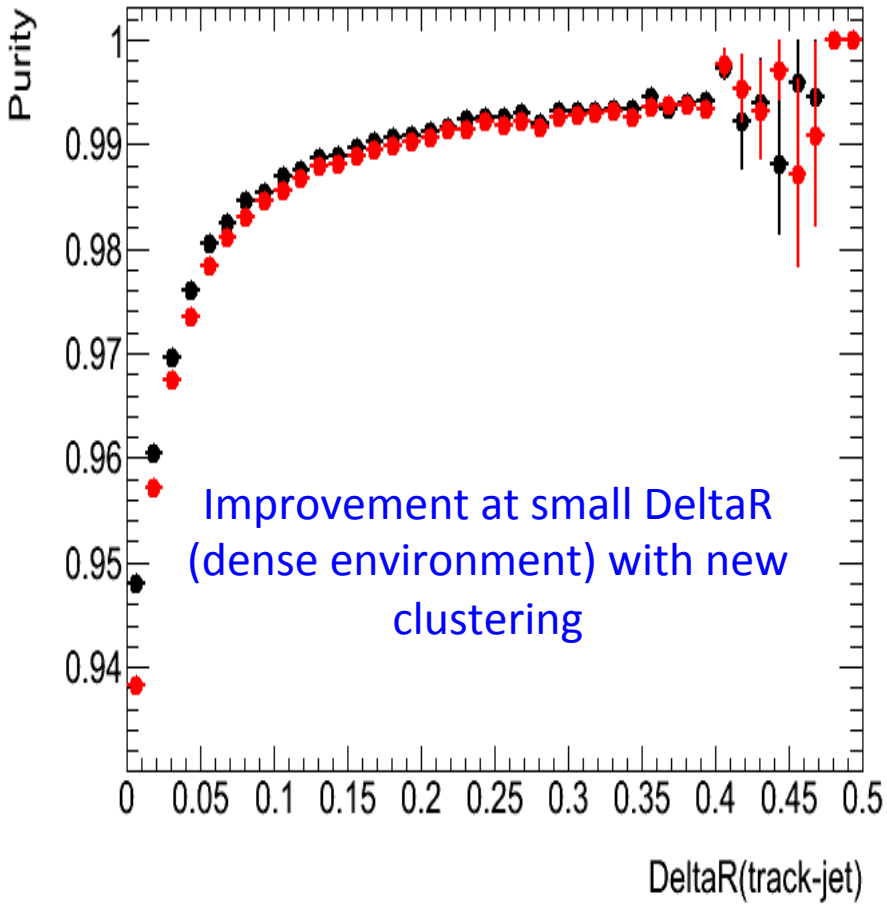
Tracking performance : purity

NN-based alg.

- Tracks passing IP b-tagging cuts in jets with p_T>30 GeV
- High p_T dijet events
- **Track in jets purity vs track p_T and DeltaR (track-jet)**
  - Purity = # reco tracks matching to true track / # reco tracks

ATLAS W/O IBL

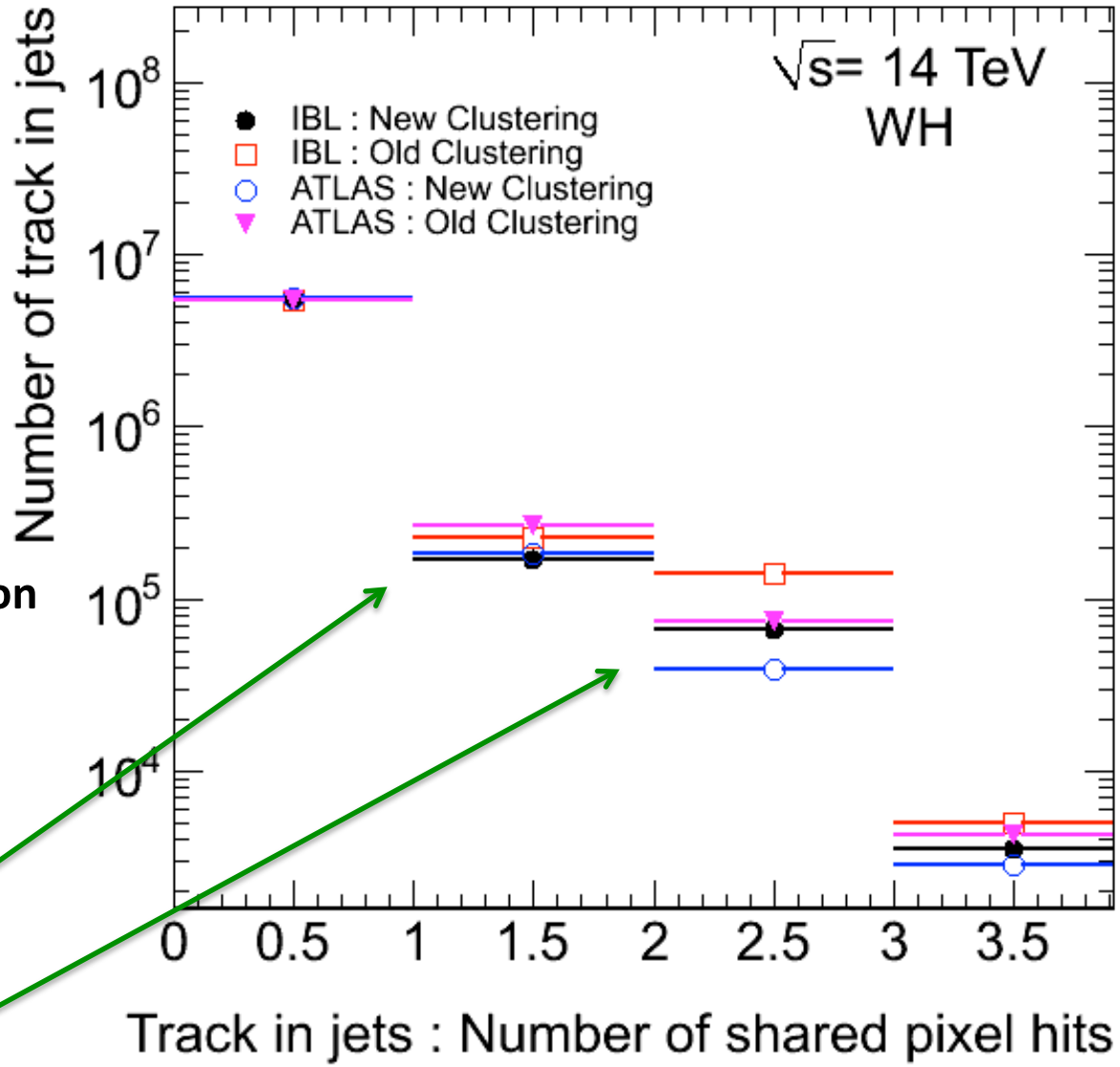




Shared pixel hits

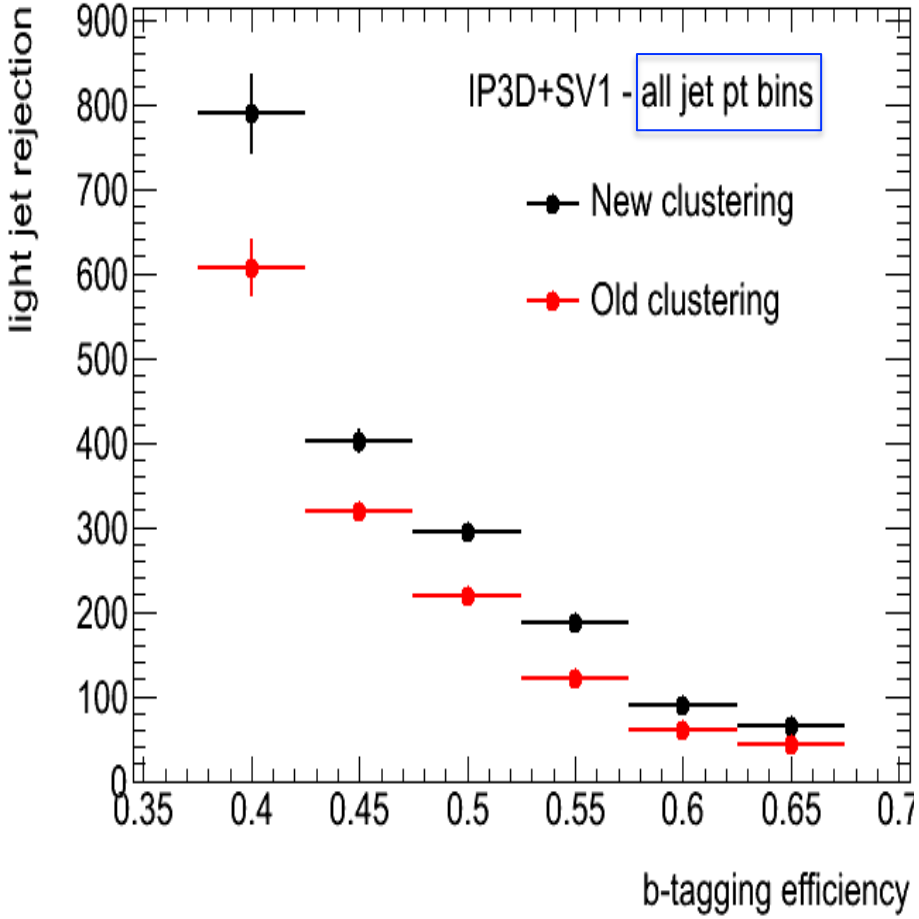
New Clustering = NN-based alg.

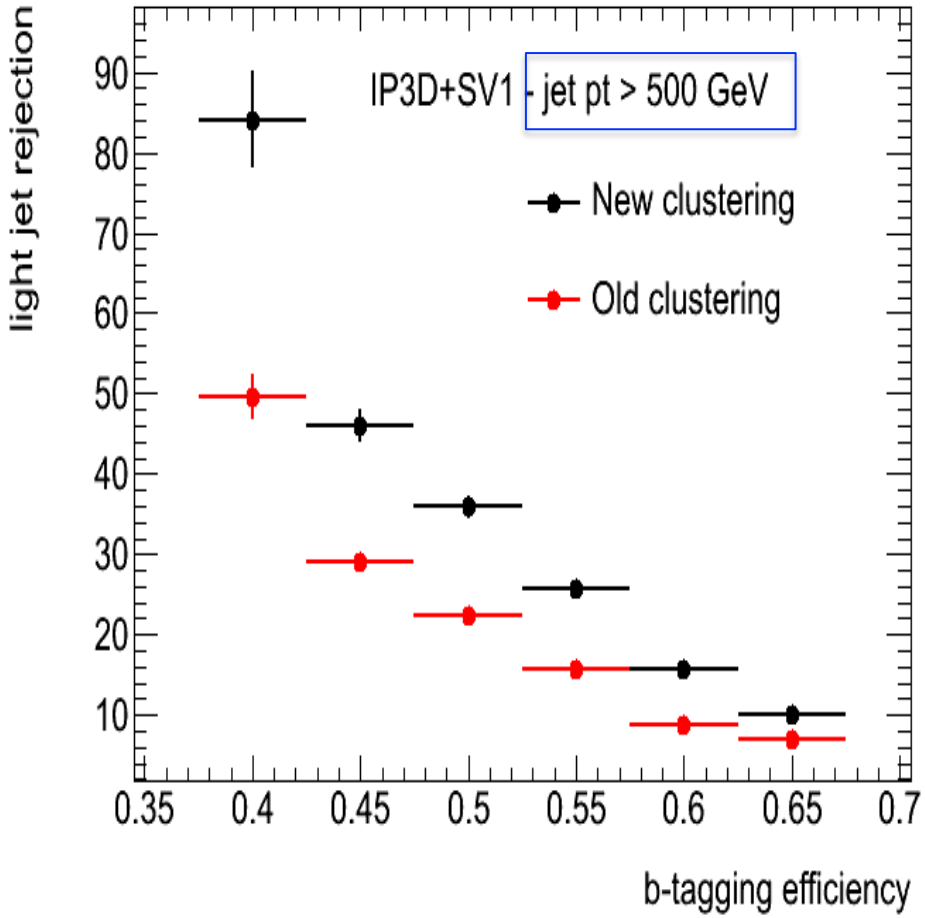
- **Number of shared pixel hits on tracks in jets**
- Without IBL, with NN clustering :
  - **~30% less tracks with 1 shared hits**
  - **~50% less tracks with 2 shared hits**



Impact on b-tagging perf.

ATLAS W/O IBL





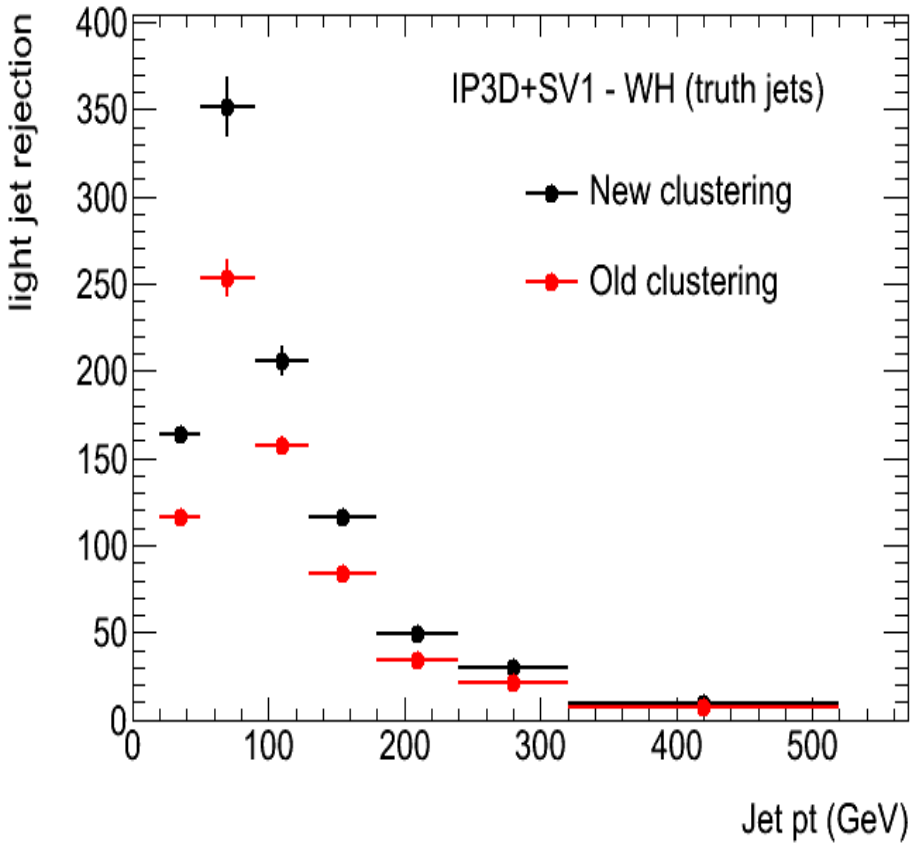
B-tagging efficiency : # of tagged true b-jets/# of true b-jets
Light jet rejection : # of true light jets/# of tagged true light jets

Perf. multiplied by ~1.5 for high p_T jets with NN-based clustering

Impact on b-tagging perf. (with IBL)

ATLAS + IBL

- **Light jet rejection vs jet p_T for a 70% b-tagging efficiency**
- **Using truth jet and IP3D+SV1 algorithm**



Performance improvement : **~40%**

Conclusion

Status

- **Significant improvement** of **cluster resolution**, **track reconstruction efficiency**, **b-tagging perf.** using the **NN-based clustering** (in both the ATLAS and ATLAS+IBL configuration)
- Another algorithm developed – performance ready to be studied

Release migration

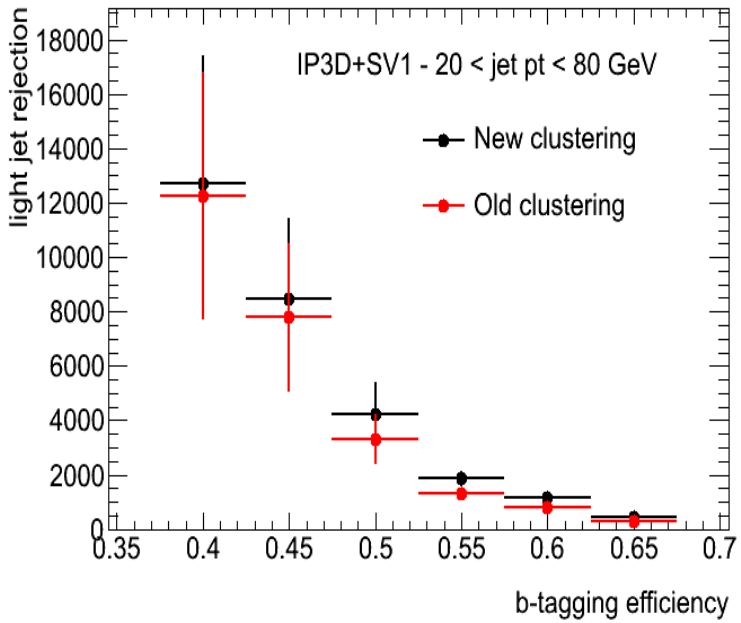
- TAGProd release **branched off from 16.6.0.1** used/in use for studies and grid deployment
- NnClusterSplitter fully integrated in TAGProd,16.6.0.1.1
  - 16.6.0.1.2 should have the first version of the TotClusterSplitter
- AmbiguitySolver being currently updated to handle these new cluster ambiguities
- Currently re-integrating the code into trunk to **devval to match the release 17.0.0 deadline** (MIG release requested)

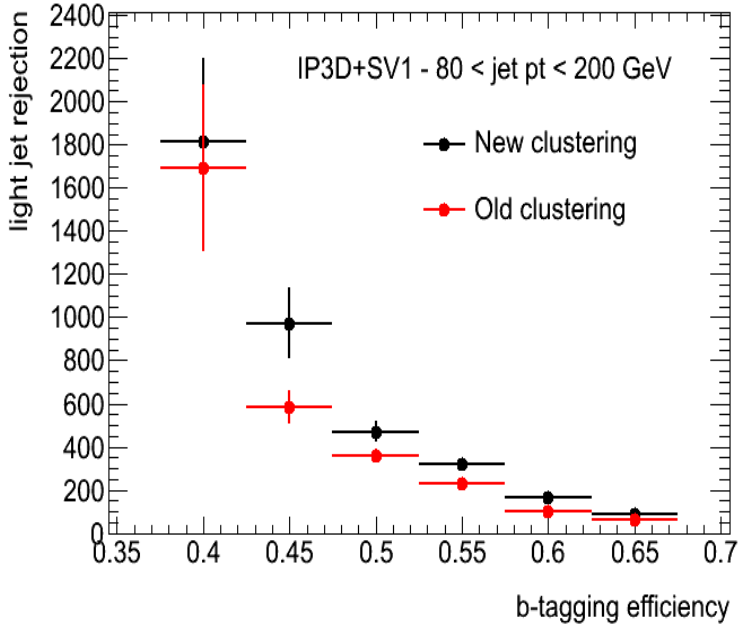
To-do

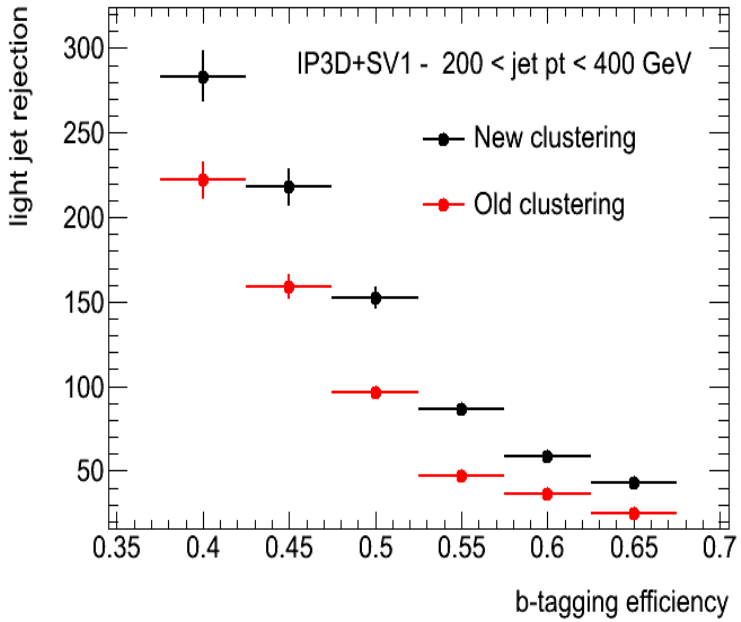
- Finish the AmbiguitySolver behavior in case of splitting
- Solve the ambiguities in truth association after splitting
- Refine implementation of NnClusterSplitter
- Test with data and specific selected events beam halo/beam splash/cosmics
- Write final report
- Full integration into 17.X.0 and correct binding to COOL for the input of the NnClusterSplitter and TotClusterSplitter

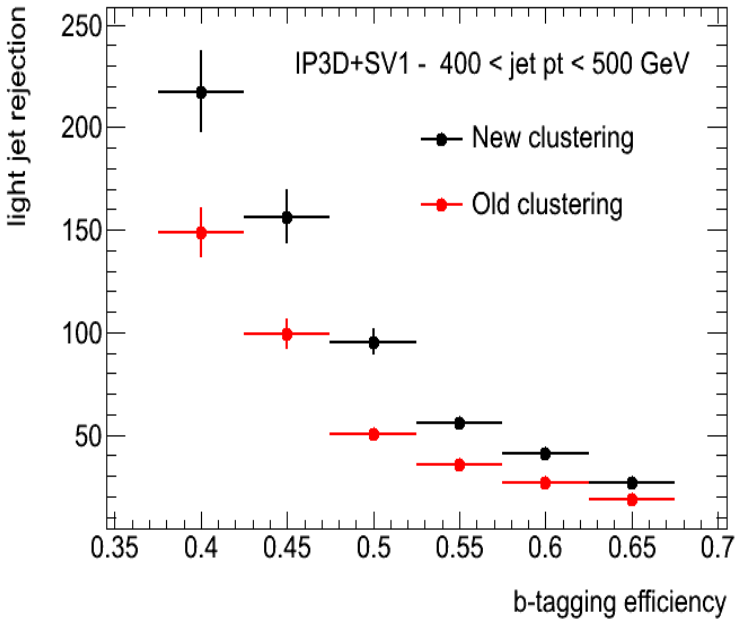
BACK-UP

Impact on b-tagging perf.



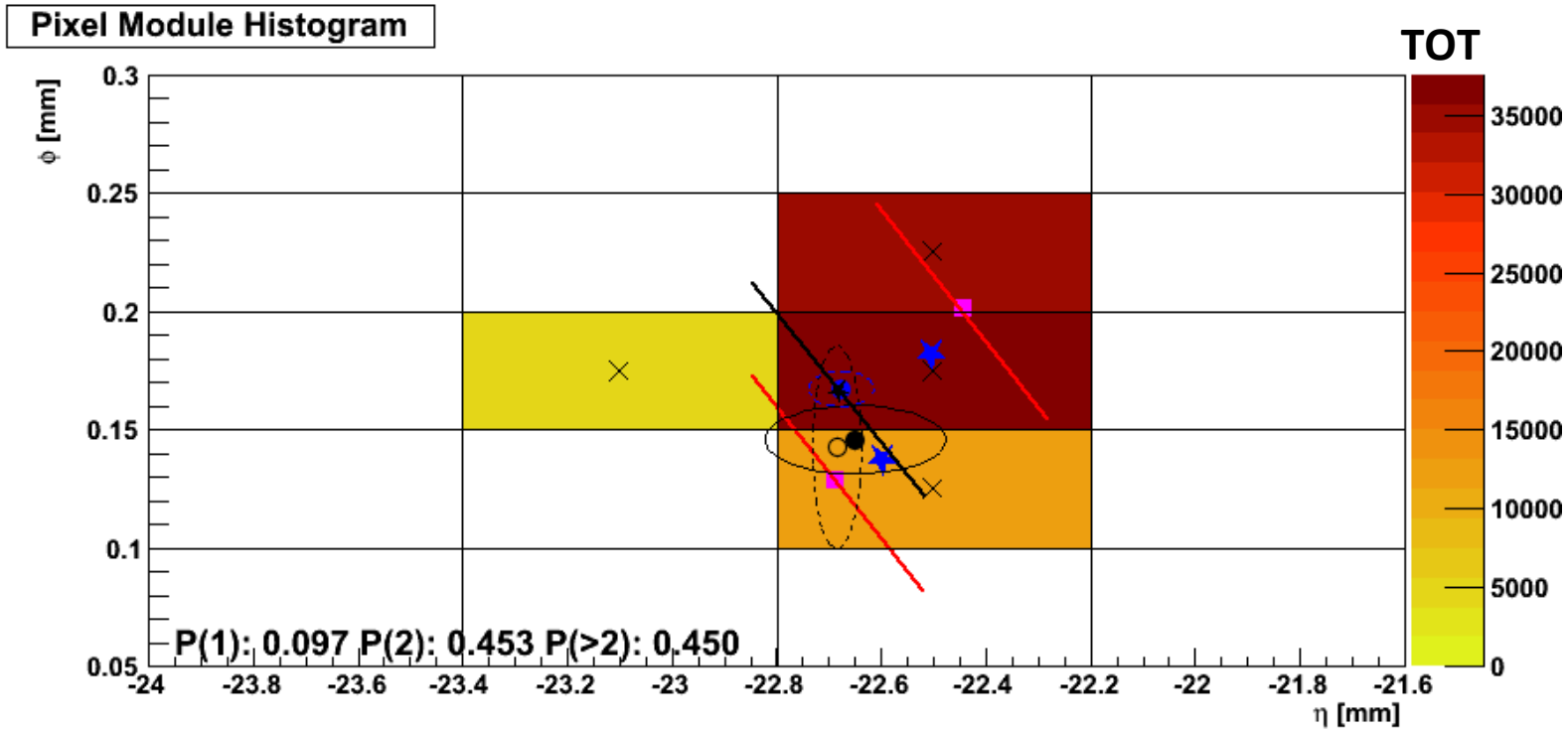






NN-based alg. : ToT map example

Giacinto & Andi



pred. truth impact (2 tracks)

pred. track impact (with path)

PRD with error

ROT with error

Reconstruction with old algorithm (1 cluster)

extrapolated track

Reconstruction with new algorithm (2 clusters)