

---

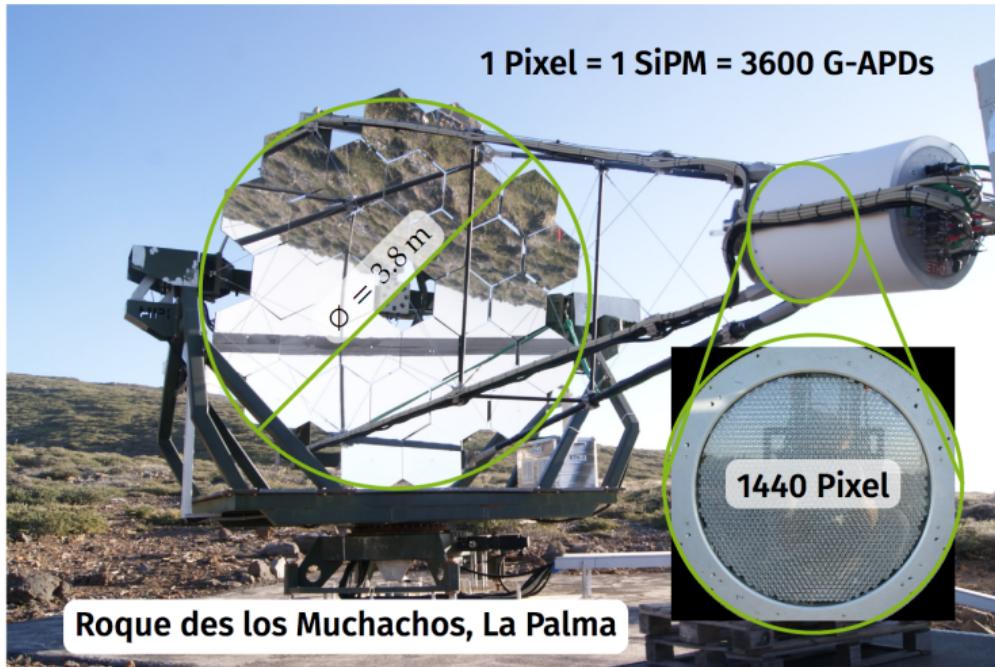
# Analysis Of The Crab Nebula Using FACT's Photon Stream Data

---

Kevin Sedlaczek for the FACT-Collaboration

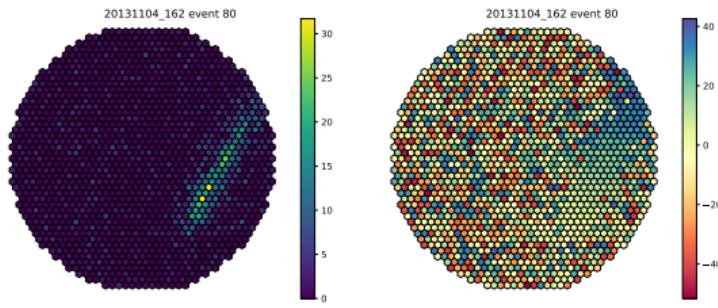
October 29, 2018

## The First G-APD Cherenkov Telescope



# data formats

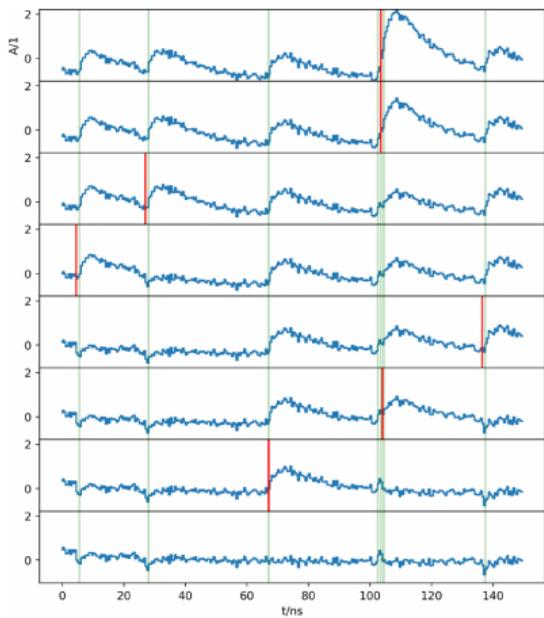
## Largest pulse representation



- FACT records data in format close to readout hardware
- superposition of multiple photon signals in time series
- data describing electric pulses instead of photons
- 2 attributes:
  - **c** charge integral / p.e.
  - **t** arrival times / ns
- shower: largest pulse along **t**-axis

→interpretation of sensor response instead of photons

## Photon Stream



- extracting single photons from time series of sensor responses
- list of arrival times of **single** photons / pixel
- 3 attributes:
  - $c_x$  first spatial coordinate
  - $c_y$  second spatial coordinate
  - $c_t$  arrival time
- shower: extraction from 3-dimensional objects
- decoupled from hardware

→interpretation of observables of single photons

## Photon Stream

$\left[ \begin{array}{c} [59, 84] \\ [93, 102, 103] \\ [58] \\ [65, 79, 97] \\ [] \\ [43, 68, 125] \\ [102] \\ [68, 100, 123] \\ [77, 88, 99, 111] \\ [42, 142, 157] \\ [99] \\ [111, 132] \\ \dots \end{array} \right]$

- extracting single photons from time series of sensor responses
- list of arrival times of **single** photons / pixel
- 3 attributes:
  - $c_x$  first spatial coordinate
  - $c_y$  second spatial coordinate
  - $c_t$  arrival time
- shower: extraction from 3-dimensional objects
- decoupled from hardware

→interpretation of observables of single photons

# data preparation / cleaning

## Largest pulse

1. search for high photon density along time axis (large pulses)  $t$
2. integrate time series to estimate number of photons
3. estimate arrival time from fit of polynomial to rising edge

## Cleaning

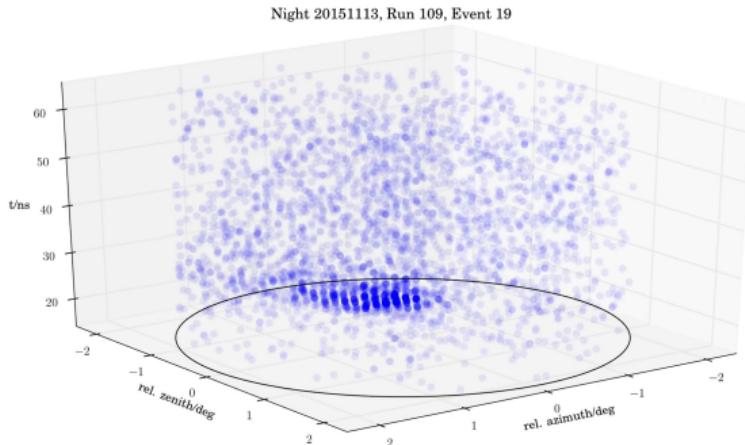
1. Find pixels containing more photons than an upper threshold  $t_1$  (5 p.e.)
2. Remove pixels with less than 2 neighbors above  $t_1$
3. Add neighbors of remaining pixels that are above lower threshold  $t_2$  (2.5 p.e.)
4. Remove pixels that have less than 2 neighbors arriving in 5 ns time window
5. Remove single pixels with less than 2 neighbors
6. Remove pixels that have less than 2 neighbors inside 5 ns time window

## Photon Stream

**density based algorithm for discovering clusters with noise** (DBSCAN)

- assign each photon to night-sky-background or cluster
- multiple clusters can be found
- no assumption of shape of clusters
- 2 parameters:
  - minimal number of photons a dense cluster must contain (20)
  - maximum distance between each photon to be considered dense ( $0,45^\circ$ )
- all observables taken into account simultaneously

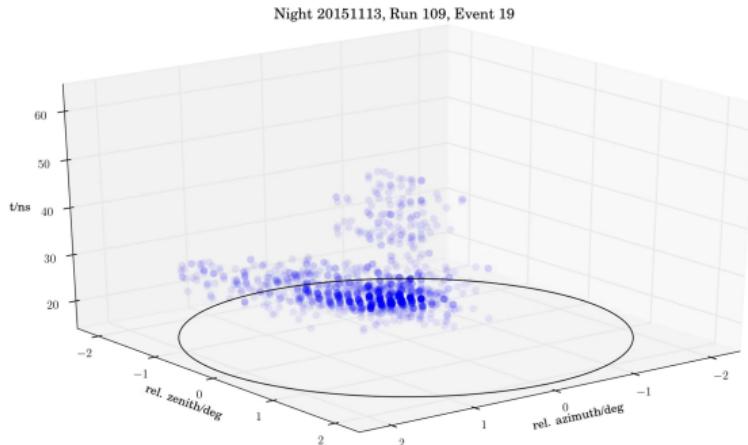
## DBSCAN



Algorithm:

1. loop over photons to find a dense cluster ( $\mathbf{x}, \mathbf{y}, t$ )
2. add photons within  $\varepsilon$  or via *dense chain*
3. if nothing to add to cluster continue with left over photons
4. photons without cluster assignment: background

# DBSCAN



Algorithm:

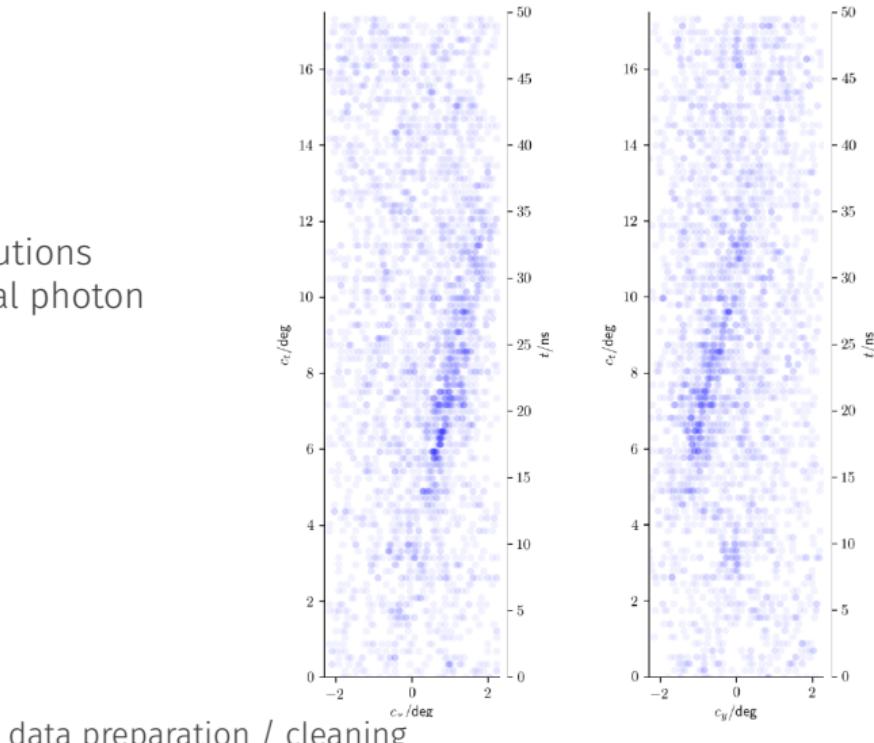
1. loop over photons to find a dense cluster ( $\mathbf{x}, \mathbf{y}, t$ )
2. add photons within  $\varepsilon$  or via *dense chain*
3. if nothing to add to cluster continue with left over photons
4. photons without cluster assignment: background

## DBSCAN

What about the metric?

- $c_t = \alpha \cdot t$
- chosen to have similar distributions along all dimensions for typical photon rates
- $\alpha = 0,35 \cdot 10^9 \text{ } \circ/\text{s}$

In total 3 parameters:  $\alpha, \varepsilon, m$



## Comparison of methods: Number of found air-shower

	Gamma-rays			Protons		
	DBSCAN	LP	DBSCAN $\in$ LP	DBSCAN	LP	DBSCAN $\in$ LP
After triggers	8725	8725	8725	2134	2134	2134
found air-shower	8300	5103	5078	1668	739	726

Table: Number of found air-showers for different cleanings on Proton and Gamma simulations.

- 63 % more gamma events in DBSCAN
- 126 % more proton events in DBSCAN

## Comparison of methods

$$\delta_{1,2} = \arccos \left( \frac{I_1 \cdot I_2}{\|I_1\| \cdot \|I_2\|} \right) \quad D_{1,2} = \|I_1 - I_2\|$$

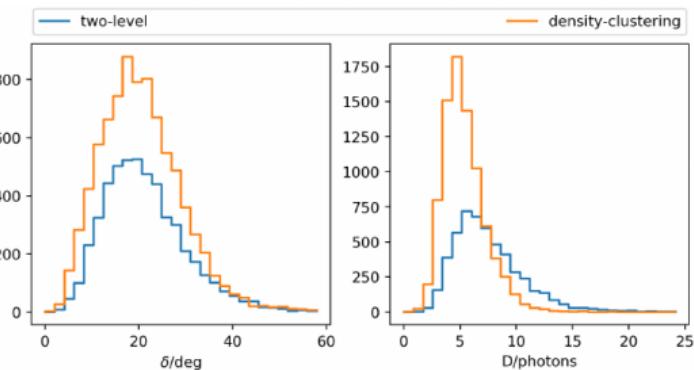


Figure: Gammas

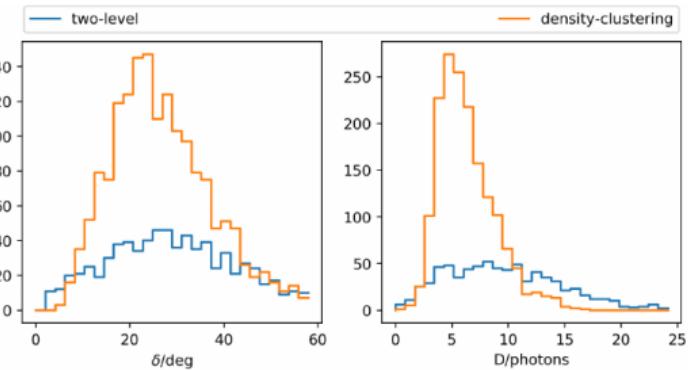


Figure: Protons

## Comparison of methods

$$\delta_{1,2} = \arccos \left( \frac{\mathbf{I}_1 \cdot \mathbf{I}_2}{\|\mathbf{I}_1\| \cdot \|\mathbf{I}_2\|} \right) \quad D_{1,2} = \|\mathbf{I}_1 - \mathbf{I}_2\|$$

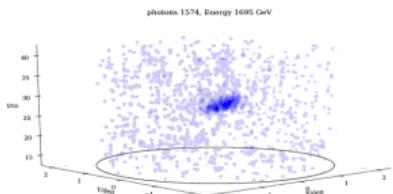
metric	Gamma-rays		Protons	
	DBSCAN	LP	DBSCAN	LP
mean $\delta/\text{deg}$	22.6	23.3	29.3	31.4
mean $D/\text{p.e.}$	6.2	8.4	7.2	10.5
<b>on overlap events</b>				
mean $\delta/\text{deg}$	18.6	23.3	22.8	31.3
mean $D/\text{p.e.}$	6.8	8.4	8.0	10.6

# Data Set

# The Data set: FACT open data crab sample

## Photon-Stream (phs)

Single photon event representation for Imaging Atmospheric Cherenkov Telescopes (IACTs)



### Public

#### Crab Nebula

High quality observations of the Crab Nebula in November 2013. Free to use for everyone. Please cite our [photon-4](#) on [GitHub](#).

#### Simulations (sim)

The FACT responses to simulated air-shower observations. This is the diffuse [ceres\\_12\\_rc](#) simulation pass.

### Private

#### Observations (obs)

All observed air-shower events of FACT. About 1e9 air-shower records containing about 3e12 single photons. See [...](#)

### About

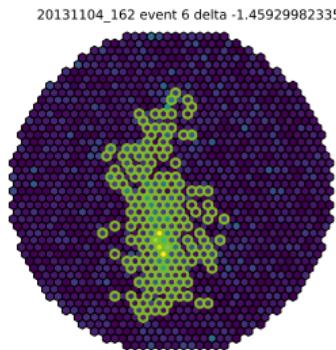
- [ICRC2017 proceeding](#)
- [photon-stream Python](#)
- [photon-stream C++](#)
- [production tools](#)

FACT collaboration, Sebastian A. Mueller, 2017

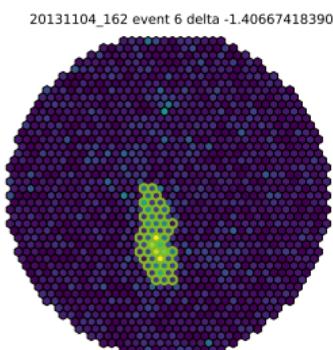
- <https://fact-project.org/data>
- Crab Nebula observations from November 2013
- including gamma-ray and proton simulations
- 17.7 hours of observations

# Images on data

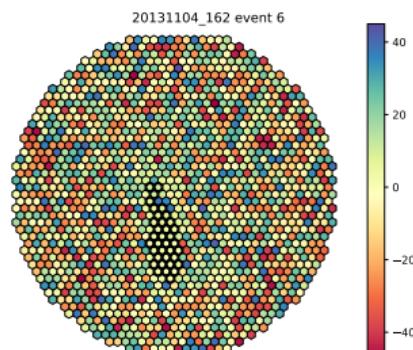
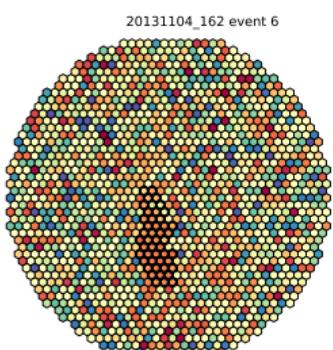
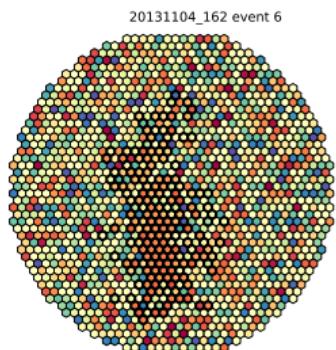
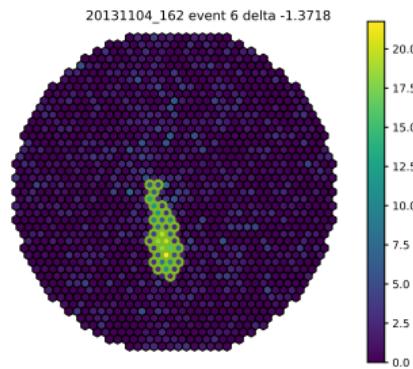
## DBSCAN



## Thresholds on phs

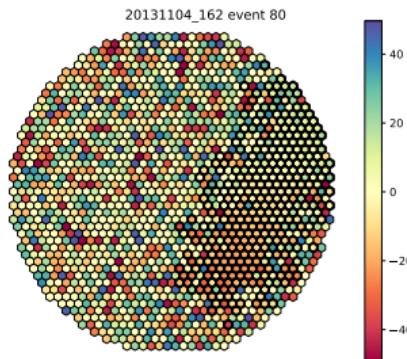
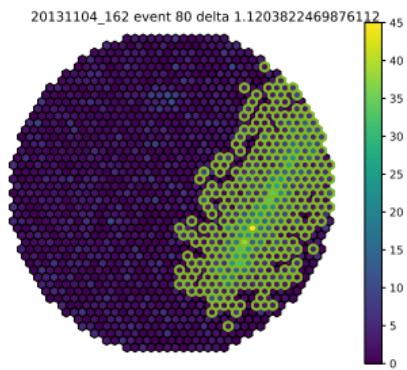


## FACTtools

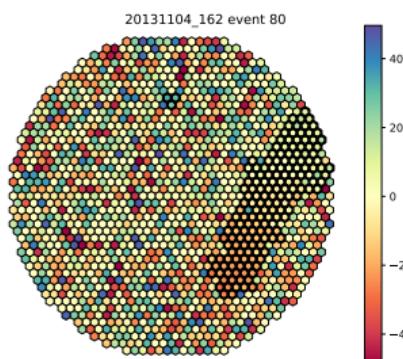
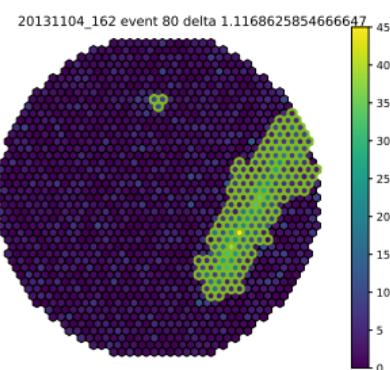


Images on data

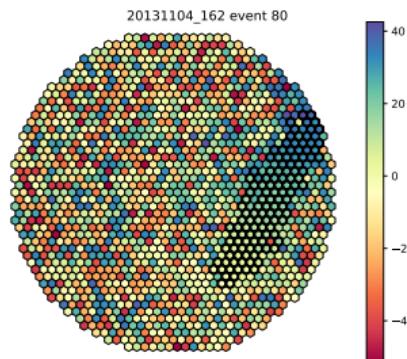
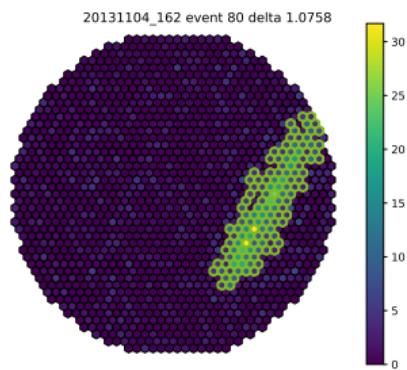
## DBSCAN



## Thresholds on phs

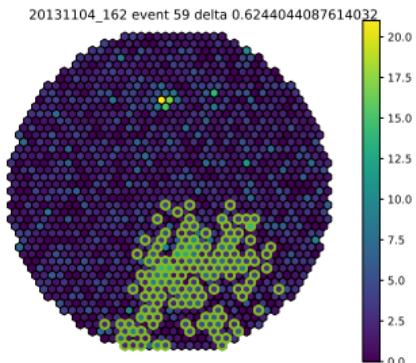


## FACTtools

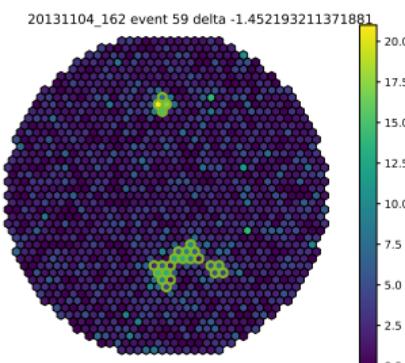


Images on data

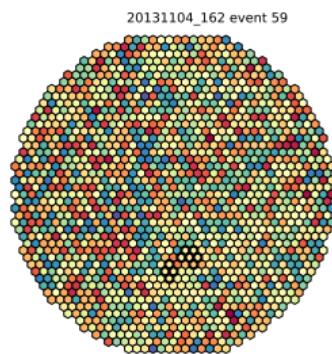
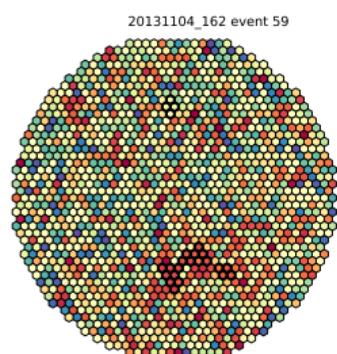
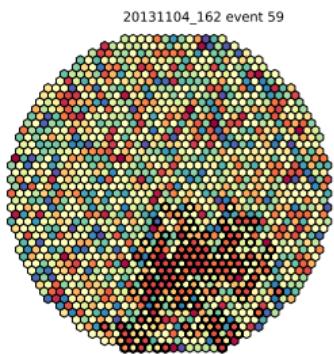
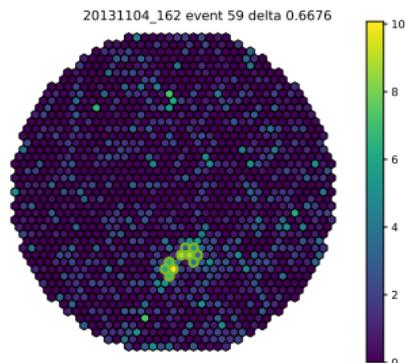
## DBSCAN



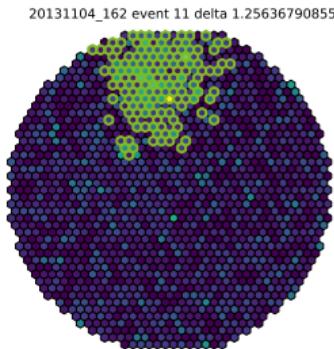
## Thresholds on phs



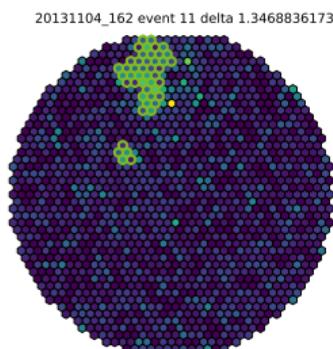
## FACTtools



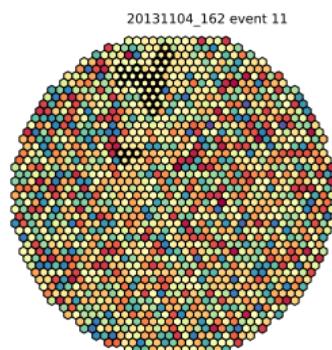
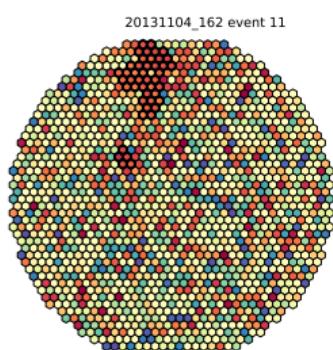
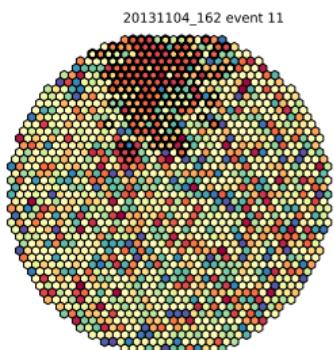
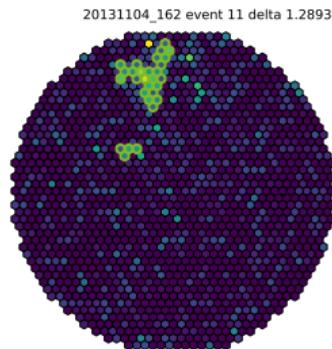
## DBSCAN



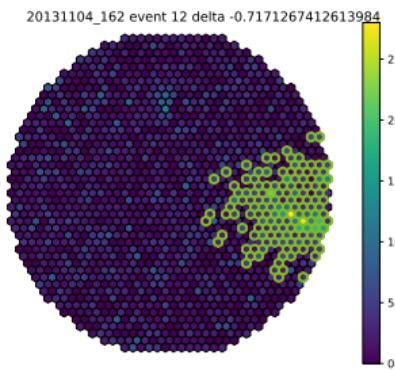
## Thresholds on phs



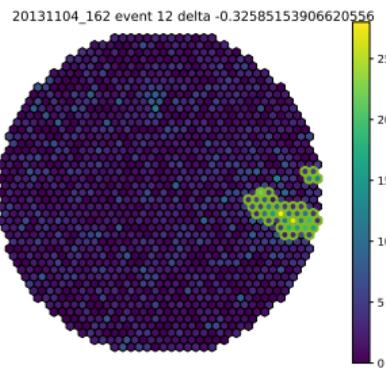
## FACTtools



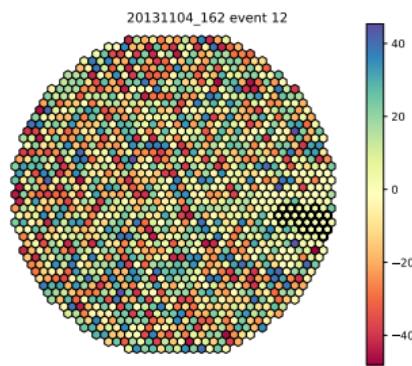
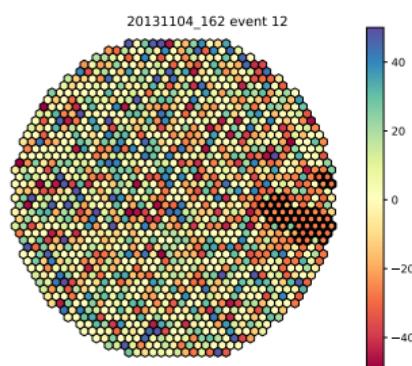
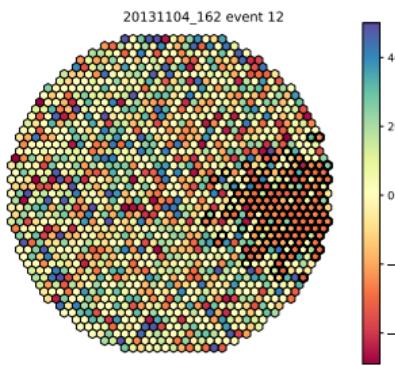
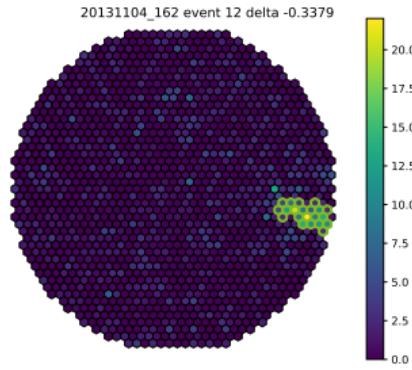
### DBSCAN



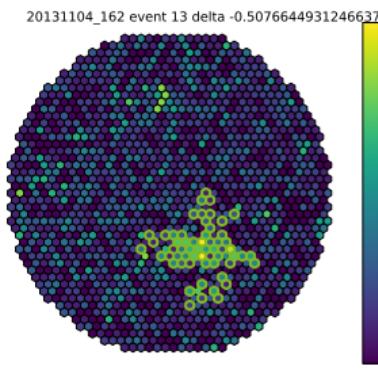
### Thresholds on phs



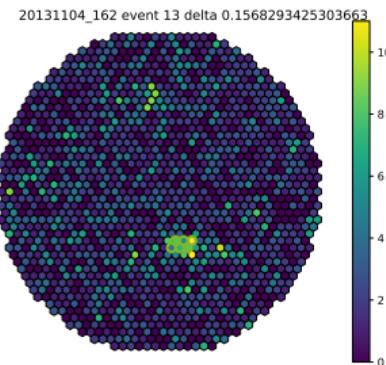
### FACTtools



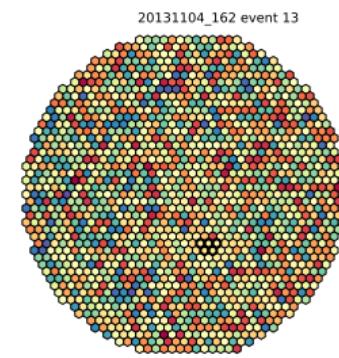
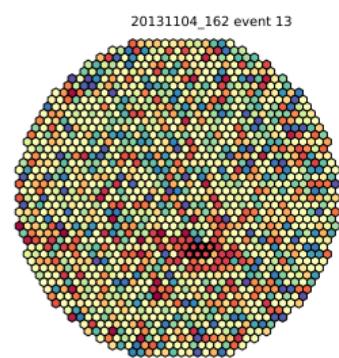
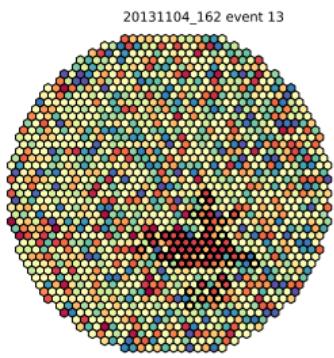
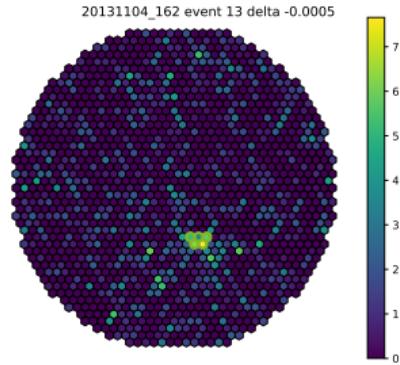
## DBSCAN



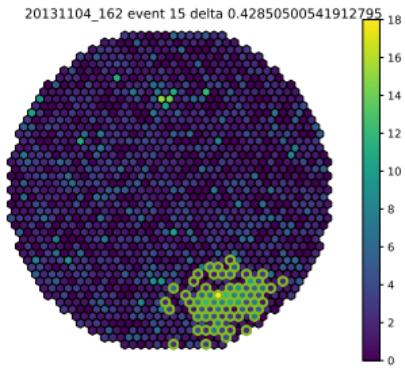
## Thresholds on phs



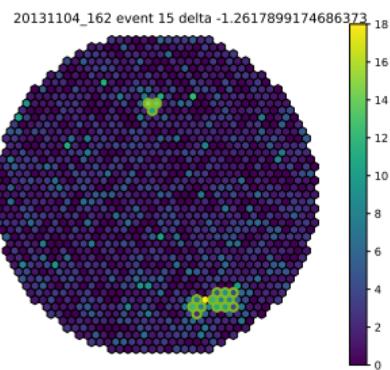
## FACTtools



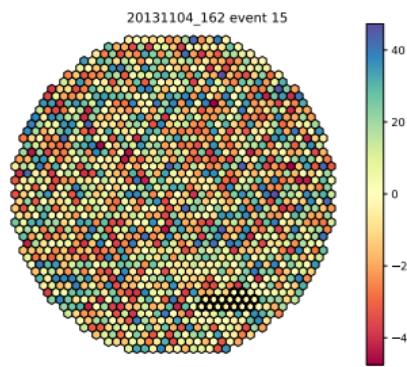
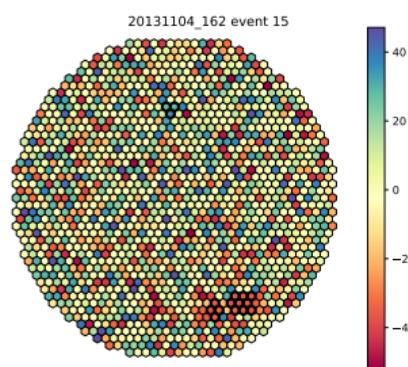
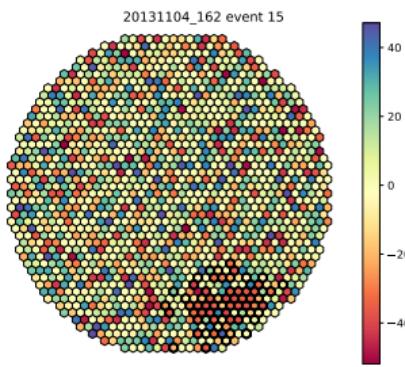
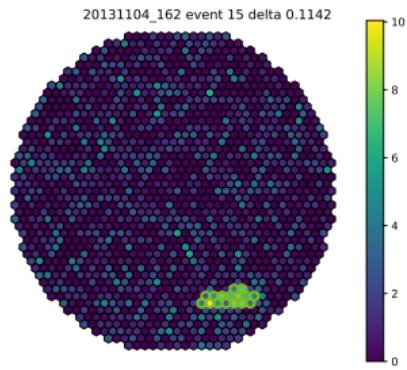
### DBSCAN



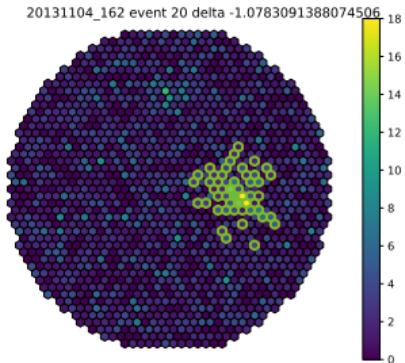
### Thresholds on phs



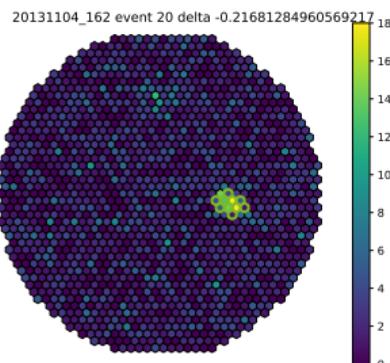
### FACTtools



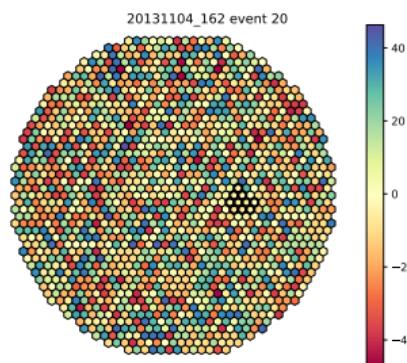
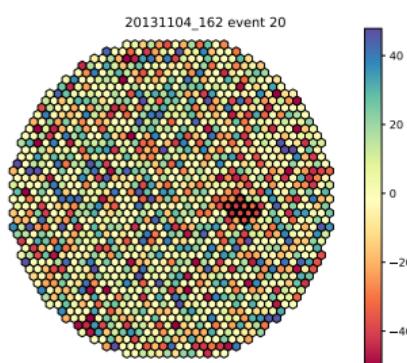
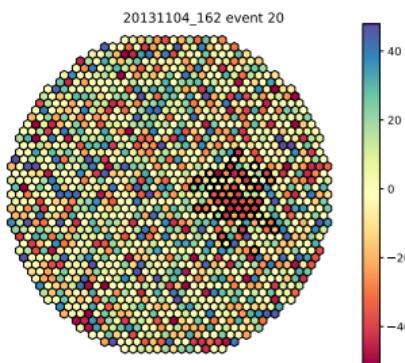
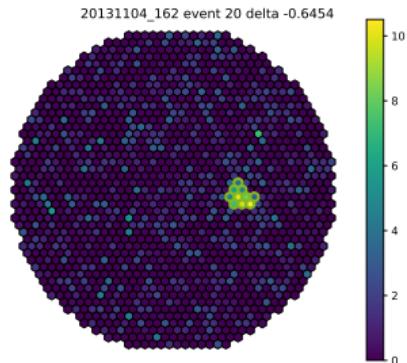
### DBSCAN



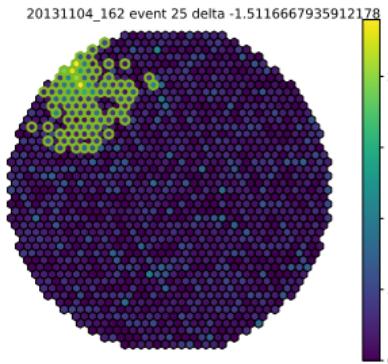
### Thresholds on phs



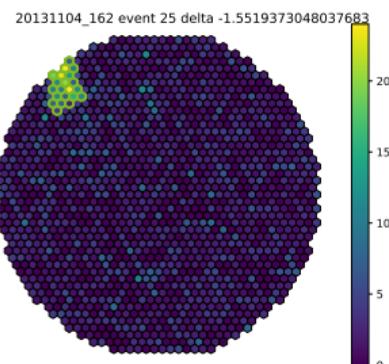
### FACTtools



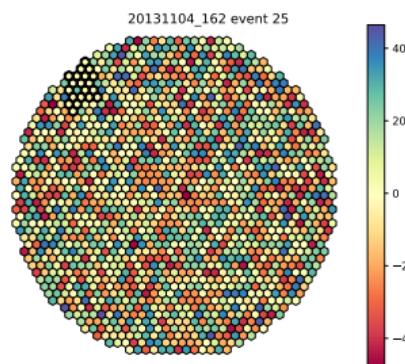
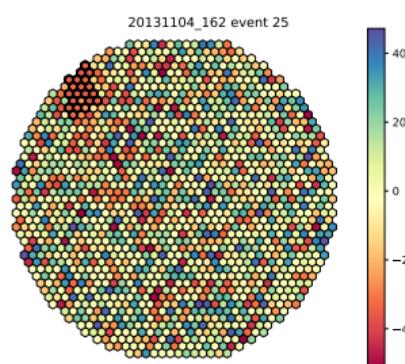
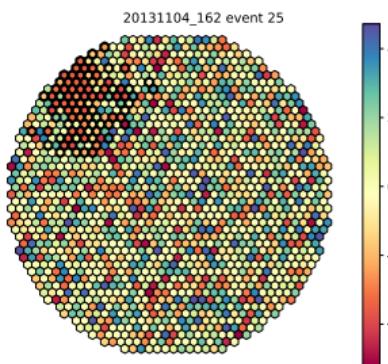
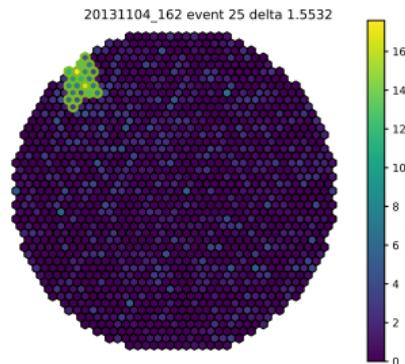
## DBSCAN



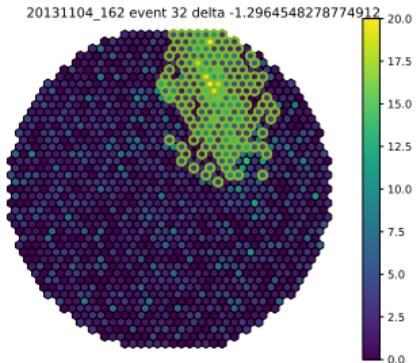
## Thresholds on phs



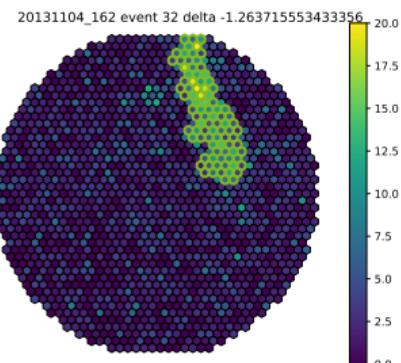
## FACTtools



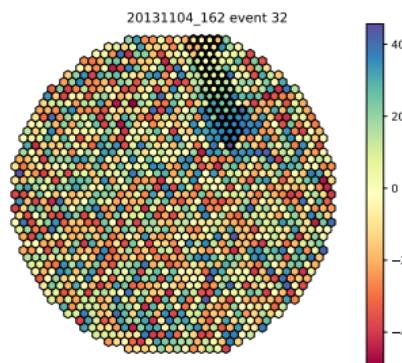
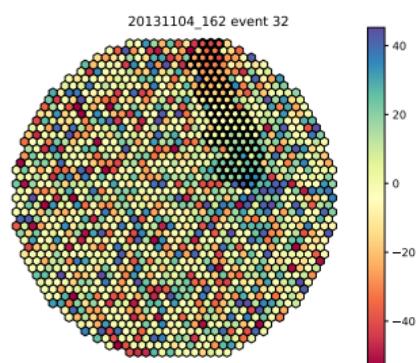
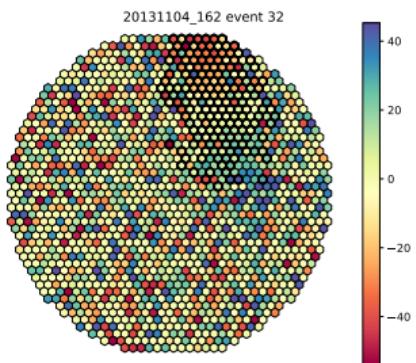
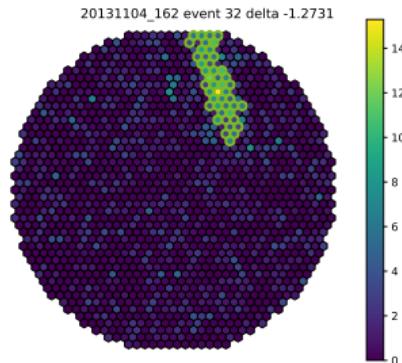
### DBSCAN



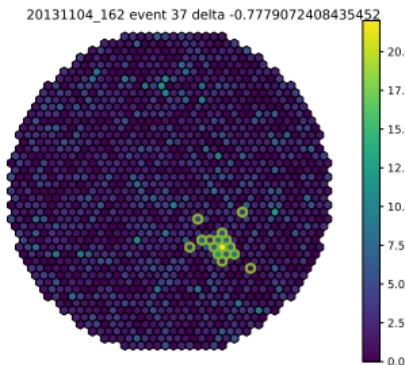
### Thresholds on phs



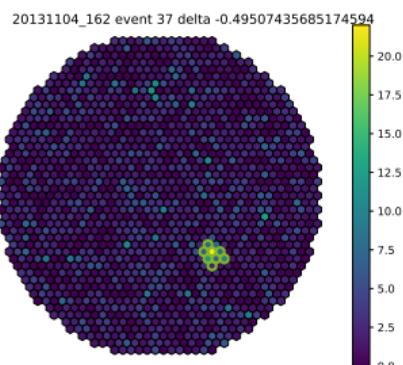
### FACTtools



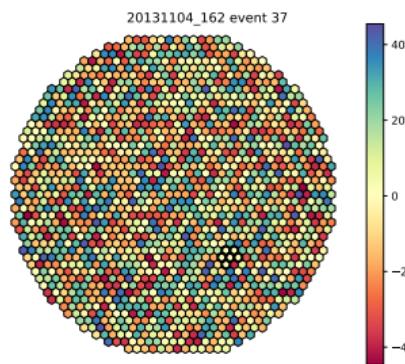
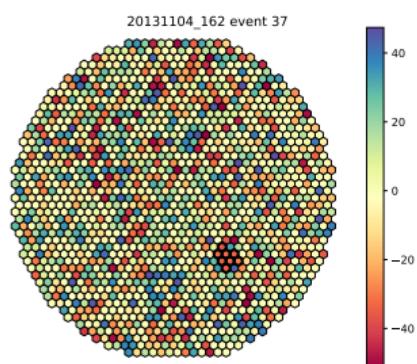
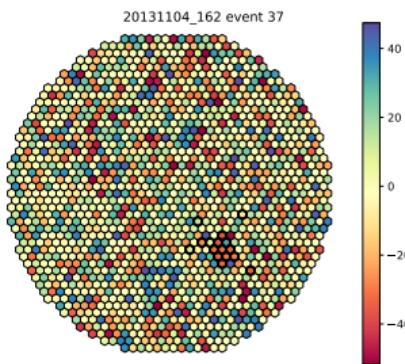
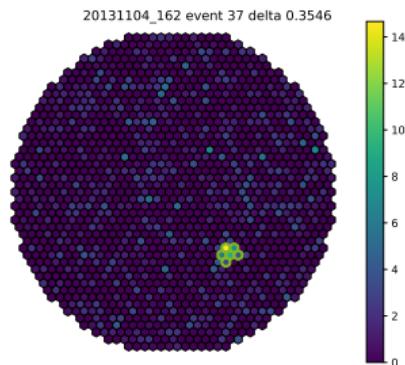
## DBSCAN



## Thresholds on phs

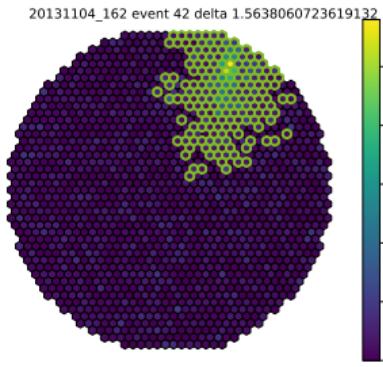


## FACTtools

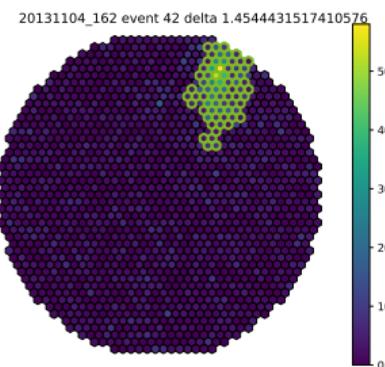


Images on data

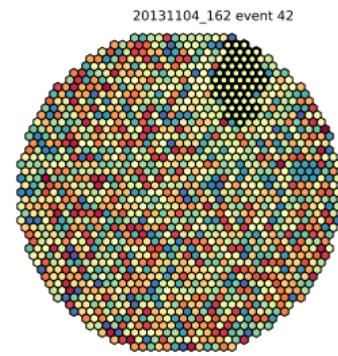
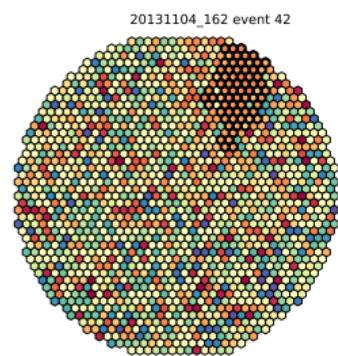
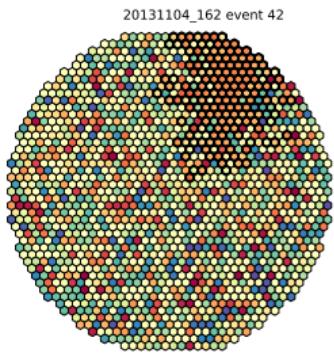
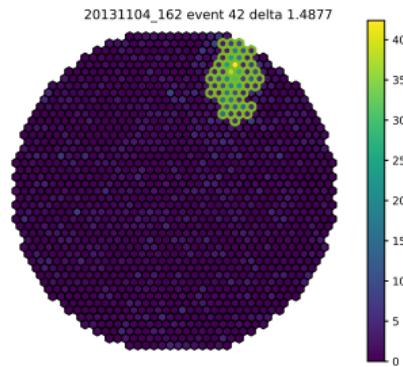
### DBSCAN



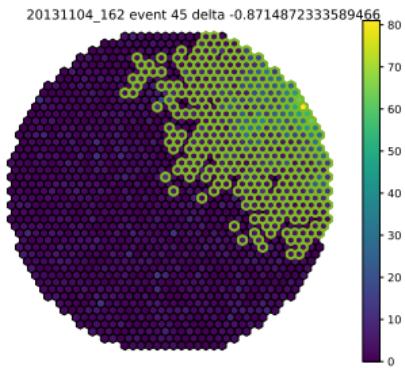
### Thresholds on phs



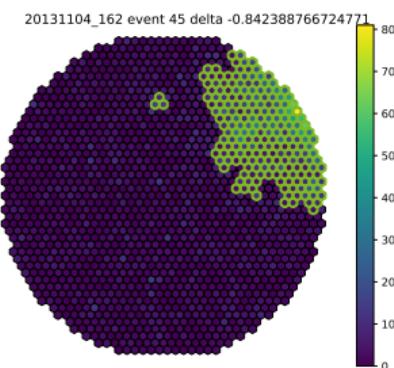
### FACTtools



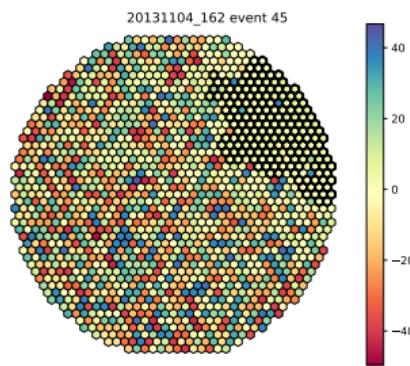
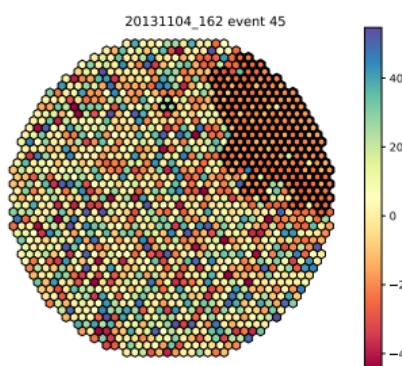
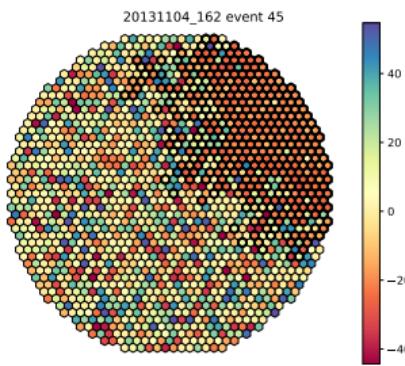
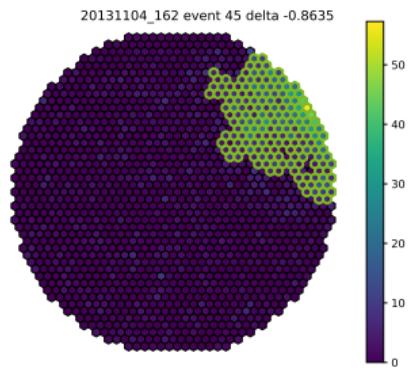
### DBSCAN



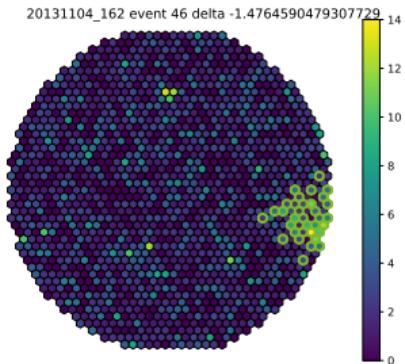
### Thresholds on phs



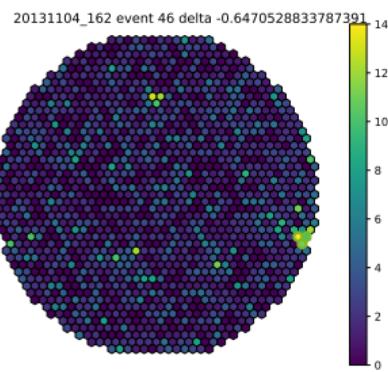
### FACTtools



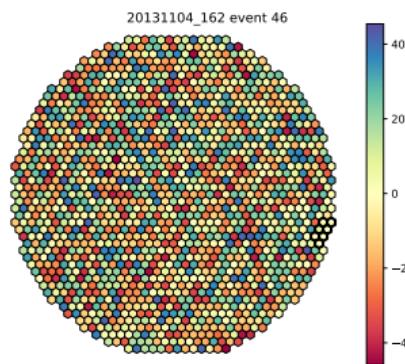
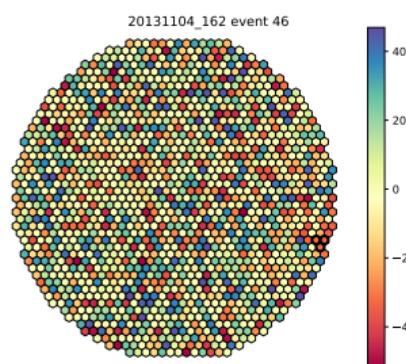
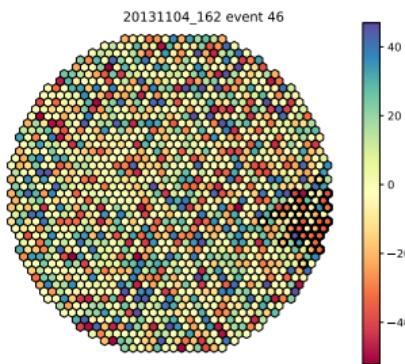
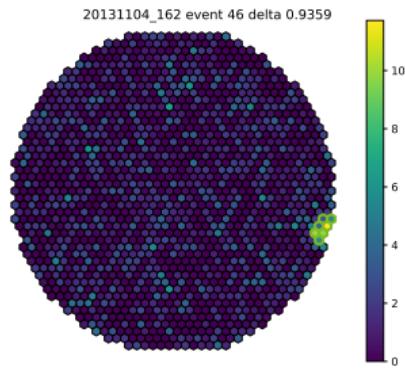
### DBSCAN



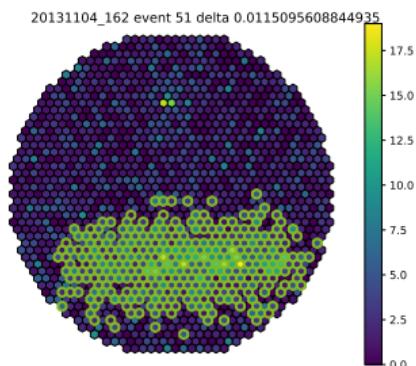
### Thresholds on phs



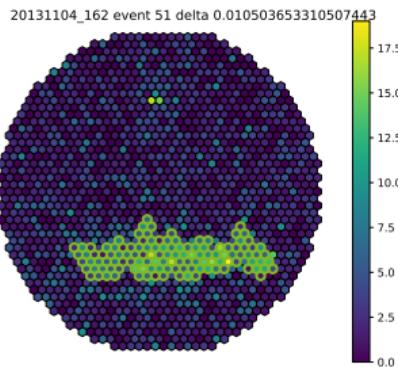
### FACTtools



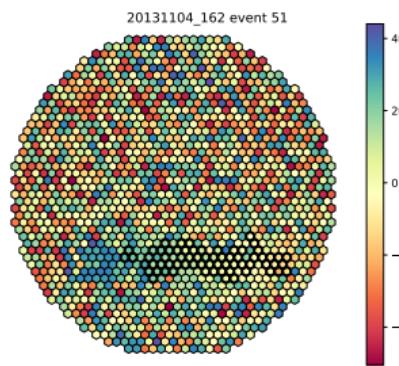
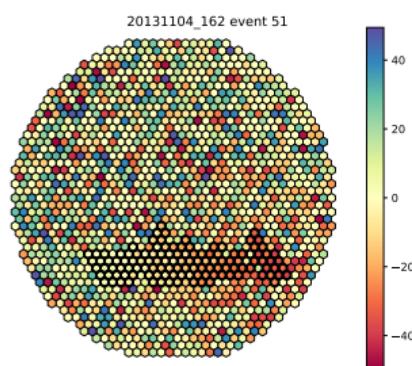
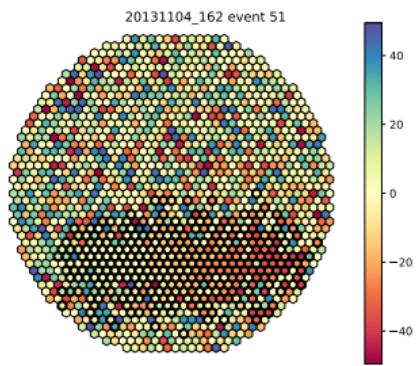
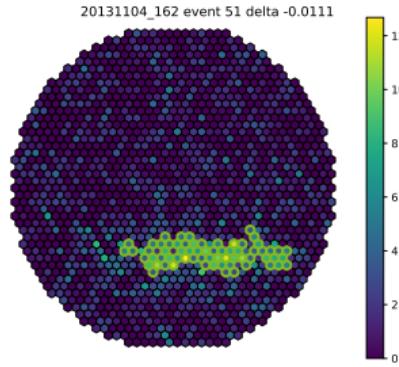
### DBSCAN



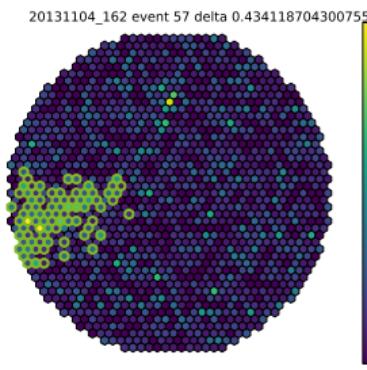
### Thresholds on phs



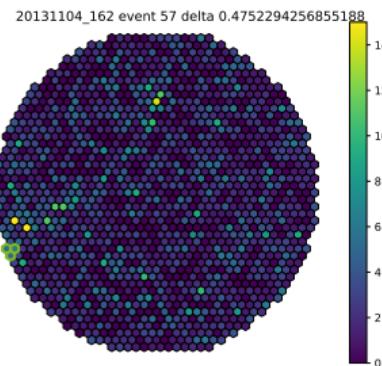
### FACTtools



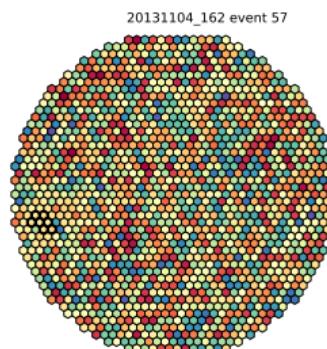
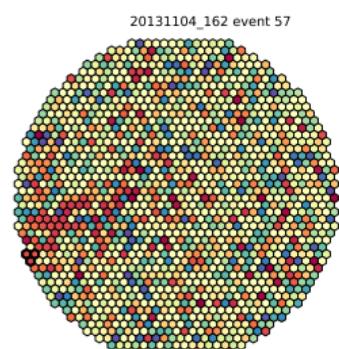
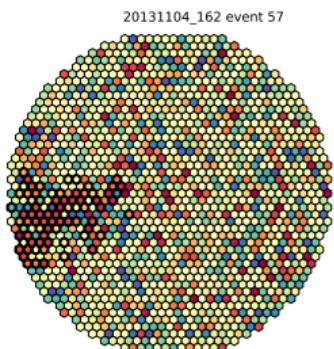
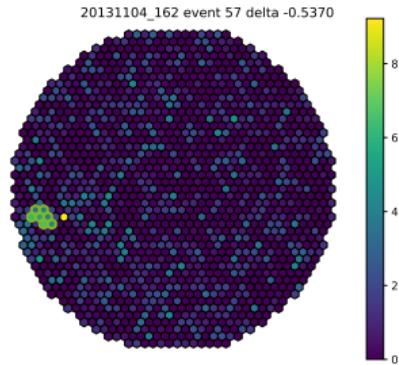
### DBSCAN



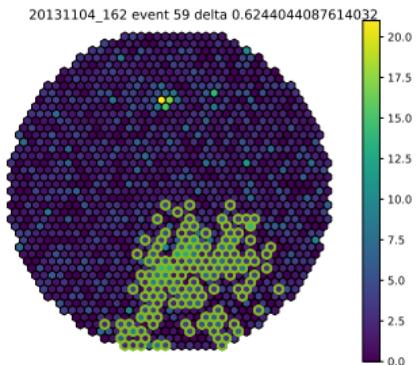
### Thresholds on phs



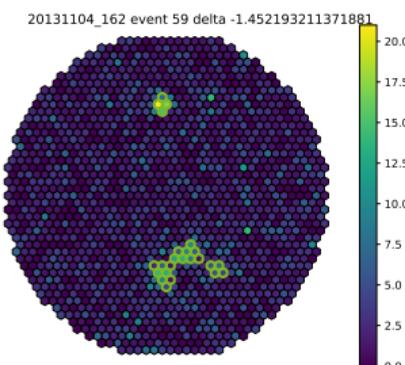
### FACTtools



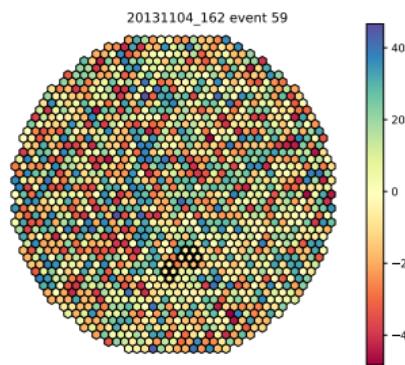
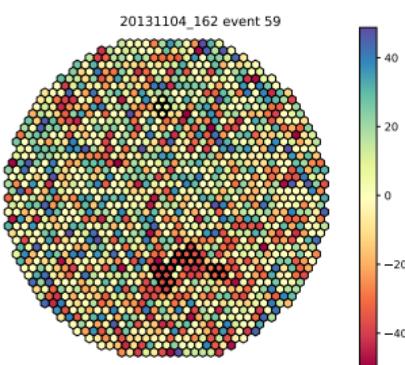
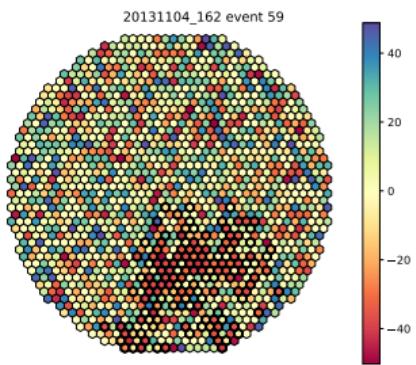
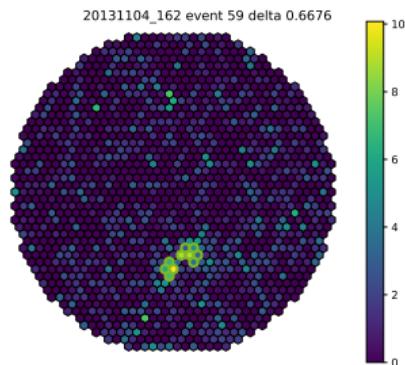
### DBSCAN



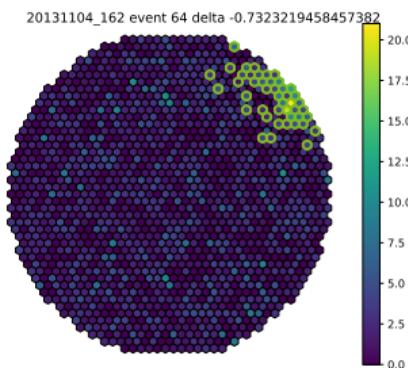
### Thresholds on phs



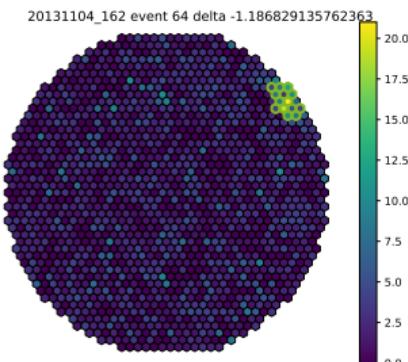
### FACTtools



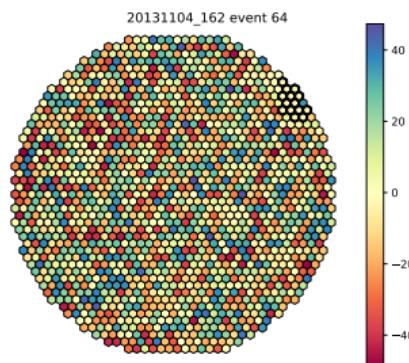
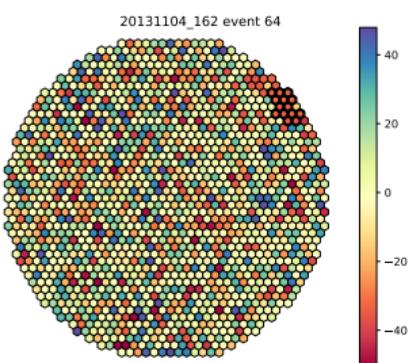
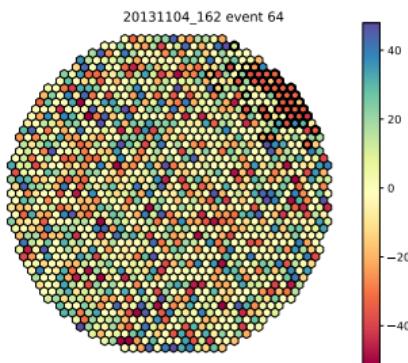
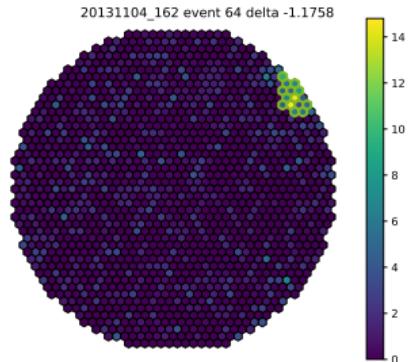
### DBSCAN



### Thresholds on phs

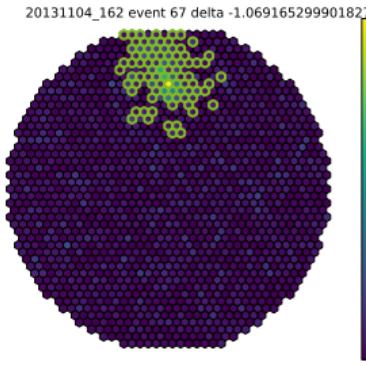


### FACTtools

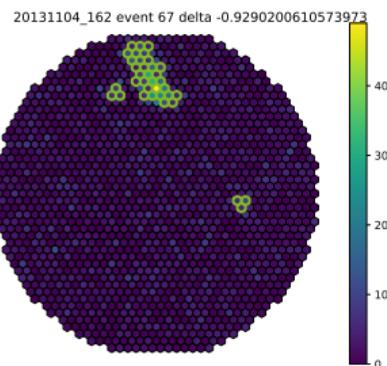


Images on data

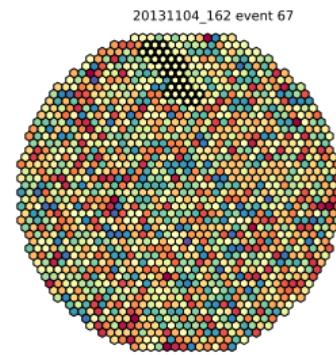
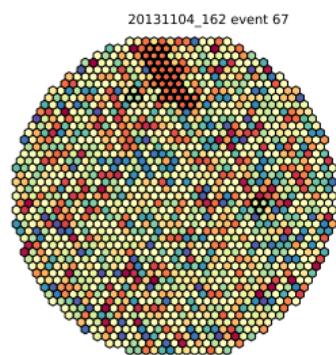
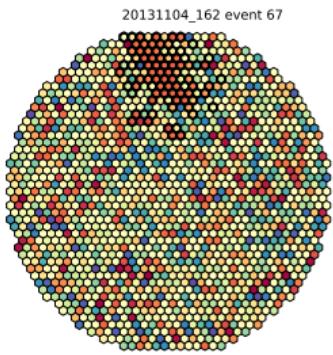
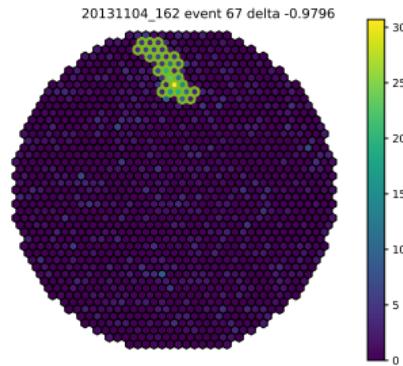
### DBSCAN



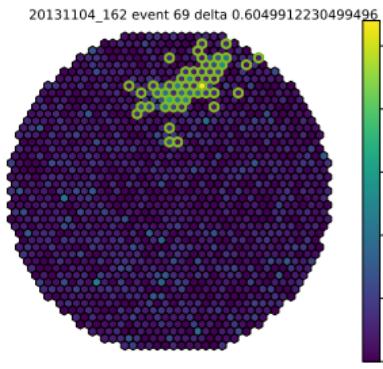
### Thresholds on phs



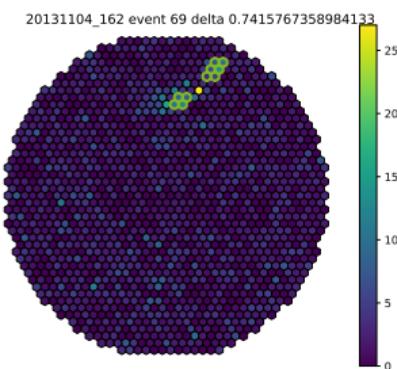
### FACTtools



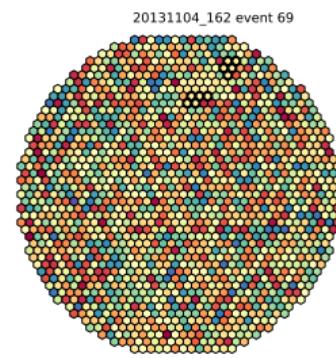
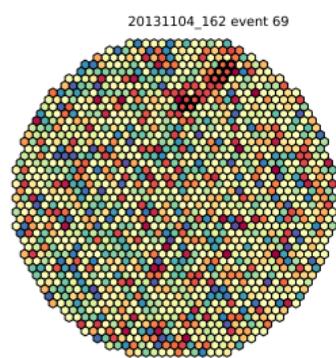
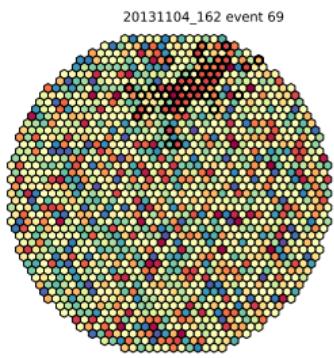
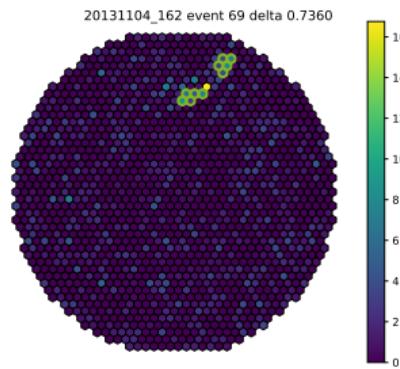
### DBSCAN



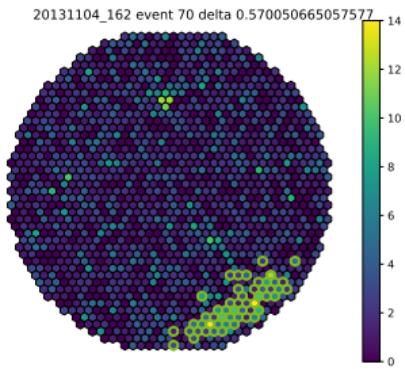
### Thresholds on phs



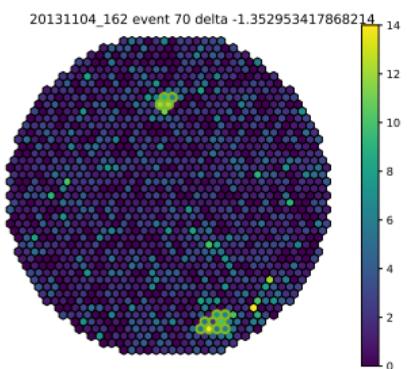
### FACTtools



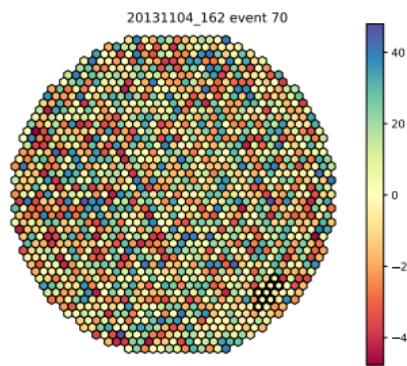
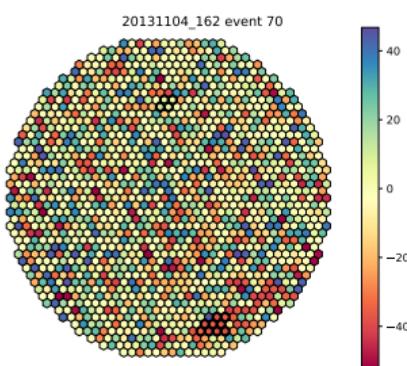
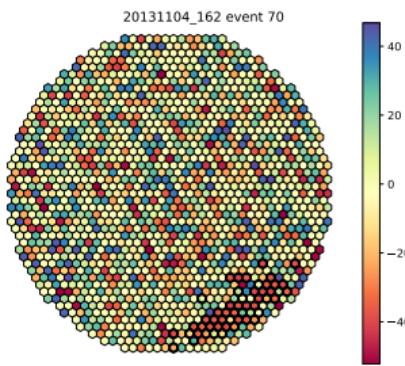
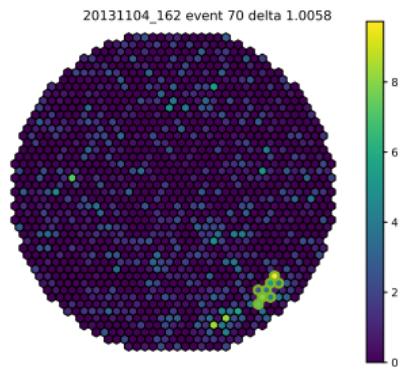
### DBSCAN



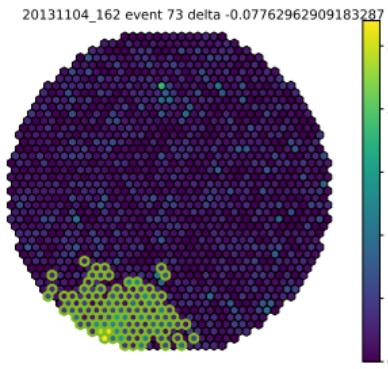
### Thresholds on phs



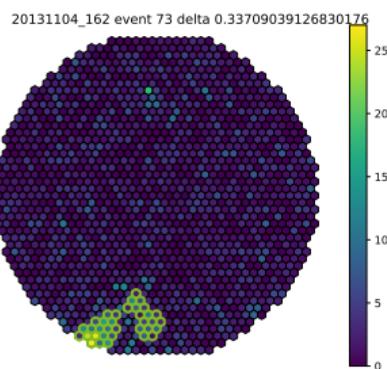
### FACTtools



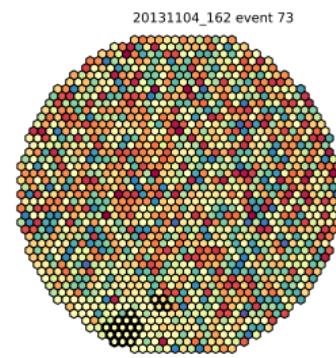
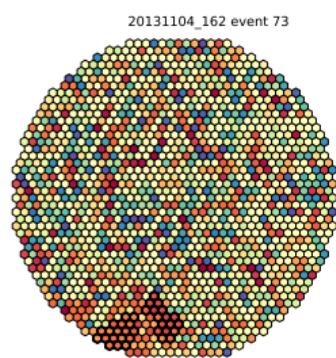
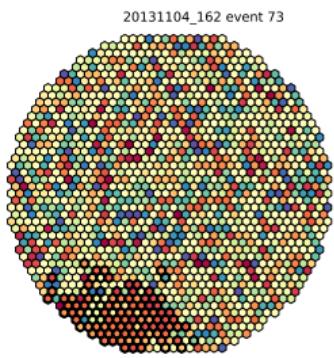
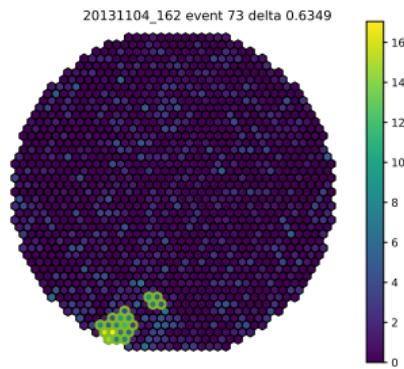
### DBSCAN



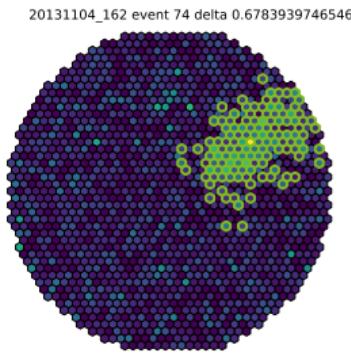
### Thresholds on phs



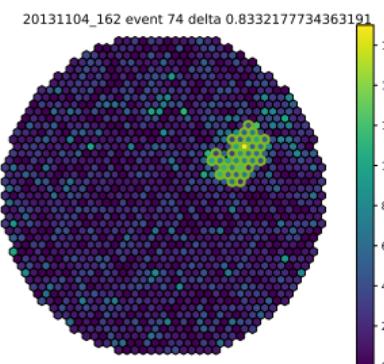
### FACTtools



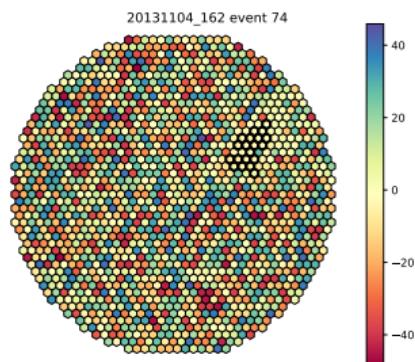
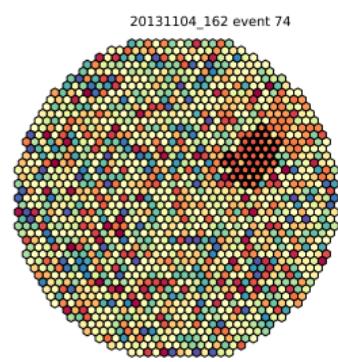
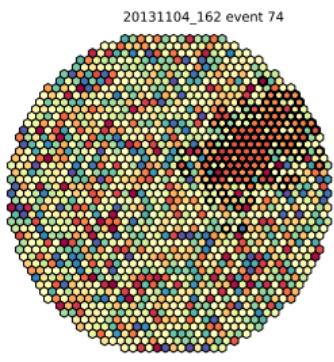
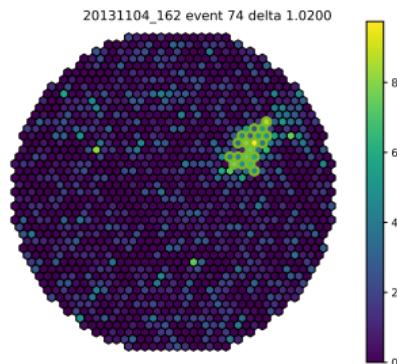
### DBSCAN



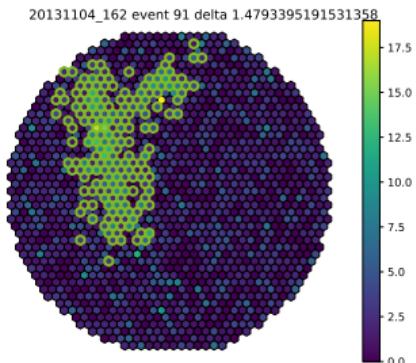
### Thresholds on phs



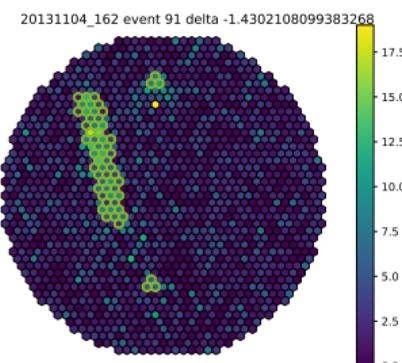
### FACTtools



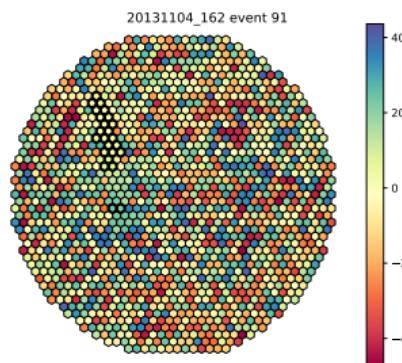
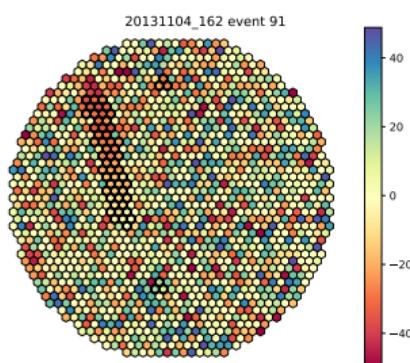
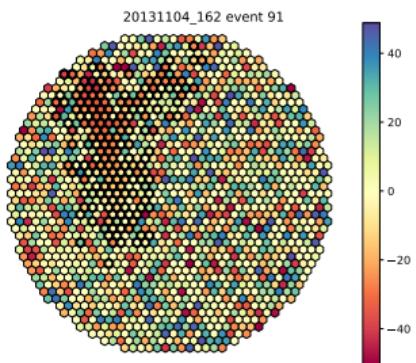
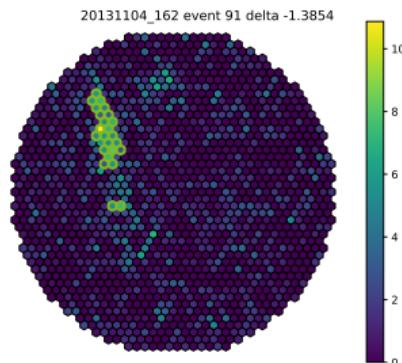
### DBSCAN



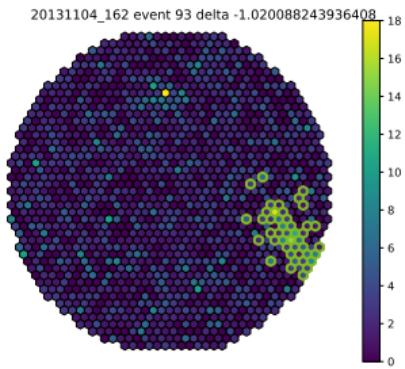
### Thresholds on phs



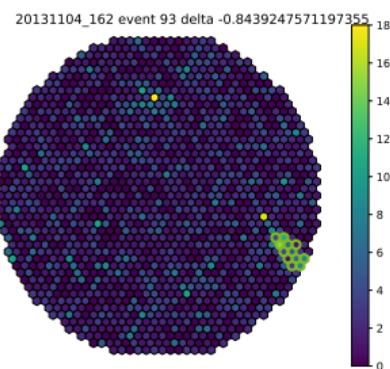
### FACTtools



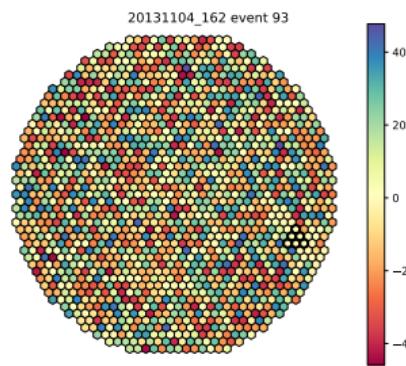
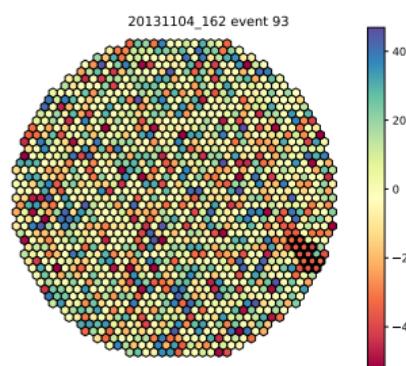
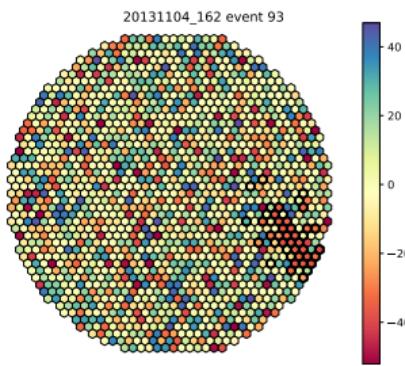
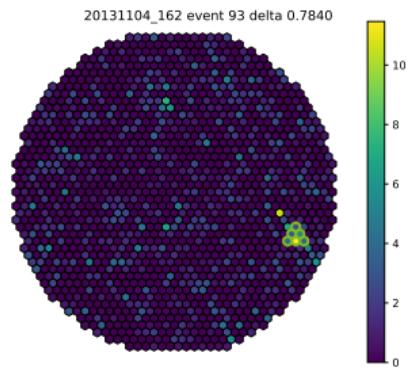
### DBSCAN



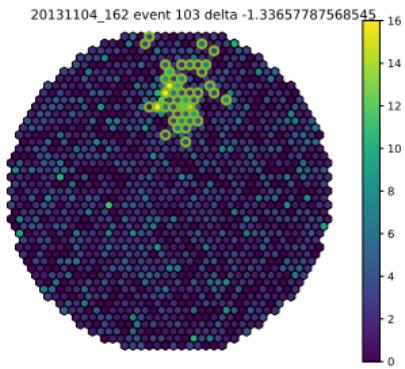
### Thresholds on phs



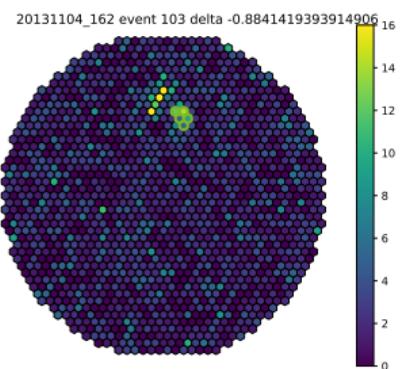
### FACTtools



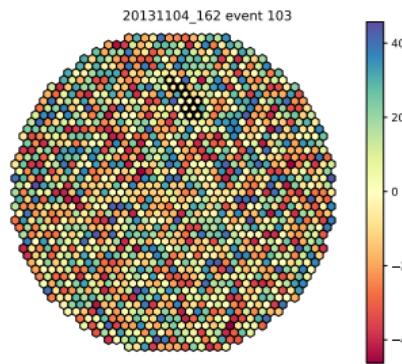
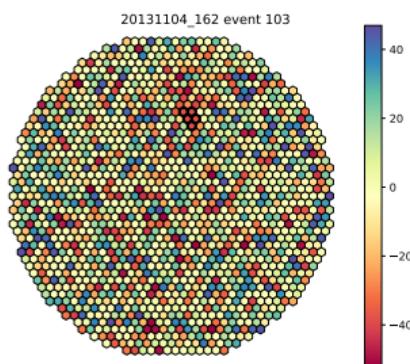
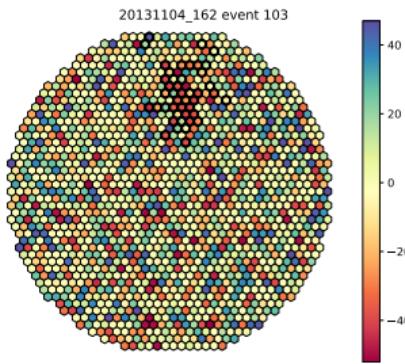
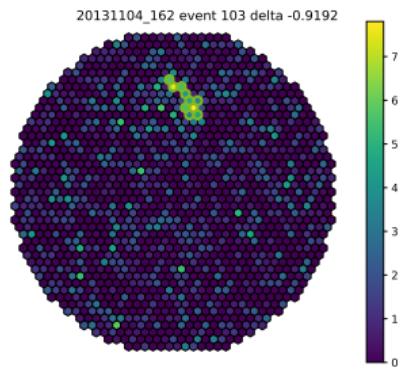
### DBSCAN



### Thresholds on phs

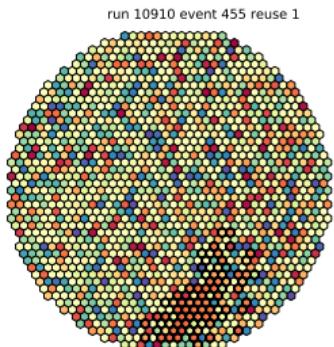
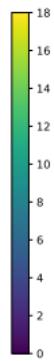
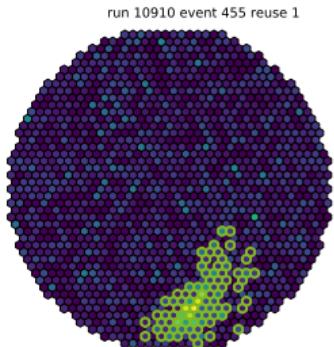


### FACTtools

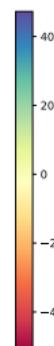
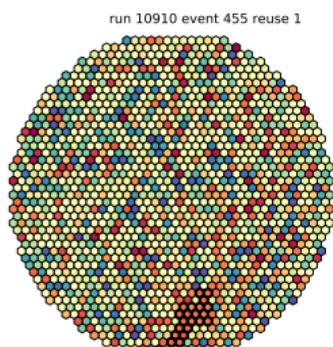
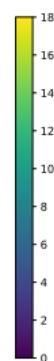
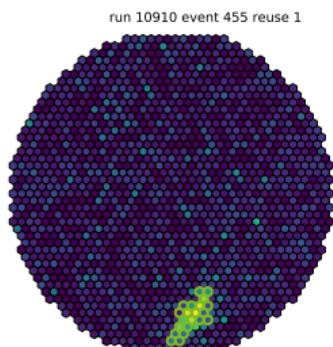


# Images on Gamma MC

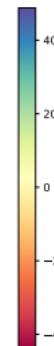
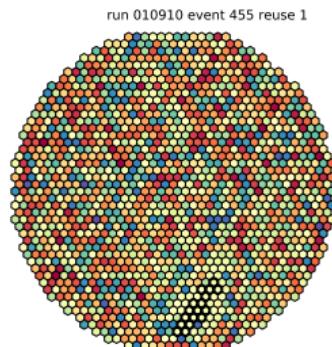
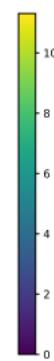
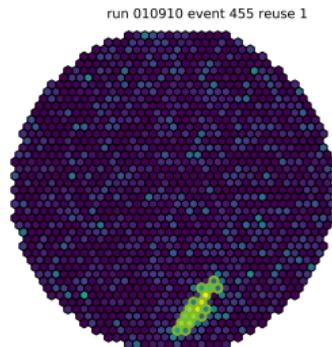
## DBSCAN



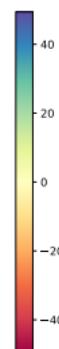
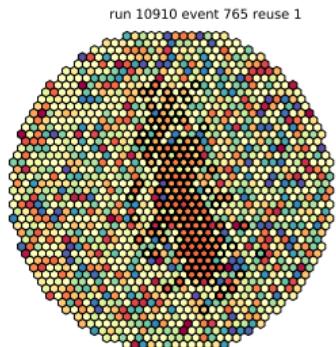
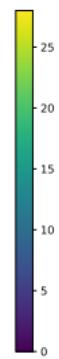
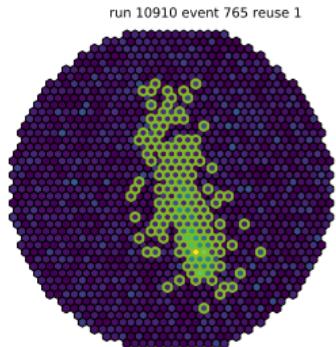
## Thresholds on phs



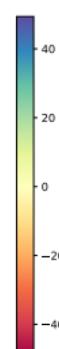
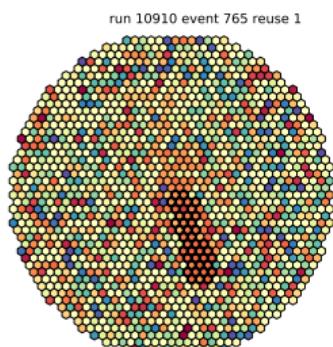
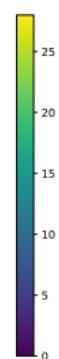
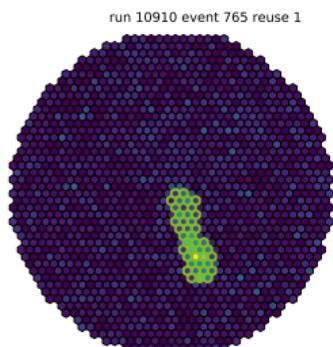
## FACTtools



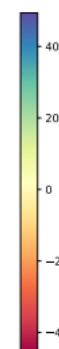
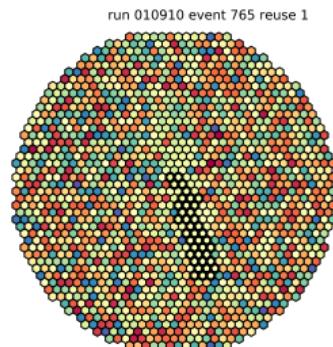
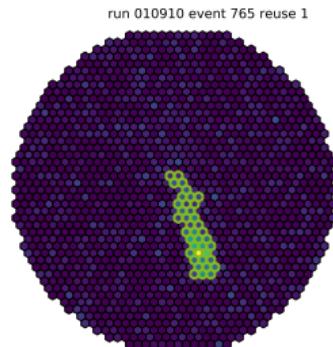
## DBSCAN



## Thresholds on phs

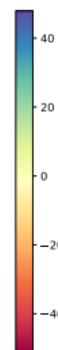
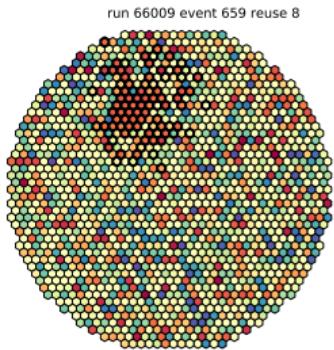
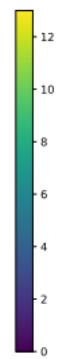
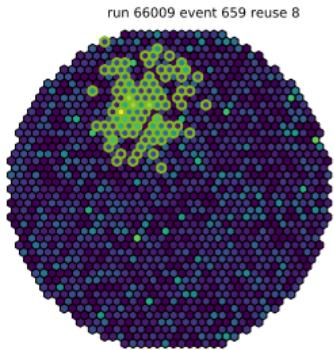


## FACTtools

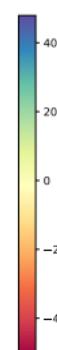
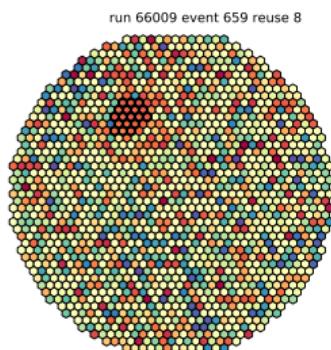
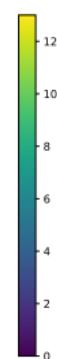
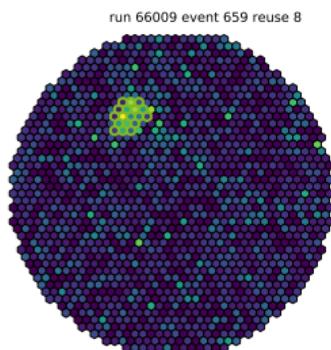


# Images on Proton MC

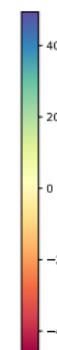
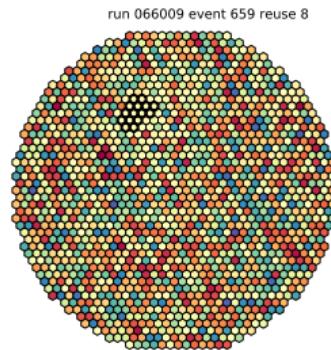
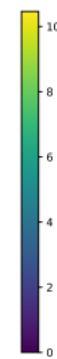
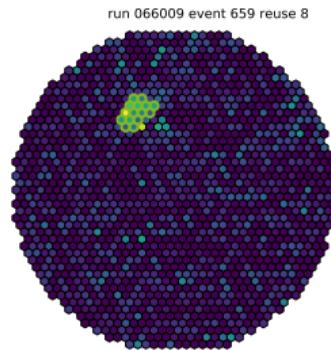
### DBSCAN



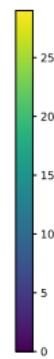
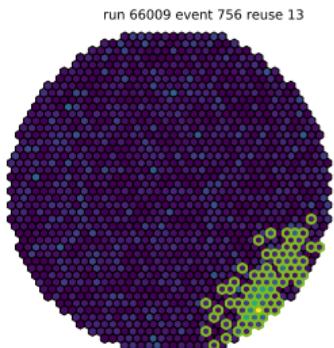
### Thresholds on phs



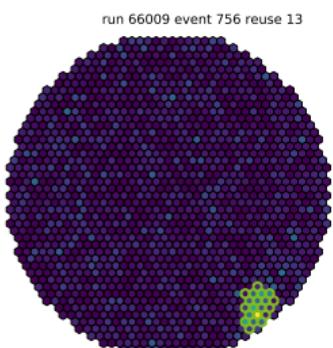
### FACTtools



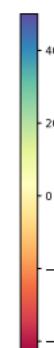
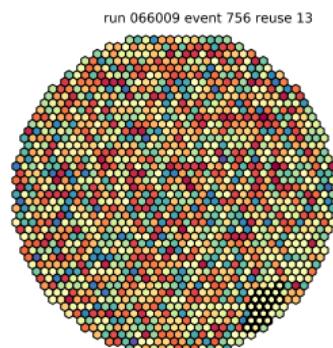
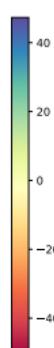
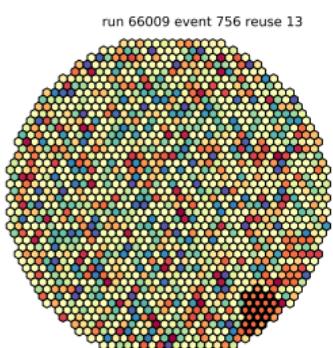
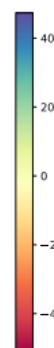
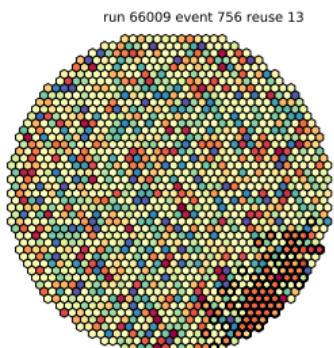
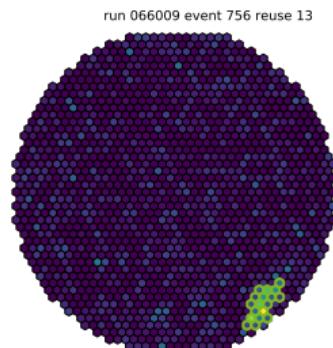
## DBSCAN



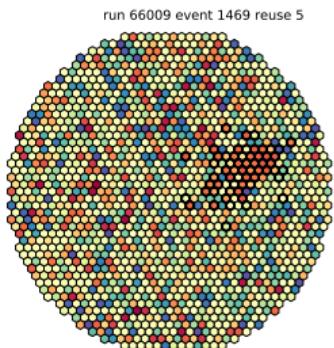
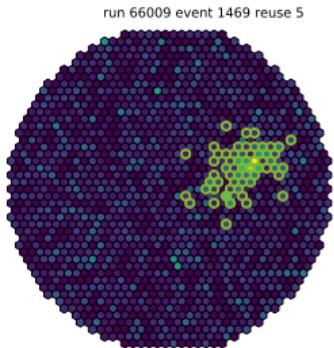
## Thresholds on phs



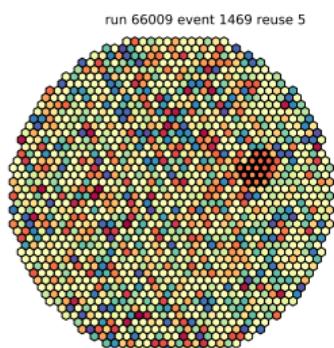
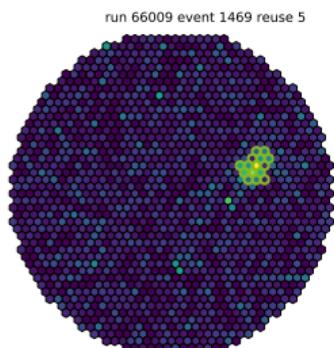
## FACTtools



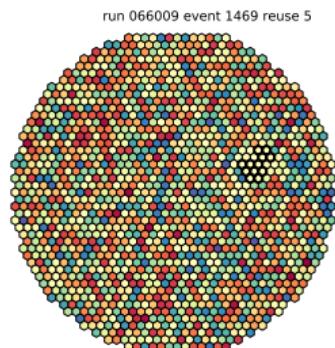
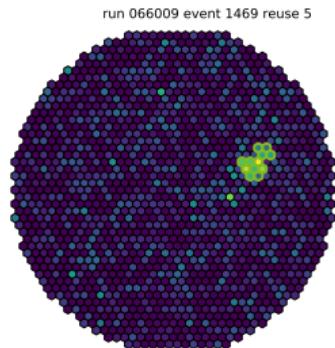
## DBSCAN



## Thresholds on phs



## FACTtools



# Analysis

## Analysis

**aim** proof of concept: generate physics results

**Crab Nebula** well measured source of cosmic gamma rays → standard candle analysis

### Photon Stream Analysis chain

**calibration** extracting single photons

**image cleaning** DBSCAN

**parametrization** parameter set A

**separation** IACT tools

**reconstruction** IACT tools

### Standard Analysis chain

**calibration** identifying large pulses

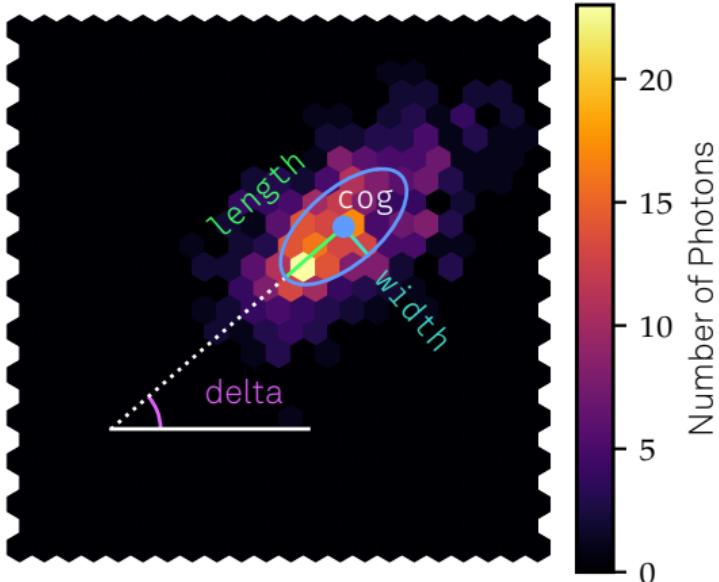
**image cleaning** time and pixel thresholds

**parametrization** parameter set B

**separation** IACT tools

**reconstruction** IACT tools

## Parameterization

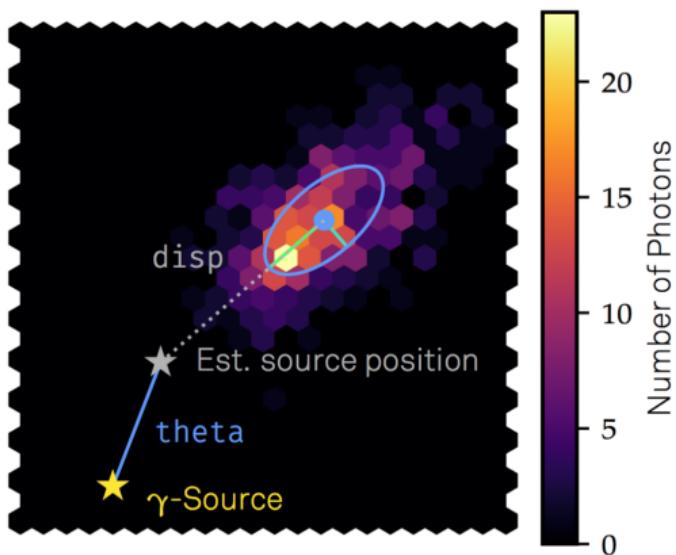


Analysis

Hillas parameters (projected back to 2D):

- size** number of photons in cluster
- length** std. dev. along long half-axis
- width** std. dev. along short half-axis
- delta** angle between length and disp
- skewness/ kurtosis** higher order statistical moments along half-axes in cluster system

## Parameterization



Source position reconstruction via  
disp-method:

**|displ|** distance from centre of gravity to  
target

**sgn(displ)** Head/Tail-Disambiguation

**theta** distance between reconstructed  
and true origin

## IACT tools

Cuts: size > 30

Machine learning with Random Forrests, using 5-fold cross validation.

### Energy estimation:

- random forest regressor (**size, length, width, n\_pixel, delta, kurtosis\_long, kurtosis\_trans, skewness\_long, skewness\_trans**)

### Gamma-hadron-separation:

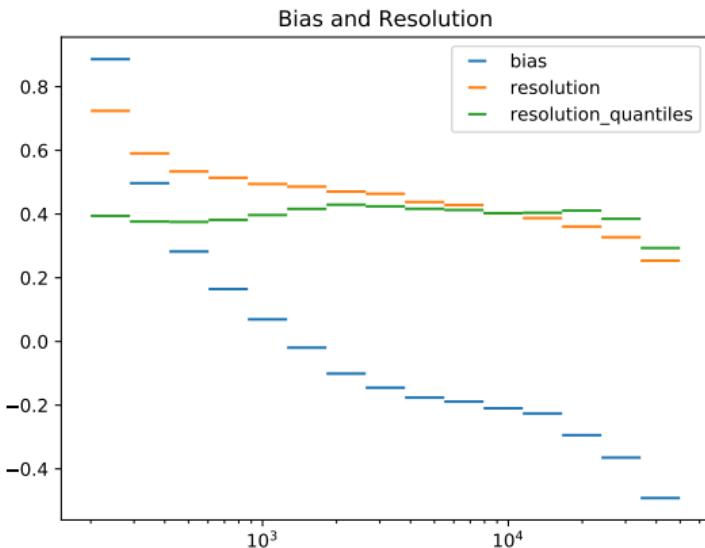
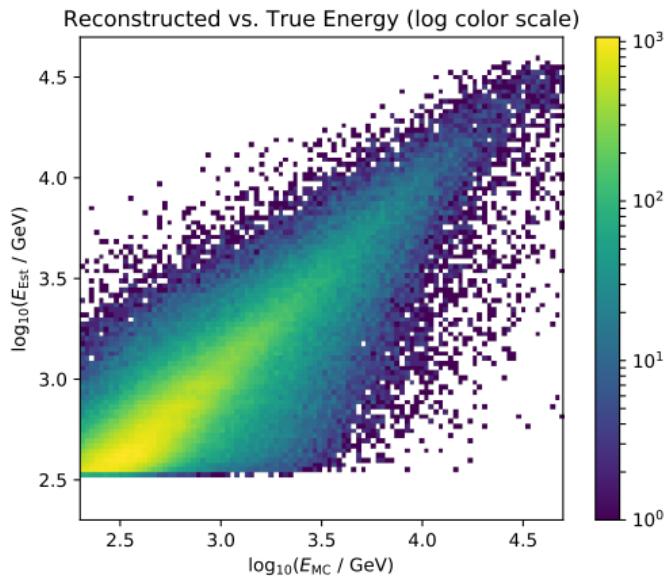
- random forest classifier (**clusters, cluster\_size\_ratio, n\_pixel, size, length, width, kurtosis\_long, kurtosis\_trans, skewness\_long, skewness\_trans**)

### Origin reconstruction:

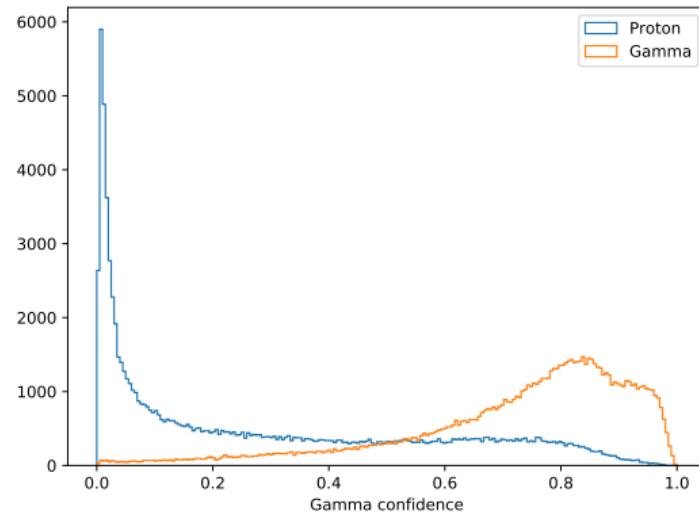
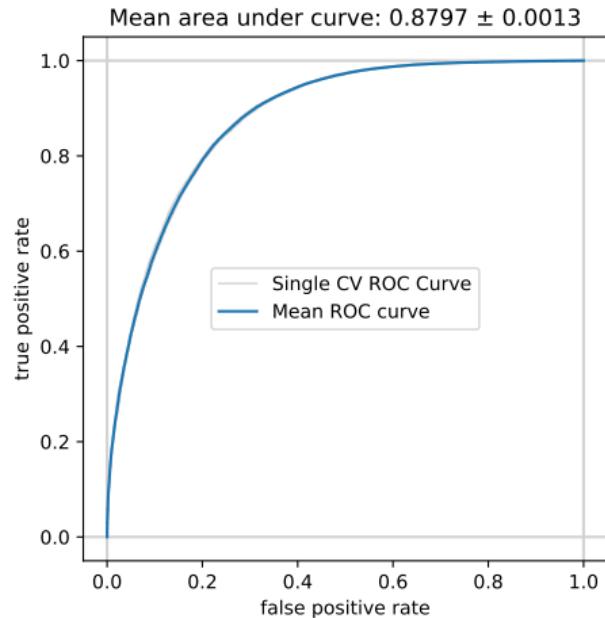
- two step task: regression of  $|disp|$  and classification of  $\text{sgn}(disp)$
- random forest regressor and classifier (**length, width, skewness\_long, kurtosis\_long**)

# Results

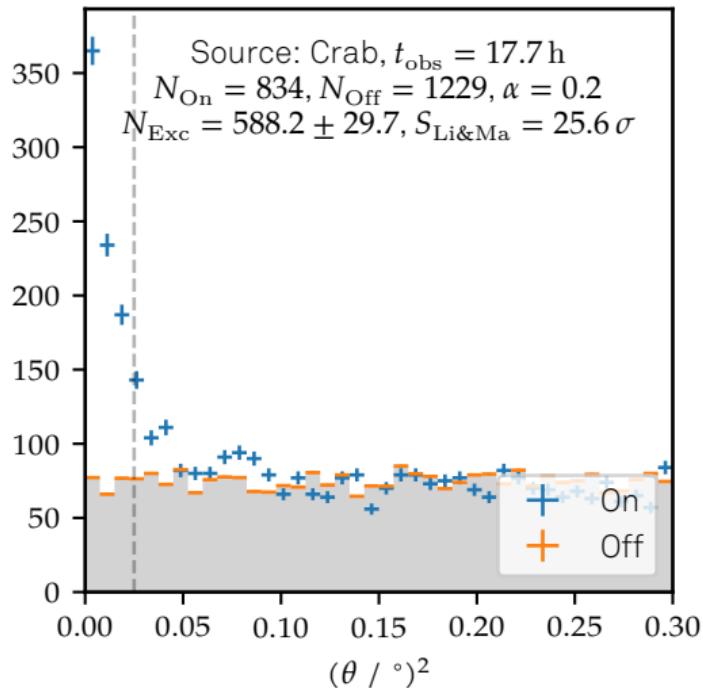
## Energy estimation $R^2$ -score = 0.740



## Separation

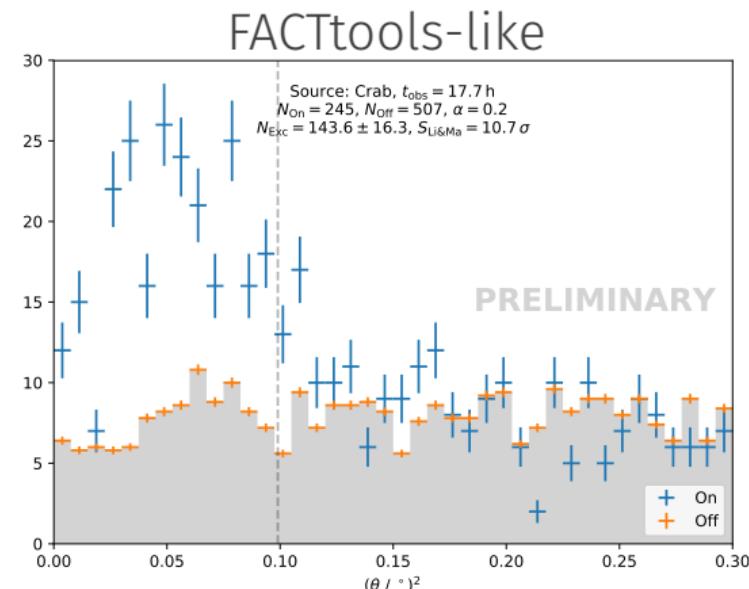
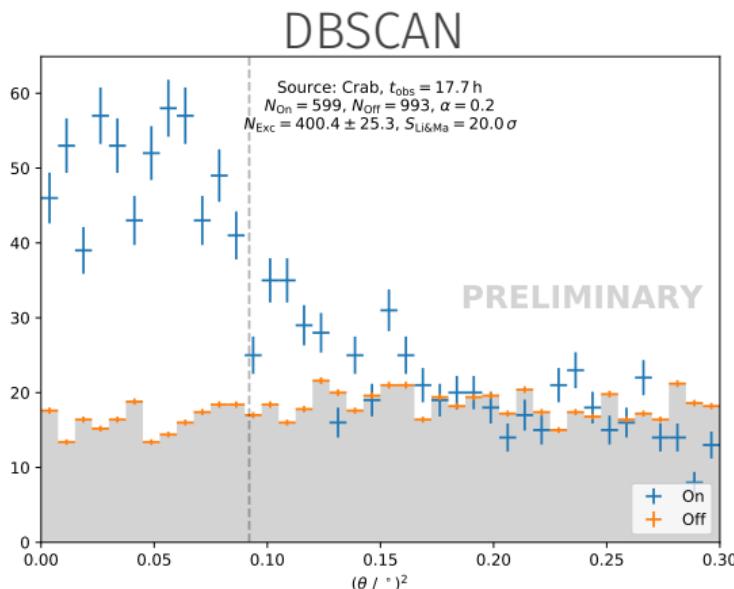


## Origin reconstruction (FACTtools)

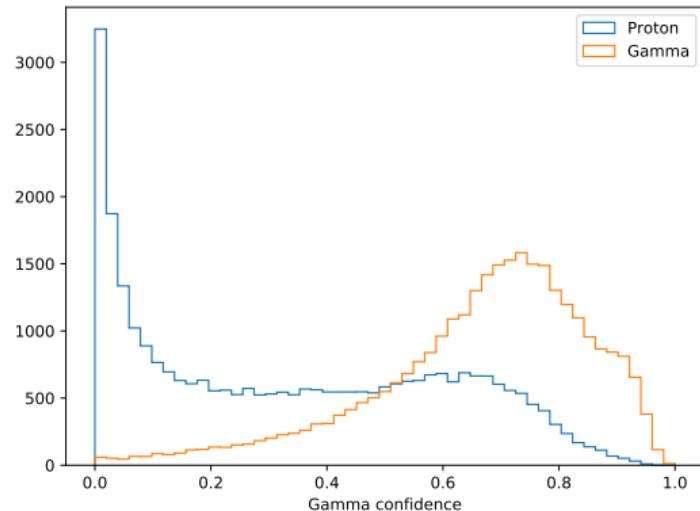
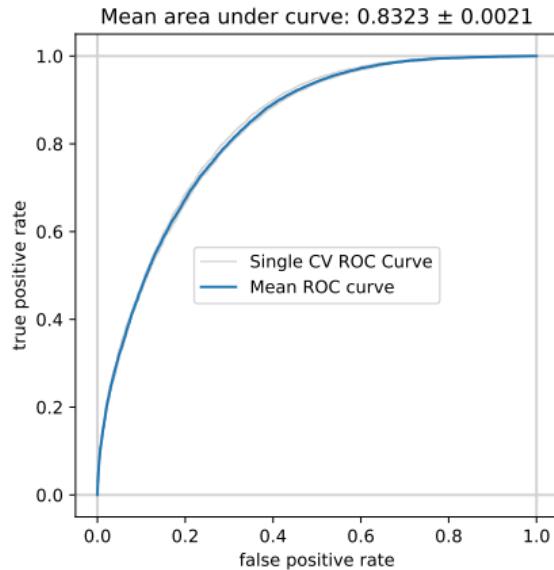


## Results

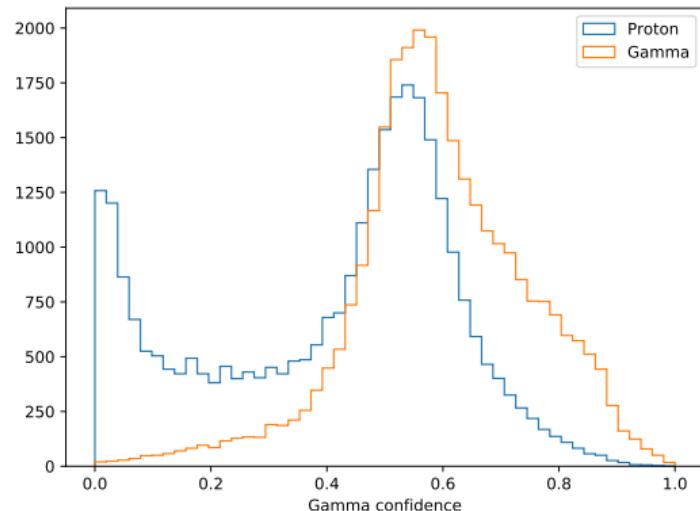
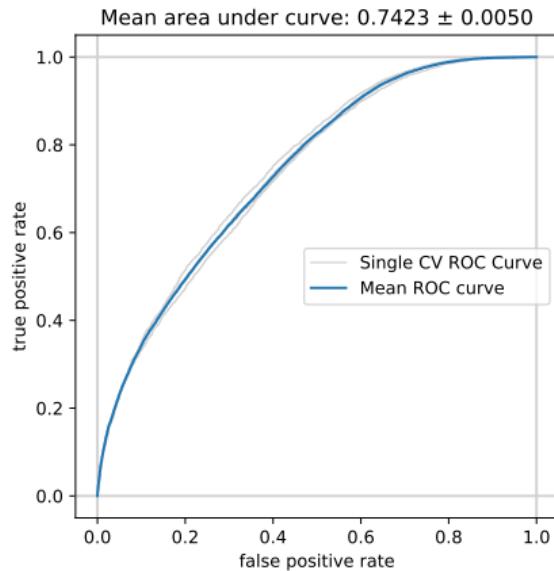
## Origin reconstruction (PhotonStream)



## Separation (DBSCAN)

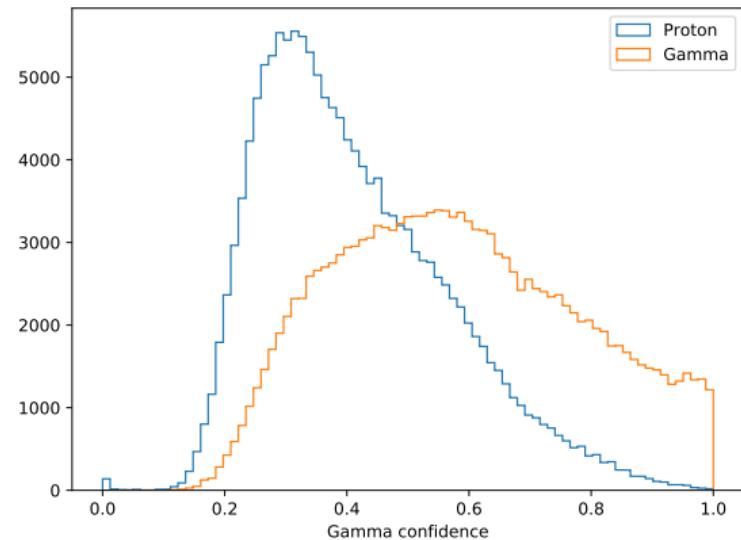
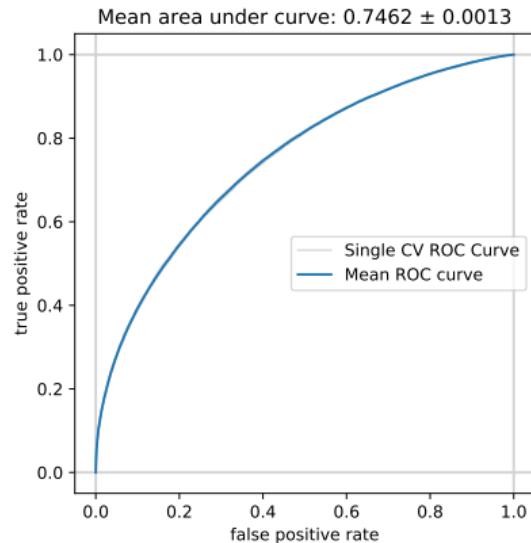


## Separation performance (pixel based)

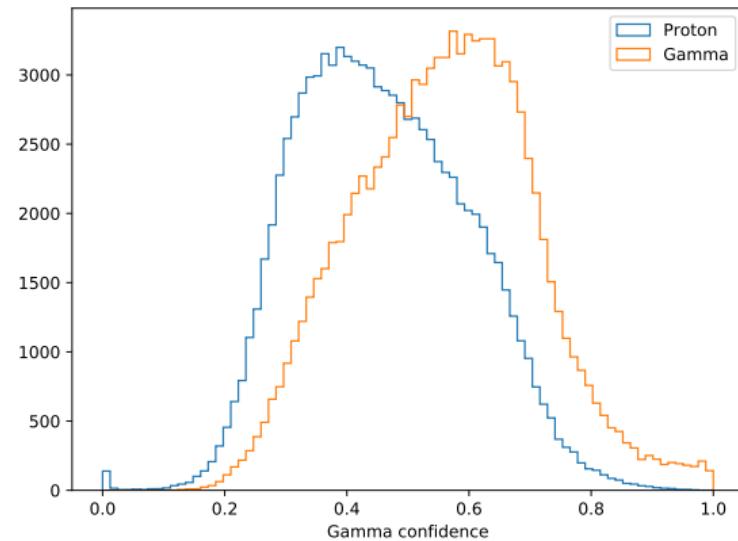
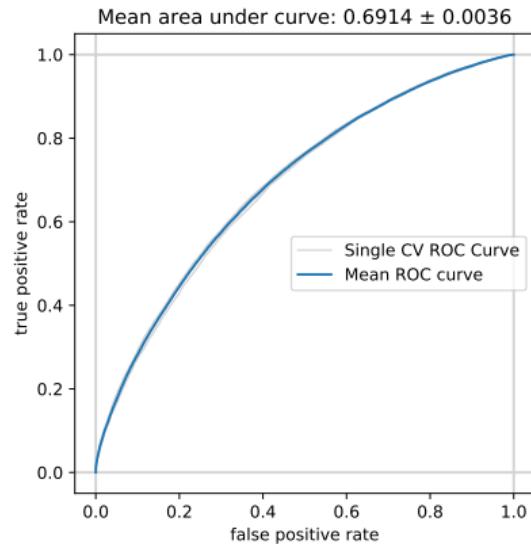


# Data-Simulation mismatches

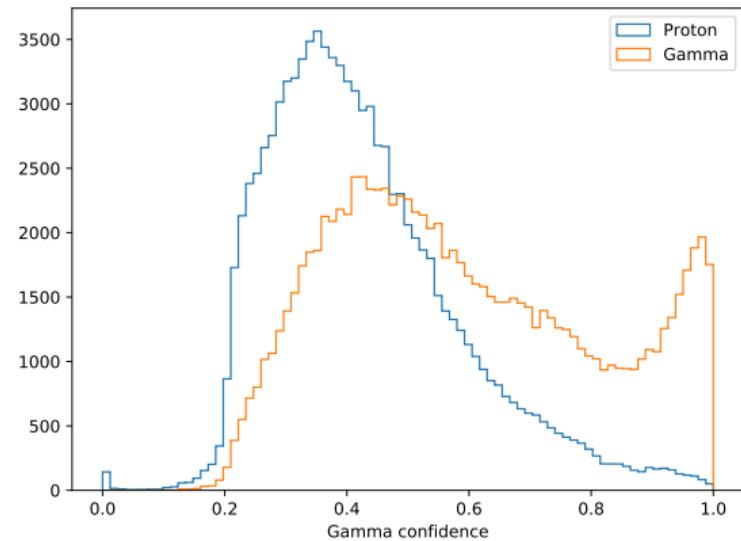
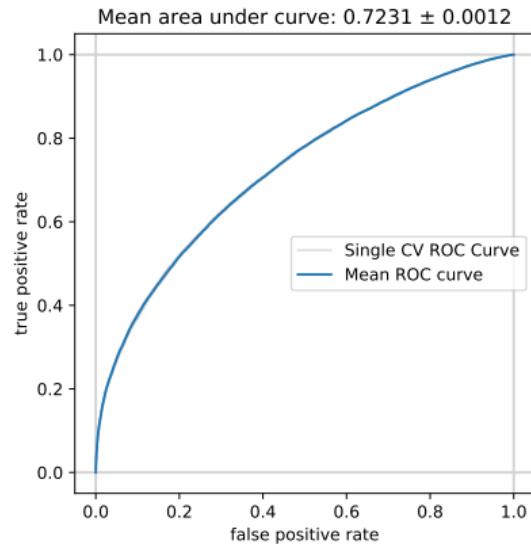
## Data-Simulation mismatches: Random Forest (DBSCAN)



## Data-Simulation mismatches: Random Forest (DBSCAN cuts)

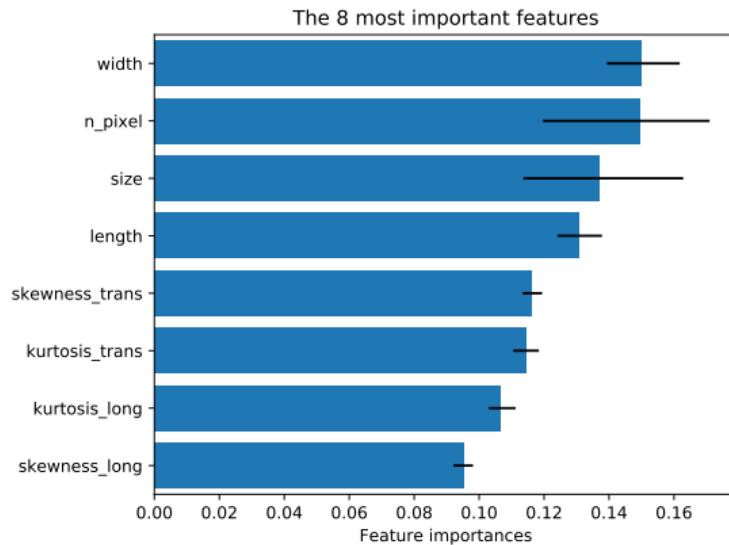


## Data-Simulation mismatches: Random Forest (thresholds)

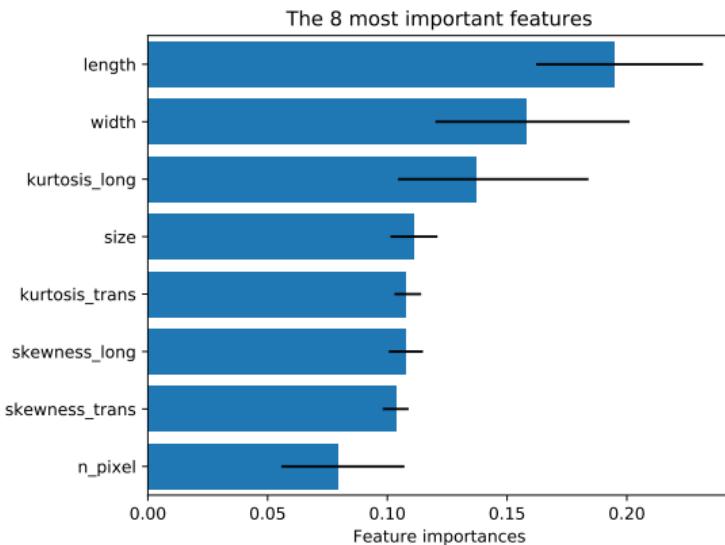


## Data-Simulation mismatches: Random Forest

DBSCAN

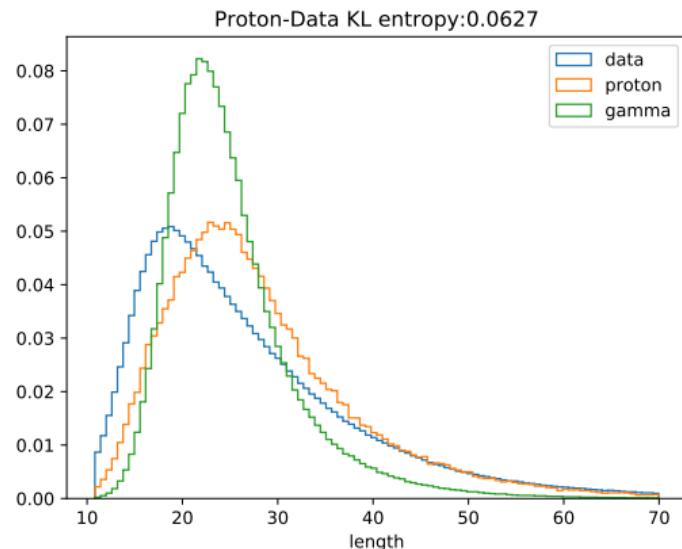
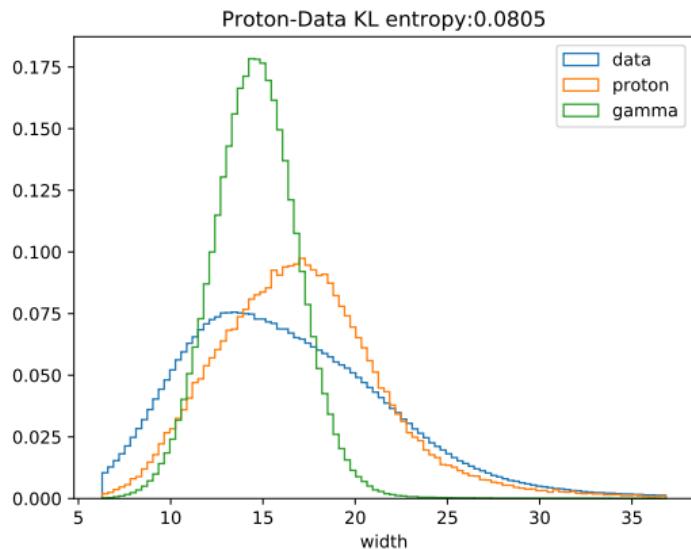


FACTtools



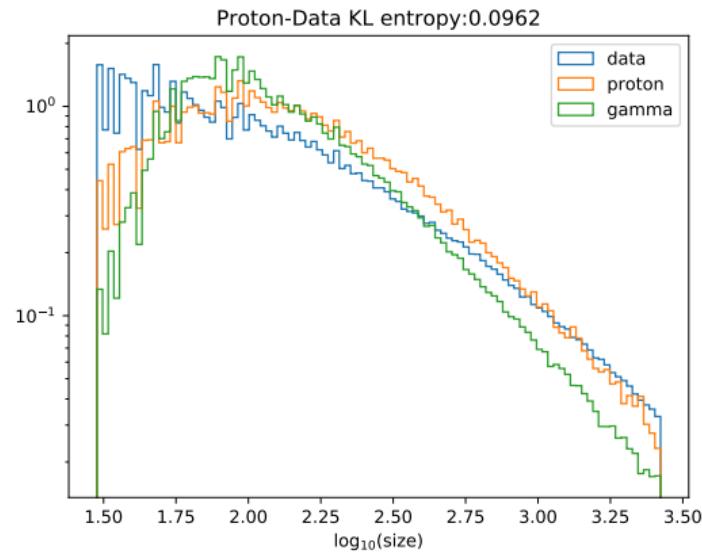
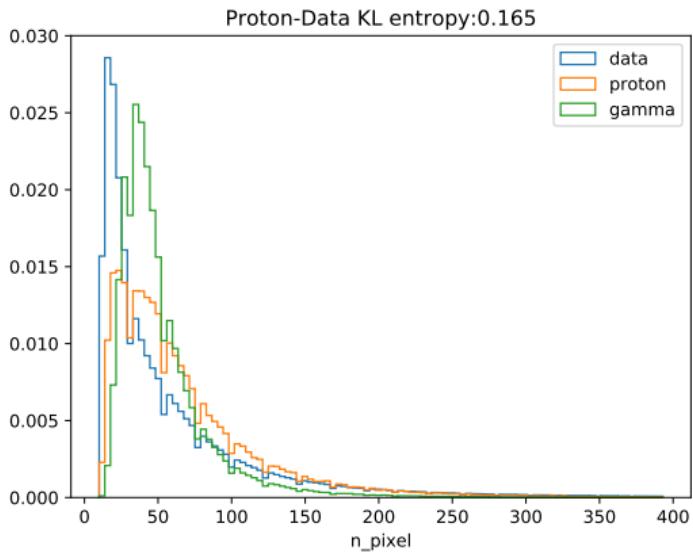
## Data-Simulation mismatches

DBSCAN



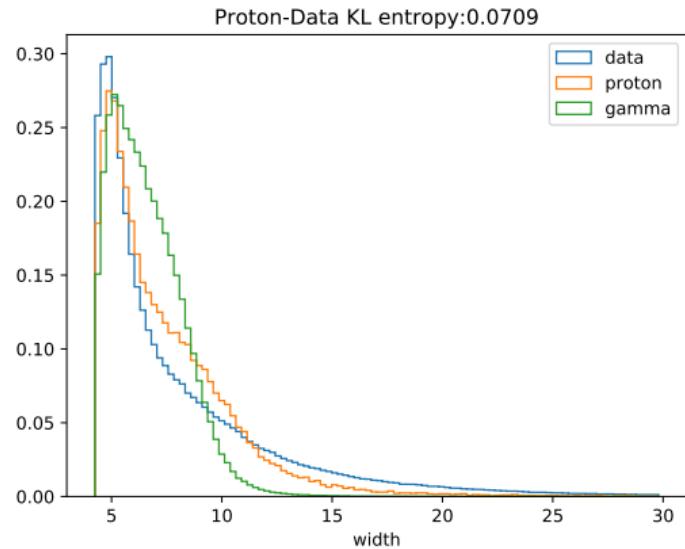
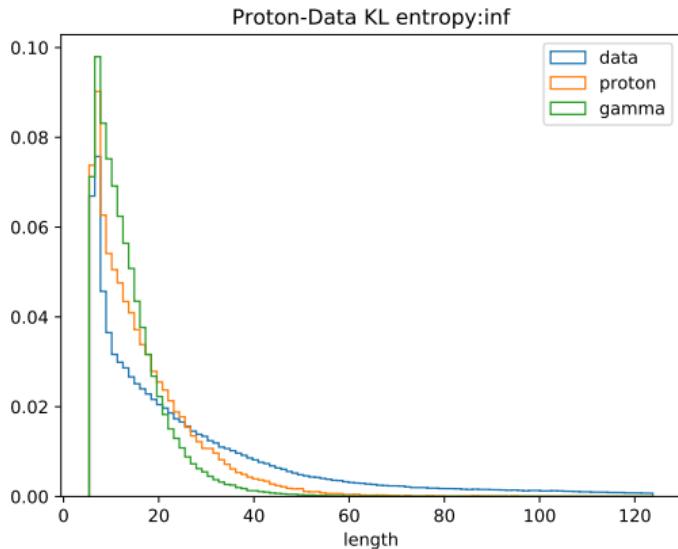
## Data-Simulation mismatches

DBSCAN



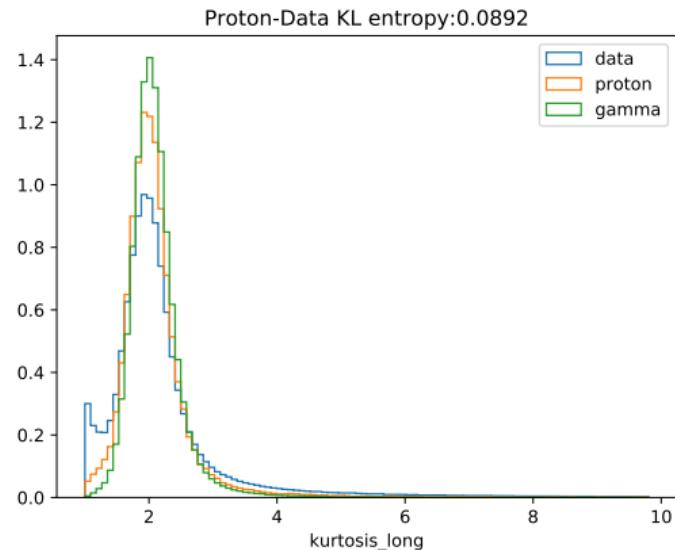
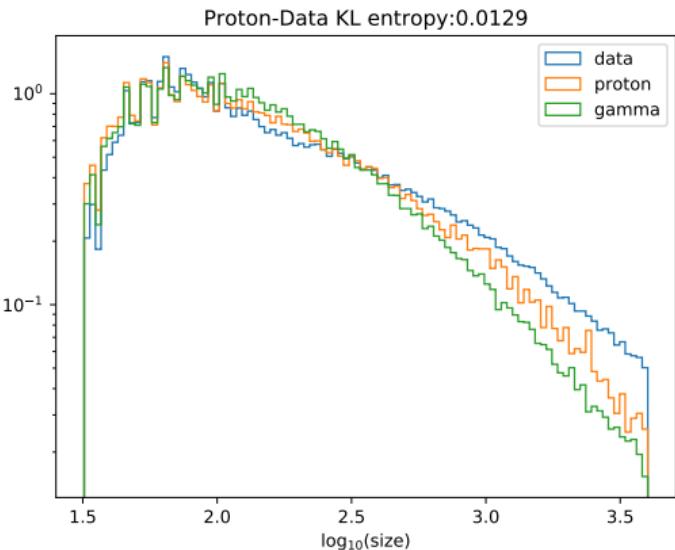
## Data-Simulation mismatches

FACTtools cleaning



## Data-Simulation mismatches

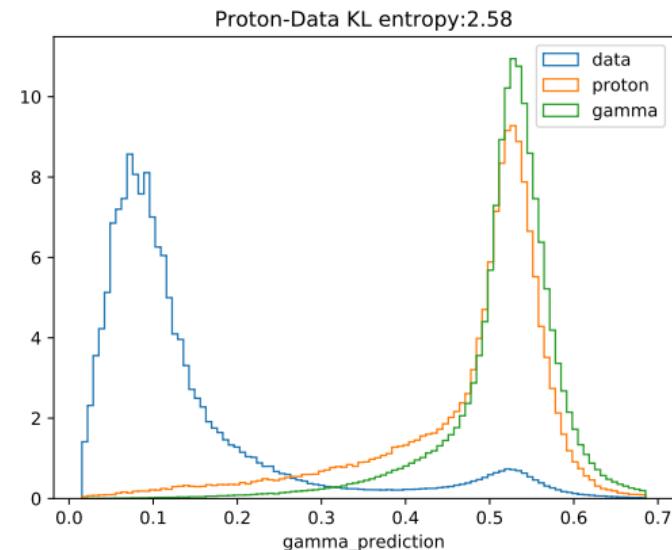
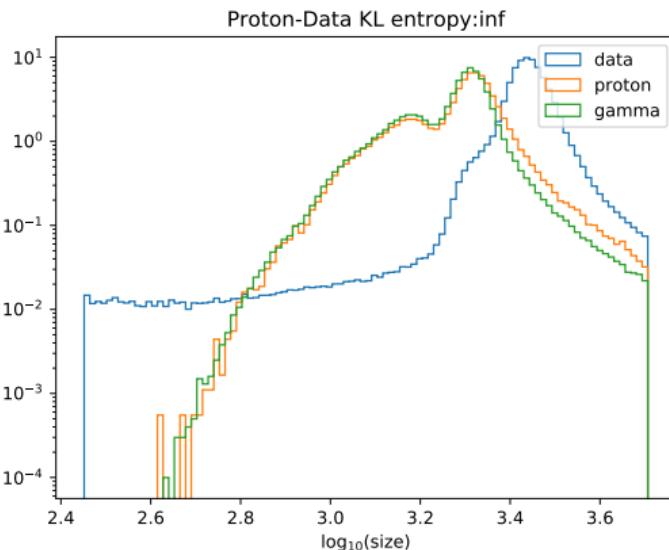
FACTtools cleaning



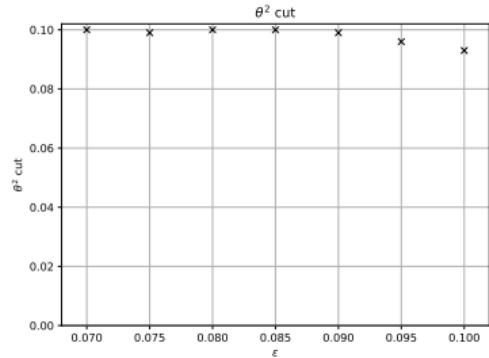
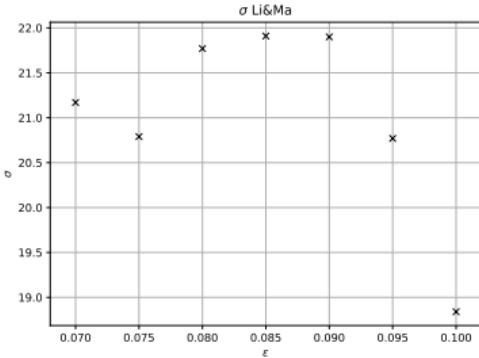
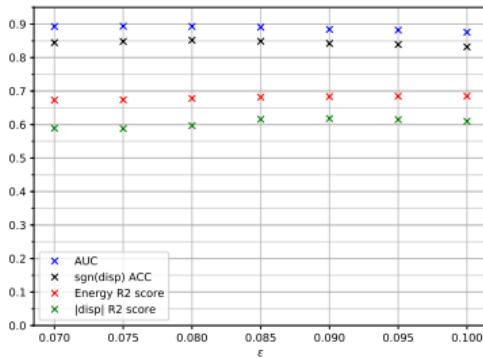
## Effects of DBSCAN parameters

$\varepsilon$  maximum distance between each photon to be considered dense

$$0.1 \rightarrow 0.175$$



## Effects of DBSCAN parameters

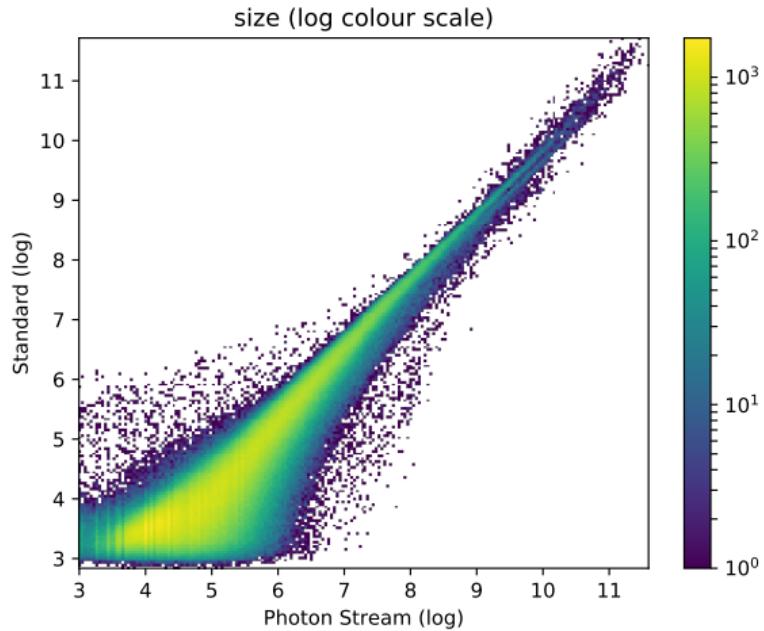




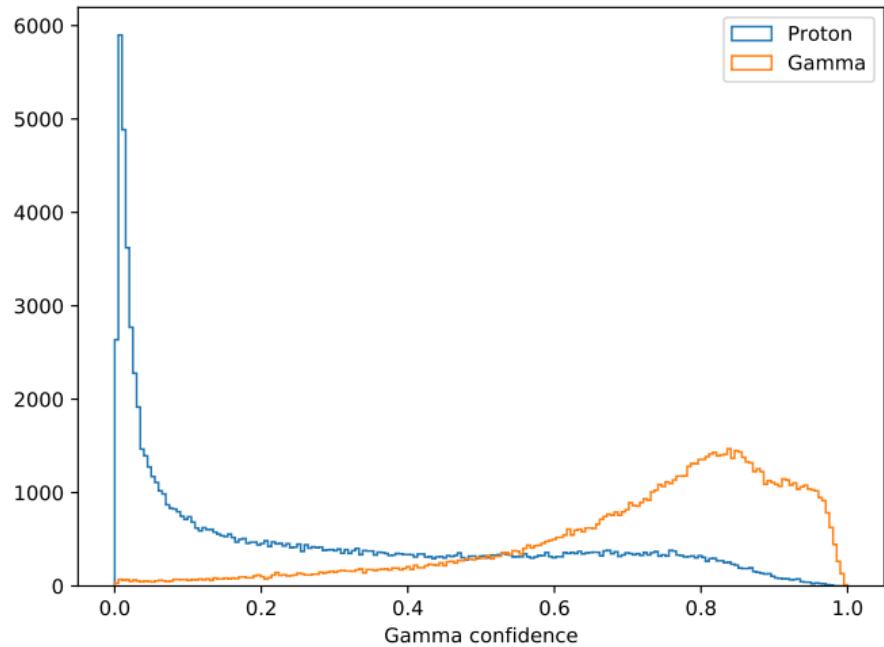
# Back up

Rare photo of an experiment with no issues  
in their Monte Carlo event generator

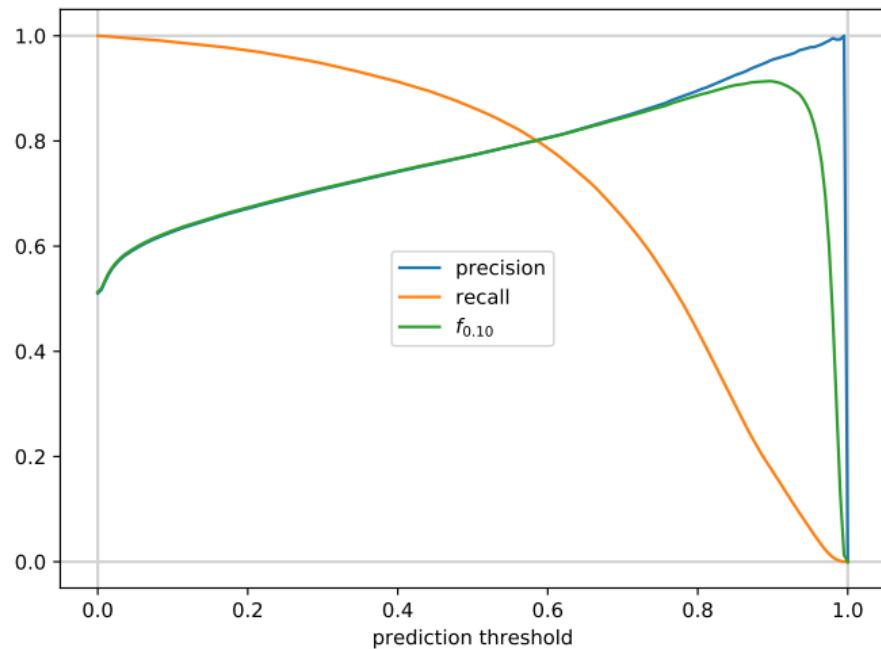




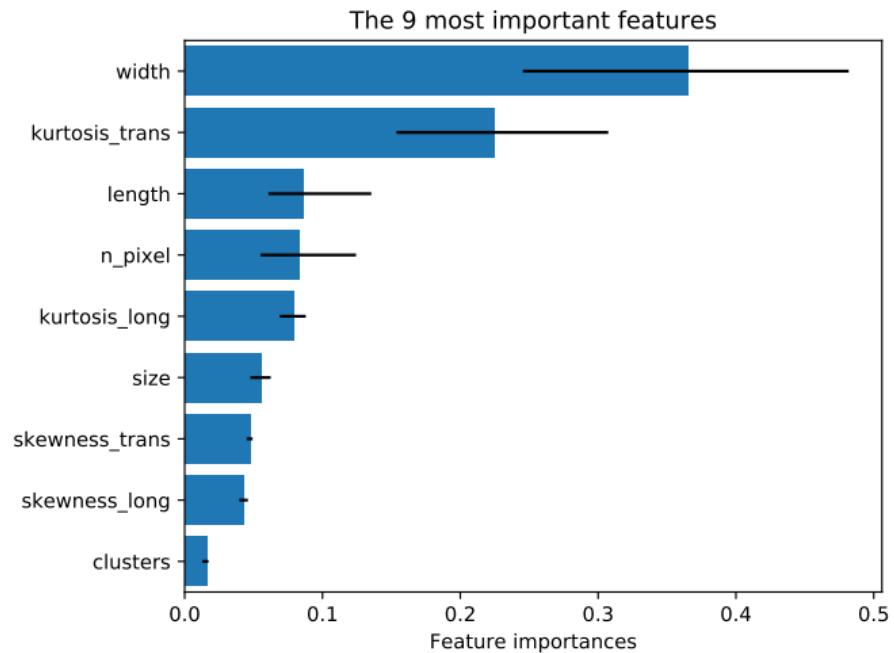
## Separation



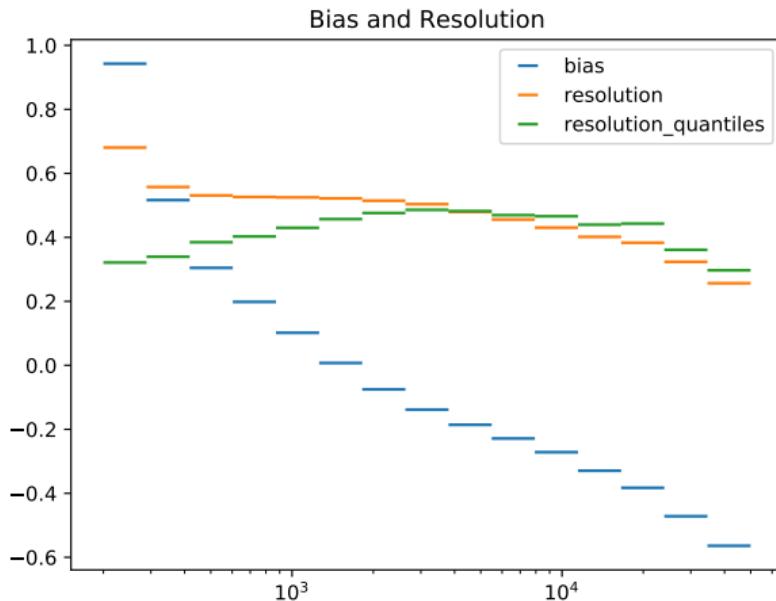
## Separation



## Separation



# Energy



# Energy

