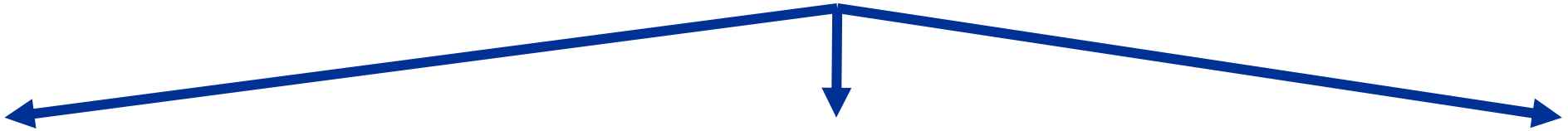
Fully connected neural network for background suppression in $\overline{He}$ research

Francesco Rossi

Sources of charged confused events

Using He Monte Carlo B1236 L1-focused and L1-L9 focused, and selecting the **reconstructed events with $\mathbf{R < 0}$**, we identified **three sources of charge confusion**



Spillover

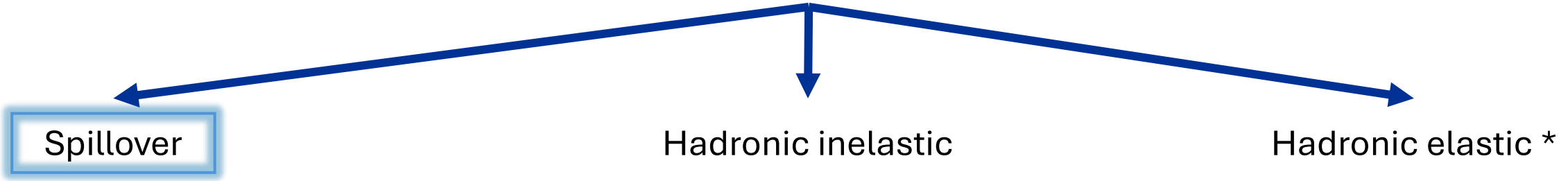
Hadronic inelastic

Hadronic elastic *

* Large angle scattering

Sources of charged confused events

Using He Monte Carlo B1236 L1-focused and B1236 L1-L9 focused, and selecting the **reconstructed events with $\mathbf{R < 0}$**, we identified **three sources of charge confusion**

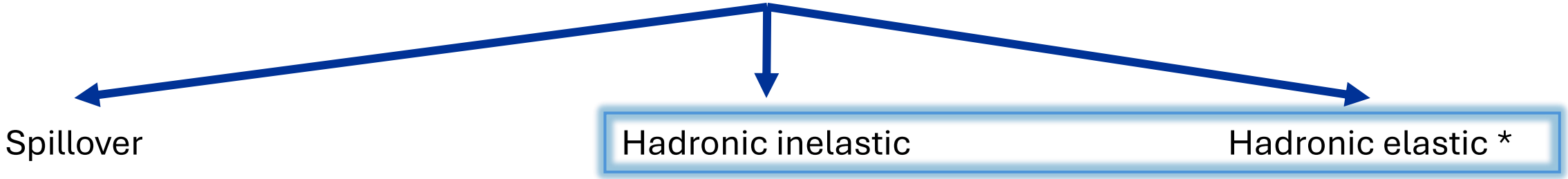


Silicon tracker finite resolution

* Large angle scattering

Sources of charged confused events

Using He Monte Carlo B1236 L1-focused and B1236 L1-L9 focused, and selecting the **reconstructed events with $\mathbf{R < 0}$**, we identified **three sources of charge confusion**

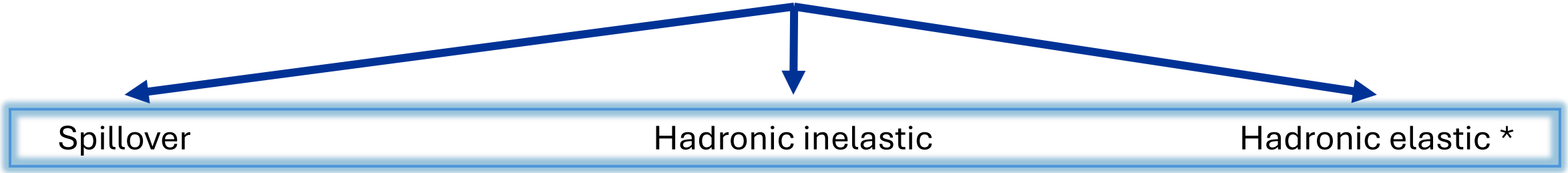


Interactions within the detector

* Large angle scattering

Sources of charged confused events

Using He Monte Carlo B1236 L1-focused and B1236 L1-L9 focused, and selecting the **reconstructed events with $\mathbf{R < 0}$**, we identified **three sources of charge confusion**

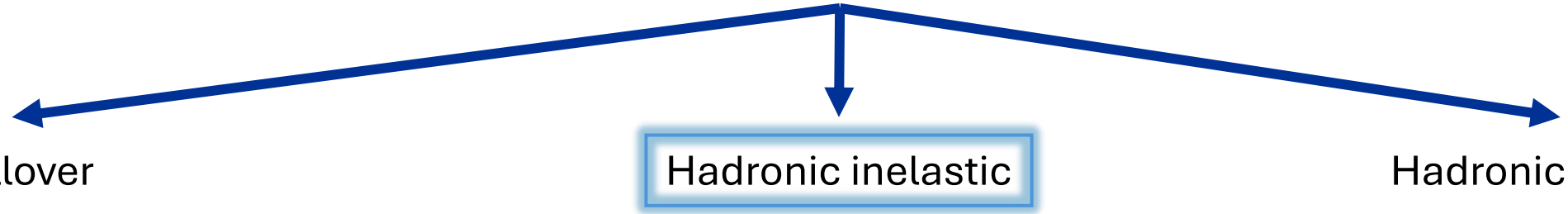


For each source, we select a sample to be used in the training of a Fully Connected Neural Network

* Large angle scattering

Sources of charged confused events

Using He Monte Carlo B1236 L1-focused and B1236 L1-L9 focused, and selecting the **reconstructed events with $\mathbf{R < 0}$**, we identified **three sources of charge confusion**

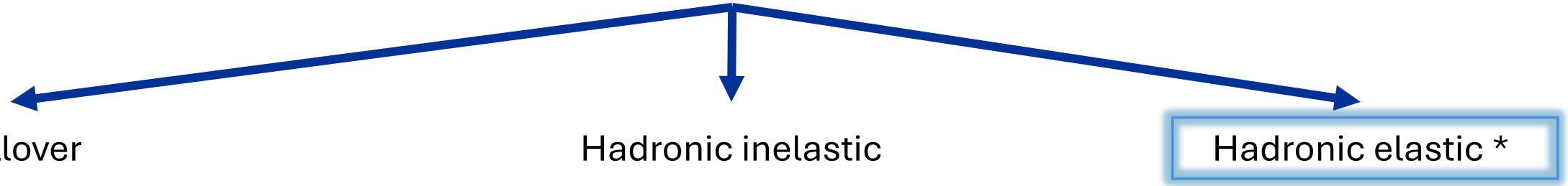


Search in the secondaries list looking for inelastic interaction products inside the inner tracker

* Large angle scattering

Sources of charged confused events

Using He Monte Carlo B1236 L1-focused and B1236 L1-L9 focused, and selecting the **reconstructed events with $\mathbf{R < 0}$**, we identified **three sources of charge confusion**



P. Zuccon added a MC flag for events interacting elastically at gBatch level.

The tag is added every time that a discrete elastic scattering is called.
It stores info about the scattering nuclei, scattering angle, initial and final momentum of primary particle.

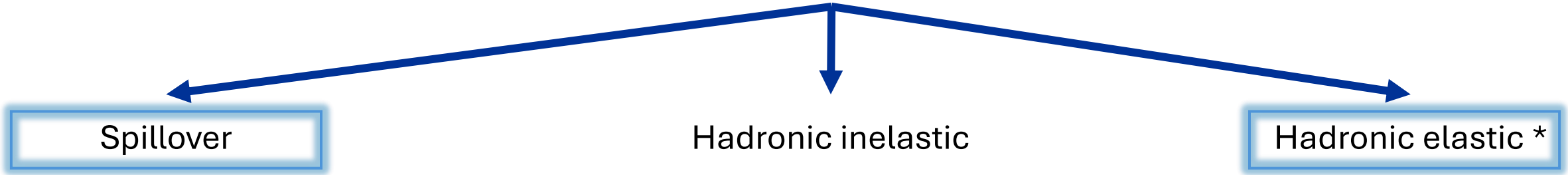
More info available in NAIA pzv1.2.0 on Gitlab.

1'000 naia ntuple produced on Trento AMS server, $R_{true} \leq 200$ [GV].

* Large angle scattering

Sources of charged confused events

Using He Monte Carlo B1236 L1-focused and B1236 L1-L9 focused, and selecting the **reconstructed events with $\mathbf{R < 0}$**, we identified **three sources of charge confusion**



Search in the secondaries list looking for secondaries produced inside the inner tracker (HasSecondary)

Events are not Had. Inel. Interactions and primary nuclei reaches L2

Propagation of two tracks: $R_{true}(L2)$ and $R_{inner}(< 0)$

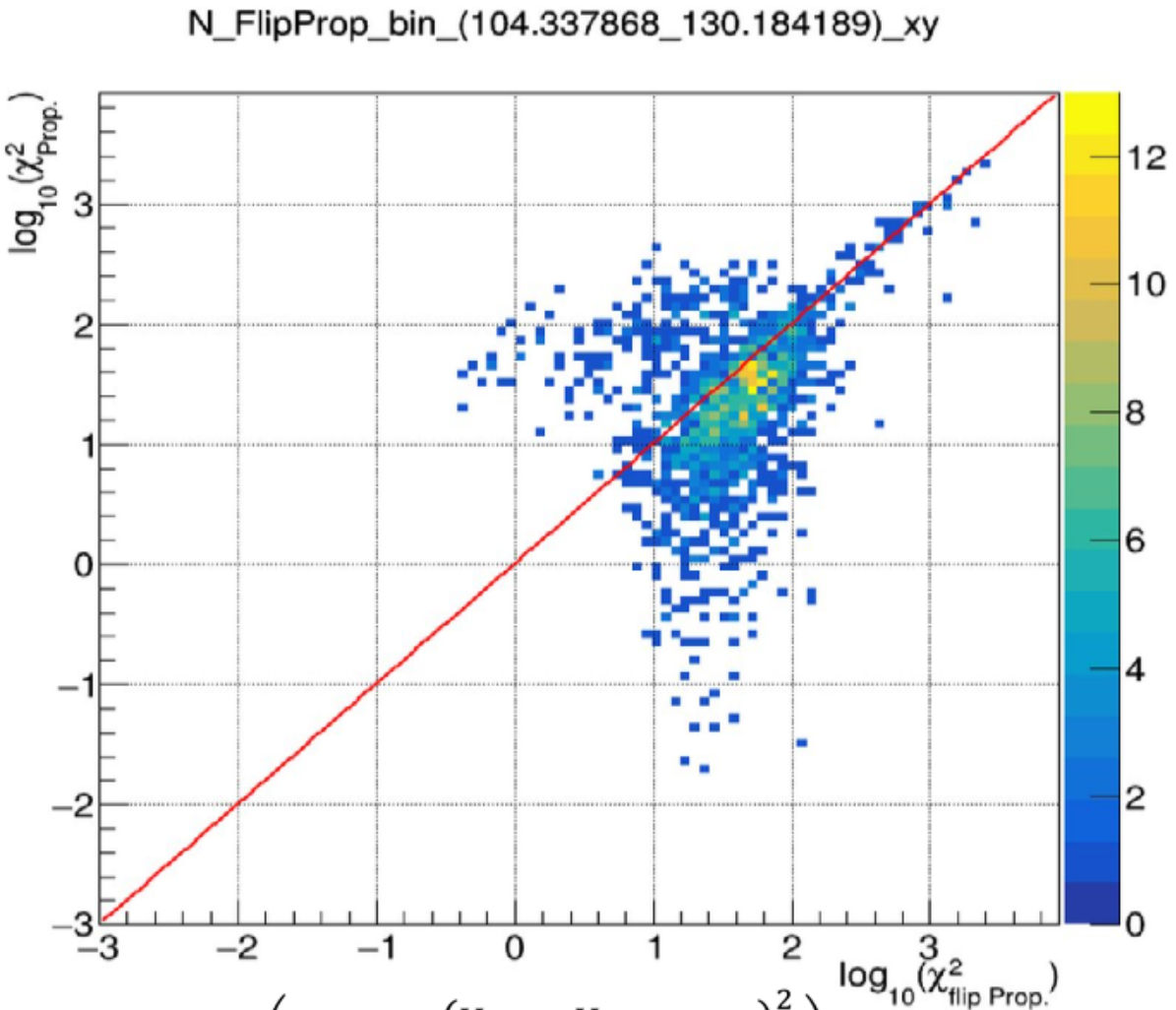
Build two χ^2 comparing y coordinate with MC true info on each layer.

$$\frac{\chi^2_{R_{inner}}}{\chi^2_{R_{true}}} \leq 0.85 \quad \rightarrow \text{El. scat.}$$

$$1.15 \leq \frac{\chi^2_{R_{inner}}}{\chi^2_{R_{true}}} \quad \rightarrow (\text{HasSecondary})\ ?\ \text{Other}\ :\ \text{Spillover}$$

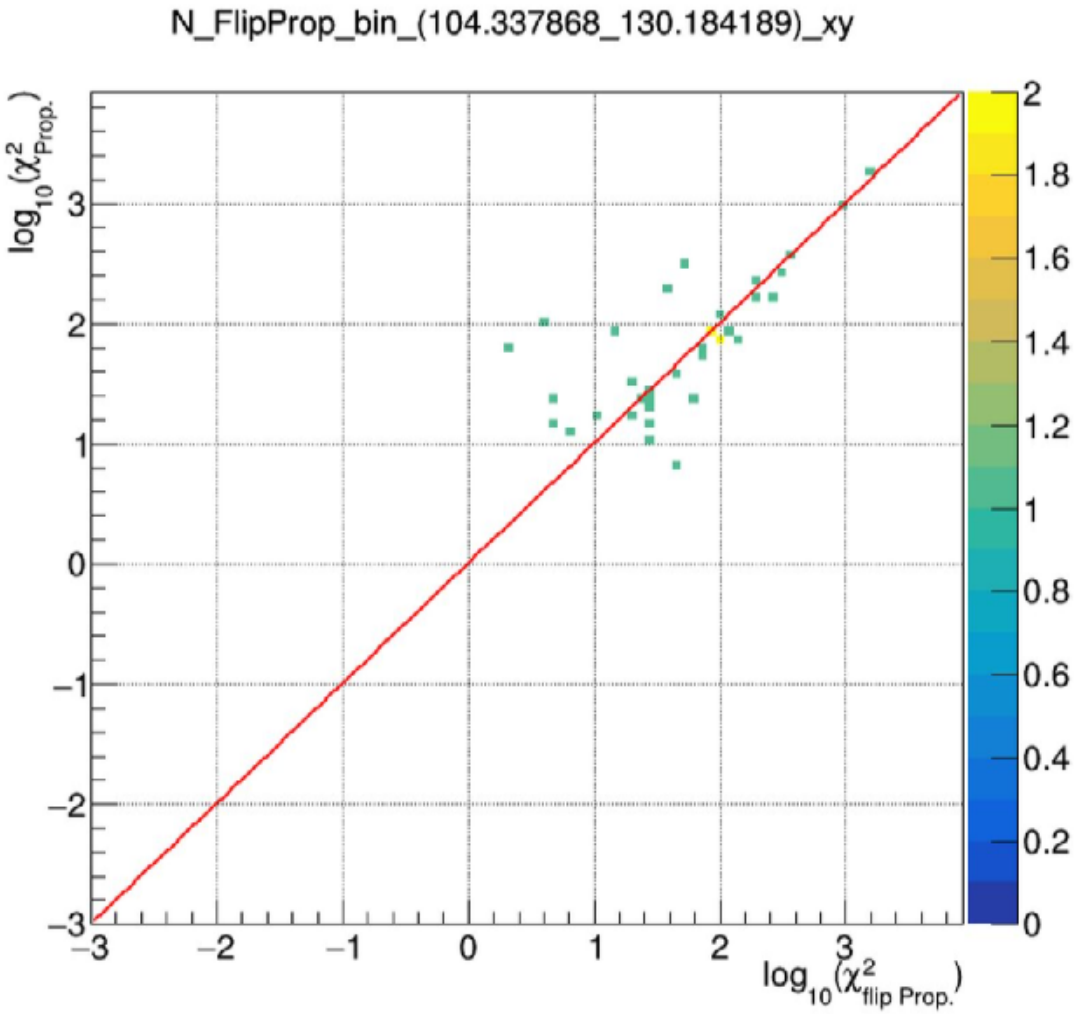
$$0.85 < \frac{\chi^2_{R_{inner}}}{\chi^2_{R_{true}}} \leq 1.15 \rightarrow (\text{HasSecondary})\ ?\ \text{Other}\ :\ \text{Spillover}$$

$R_{inner} < 0$ (no secondaries within inner tracker)



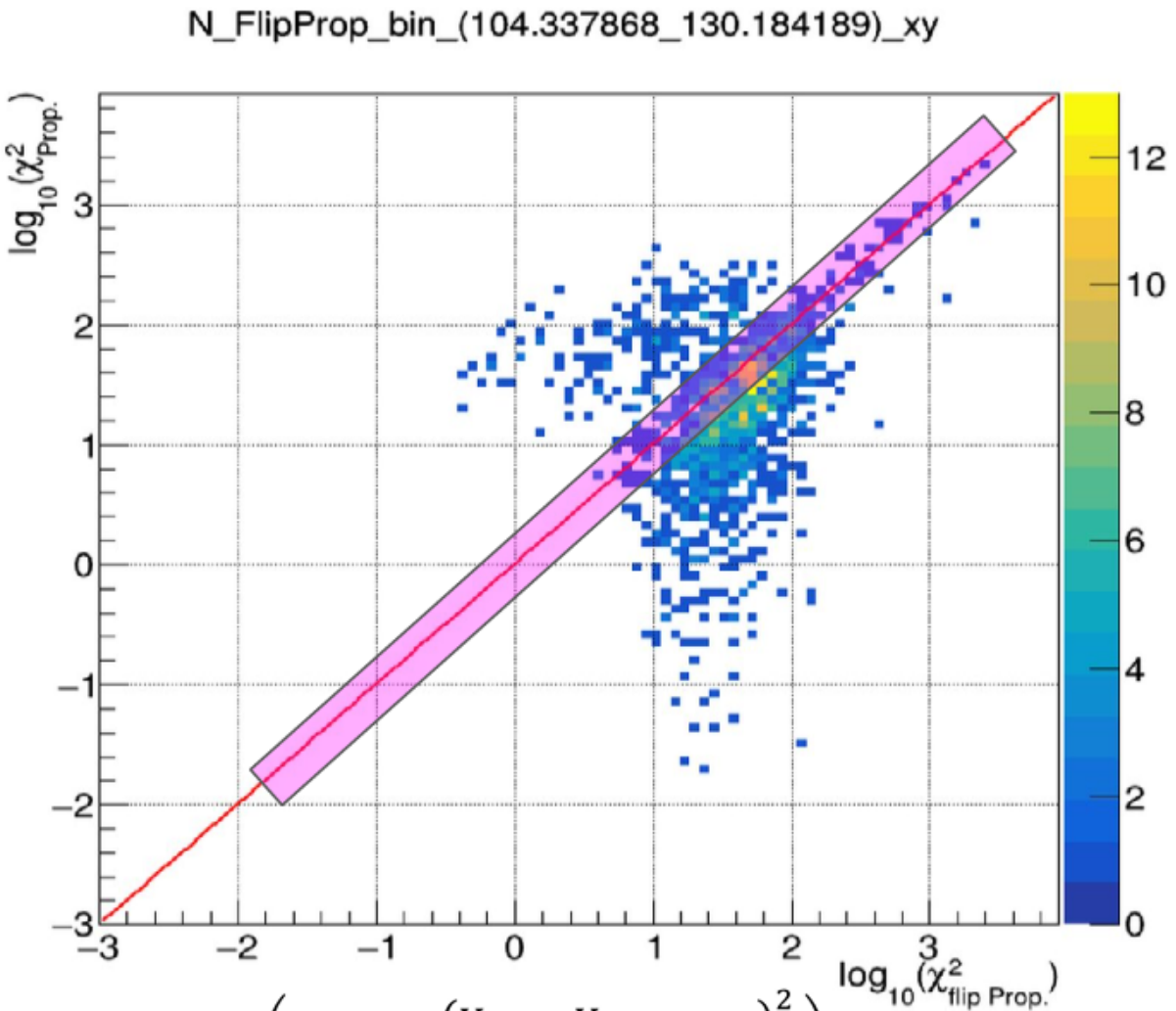
$$\chi^2_{prop} = \frac{\left(\sum_{i-layer} \frac{(Y_{MC} - Y_{prop.track})^2}{\sigma^2_{hit}}\right)}{Ndof}$$

$R_{inner} < 0$ (elastic scattering within inner tracker)



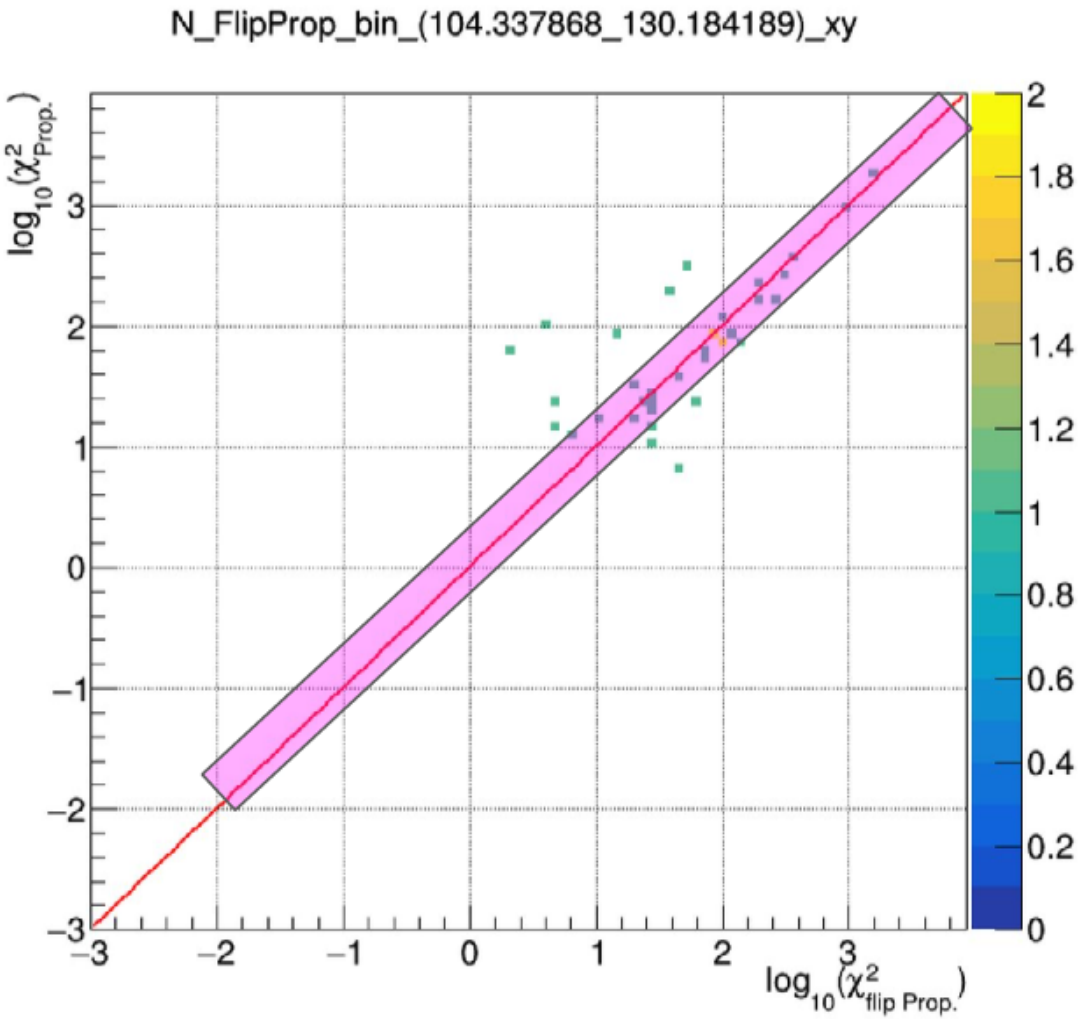
$$\sigma^2_{hit} = 15\ \mu m$$

$R_{inner} < 0$ (no secondaries within inner tracker)



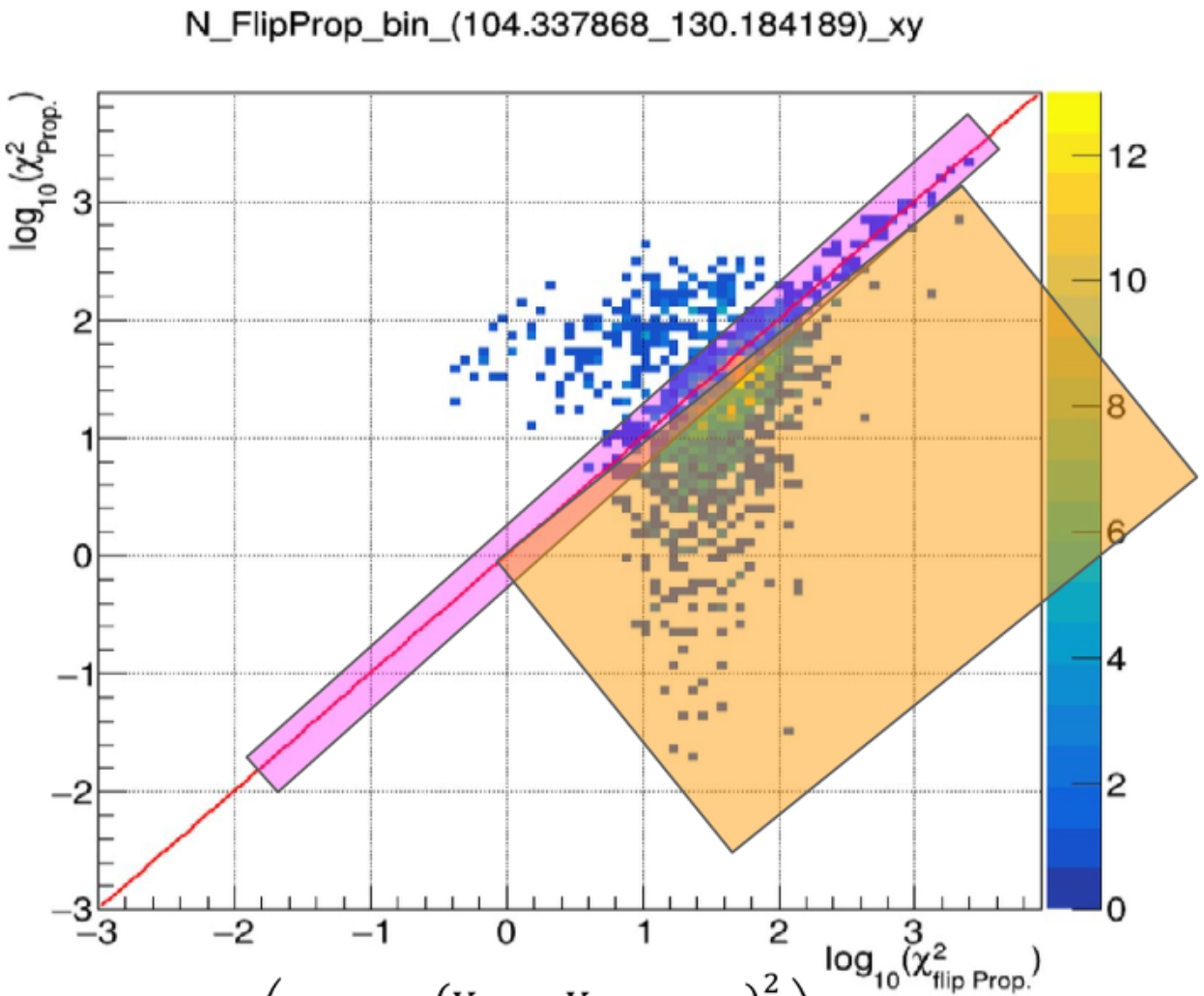
$$\chi^2_{prop} = \frac{\left(\sum_{i-layer} \frac{(Y_{MC} - Y_{prop.track})^2}{\sigma^2_{hit}}\right)}{Ndof}$$

$R_{inner} < 0$ (elastic scattering within inner tracker)



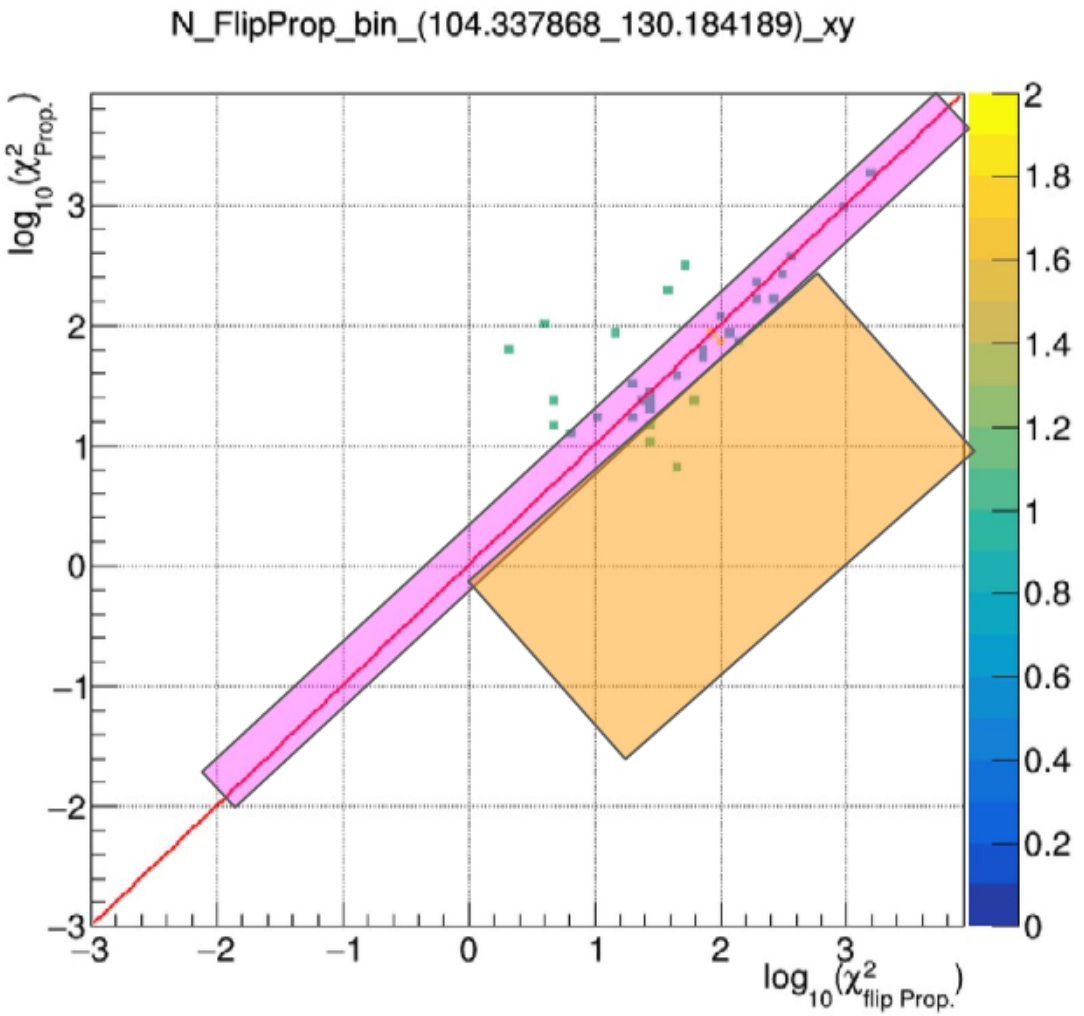
$$\sigma^2_{hit} = 15\ \mu m$$

$R_{inner} < 0$ (no secondaries within inner tracker)



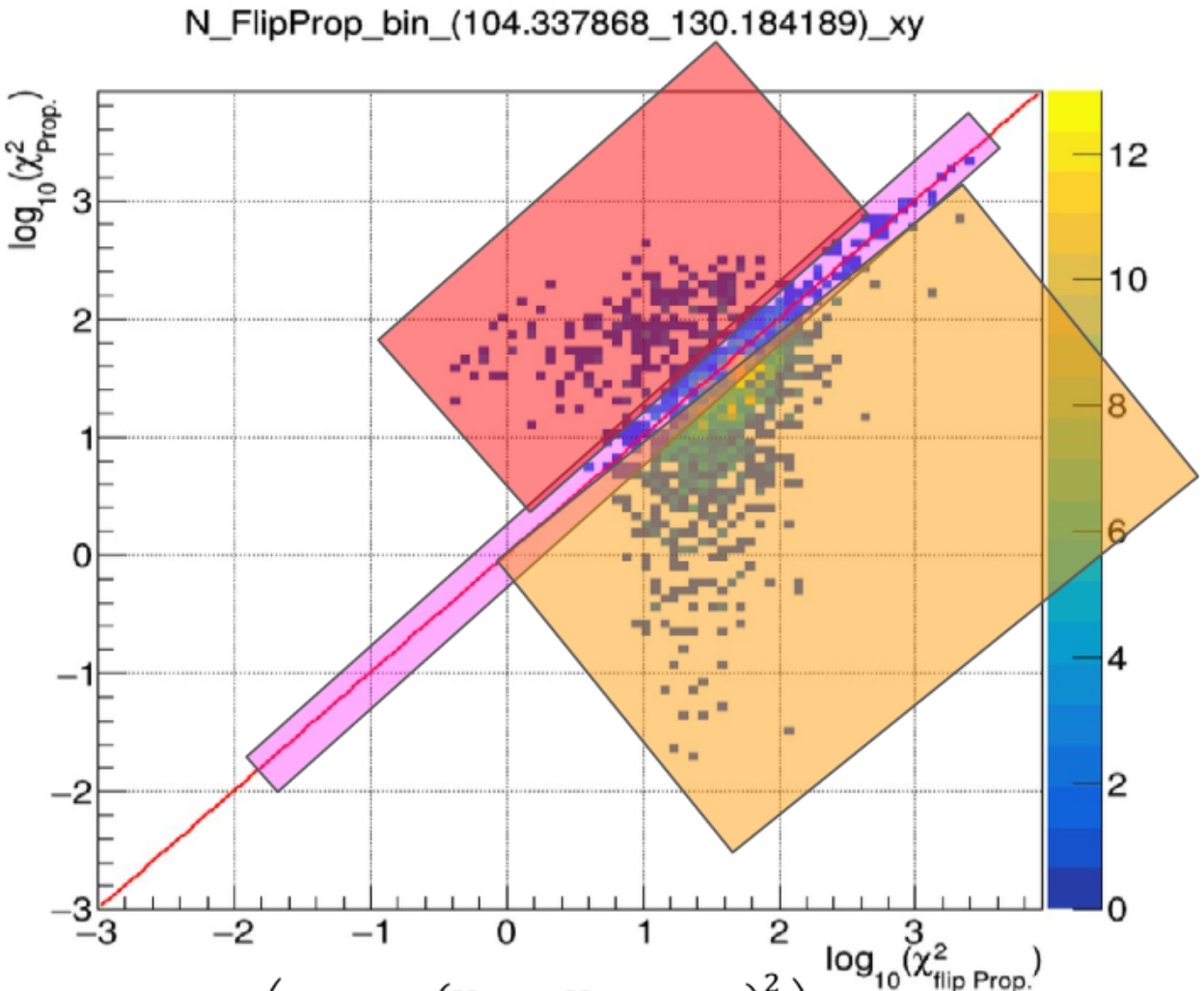
$$\chi^2_{prop} = \frac{\left(\sum_{i-layer} \frac{(Y_{MC} - Y_{prop.track})^2}{\sigma^2_{hit}}\right)}{Ndof}$$

$R_{inner} < 0$ (elastic scattering within inner tracker)



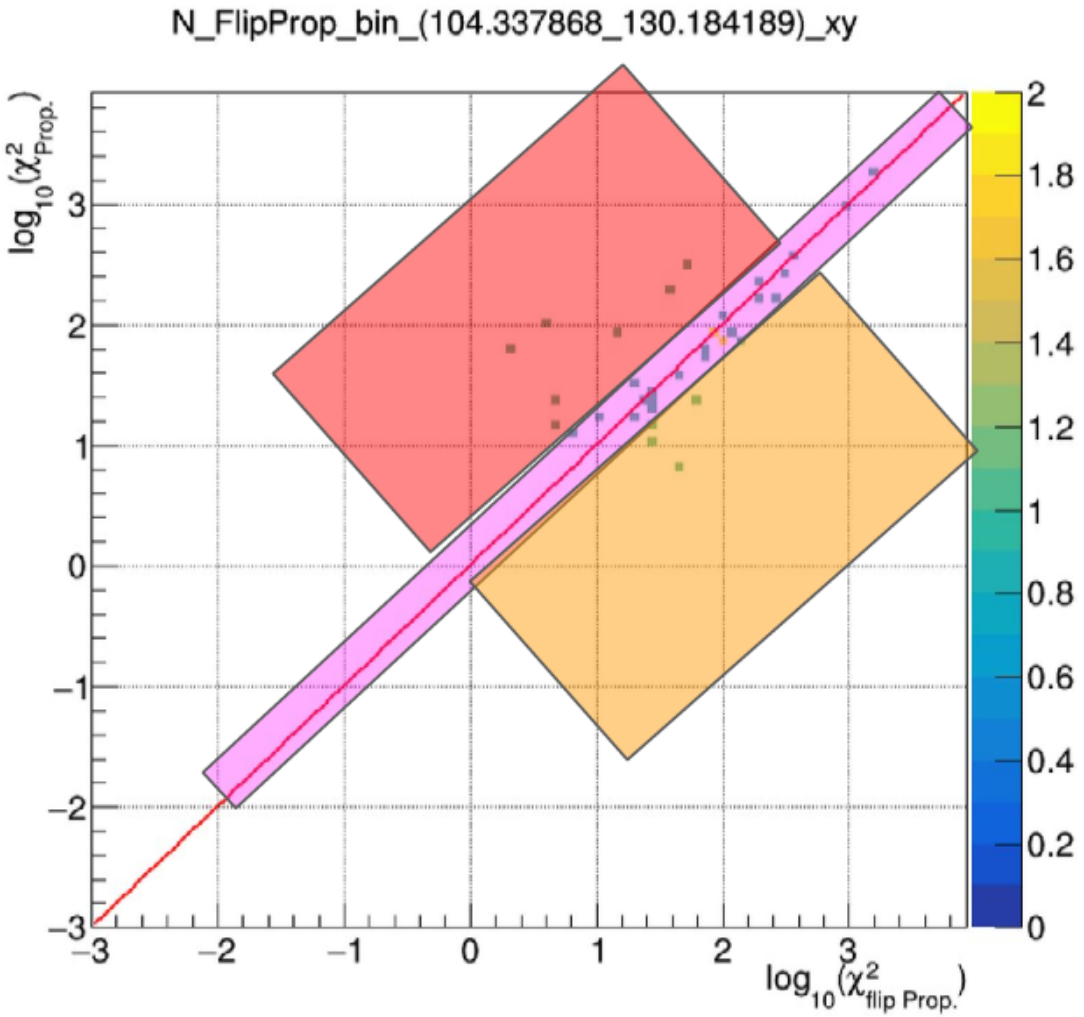
$$\sigma^2_{hit} = 15\ \mu m$$

$R_{inner} < 0$ (no secondaries within inner tracker)



$$\chi^2_{prop} = \frac{\left(\Sigma_{i-layer} \frac{(Y_{MC} - Y_{prop.track})^2}{\sigma^2_{hit}}\right)}{Ndof}$$

$R_{inner} < 0$ (elastic scattering within inner tracker)



$$\sigma^2_{hit} = 15\ \mu m$$

Monte carlo selection (He B1236 L1-focused and L1-L9 focused)

IsPhysicsTrigger

TOF hits ≥ 4
$\beta > 0$
$\chi^2_{Coo} < 4$

Track number ≥ 1
track pattern L1&L2&(L3|L4)&(L5|L6)&(L7|L8) (≥ 5)

charge STD (inner) $\in [1.7, 2.4]$
charge STD (L1) $\in [1.6, 3.0]$
Inner fiducial volume
L1 fiducial volume

$\chi^2_Y < 10$
charge UTOF $\in [1.5, 3.0]$
charge LTOF ≥ 1.5

Sign inner upper half (UH) and lower half (LH)

Tracker fiducial volume cut:
- L1: |R|<62cm, |Y|<47cm
- L2: |R|<62cm, |Y|<40cm
- L3: |R|<46cm, |Y|<44cm
- L4: |R|<46cm, |Y|<44cm
- L5: |R|<46cm, |Y|<36cm
- L6: |R|<46cm, |Y|<36cm
- L7: |R|<46cm, |Y|<44cm
- L8: |R|<46cm, |Y|<44cm

UH = Rig. [UH-inner]
LH = Rig. [LH-inner]
Inner = Rig. [inner]

If(Inner > 0) $\rightarrow (UH > 0, LH > 0)$
If(Inner < 0) $\rightarrow (UH < 0, LH < 0)$

Monte carlo selection (He B1236 L1-focused and L1-L9 focused)

IsPhysicsTrigger

TOF hits ≥ 4
$\beta > 0$
$\chi^2_{Coo} < 4$

Track number ≥ 1
track pattern L1&L2&(L3|L4)&(L5|L6)&(L7|L8) (≥ 5)

charge STD (inner) $\in [1.7, 2.4]$
charge STD (L1) $\in [1.6, 3.0]$
Inner fiducial volume
L1 fiducial volume

$\chi^2_Y < 10$
charge UTOF $\in [1.5, 3.0]$
charge LTOF ≥ 1.5

Sign inner upper half (UH) and lower half (LH)

Composition:

76% Spillover,
21% El. Scat.
2% Had. Inel
0.04% Other

~ $0.9 \cdot 10^6$ training

~ $0.4 \cdot 10^6$ testing

Input features for Fully Connected Neural Network (FCNN)

- Distance between track hit and strip with max energy deposit on L8.
- Number of hits in L3-L8 inner tracker Y side.
- Track hit energy deposition on L8.
- Total energy deposition of tracker clusters on L8 Y-side.
- Ratio between cluster amplitude and its neighbouring 10 strips (L8)
- Max cluster energy deposit on L8 Y-side.
- ACC clusters
- Number of hits in L3-L8 inner tracker X side.
- Track hit energy deposition on L7.
- Total energy deposition of tracker clusters on L7 Y-side.
- Ratio between cluster amplitude and its neighbouring 10 strips (L2)
- Max cluster energy deposit on L7 Y-side.

Input features for Fully Connected Neural Network (FCNN)

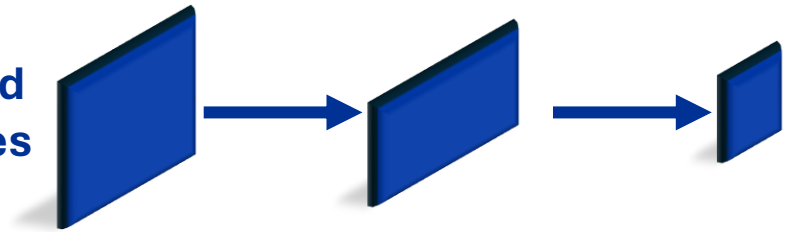
- Distance between track hit and strip with max energy deposit on L8.
- Number of hits in L3-L8 inner tracker Y side.
- Track hit energy deposition on L8.
- Total energy deposition of tracker clusters on L8 Y-side.
- Ratio between cluster amplitude and its neighbouring 10 strips (L8)
- Max cluster energy deposit on L8 Y-side.

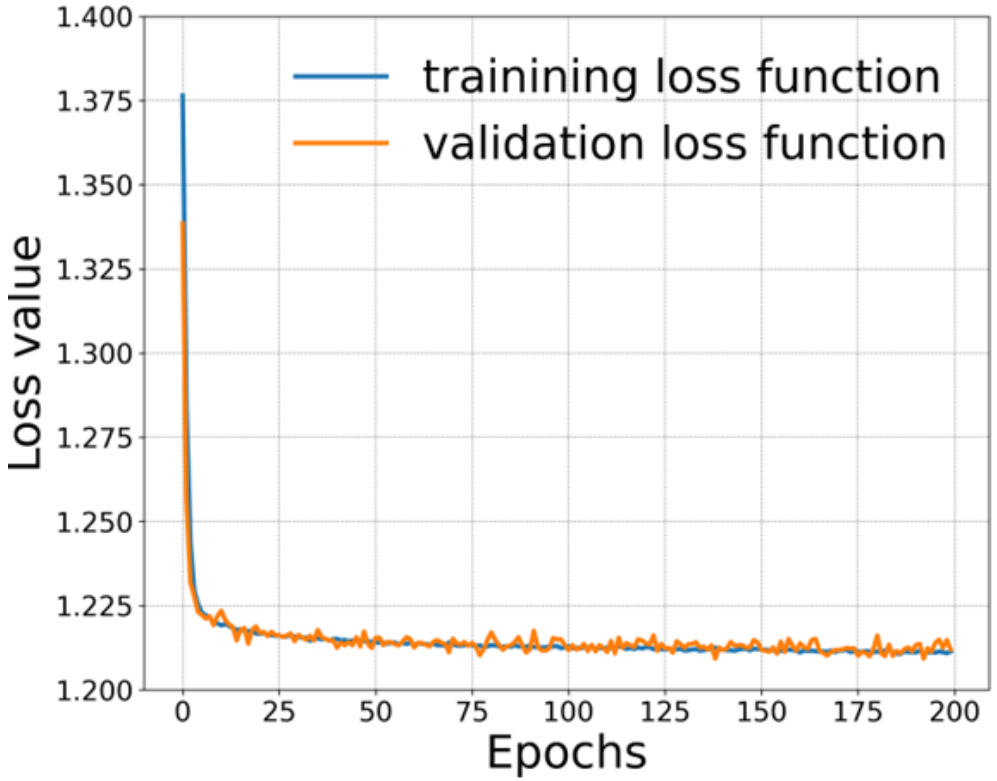
- ACC clusters

- Number of hits in L3-L8 inner tracker X side.
- Track hit energy deposition on L7.
- Total energy deposition of tracker clusters on L7 Y-side.
- Ratio between cluster amplitude and its neighbouring 10 strips (L2)
- Max cluster energy deposit on L7 Y-side.

Fully Connected Neural Network (FCNN) for classification

- FCNN **structure**:
  - PyTorch
  - Five linear layers: [12, 10, 10, 10, 4].

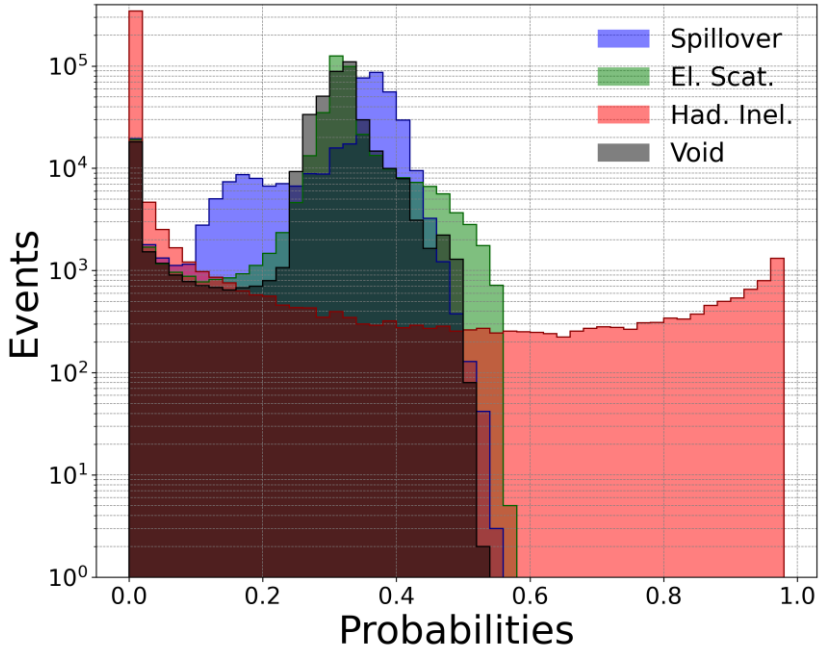




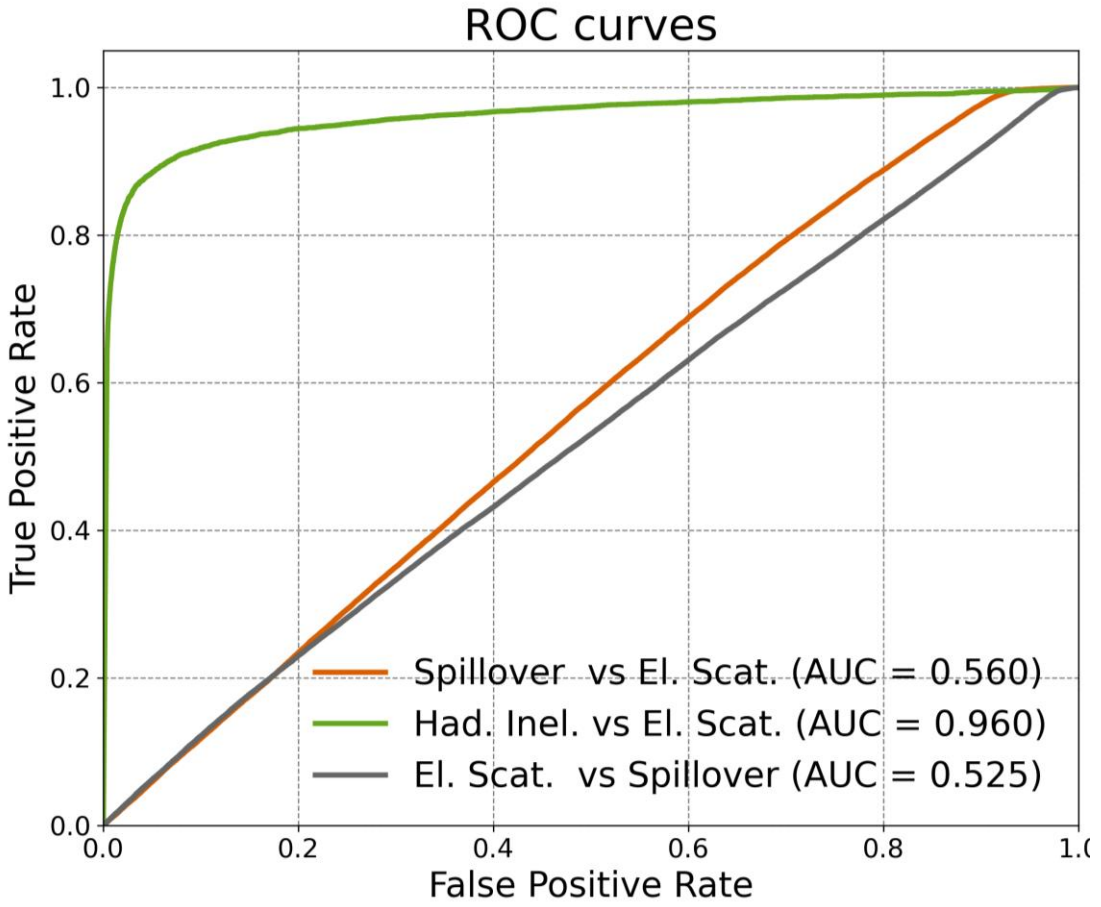
- FCNN **hyperparameters**:
  - Optimizer: Adam
  - Learning rate: $1.0 \cdot 10^{-3}$
  - Batch size: $5.0 \cdot 10^{3}$
  - Loss function: Cross Entropy

Fully Connected Neural Network (FCNN) for classification

- Training (200) epochs on $9 \cdot 10^5$ events
- Validation on $4 \cdot 10^5$ events
- **Training** and **validation** dataset **are unbalanced**



The classifier predict random probabilities, maybe due to samples unbalance.



The network is not able to distinguish Spillover and El. Scat.

Discriminants

- The network returns a vector of three elements.
  - Each element corresponds to the probability that the current event is belonging to the 1st , 2nd , 3rd or 4th class:

$$\text{FCNN output} = \left(p_{spillover},\ p_{Had.Inel}\ , p_{El.Scat.}, p_{Other}\right)$$

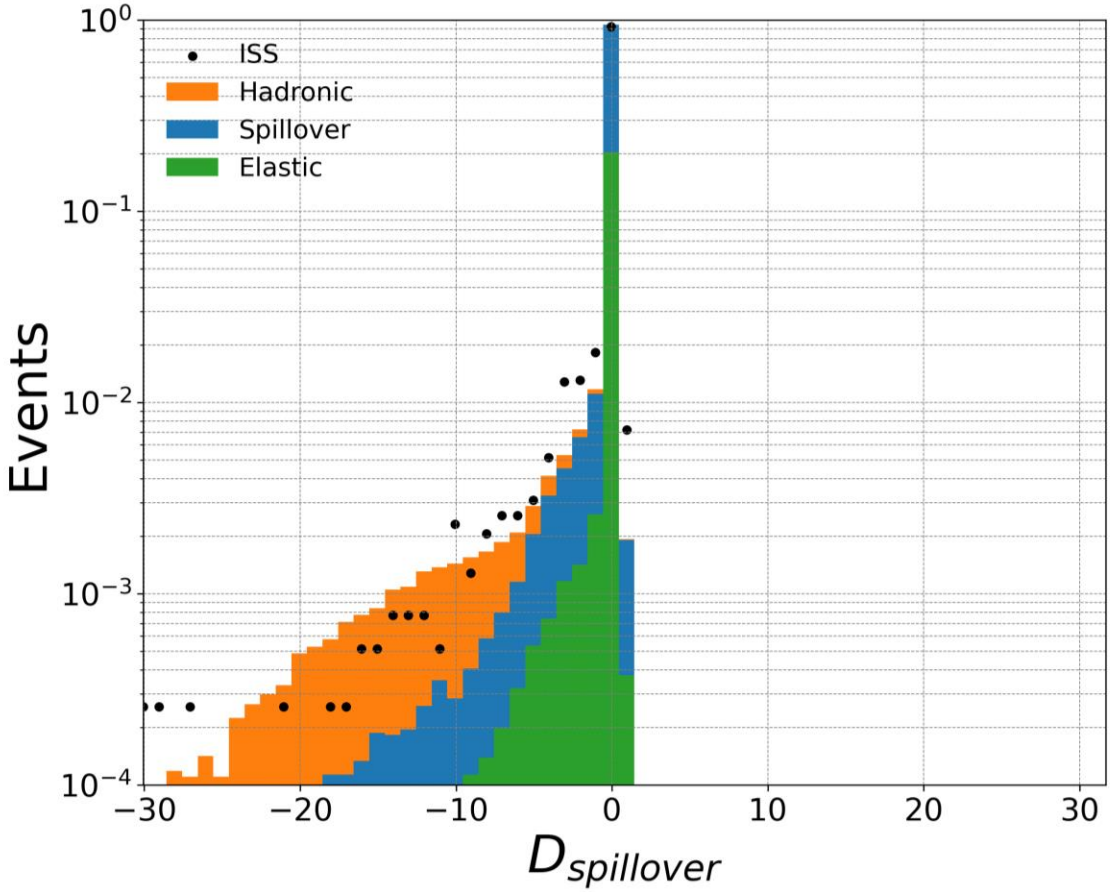
- The fraction of each class is defined as:

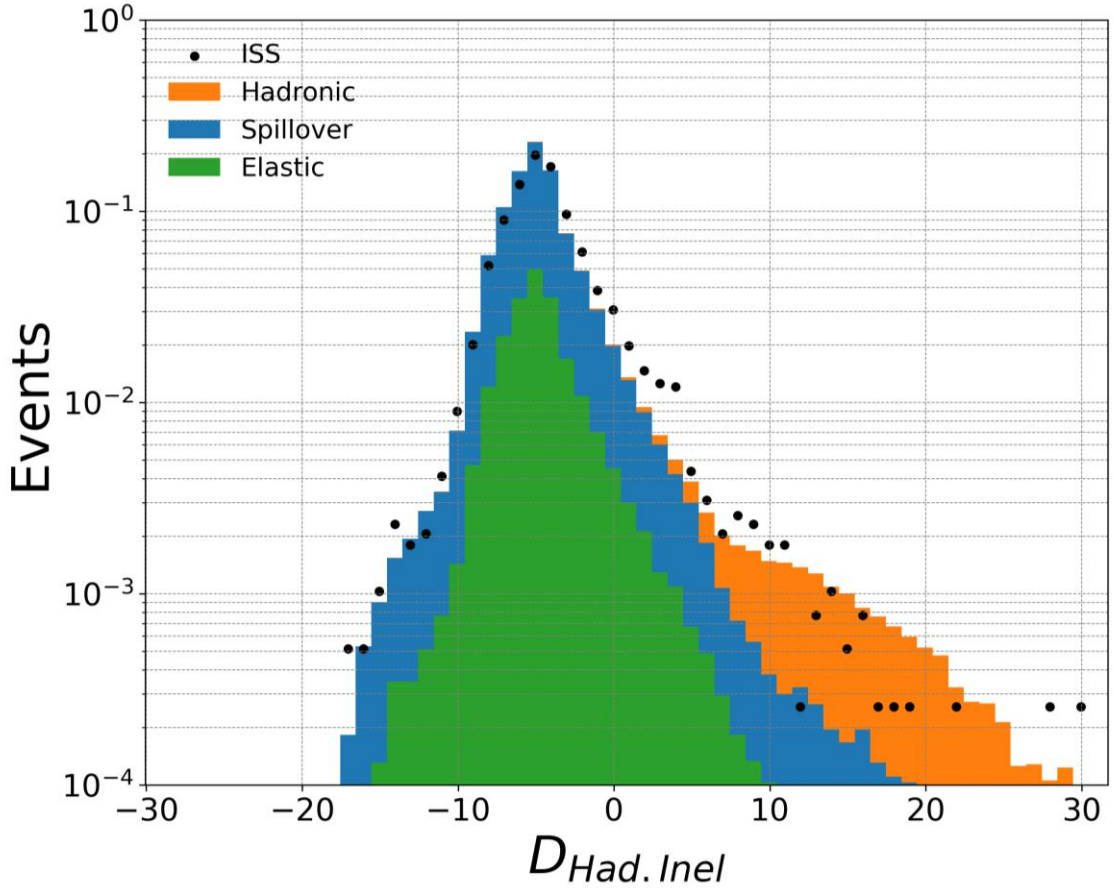
$$f_{Had.inel.} = \frac{\#Had.inel.}{\#Spillover + \#Had.Inel + \#El.scat. + \#Other}$$

- The discriminant is defined as

$$D_{Spillover} = \log_{10}\left(\frac{p_{Spillover}}{f_{Had.inel.} \cdot p_{Had.inel.} + f_{El.scat.} \cdot p_{El.scat.} + f_{Other} \cdot p_{Other}}\right)$$

Discriminants for spillover and had. Inel.

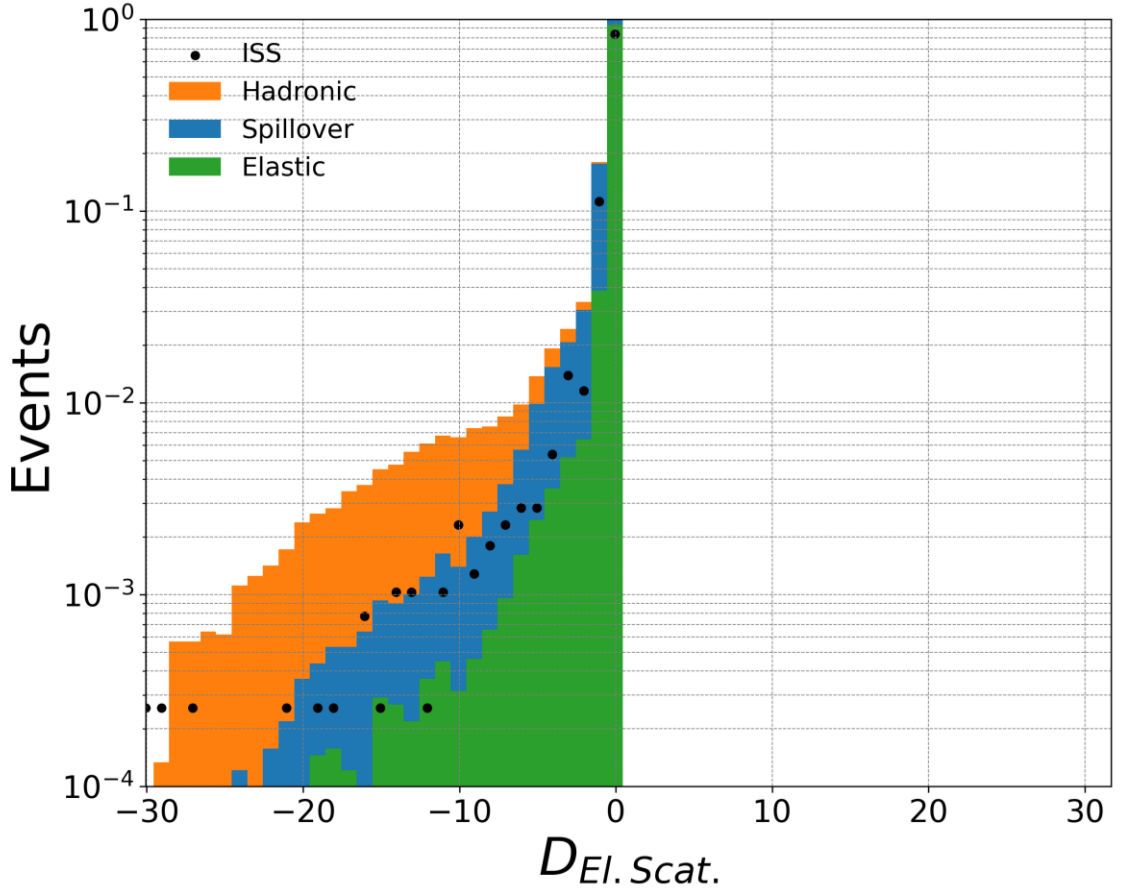




$$D_{Spillover} = \log_{10}\left(\frac{p_{Spillover}}{f_{Had.inel.} \cdot p_{Had.inel.} + f_{El.scat.} \cdot p_{El.scat.} + f_o \cdot p_o}\right)$$

$$D_{El.scat} = \log_{10}\left(\frac{p_{Had.inel.}}{f_{Spillover} \cdot p_{Spillover.} + f_{El.scat.} \cdot p_{El.scat.} + f_o \cdot p_o}\right)$$

Discriminant for el. scat.



$$D_{El.scat} = \log_{10}\left(\frac{p_{El.scat}}{f_{Had.inel.} \cdot p_{Had.inel.} + f_{Spillover} \cdot p_{Spillover} + f_O \cdot p_O}\right)$$

Prospects

- Propagator does not take into account energy losses and multiple scattering.
- Study alternative label definition in the MC sample.
- Investigate new input features to discriminate between spillover and elastic scattering.
- Rigidity dependence of the input variables.

- Reweight the MC; more on checks on agreement between MC and data
- Train the network on balanced dataset and validate on unbalanced sample.
- Improve data statics (~ 40 bartel rotations to be analyzed).

- MC B1306 is not focused, interesting to integrate the P.Z. flag for elastic interactions.
- Modify NAIA to save each secondary produced by hadronic (inelastic and elastic) interactions.

- **Study possible alternative ML techniques as anomaly detection.**

Prospects

- Propagator does not take into account energy losses and multiple scattering.
- Study alternative label definition in the MC sample.
- Investigate new input features to discriminate between spillover and elastic scattering.
- Rigidity dependence of the input variables.

- Reweight the MC; more on checks on agreement between MC and data
- Train the network on balanced dataset and validate on unbalanced sample.
- Improve data statics (~ 40 bartel rotations to be analyzed).

Thank you for your attention!

- MC B1306 is not focused, interesting to integrate the P.Z. flag for elastic interactions.
- Modify NAIA to save each secondary produced by hadronic (inelastic and elastic) interactions.

- **Study possible alternative ML techniques as anomaly detection.**

Backup

Monte Carlo label defintion

1. Check for secondary production due to the primary particle within the inner tracker (HasSecondary)

2. Check for inelastic interactions (HasHadInel).
    1. If true the event is tagged as Had. Inel.
1. Otherwise:
    1. If primary reaches L2 propagate two different tracks inside the inner tracker: $R_{true}(L2)$ and $R_{inner}(< 0)$.
        1. Build two χ^2 comparing y coordinate with MC true info on each layer.
        2. Check the χ^2 ratio, i.e.:

$$\frac{\chi^2_{R_{inner}}}{\chi^2_{R_{true}}} \leq 0.85 \rightarrow \text{El. scat.}$$

$$1.15 \leq \frac{\chi^2_{R_{inner}}}{\chi^2_{R_{true}}} \rightarrow (\text{HasSecondary})\ ?\ \text{Other}\ :\ \text{Spillover}$$

$$0.85 < \frac{\chi^2_{R_{inner}}}{\chi^2_{R_{true}}} \leq 1.15 \rightarrow (\text{HasSecondary})\ ?\ \text{Other}\ :\ \text{Spillover}$$

Known problems: no MS and E losses in propagator

    2. Otherwise, tagged the event as Other

Monte carlo selection (He B1236 L1-focused and L1-L9 focused)

IsPhysicsTrigger

TOF hits ≥ 4
$\beta > 0.2$
$\chi^2_{Coo} < 4$

Track number ≥ 1
track pattern L1&L2&(L3|L4)&(L5|L6)&(L7|L8) (≥ 5)

charge STD (inner) $\in [1.7, 2.4]$
charge STD (L1) $\in [1.6, 3.0]$
Inner fiducial volume
L1 fiducial volume

$\chi^2_Y < 10$
charge UTOF $\in [1.5, 3.0]$
charge LTOF ≥ 1.5

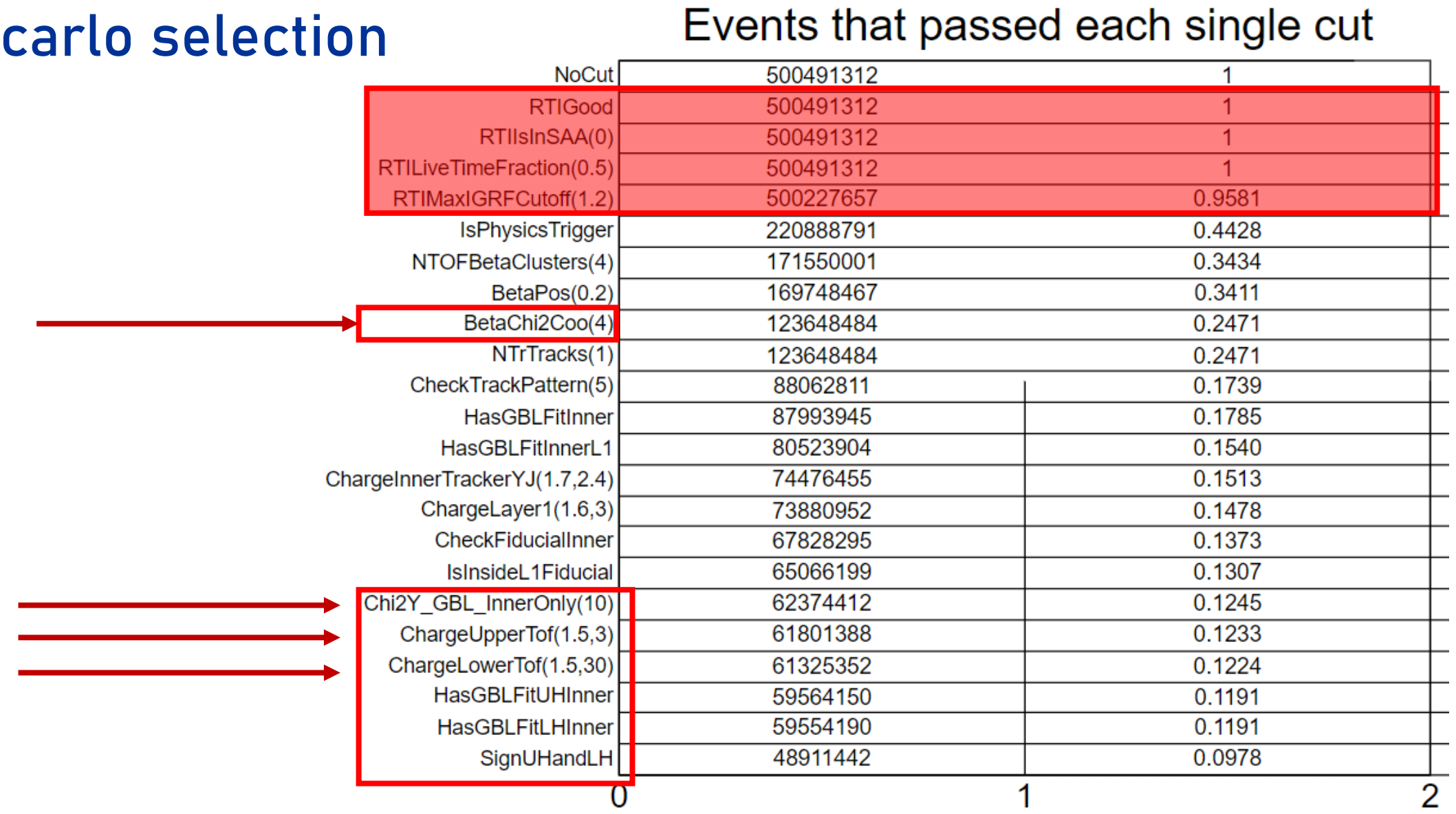
Sign inner upper half (UH) and lower half (LH)

Tracker fiducial volume cut:
- L1: |R|<62cm, |Y|<47cm
- L2: |R|<62cm, |Y|<40cm
- L3: |R|<46cm, |Y|<44cm
- L4: |R|<46cm, |Y|<44cm
- L5: |R|<46cm, |Y|<36cm
- L6: |R|<46cm, |Y|<36cm
- L7: |R|<46cm, |Y|<44cm
- L8: |R|<46cm, |Y|<44cm

UH = Rig. [UH-inner]
LH = Rig. [LH-inner]
Inner = Rig. [inner]

If(Inner > 0) $\rightarrow (\mathrm{UH} > 0, \mathrm{LH} > 0)$
If(Inner < 0) $\rightarrow (\mathrm{UH} < 0, \mathrm{LH} < 0)$

Monte carlo selection



MC with large angle scattering flag

Selection

Selezione comune attualmente usata

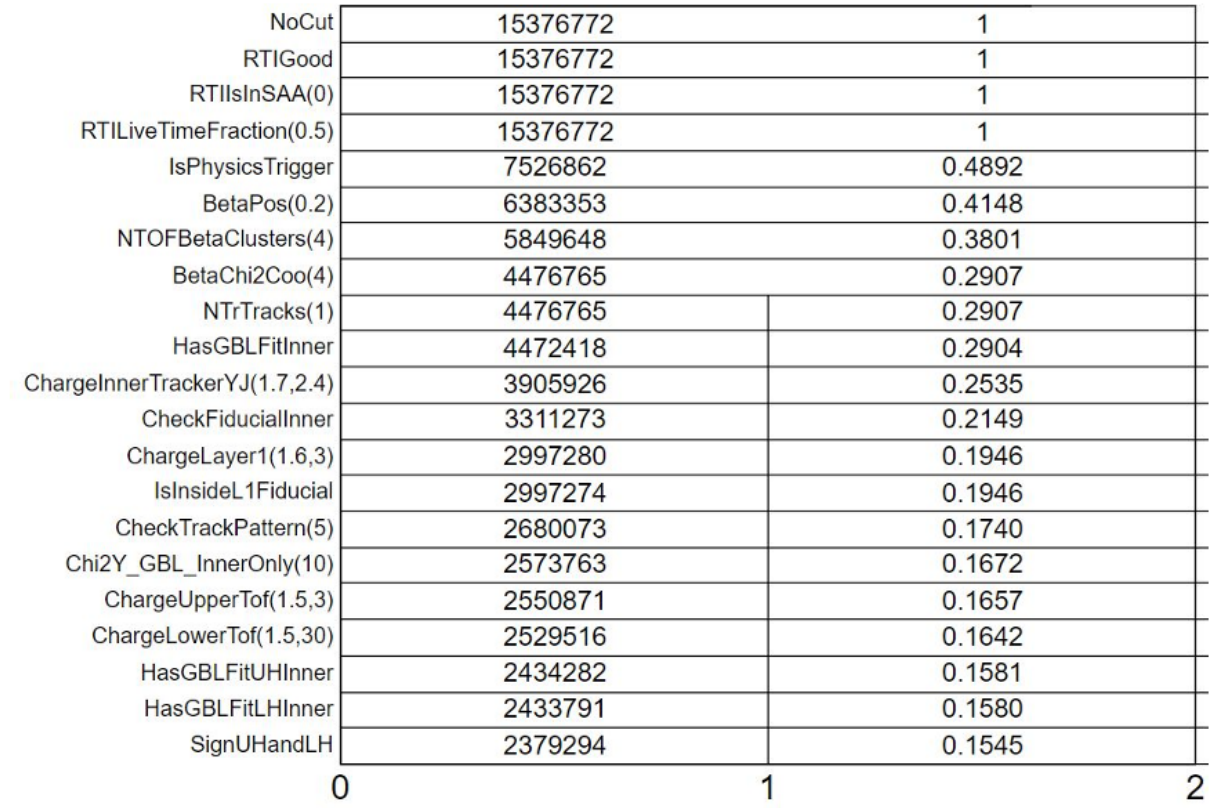
IsPhysicsTrigger

$\beta > 0$
TOF hits = 4
Chi2Coo < 4

Track number >= 1
charge YJ (inner) ϵ [1.7, 2.4]
inner fiducial volume
charge YJ (L1) ϵ [1.6, 3.0]
track pattern 5/8 (L1-inner)
$\chi^2_Y < 10$

charge (UTOF) ϵ [1.5, 3.0]
charge (LTOF) > 1.5

(R_{UH} , R_{LH} < 0) oppure (R_{UH} , R_{LH} > 0)



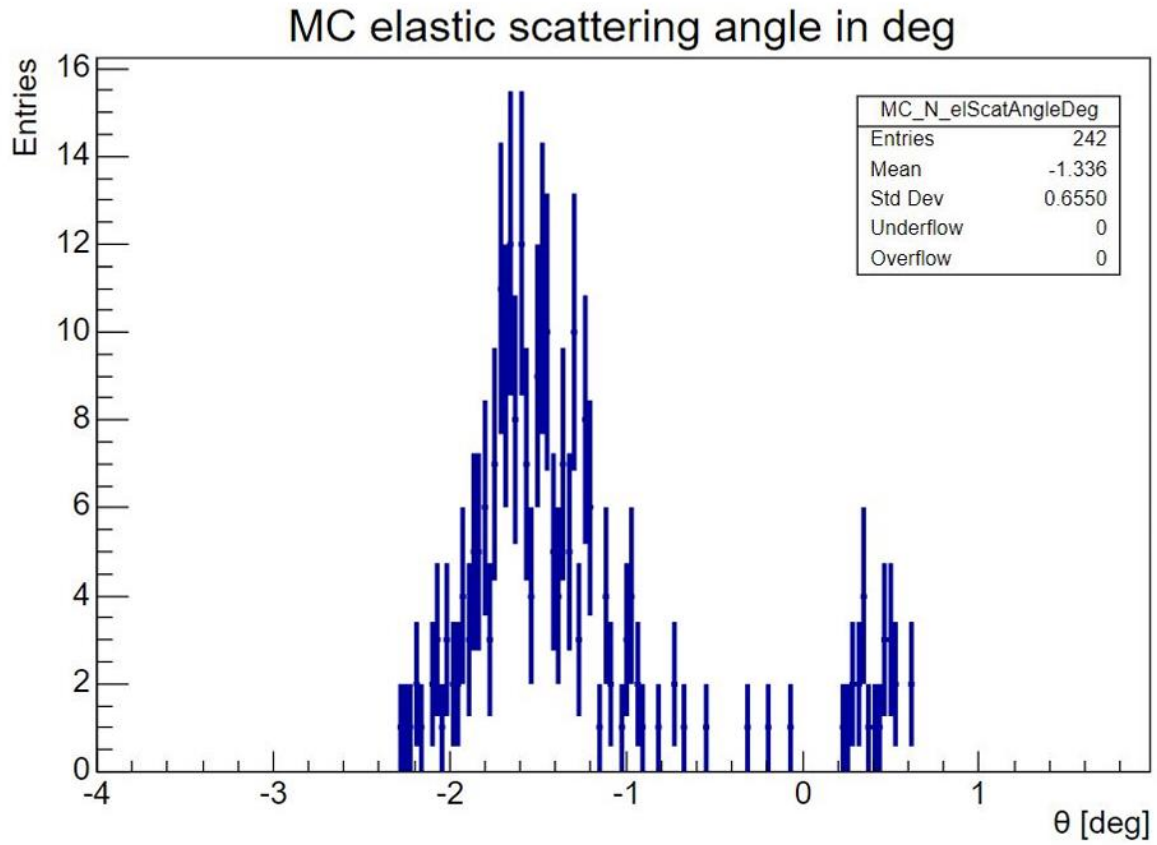
R_{inner} < 0 → 696 eventi

Distribution of the scattering angle ($R < 0$)

$R_{inner} < 0$ ⟶ 696 eventi

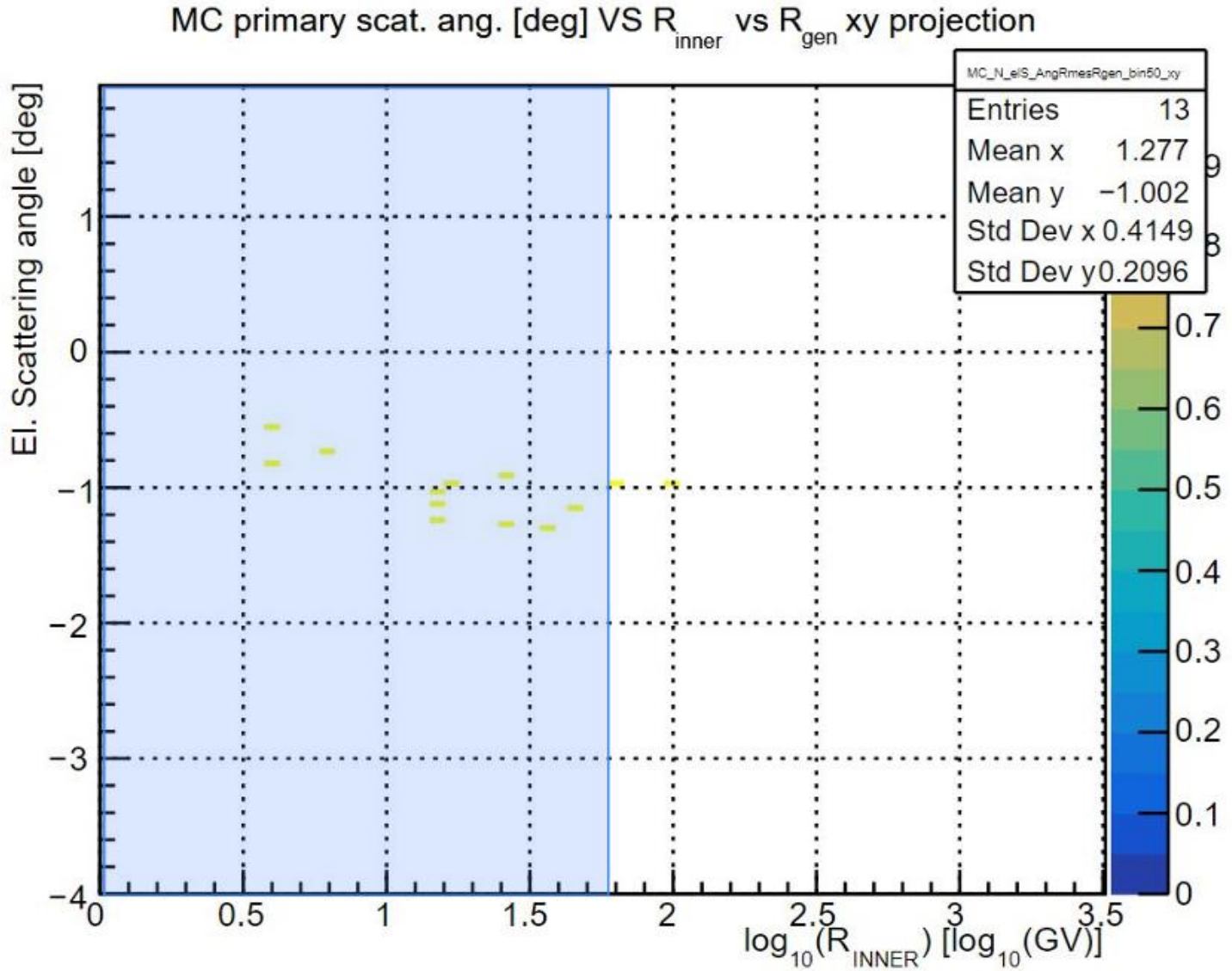
Verità MC per scattering elastic

242 eventi con scattering elastico nell'inner tracker



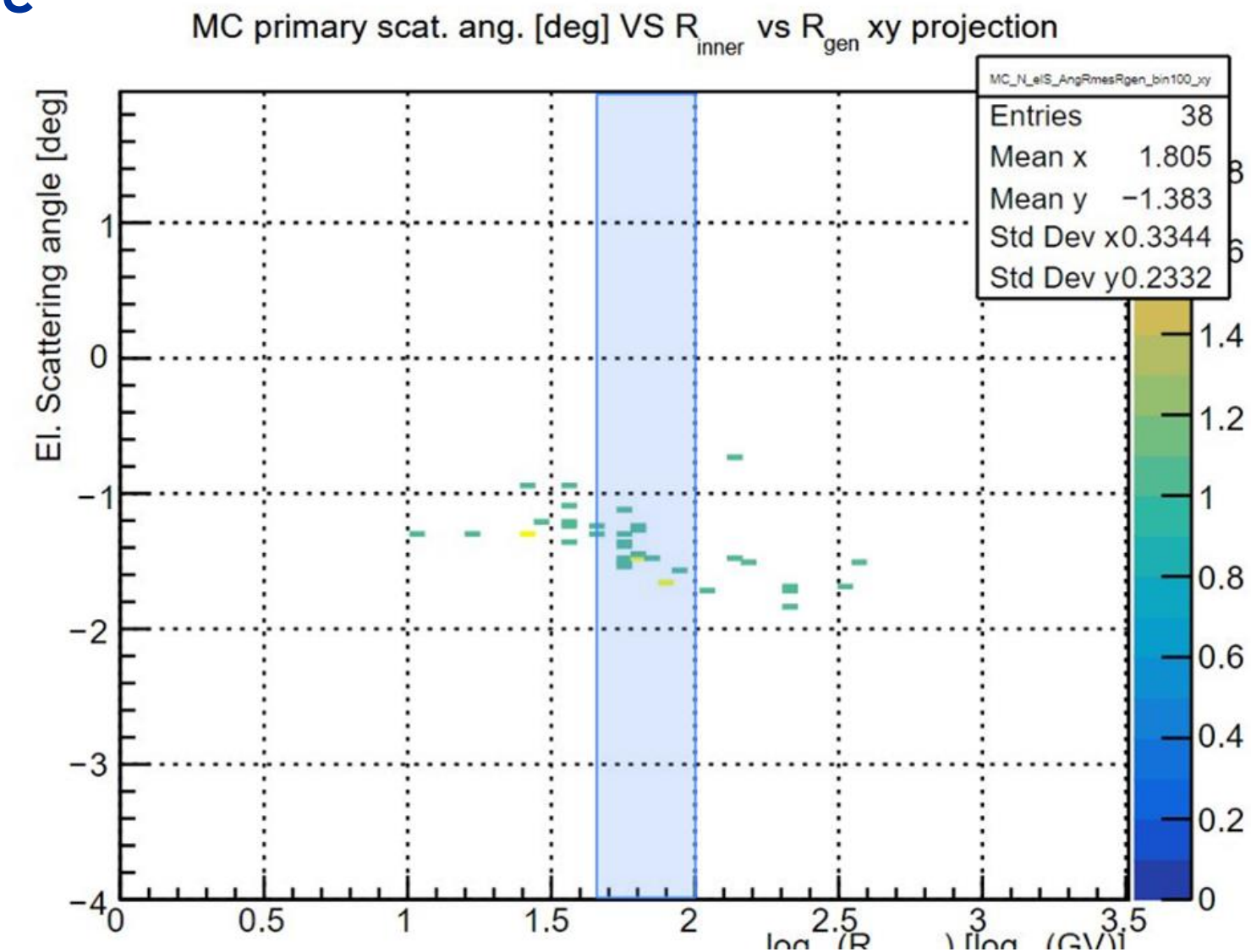
$(R_{inner} < 0)$ and elasti scattering inside the inner tracker

$$R_{gen} \in [0, 50[$$
$$\log_{10}(R_{gen}) < 1.67$$



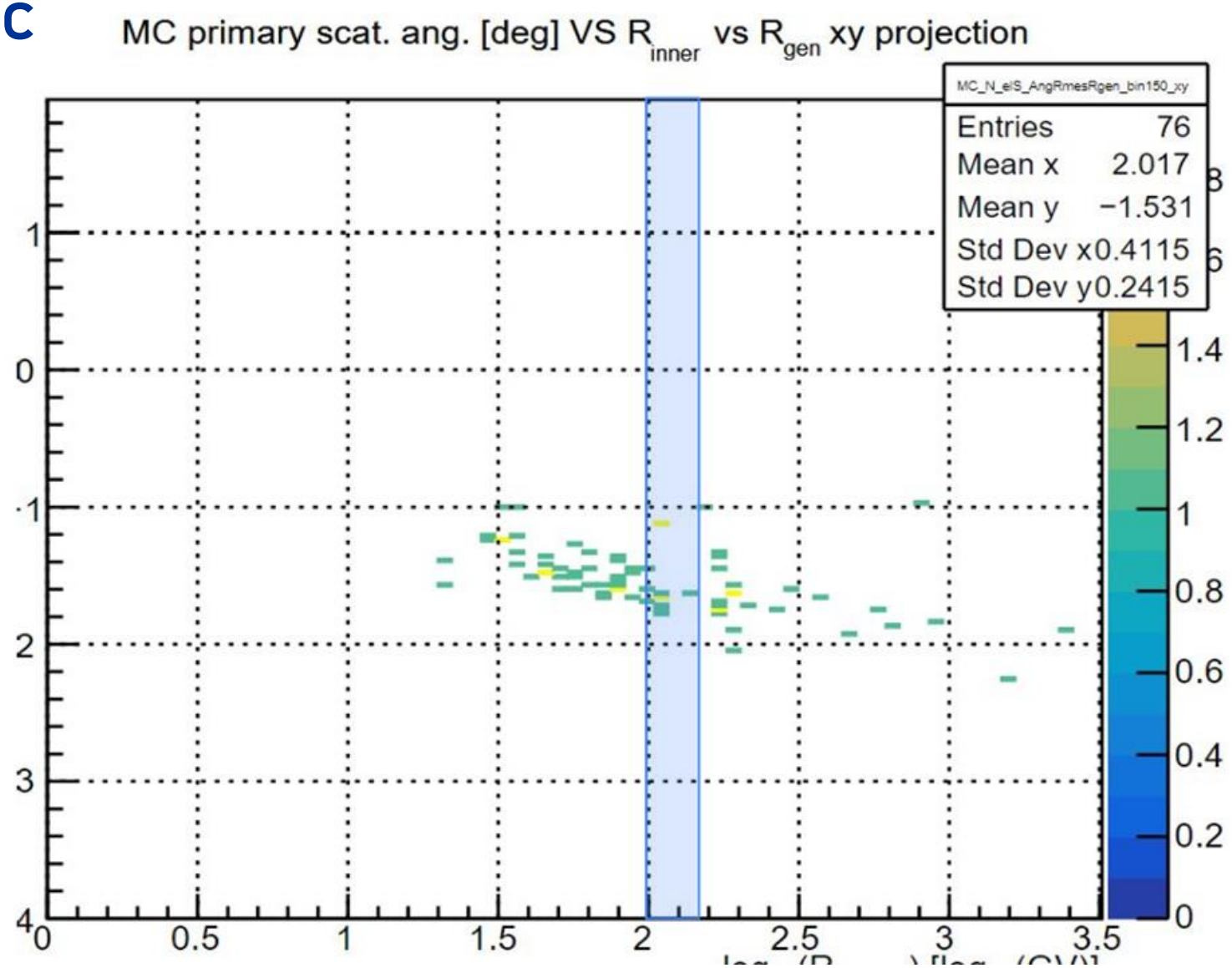
$(R_{inner} < 0)$ and elastic scattering inside the inner tracker

$$R_{gen} \in [50, 100[$$
$$\log_{10}(R_{gen}) < 2.0$$



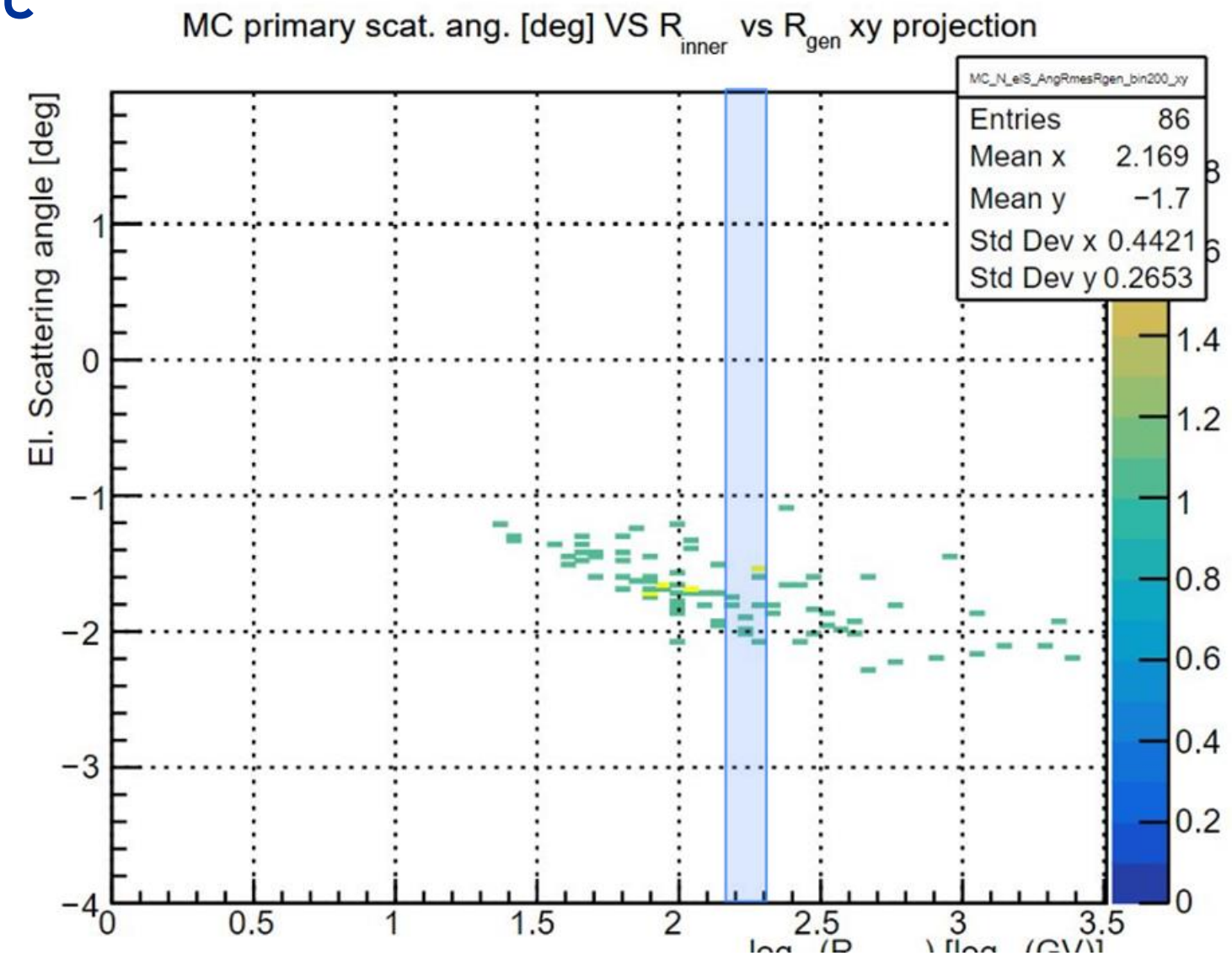
$(R_{inner} < 0)$ and elastic scattering inside the inner tracker

$$R_{gen} \in [100, 150[$$
$$\log_{10}(R_{gen}) < 2.18$$

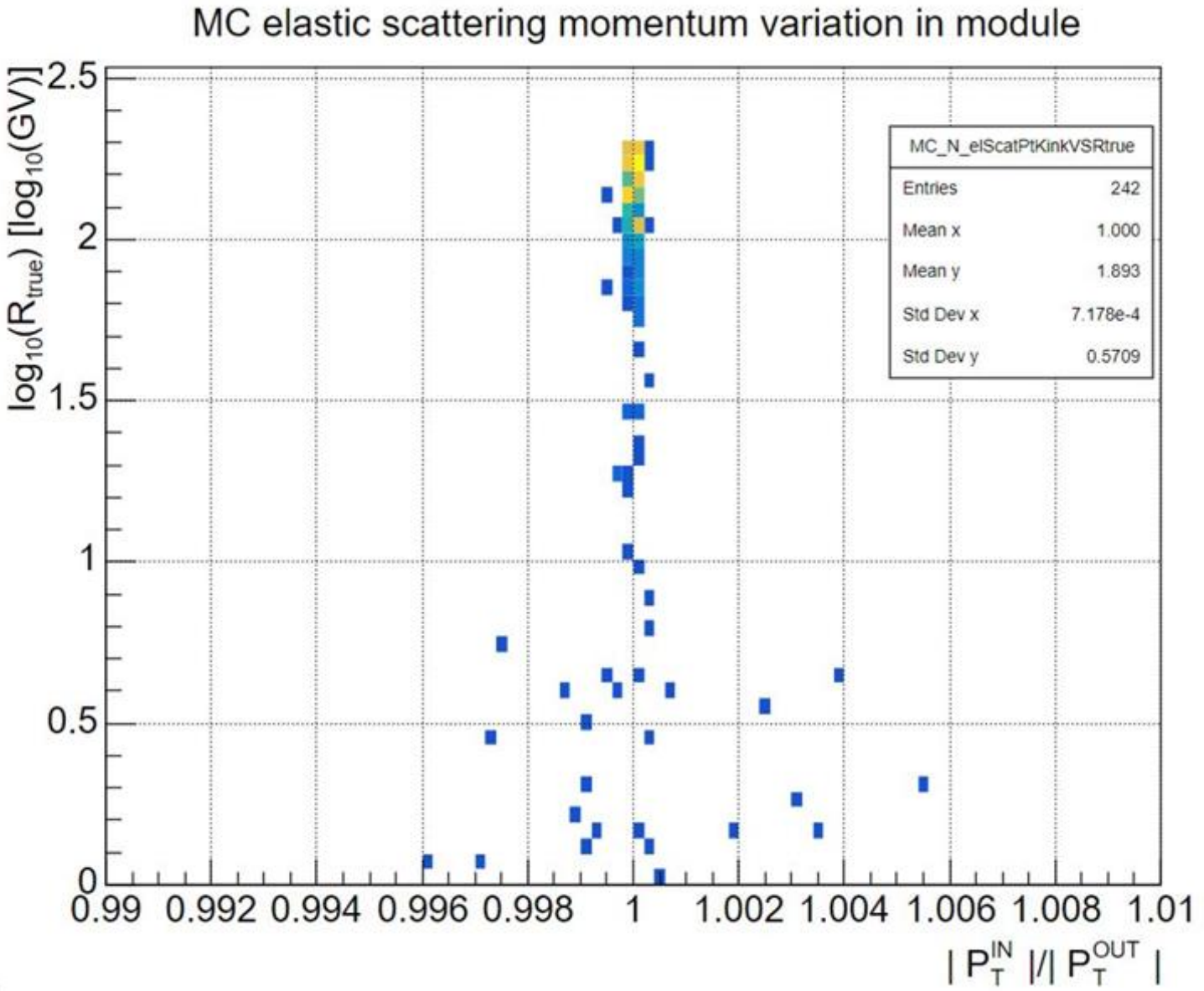


$(R_{inner} < 0)$ and elastic scattering inside the inner tracker

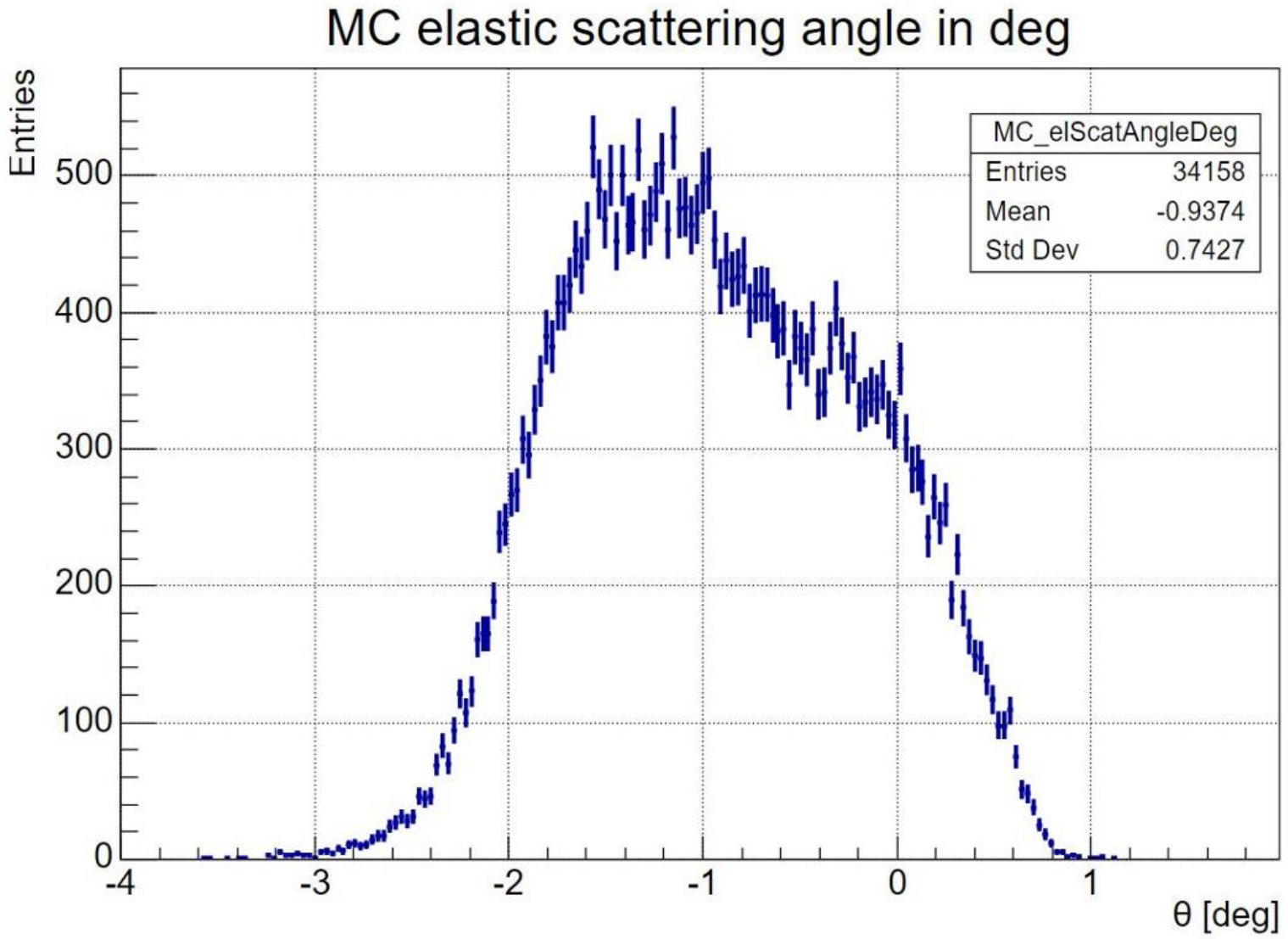
$$R_{gen} \in [150, 200[$$
$$\log_{10}(R_{gen}) < 2.30$$



$(R_{inner} < 0)$ and elastic scattering inside the inner tracker

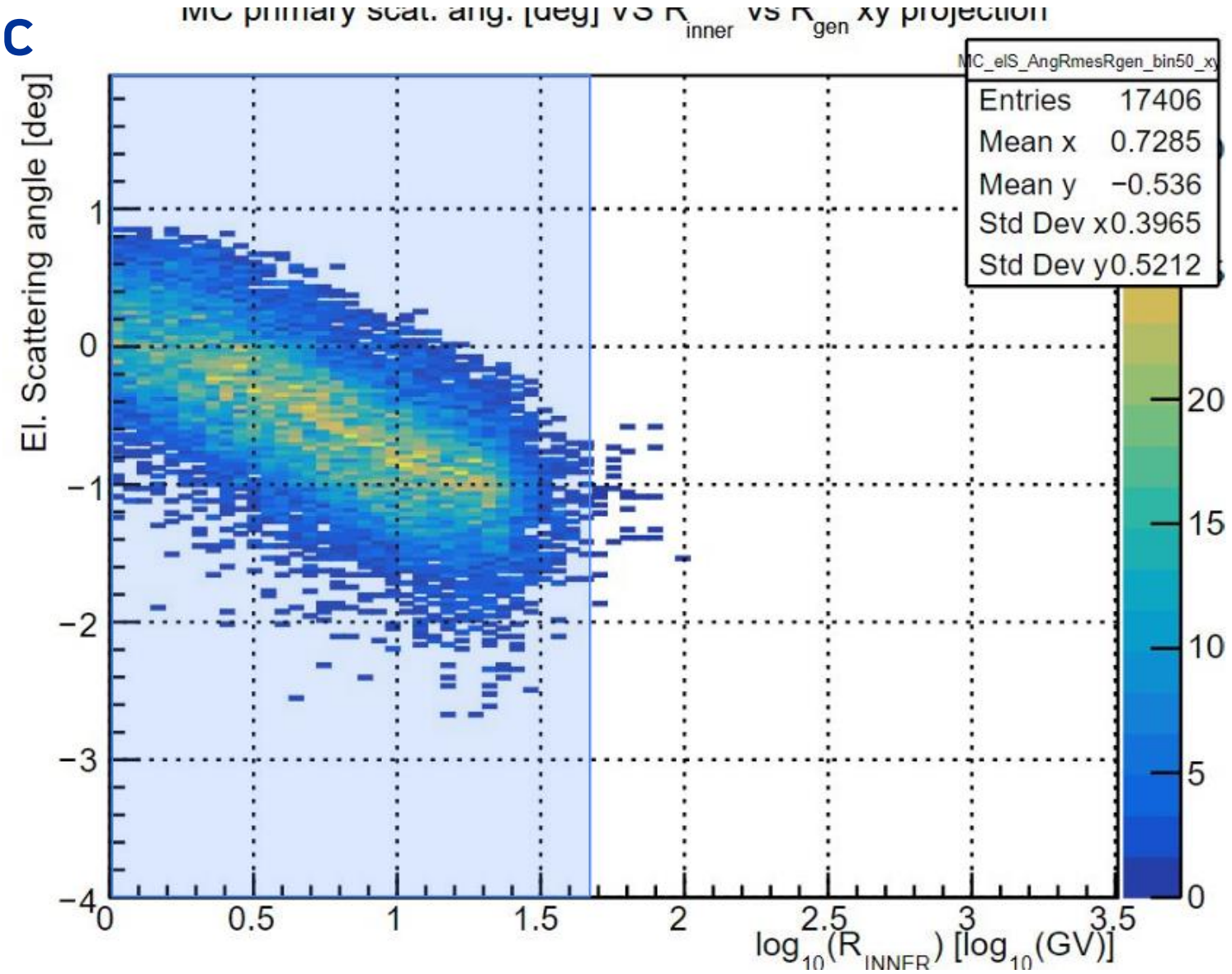


Distribution of the scattering angle ($R > 0$)



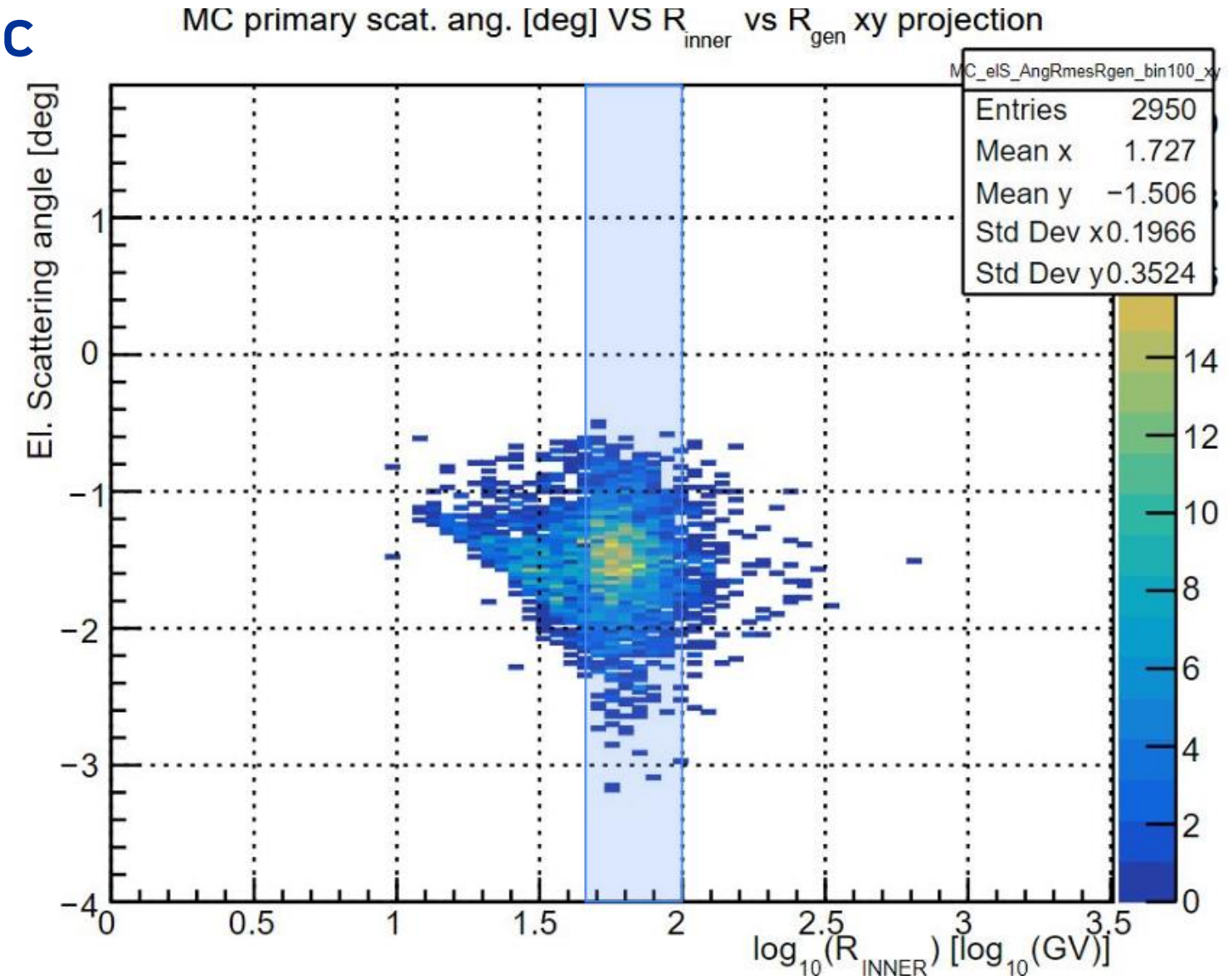
$(R_{inner} > 0)$ and elastic scattering inside the inner tracker

$$R_{gen} \in [50, 100[$$
$$\log_{10}(R_{gen}) < 2.0$$



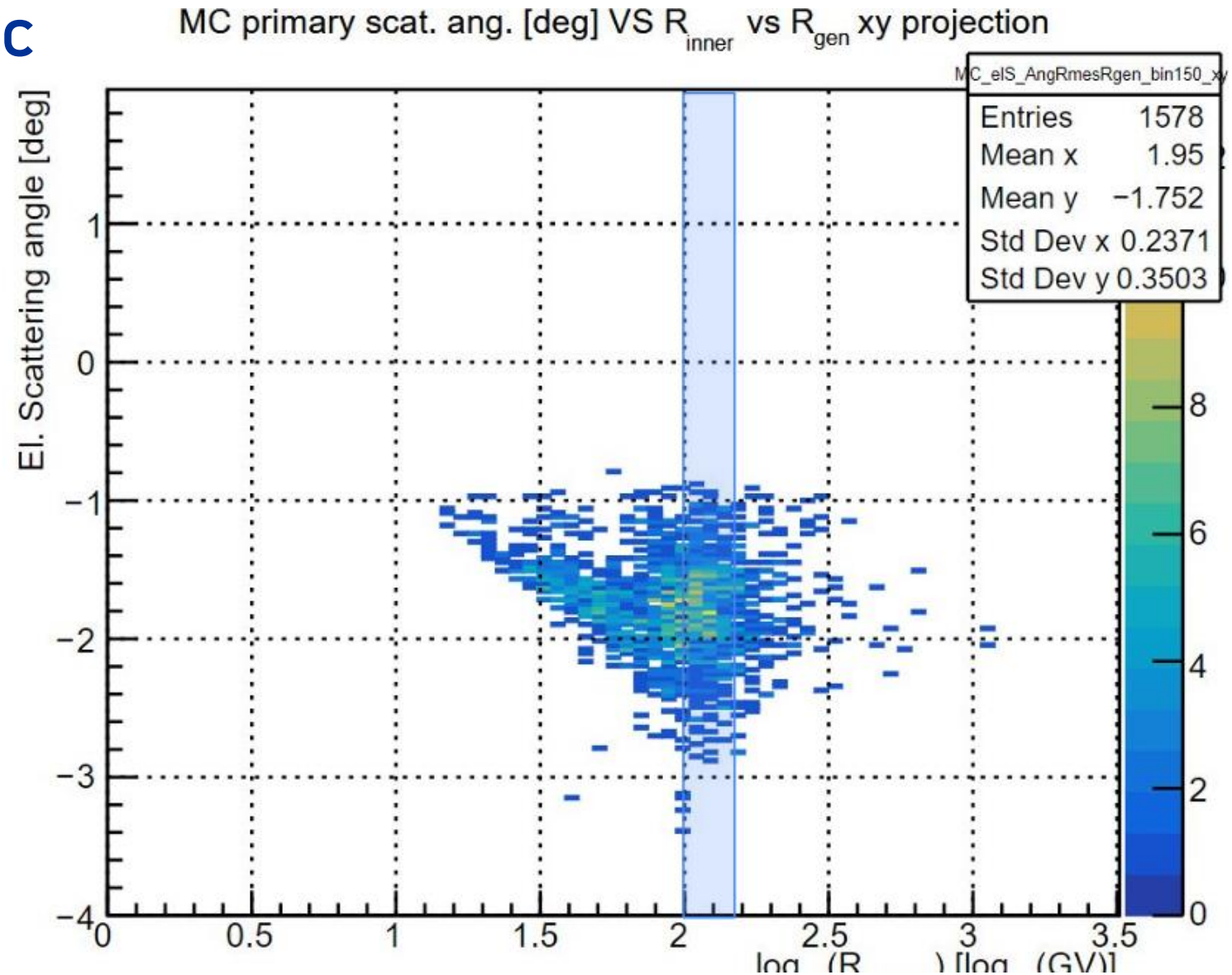
$(R_{inner} > 0)$ and elastic scattering inside the inner tracker

$$R_{gen} \in [50, 100[$$
$$\log_{10}(R_{gen}) < 2.0$$



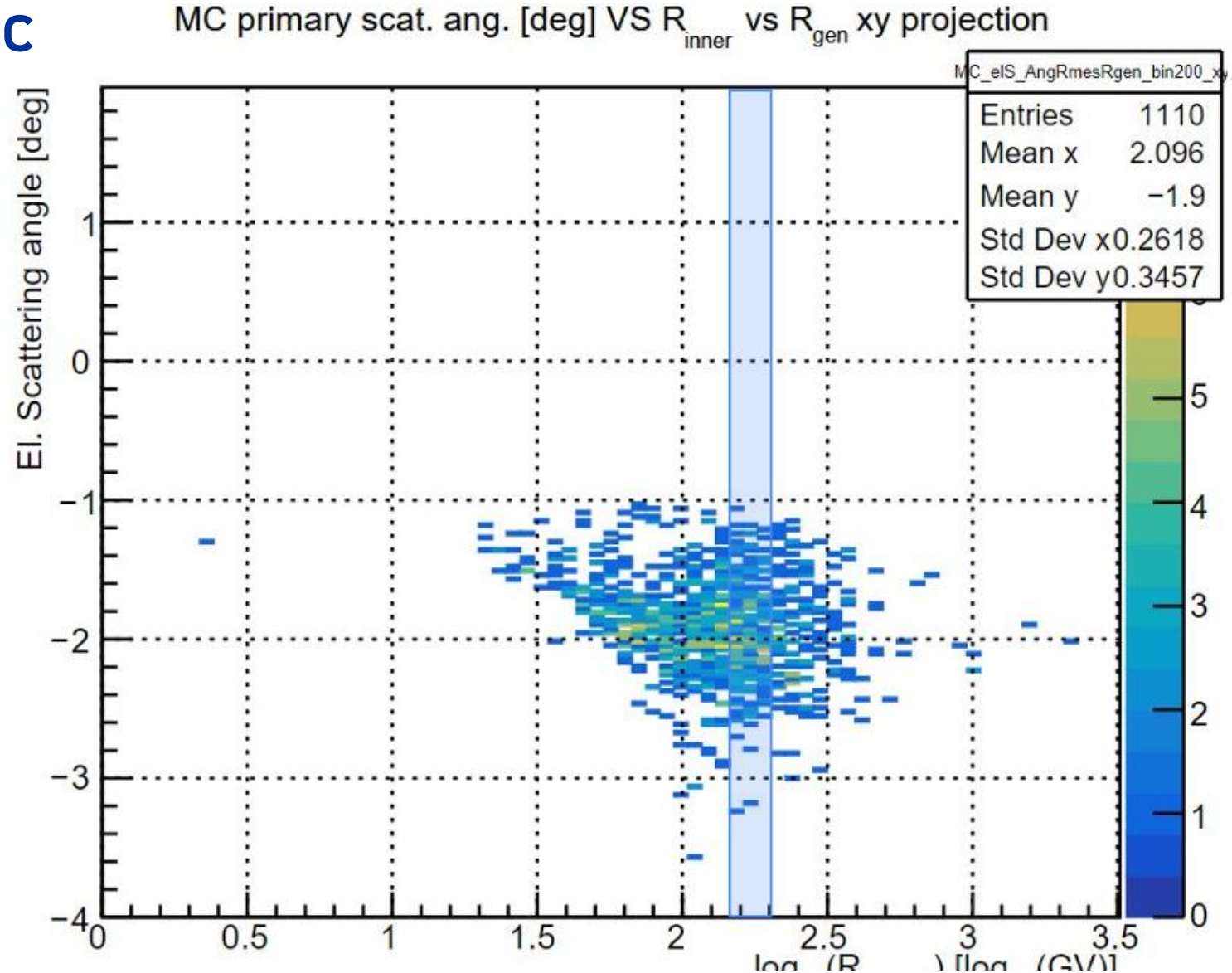
$(R_{inner} < 0)$ and elastic scattering inside the inner tracker

$$R_{gen} \in [100, 150[$$
$$\log_{10}(R_{gen}) < 2.18$$

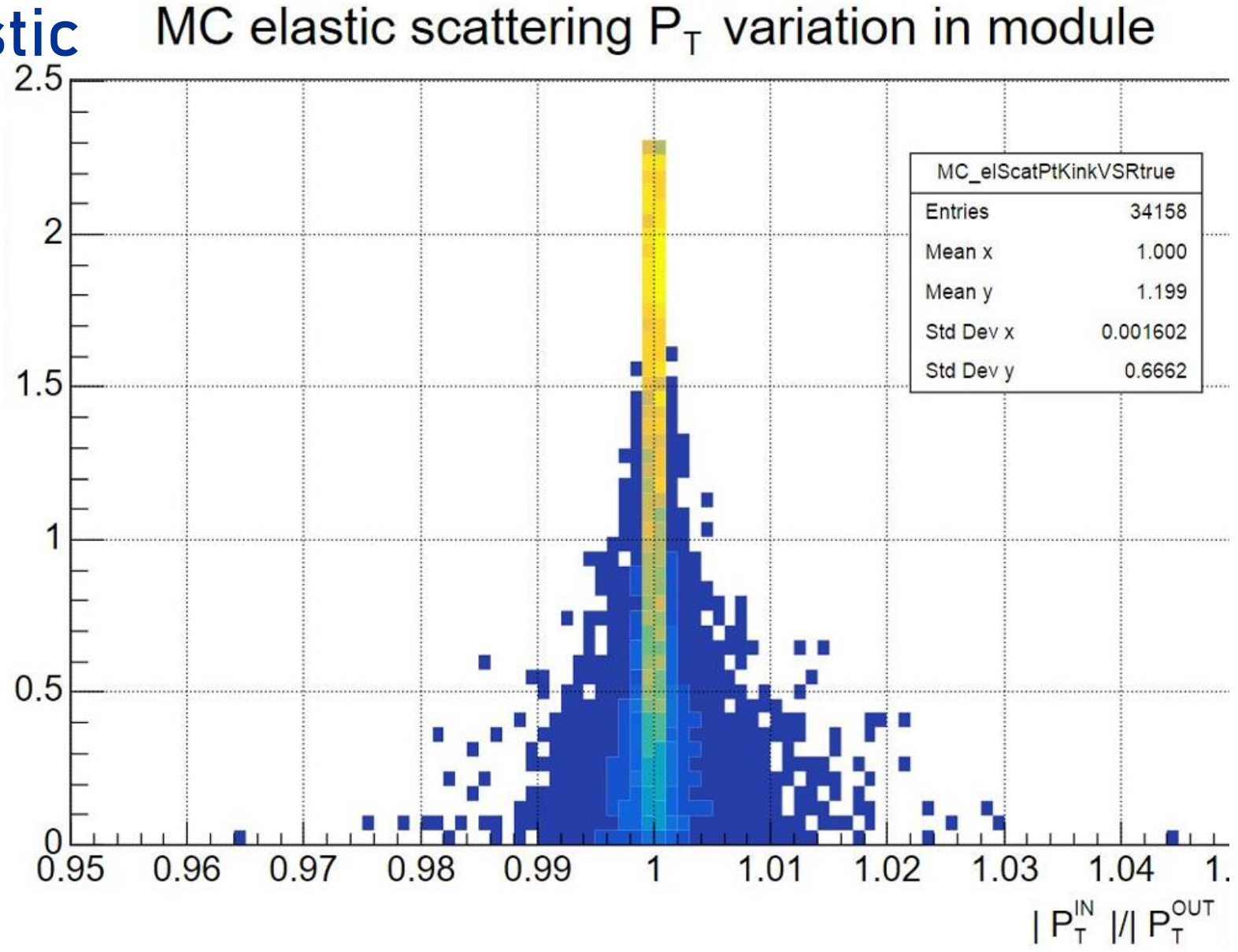


$(R_{inner} < 0)$ and elastic scattering inside the inner tracker

$$R_{gen} \in [150, 200[$$
$$\log_{10}(R_{gen}) < 2.30$$

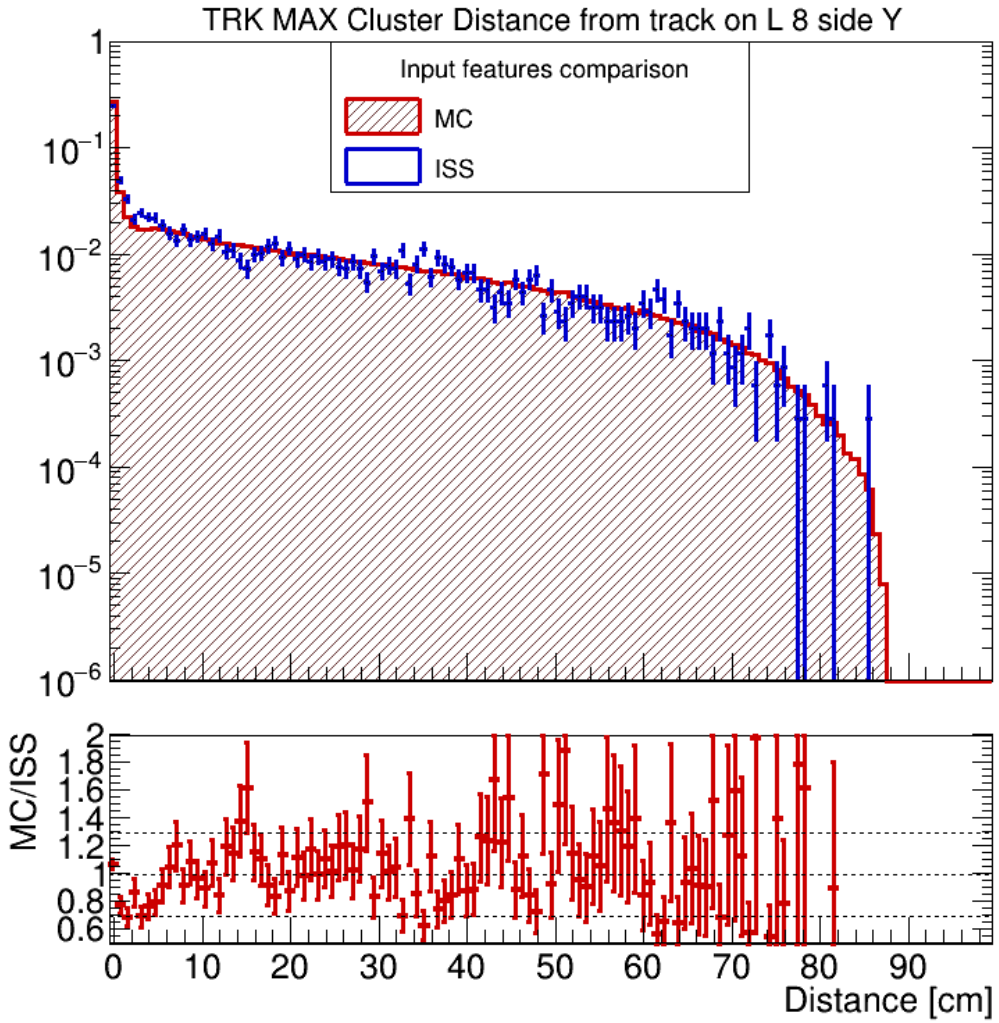


$(R_{inner} < 0)$ and elastic scattering inside the inner tracker

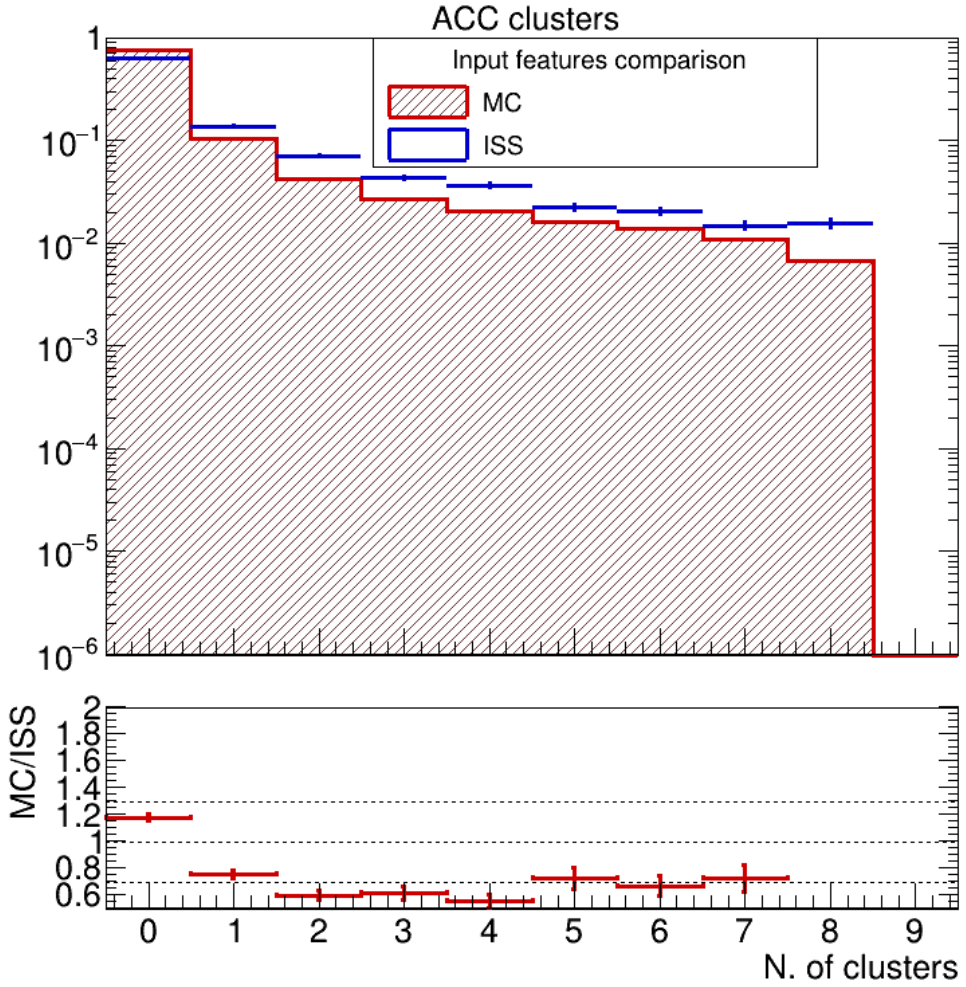


MC and ISS-data comparison

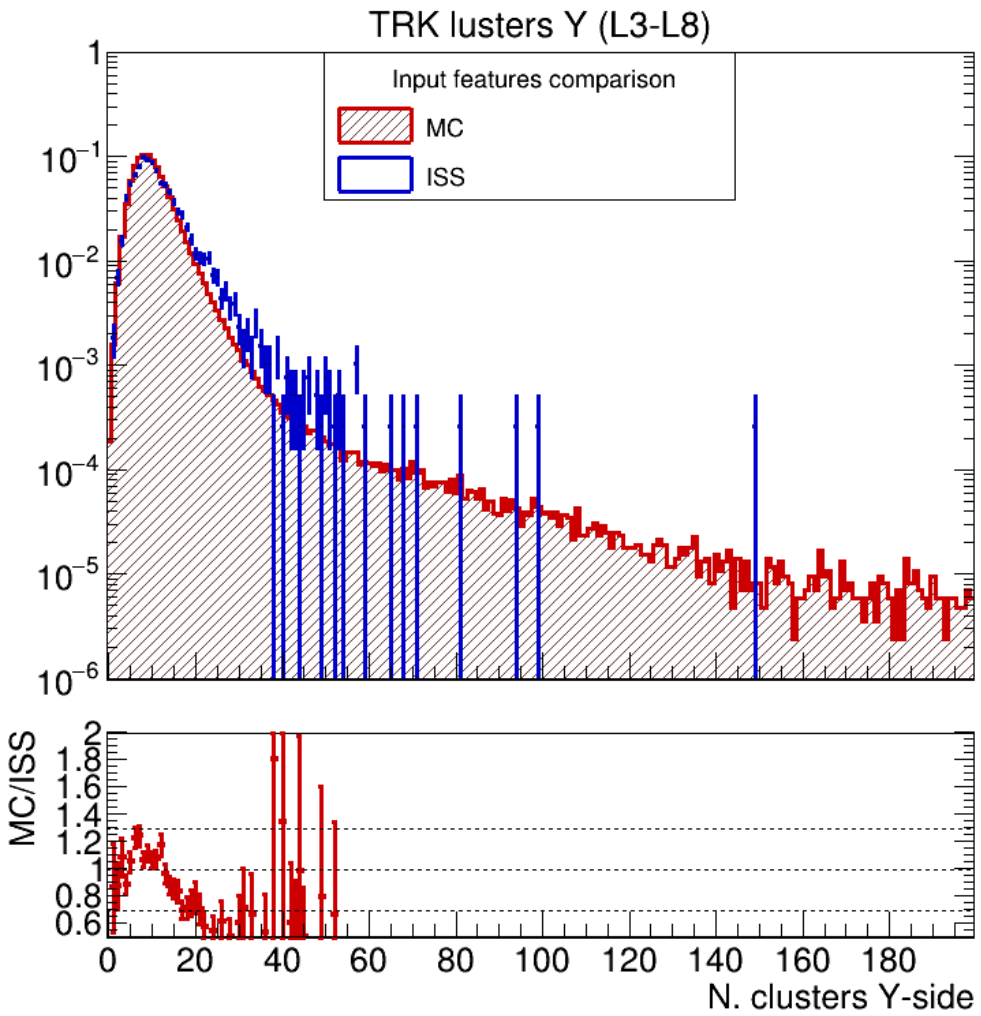
Distance between track hit and strip with max energy deposit on L8.



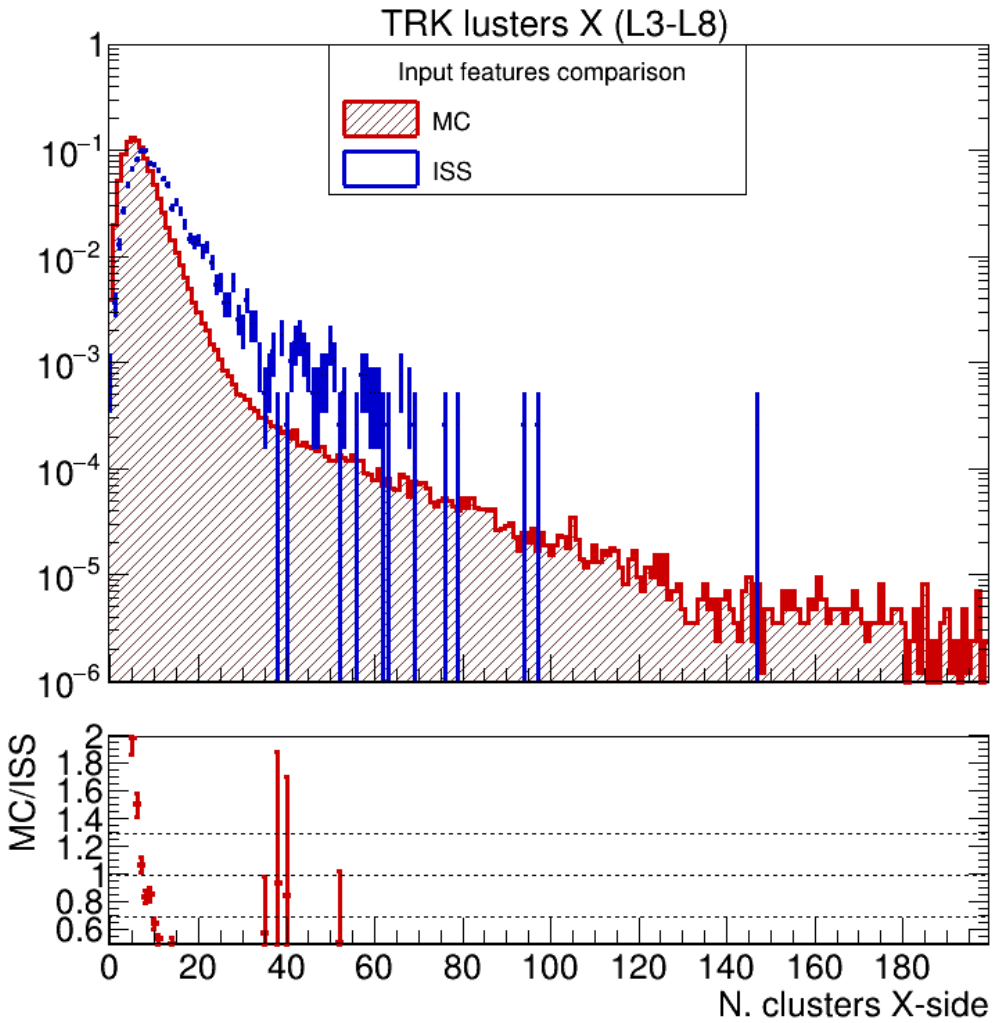
ACC clusters



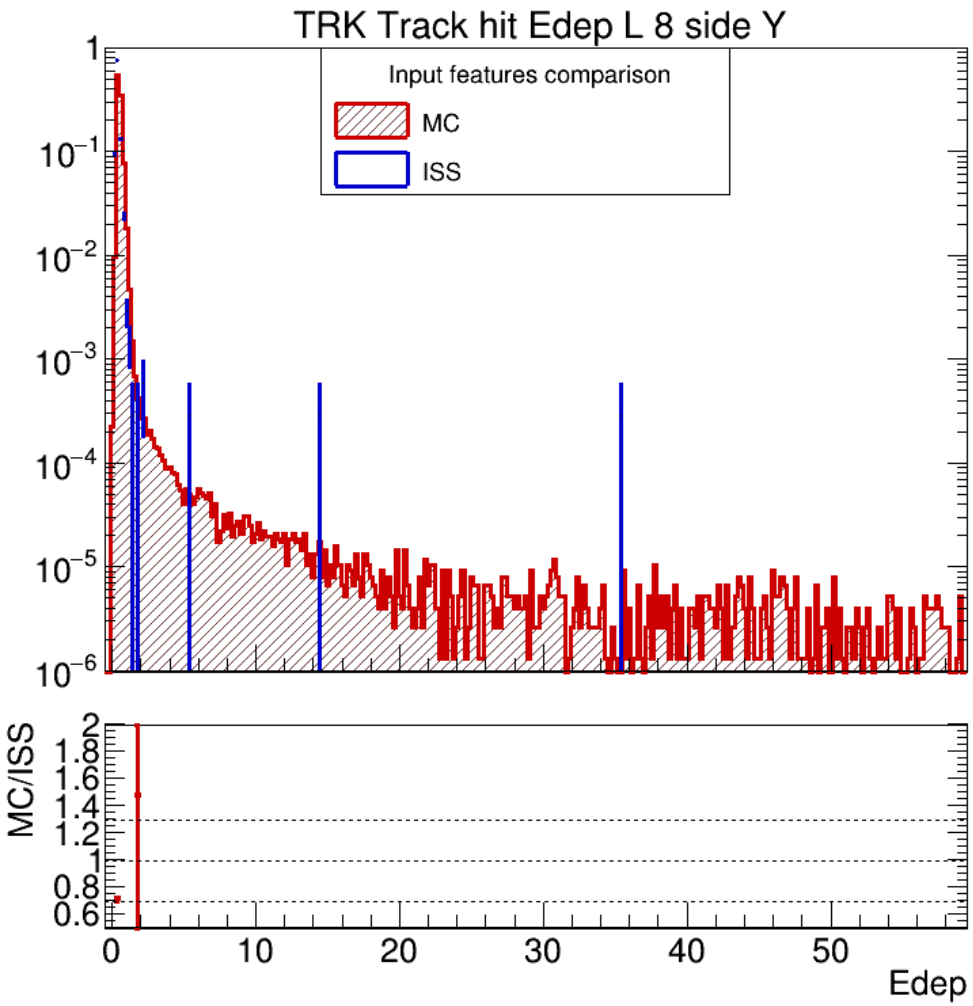
Number of hits in L3-L8 inner tracker Y side.



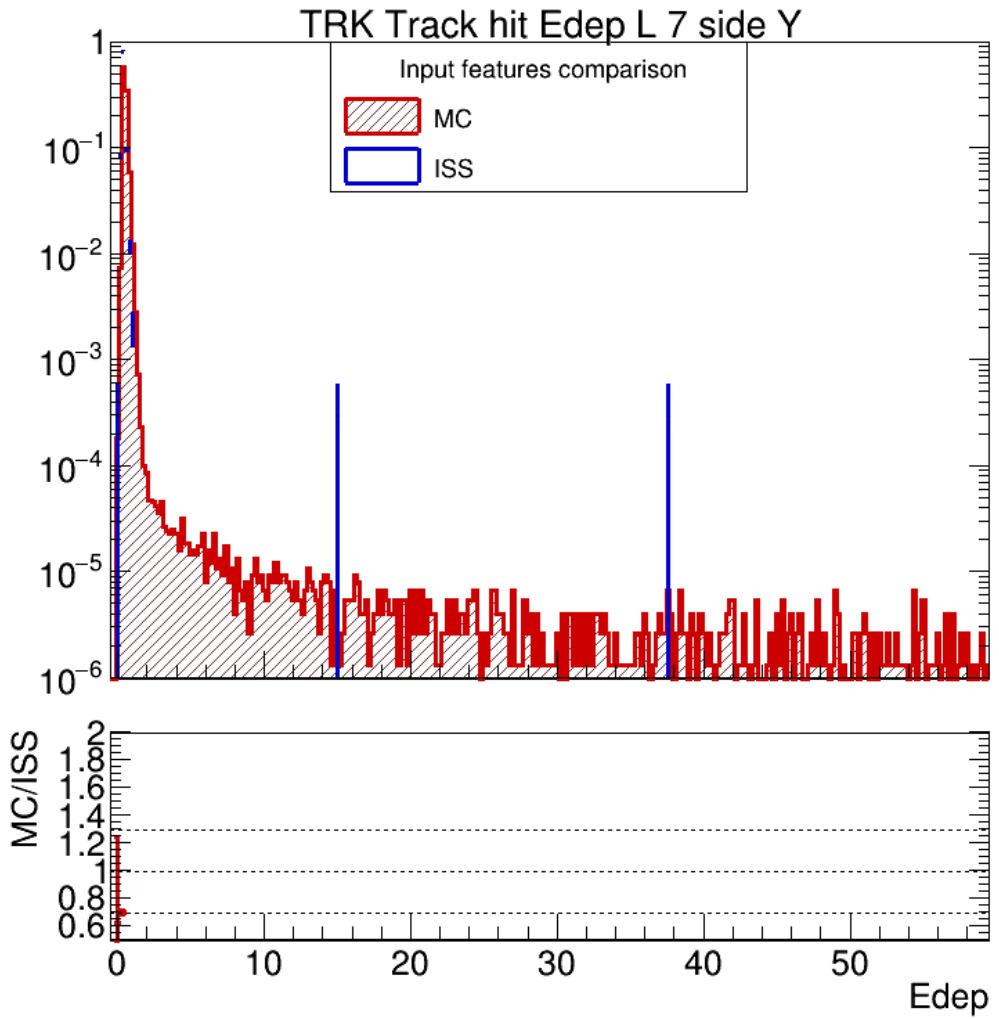
Number of hits in L3-L8 inner tracker X side.



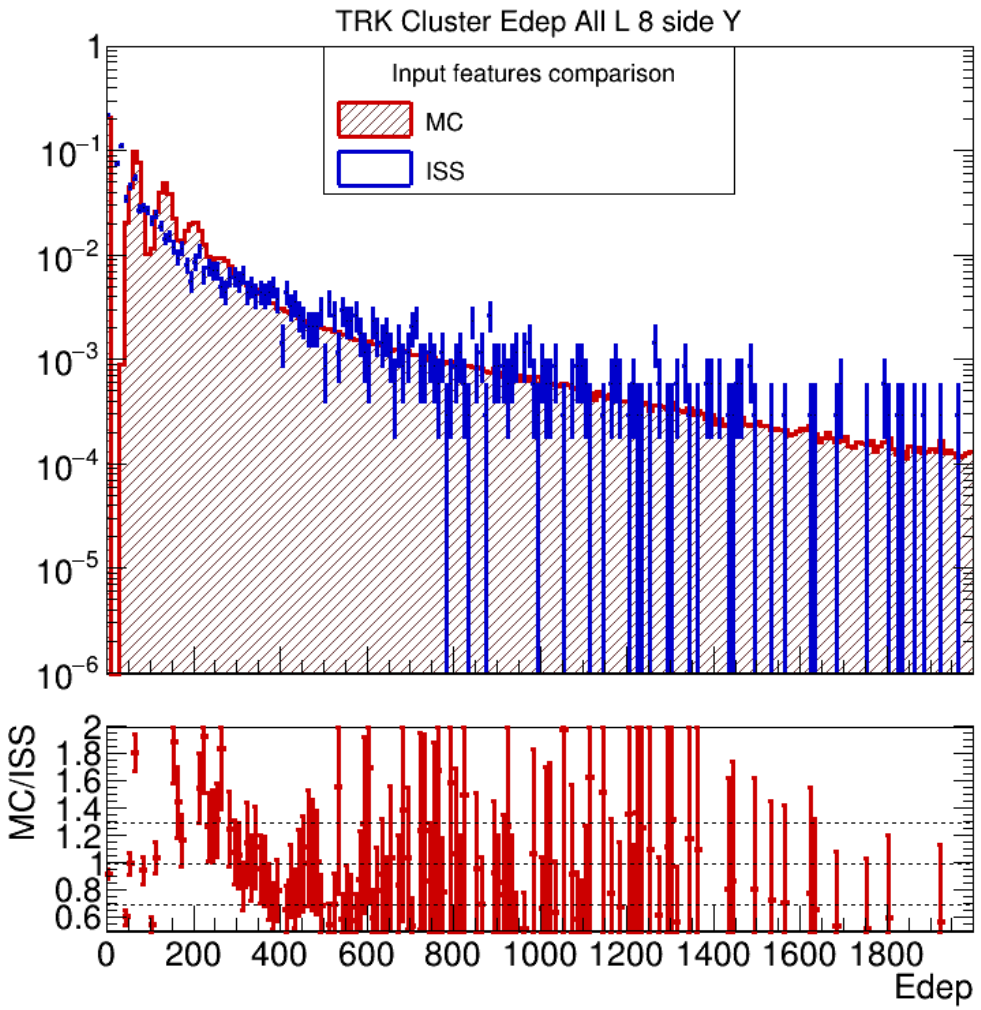
Track hit energy deposition on L8.



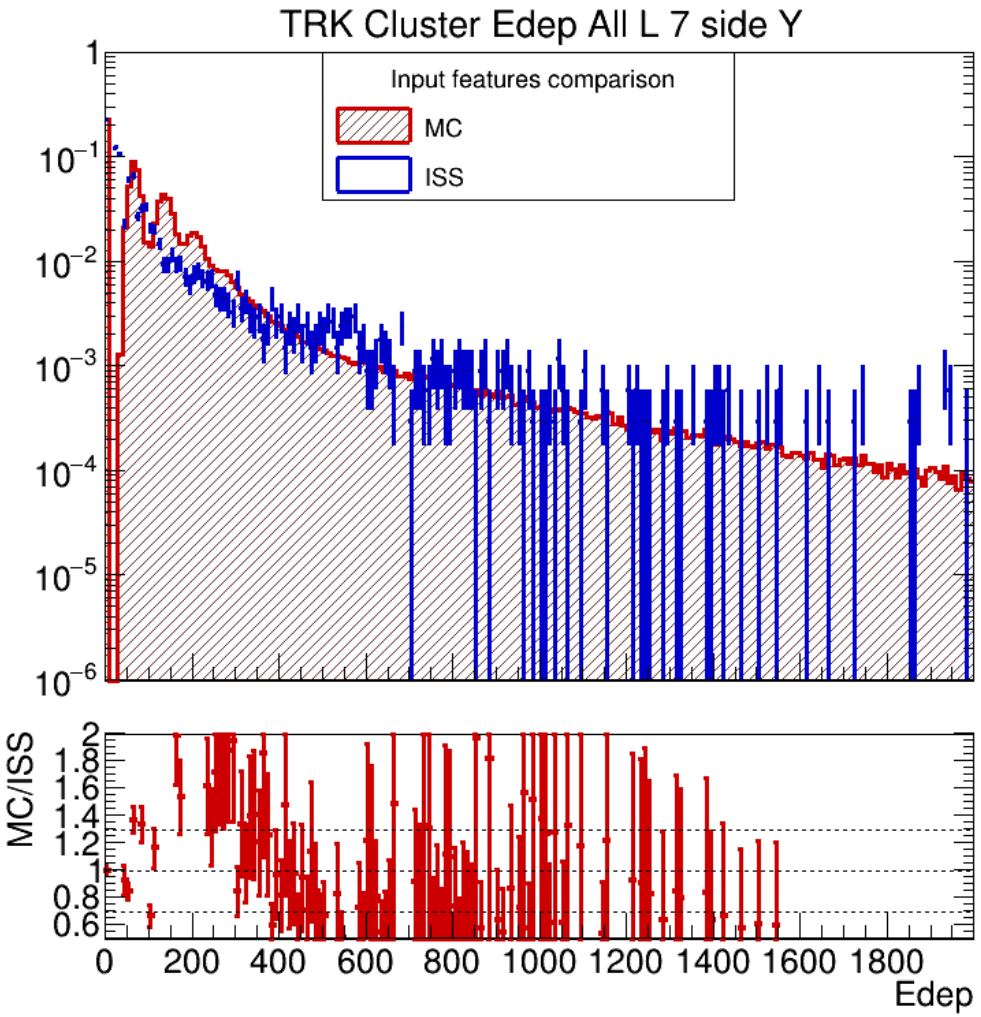
Track hit energy deposition on L7.



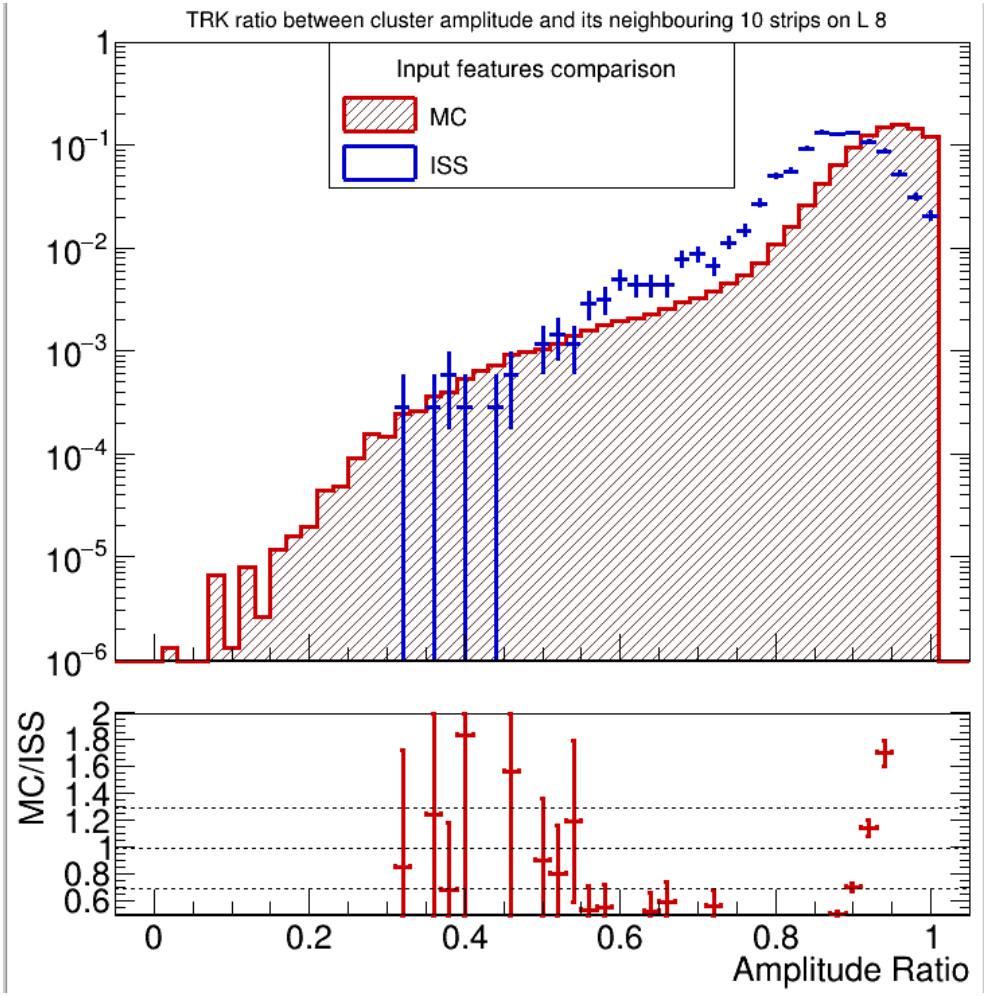
Total energy deposition of tracker clusters on L8 Y-side .



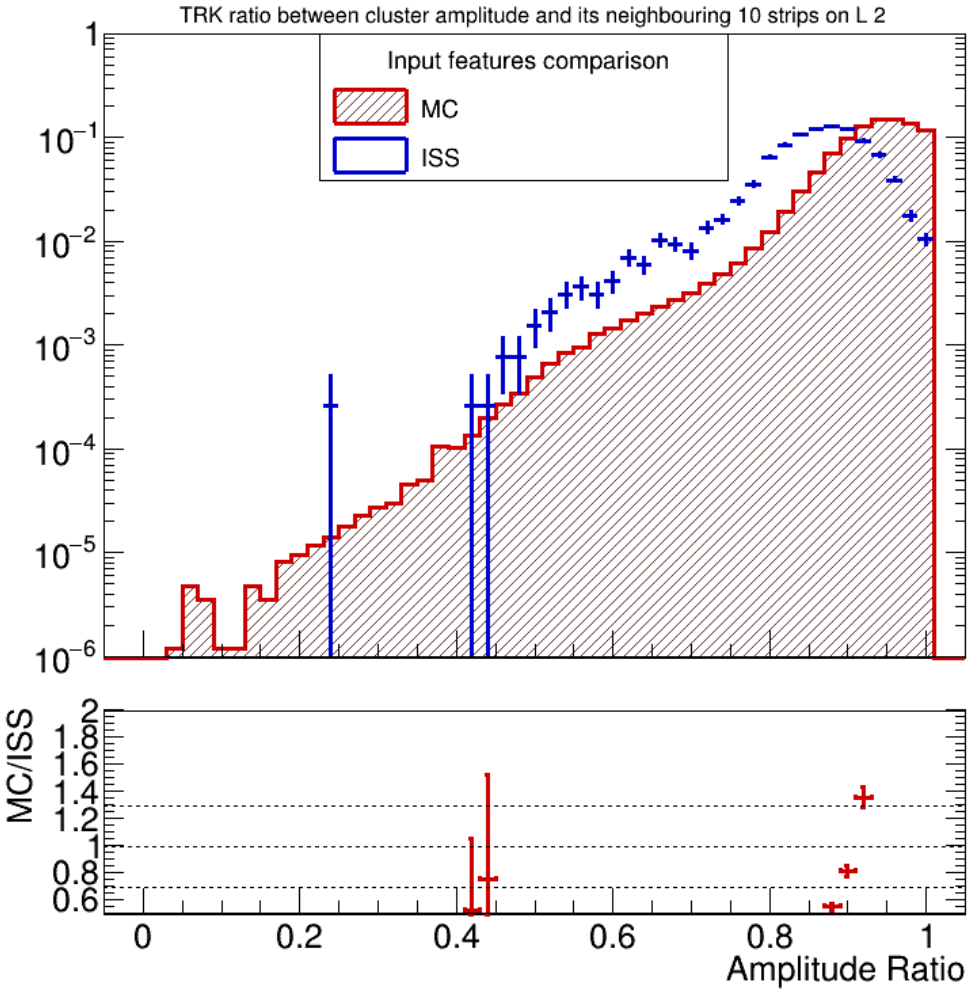
Total energy deposition of tracker clusters on L7 Y-side .



Ratio between cluster amplitude and its neighbouring 10 strips (L8)



Ratio between cluster amplitude and its neighbouring 10 strips (L2)



ISS selections for skimming and analysis

ISS-data skimming cutflow

No cut	1.3638e+10	1
RTI good	1.36258e+10	0.999109
SAA	1.28328e+10	0.941883
LiveTimeFraction > 0.5	1.27602e+10	0.936756
Zenith < 40 deg	1.27443e+10	0.93578
Physics Trigger	1.20464e+10	0.887975
NBetaClusters > 4	7.89321e+09	0.581799
Beta > 0	6.61817e+09	0.487718
BetaChi2Coo < 8	6.15594e+09	0.453648
NTracks >= 1	6.15594e+09	0.453648
Track pattern 5/8	1.65662e+09	0.122044
GBL fit Inner	1.65602e+09	0.121999
GBL fit InL1	1.41551e+09	0.104254
Inner CH [1.7, 2.4]	1.60342e+08	0.0118012
L1 CH [1.6, 3.0]	1.54294e+08	0.0113556
Inner fiducial volume	1.43048e+08	0.01053
L1 fiducial volume	1.3741e+08	0.0101151
ChiSq Inner < 20	1.32873e+08	0.00978084

0 1 2

RTIgood = 0
isInSAA
LiveTimeFraction > 0.5
Zenith < 40 deg

IsPhysicsTrigger

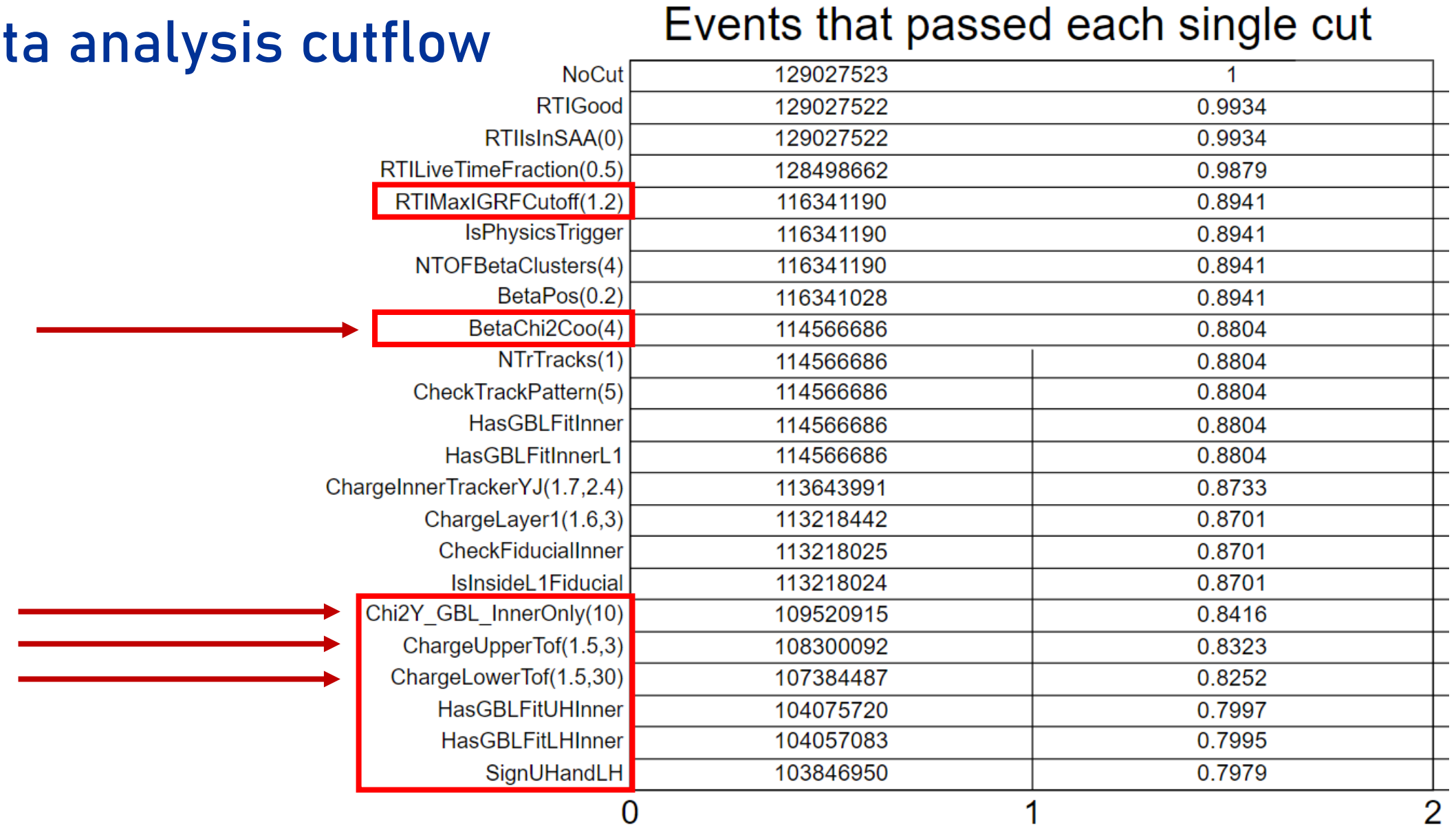
TOF hits ≥ 4
$\beta > 0$
$\chi^2_{COO} < 8$

Track number ≥ 1
charge STD (inner) $\in [1.7, 2.4]$
Inner fiducial volume
charge STD (L1) $\in [1.6, 3.0]$
track pattern L1&L2&(L3|L4)&(L5|L6)&(L7|L8) (≥ 5)

$\chi^2_Y < 20$

Tracker fiducial volume cut:
L1: |R|<62cm, |Y|<47cm
L2: |R|<62cm, |Y|<40cm
L3: |R|<46cm, |Y|<44cm
L4: |R|<46cm, |Y|<44cm
L5: |R|<46cm, |Y|<36cm
L6: |R|<46cm, |Y|<36cm
L7: |R|<46cm, |Y|<44cm
L8: |R|<46cm, |Y|<44cm

ISS-data analysis cutflow

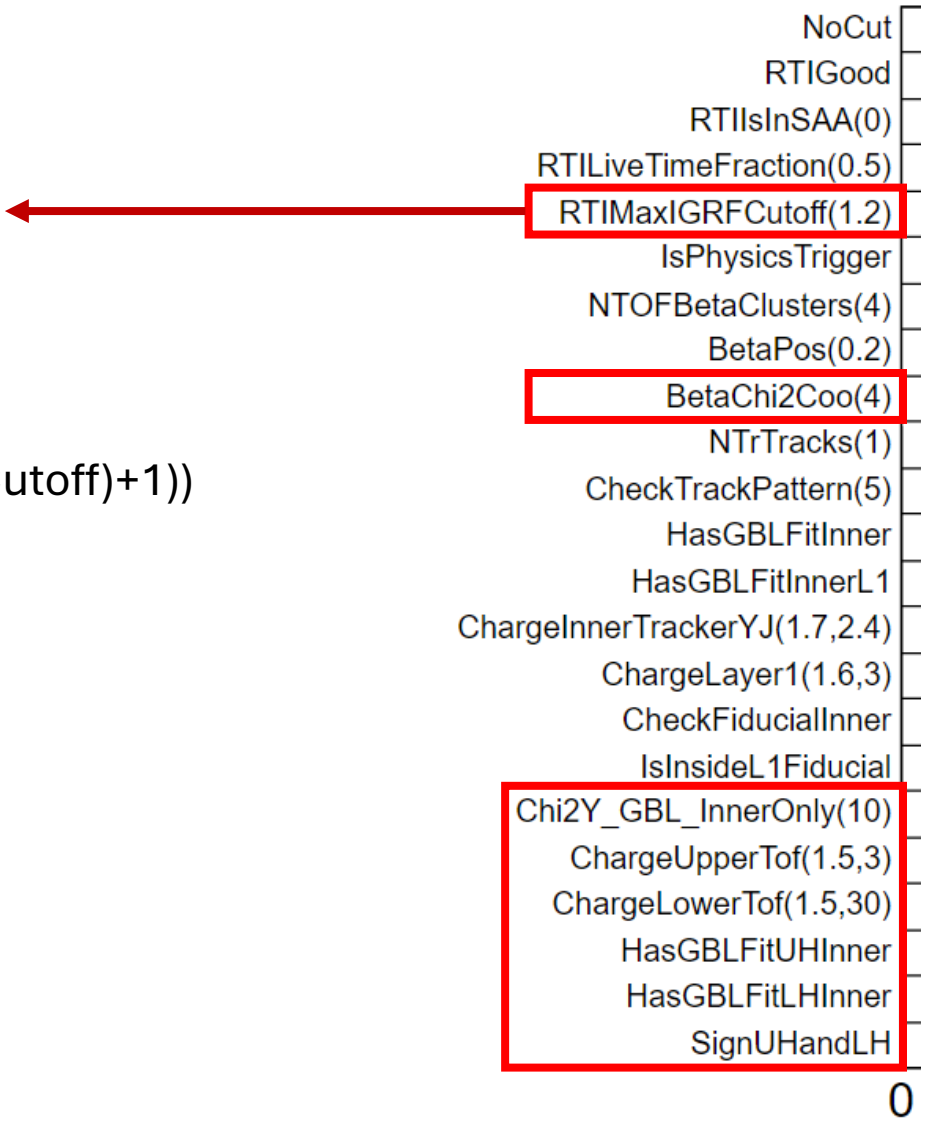


ISS-data analysis cutflow

L1Inner = Rig. [L1-inner]
IGRFCutoff = rti_info.MaxIGRFCutoff[40°][ch_sign]
Hist = Rig. Histogram with binning from He_7_5y

if (L1Inner <= hist->GetBinLowEdge(hist->FindBin(1.2 * IGRFCutoff)+1))

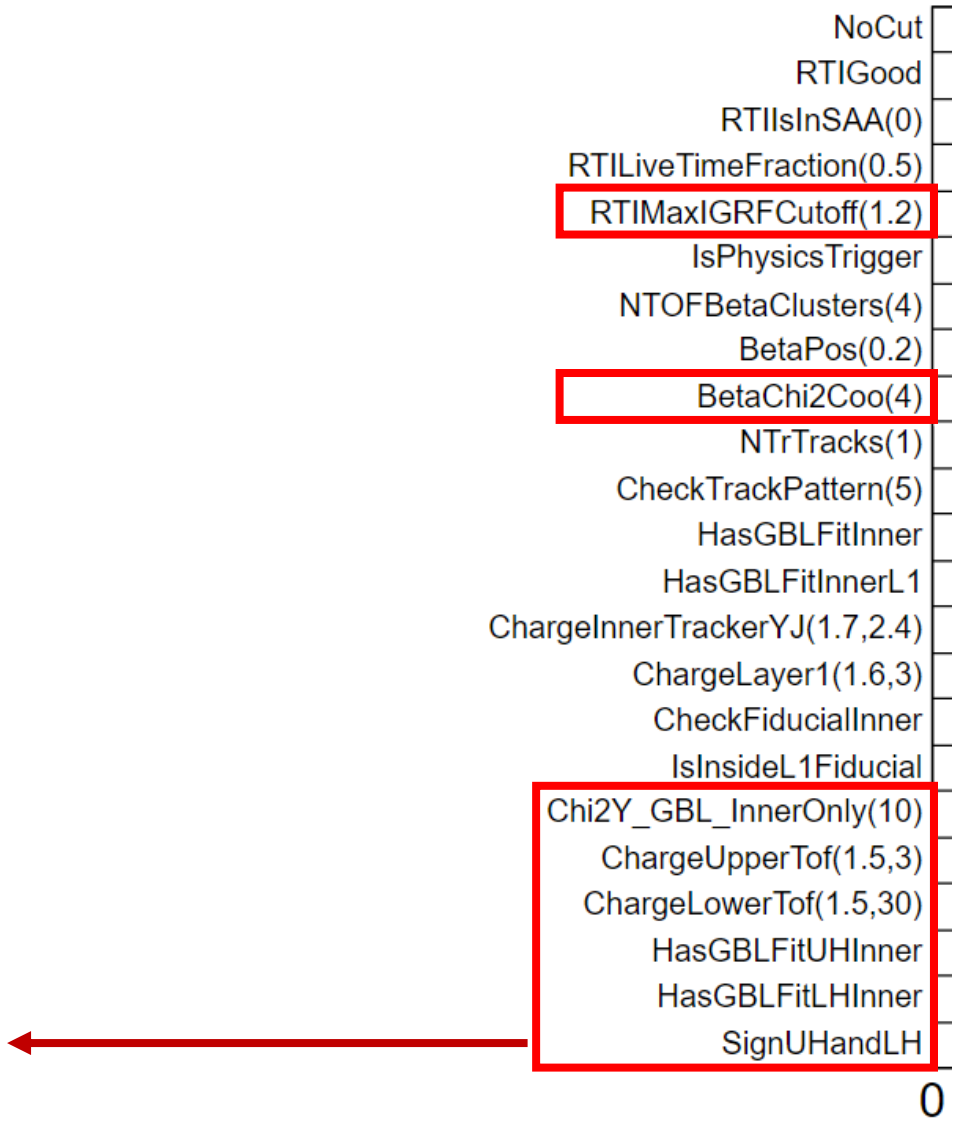
Do I have to consider positive or negative charges?



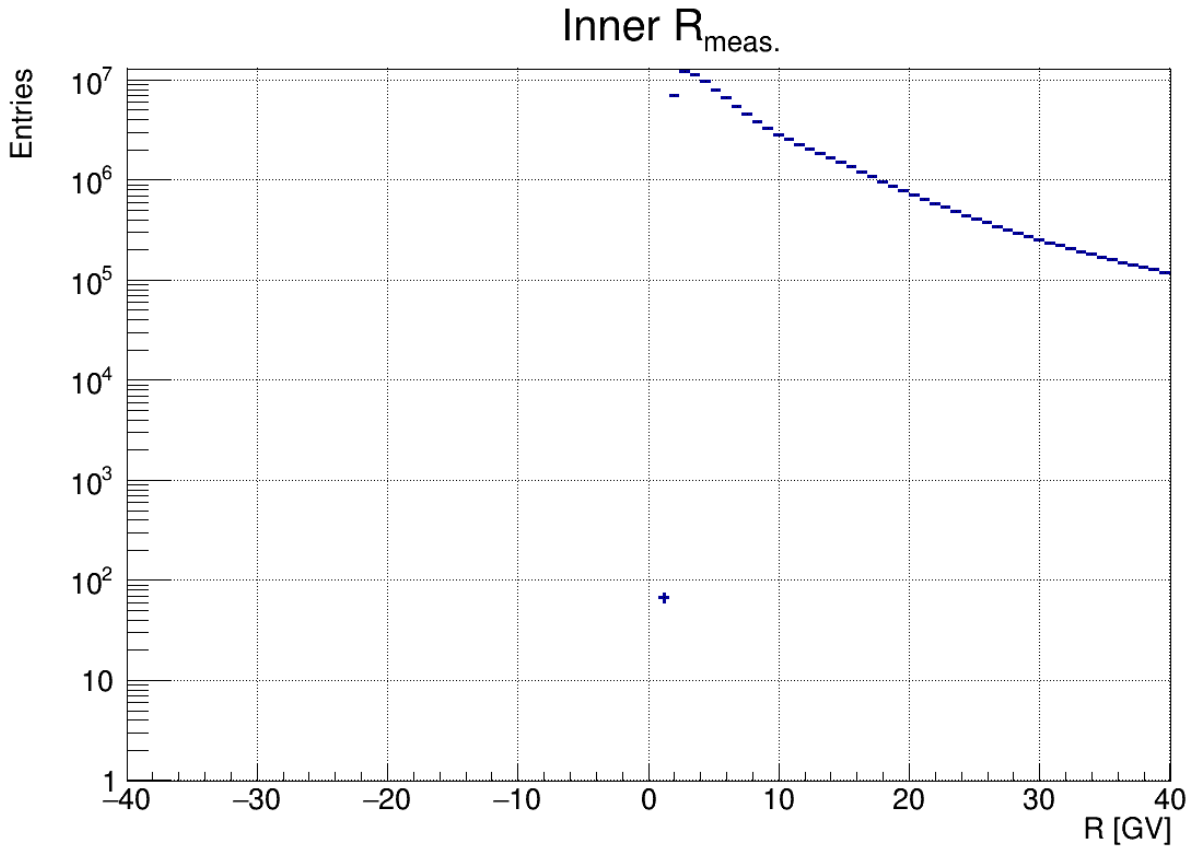
ISS-data analysis cutflow

UH = Rig. [UH-inner]
LH = Rig. [LH-inner]
Inner = Rig. [inner]

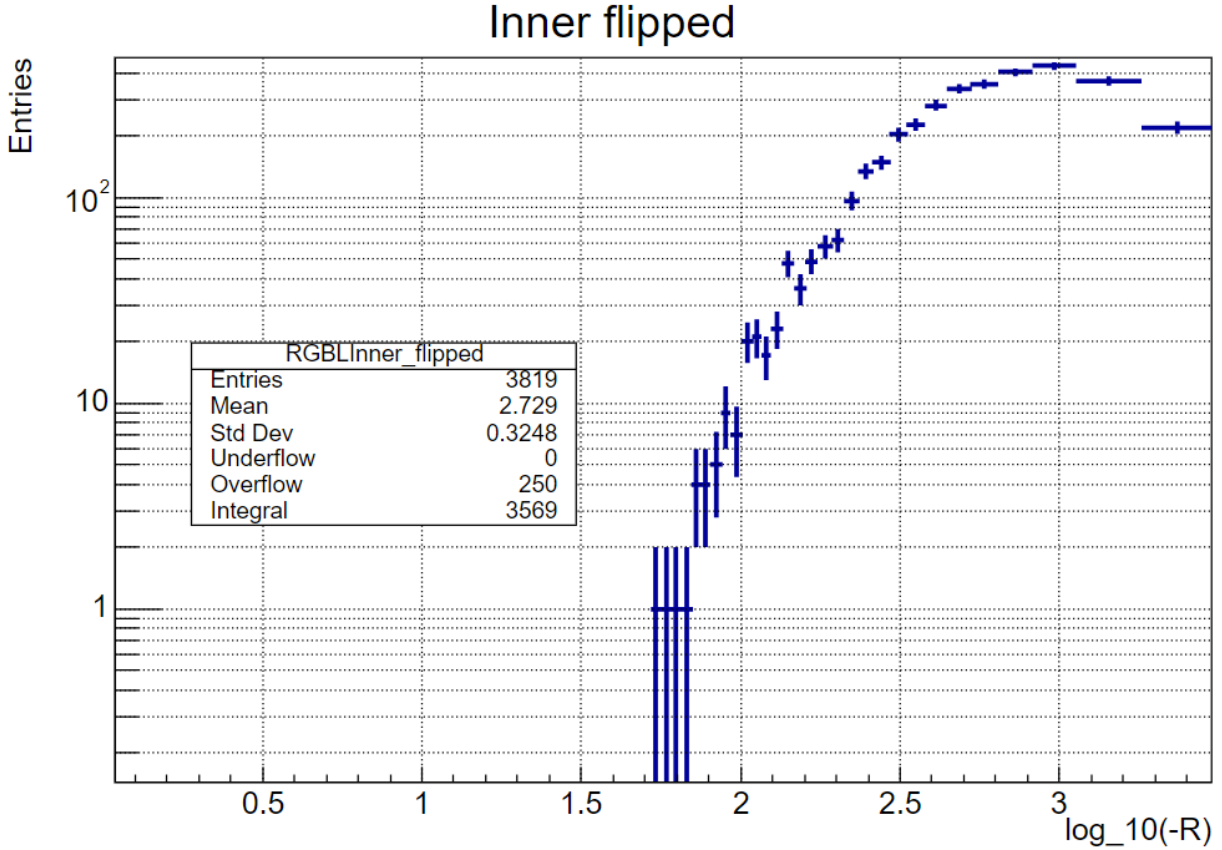
If(Inner > 0) → $(\mathrm{UH} > 0, \mathrm{LH} > 0)$
If(Inner < 0) → $(\mathrm{UH} < 0, \mathrm{LH} < 0)$



Measured inner tracker rigidity:



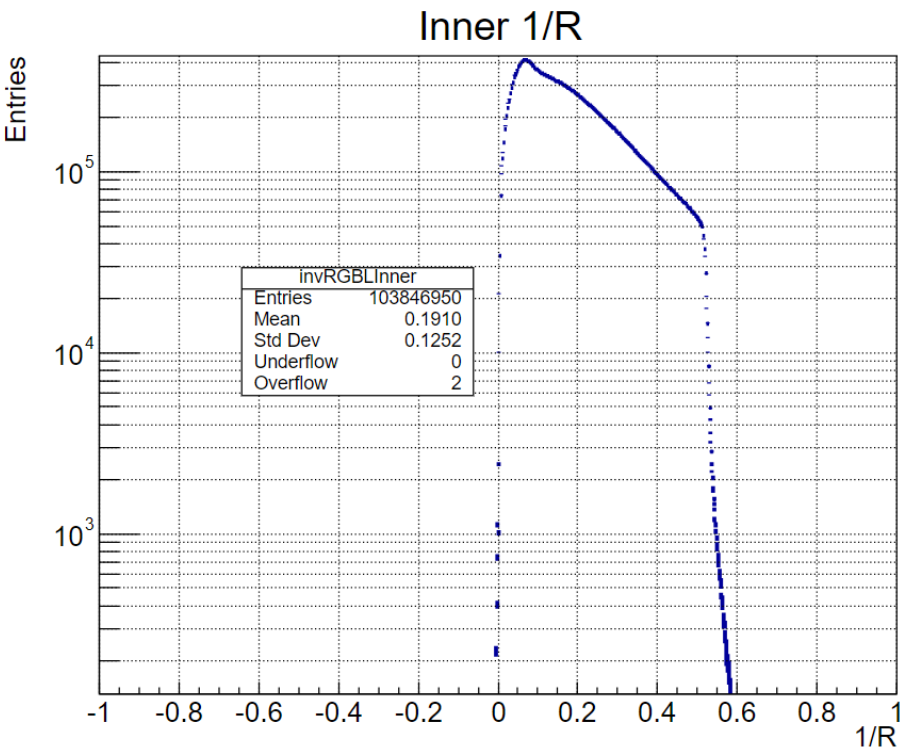
Measured inner tracker rigidity < 0:

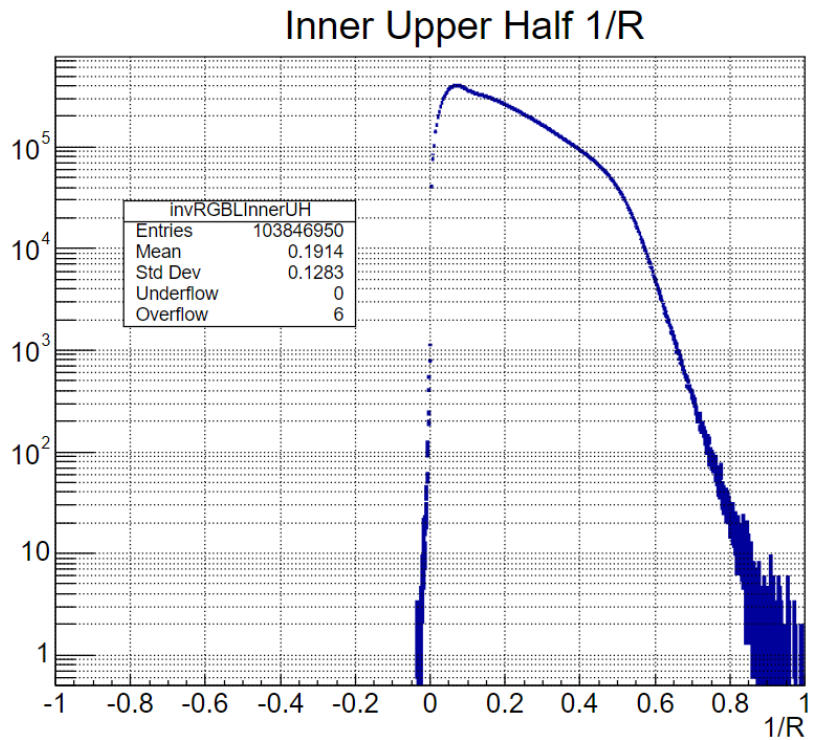


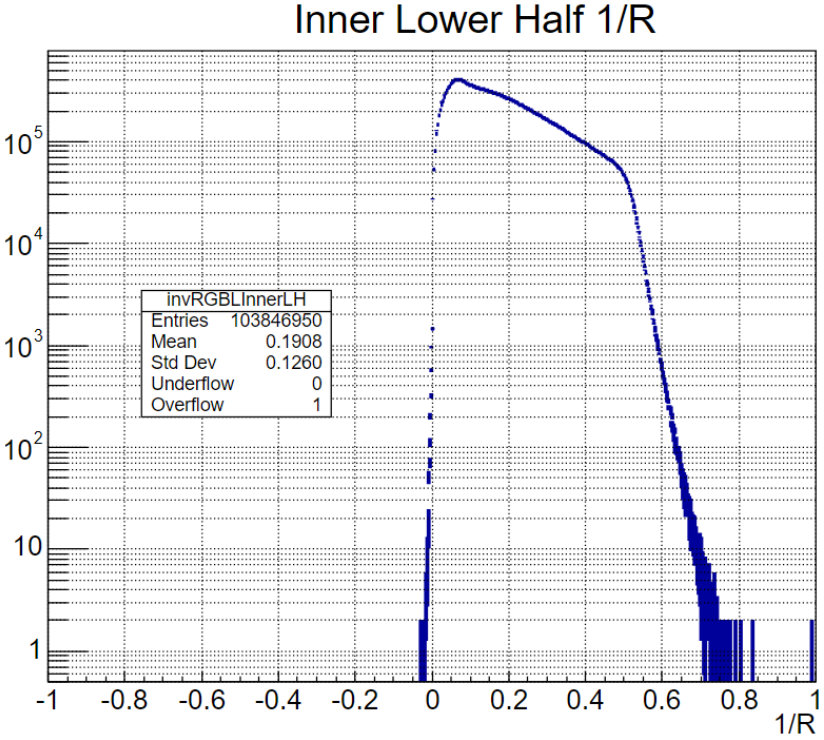
No events with $R_{inner} \in [-40, 0]$.

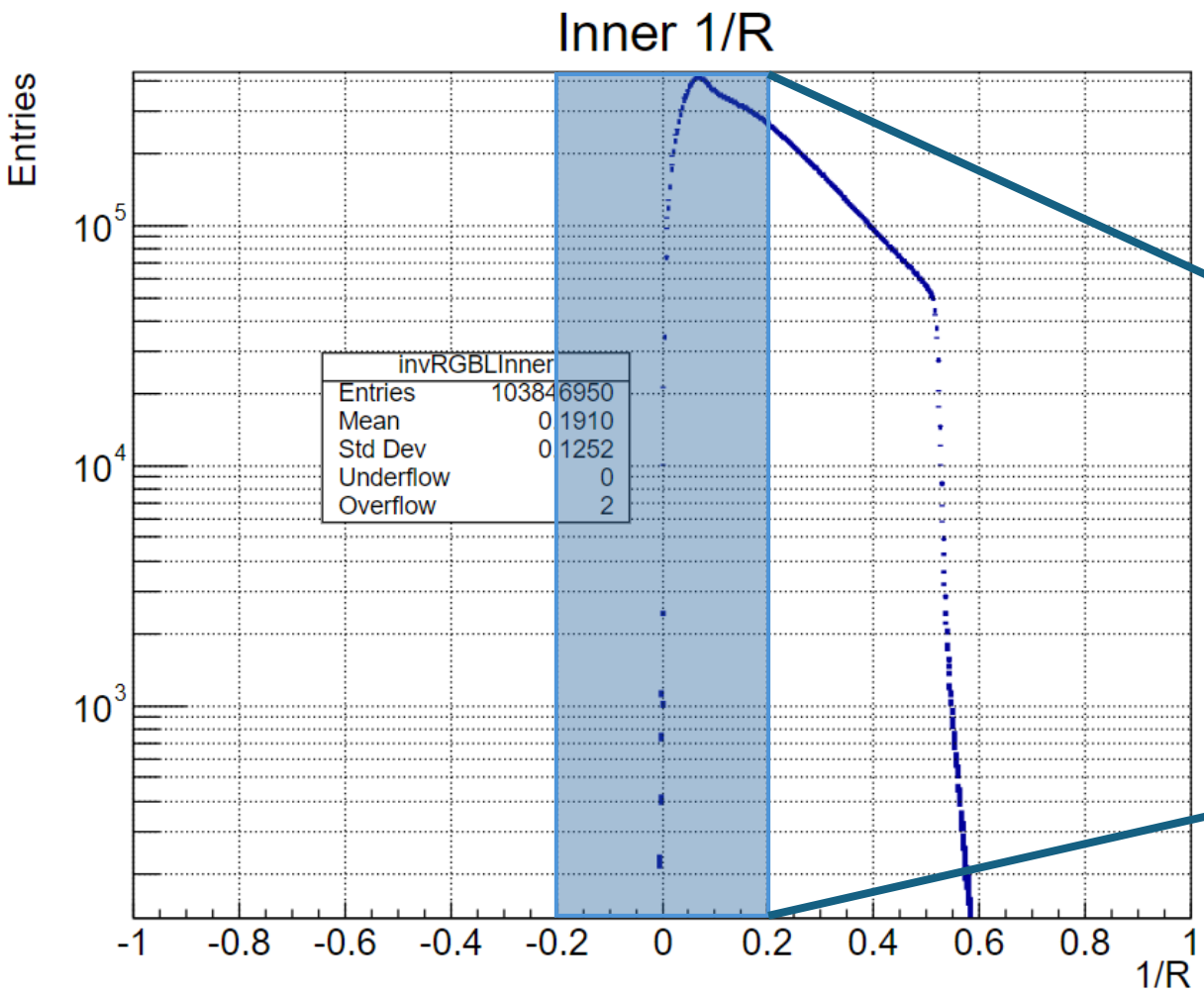
$1.3 \cdot 10^8$ events with $R_{inner} > 0$

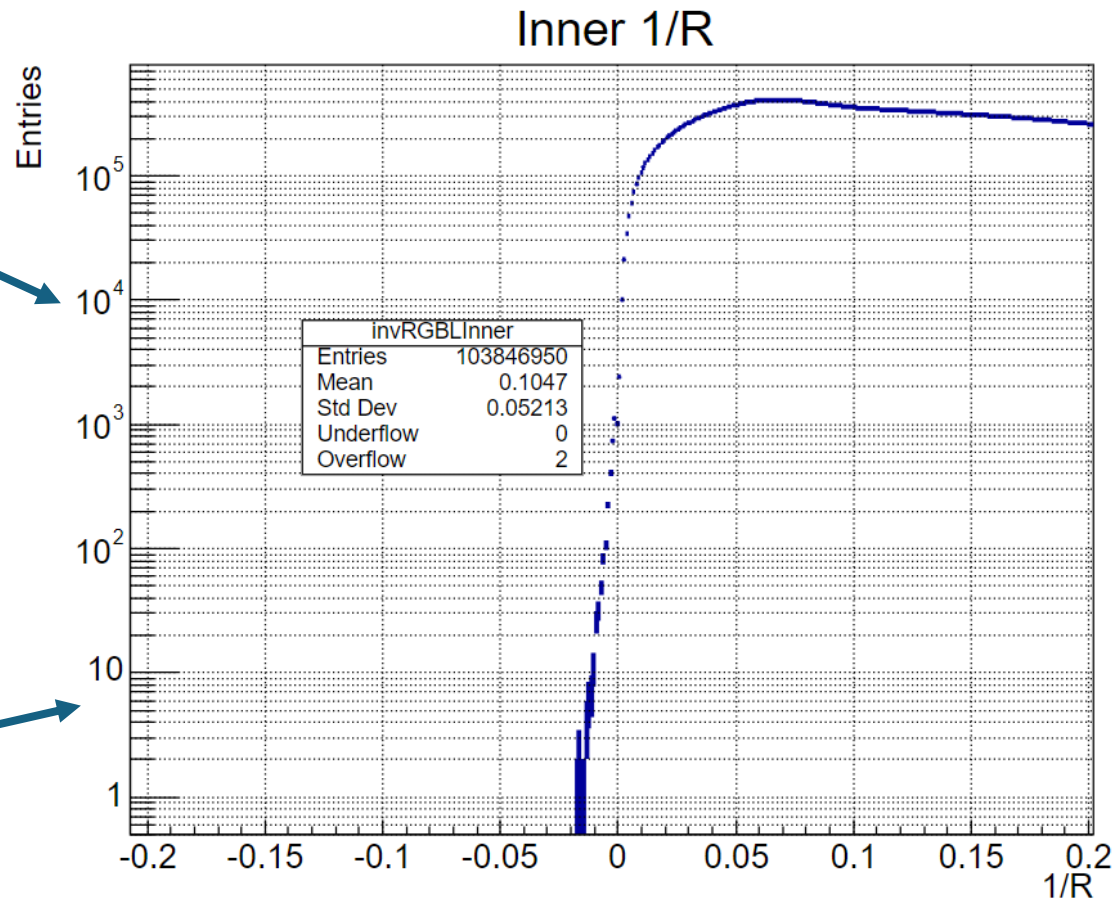
$3.8 \cdot 10^3$ events with $R_{inner} < 0$


Inner 1/R
Entries
invRGBLInner
Entries 103846950
Mean 0.1910
Std Dev 0.1252
Underflow 0
Overflow 2
1/R


Inner Upper Half 1/R
invRGBLInnerUH
Entries 103846950
Mean 0.1914
Std Dev 0.1283
Underflow 0
Overflow 6
1/R


Inner Lower Half 1/R
invRGBLInnerLH
Entries 103846950
Mean 0.1908
Std Dev 0.1260
Underflow 0
Overflow 1
1/R

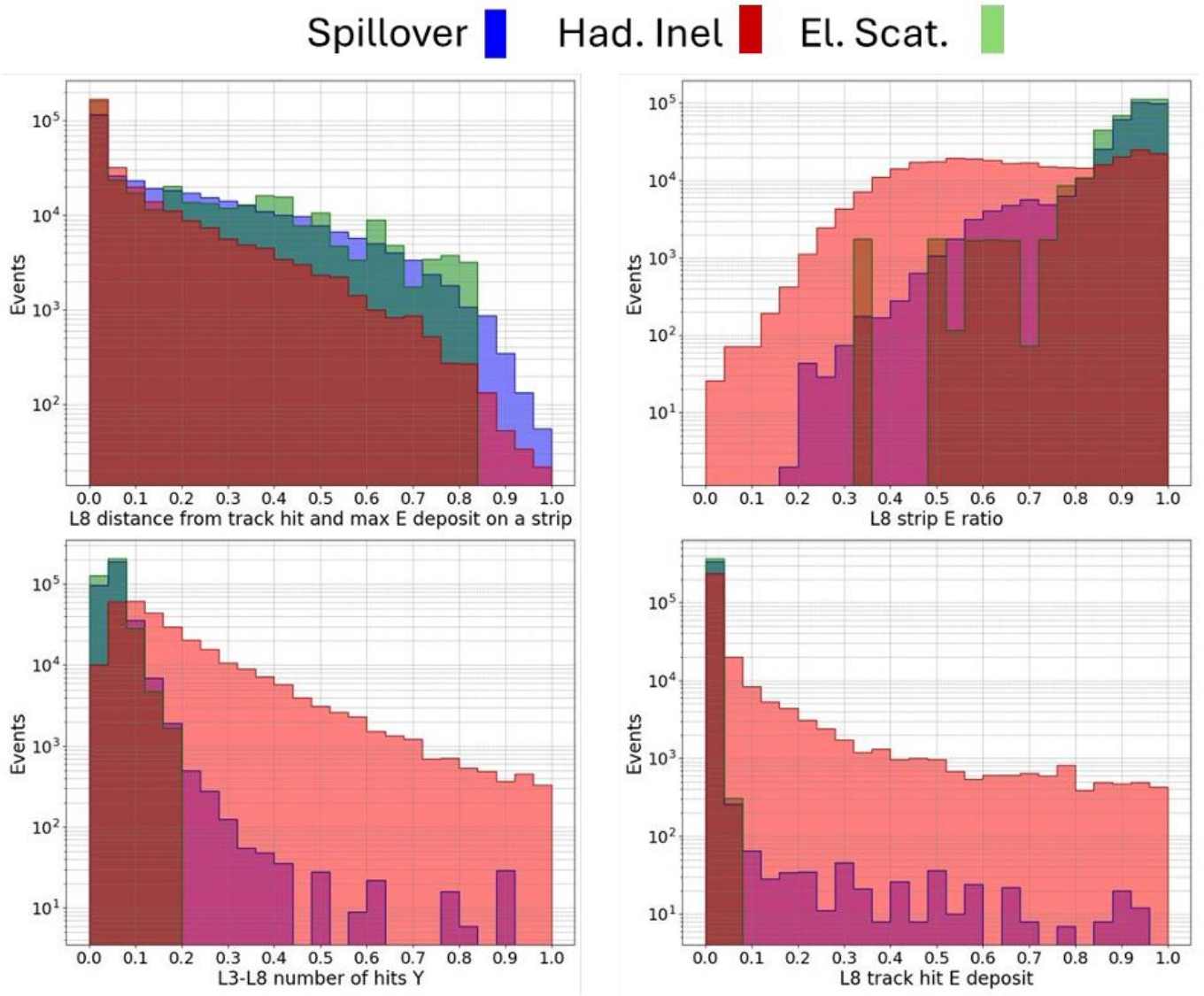

Inner 1/R
Entries
invRGBLInner
Entries 103846950
Mean 0.1910
Std Dev 0.1252
Underflow 0
Overflow 2
1/R


Inner 1/R
Entries
invRGBLInner
Entries 103846950
Mean 0.1047
Std Dev 0.05213
Underflow 0
Overflow 2
1/R

FCNN input features (old selection)

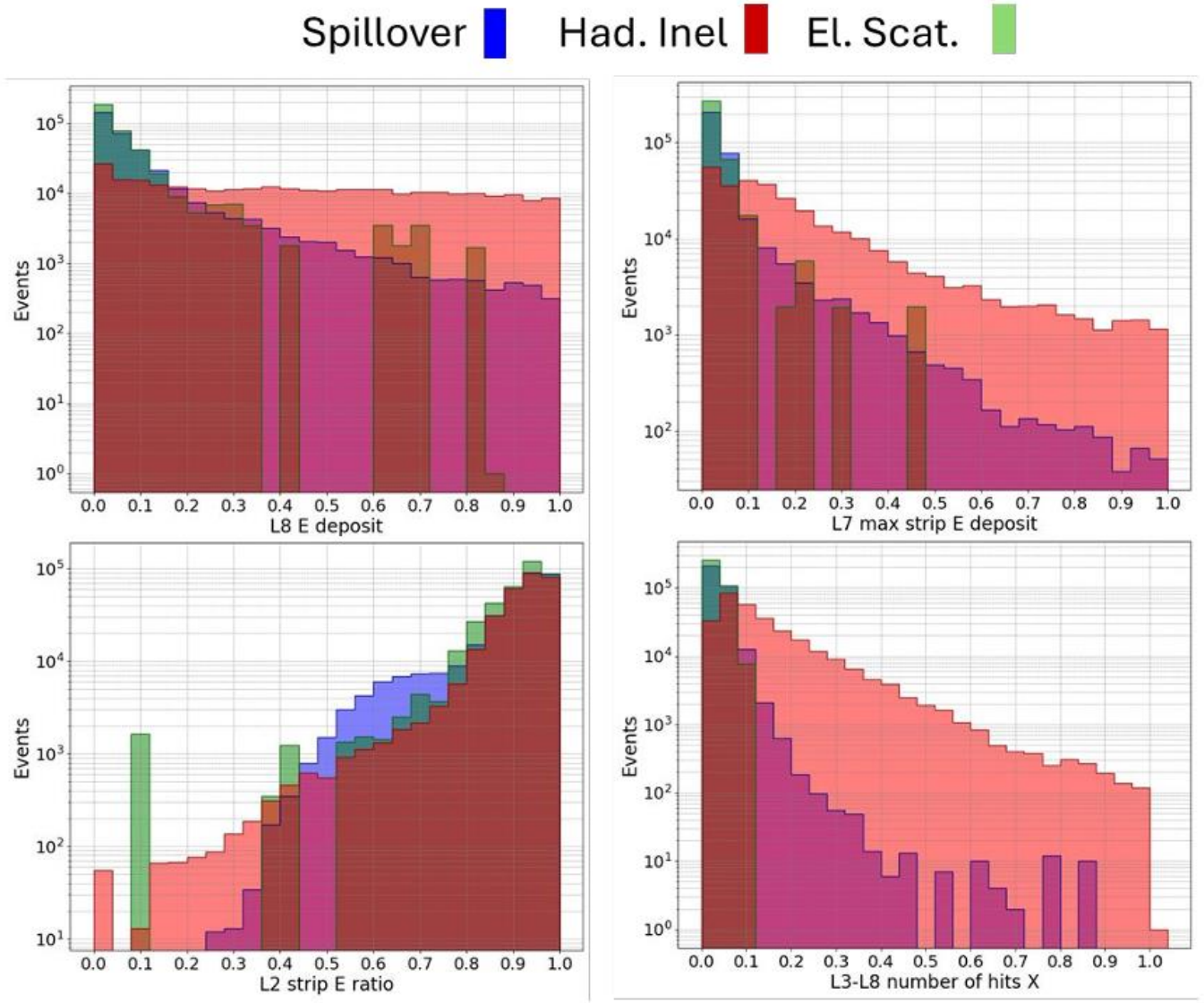
Input features 1

- Distance between track hit and strip with max energy deposit on L8.
- Ratio between strip energy deposition and and its neighbouring 10 strips, on L8.
- Number of hits in L3-L8 inner tracker Y side.
- Track hit energy deposition on L8.



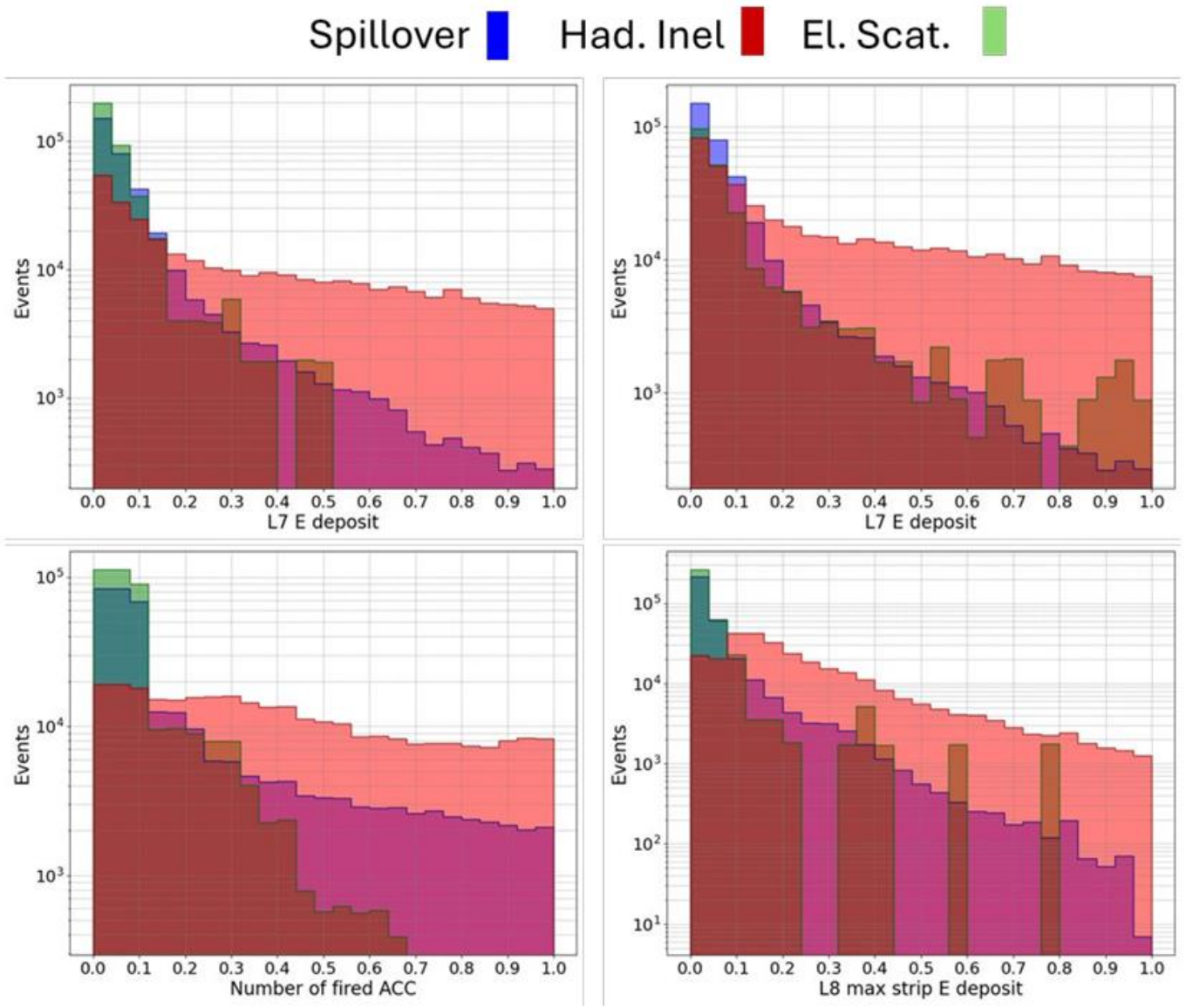
Input features 2

- Total energy deposition on L8.
- Max energy deposit on a single strip on L7.
- Ratio between strip energy deposition and and its neighbouring 10 strips, on L2.
- Number of hits in L3-L8 inner tracker X side.



Input features 3

- Track hit energy deposition on L7.
- Total energy deposition on L7.
- Number of fired ACC.
- Max energy deposit on a single strip on L8.



Correlation matrix and features ranking