



# Detector simulation and digitization

(studying the simulation with Nuclear Recoils)



P.Meloni , F. Petrucci



The University Of Sheffield.



CYGNO simulation was studied and optimized on Electron Recoils with LIME overground data.

PREPARED FOR SUBMISSION TO JINST

## Simulation of the CYGNO Gaseous TPC Optical Readout

F.D. Amaro<sup>1</sup> R. Antonietti<sup>2,3</sup> E. Baracchini<sup>4,5</sup> L. Benussi<sup>6</sup> S. Bianco<sup>6</sup> C. Capocchia<sup>6</sup> M. Caponero<sup>6,7</sup> L.G.M de Carvalho<sup>8</sup> G. Cavoto<sup>9,10</sup> I.A. Costa<sup>6</sup> A. Croce<sup>6</sup> M. D'Astolfo<sup>4,5</sup> G. D'Imperio<sup>10</sup> G. Dho<sup>6</sup> E. Di Marco<sup>10</sup> J.M.F. dos Santos<sup>1</sup> D. Fiorina<sup>4,5</sup> F. Iacoangeli<sup>10</sup> Z. Islam<sup>4,5</sup> E. Kemp<sup>11</sup> H. P. Lima Jr<sup>4,5</sup> G. Maccarrone<sup>6</sup> R.D.P. Mano<sup>1</sup> D. J. G. Marques<sup>4,5</sup> G. Mazzitelli<sup>6</sup> P. Meloni<sup>2,3</sup> A. Messina<sup>9,10</sup> C.M.B. Monteiro<sup>1</sup> R.A. Nobrega<sup>8</sup> I.F. Pains<sup>8</sup> E. Paoletti<sup>6</sup> F. Petrucci<sup>2,3</sup> S. Piacentini<sup>4,5</sup> D. Pierluigi<sup>6</sup> D. Pinci<sup>10</sup> F. Renga<sup>10</sup> A. Russo<sup>6</sup> G. Saviano<sup>6,12</sup> P.A.O.C. Silva<sup>1</sup> N. J. C. Spooner<sup>13</sup> R. Tesaro<sup>6</sup> S. Tomassini<sup>6</sup> S. Torelli<sup>4,5</sup> D. Tozzi<sup>9,10</sup>

# Status

CYGNO simulation was studied and optimized on Electron Recoils with LIME overground data.

PREPARED FOR SUBMISSION TO JINST

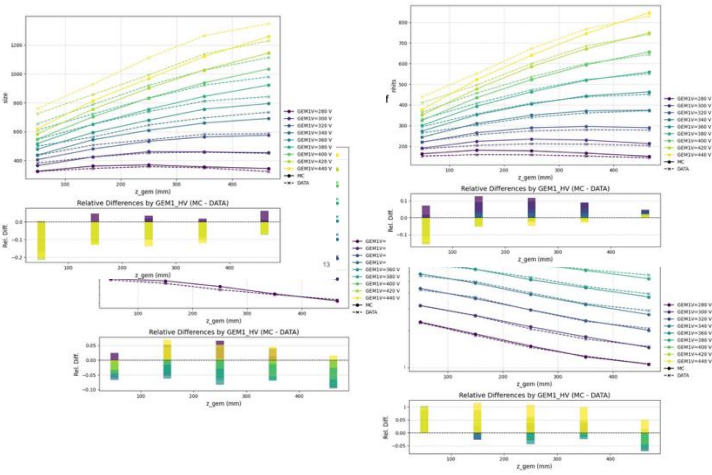
## Simulation of the CYGNO Gaseous TPC Optical Readout

F.D. Amaro<sup>1</sup> R. Antonietti<sup>2,3</sup> E. Baracchini<sup>4,5</sup> L. Benussi<sup>6</sup> S. Bianco<sup>6</sup> C. Capoccia<sup>6</sup> M. Caponero<sup>6,7</sup> L.G.M de Carvalho<sup>8</sup> G. Cavoto<sup>9,10</sup> I.A. Costa<sup>6</sup> A. Croce<sup>6</sup> M. D'Astoifo<sup>4,5</sup> G. D'Imperio<sup>10</sup> G. Dho<sup>6</sup> E. Di Marco<sup>10</sup> J.M.F. dos Santos<sup>1</sup> D. Fiorina<sup>4,5</sup> F. Iacoangeli<sup>10</sup> Z. Islam<sup>4,5</sup> E. Kemp<sup>11</sup> H. P. Lima Jr<sup>4,5</sup> G. Maccarrone<sup>6</sup> R.D.P. Mano<sup>1</sup> D. J. G. Marques<sup>4,5</sup> G. Mazzitelli<sup>6</sup> P. Meloni<sup>2,3</sup> A. Messina<sup>9,10</sup> C.M.B. Monteiro<sup>1</sup> R.A. Nobrega<sup>8</sup> I.F. Pains<sup>8</sup> E. Paoletti<sup>6</sup> F. Petrucci<sup>2,3</sup> S. Piacentini<sup>4,5</sup> D. Pierluigi<sup>6</sup> D. Pinci<sup>10</sup> F. Renga<sup>10</sup> A. Russo<sup>6</sup> G. Saviano<sup>6,12</sup> P.A.O.C. Silva<sup>1</sup> N. J. C. Spooner<sup>13</sup> R. Tesauro<sup>6</sup> S. Tomassini<sup>6</sup> S. Torelli<sup>4,5</sup> D. Tozzi<sup>9,10</sup>

Extensive validation and further optimization in underground data taking, including dedicated studies varying Vgem. Simulation closely reproduce ER tracks behaviour.

Summary with all the information in a presentation from Pietro [in a presentation in July 2025](#)

N.B.: digitization heavily depends on gain and attenuation length that NEED to be evaluated for each time period. The procedure to use digitization is there, but cannot be plug&play



# Extending simulation validation to Nuclear Recoils

It is mandatory to validate our simulation model also with Nuclear Recoils:

- some test with alphas already ongoing
- AmBe source runs allows to study NR from neutrons

AmBe data are available in Run3 and Run5.

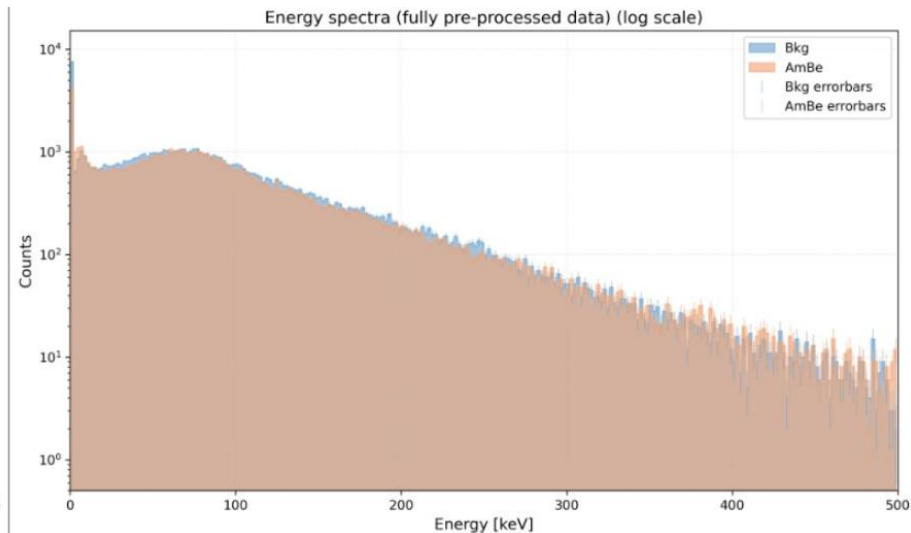
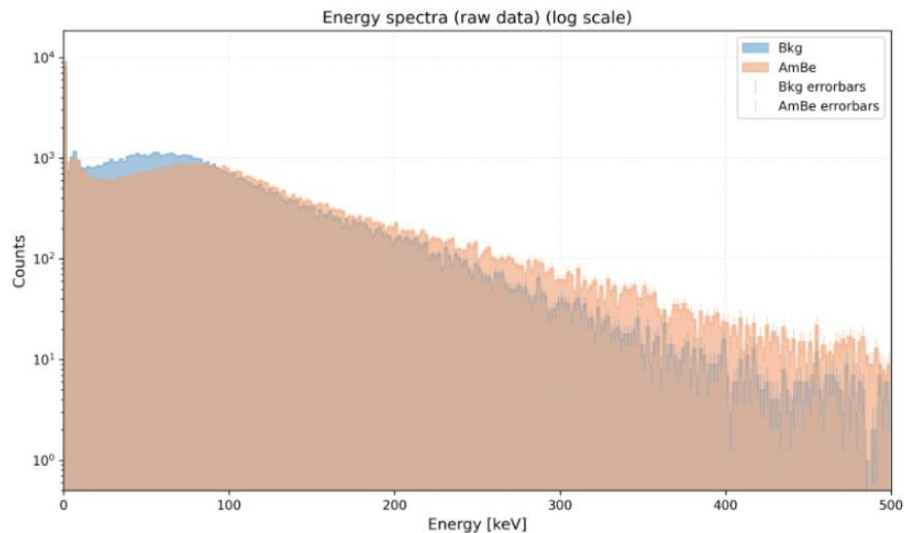
We started with the Run3 dataset.

Dataset	Run range	Date range	Effective Exposure [h]
Stability	25486–25772	Sept. 26–29	10.5
Bkg_calibration	27512	Oct. 9	–
Bkg	27322–27844	Oct. 7–10	56.6
AmBe_p1	23820–23984	Aug. 2–3	16.9
AmBe_p2	24023–24328	Aug. 3–4	31.0
AmBe + $^{55}\text{Fe}$	23988–24022	Aug. 3	2.1

# Energy spectra

Normalization: background and source data samples are normalized using the effective exposure

Pre-processing: data are normalized using calibration runs to account for instabilities (mainly light yield correction for humidity variations)



# AmBe simulation

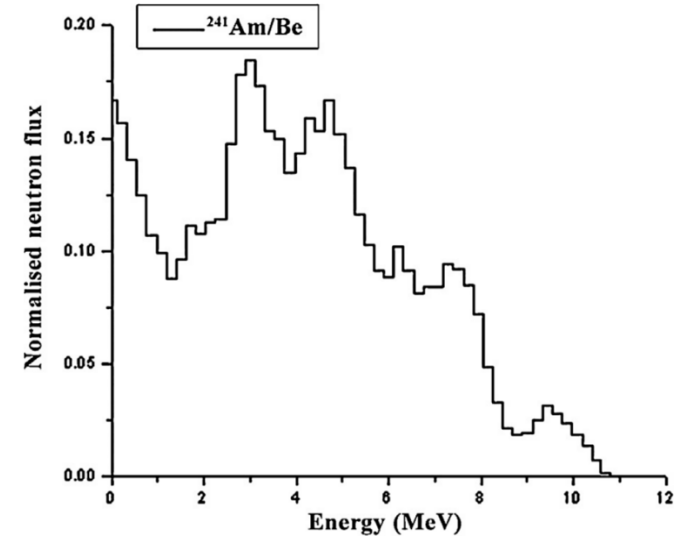
3.5  $10^5$  neutrons generated with Geant4:

- Neutron energy spectrum: 0-11 MeV (AmBe distribution)
- Quenching factor incorporated for reduced ionization yield of NRs.

Generated events are digitized using the optimized parameters and reconstructed

Only  $\sim 10^4$  events produce measurable deposits in gas:

- Low rate reflects geometrical acceptance and interaction probabilities



Large ER Background from Neutrons is found:

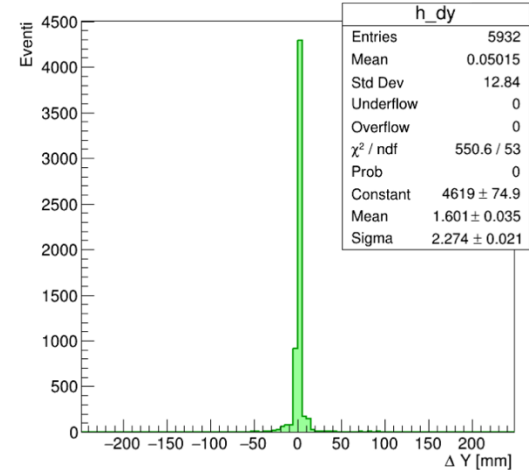
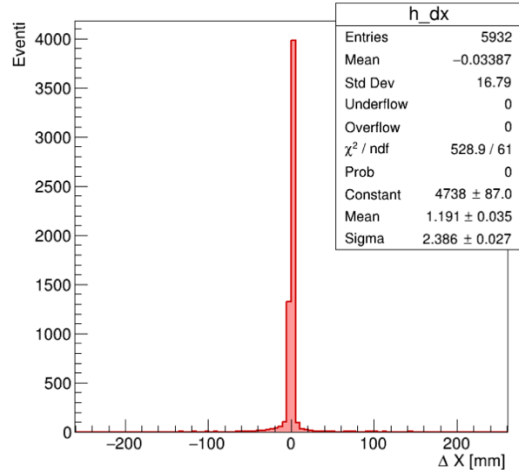
- More than half of signal events are electron recoils, not nuclear recoils;
  - Neutrons interact with PMMA vessel and Cu shielding generating Photons and electrons
- AmBe is not a "pure" NR source: **discrimination is essential** to test MC on NR

# Matching truth and reco quantities in simulation

MC truth (particle type, true energies, vertex) not preserved after reconstruction.  
Clustering may cut or merge tracks.

Reco tracks and true particles must be matched again:

- Compute barycentre (x,y) for each reconstructed cluster
- Extract all Geant4 hits in gas, group by track ID
- Calculate MC barycentre for each track ID
- Match to closest MC barycentre if distance < 10 mm



# Particle types in simulation

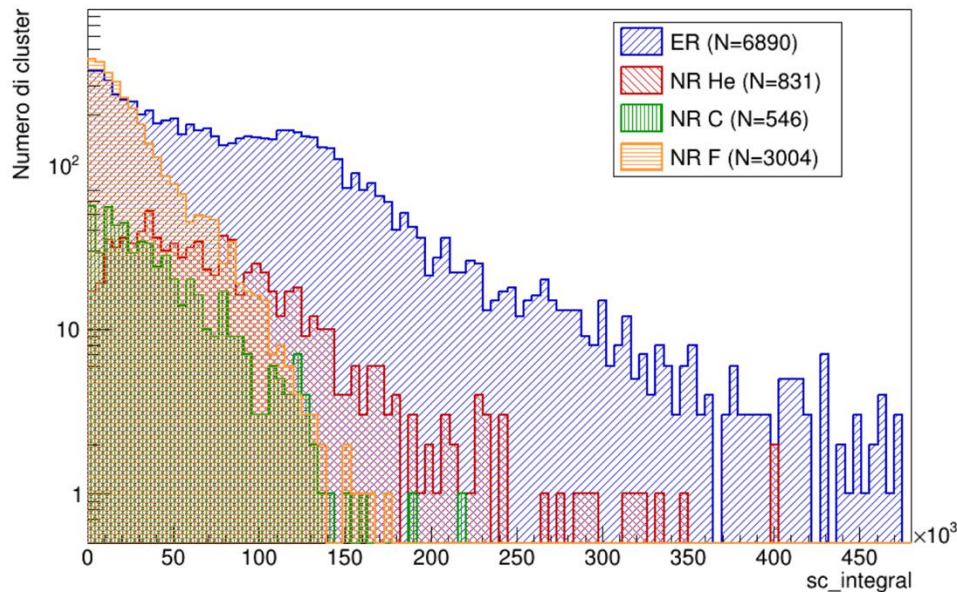
## Four Categories Based on Interacting Particle

- Electron Recoils (ER): e<sup>-</sup>, e<sup>+</sup>, or photons (6,890 events, 61%)
- Helium Recoils (He-NR): 831 events (7.4%)
- Carbon Recoils (C-NR): 546 events (4.8%)
- Fluorine Recoils (F-NR): 3,004 events (26.7%)

## Interpretation:

- F-NR prevalence reflects CF<sub>4</sub> composition and higher mass
- Total NRs: 4,381 events (38.9% of signal)
- ERs dominate: neutron interactions with structural materials

# Tracks total light integral

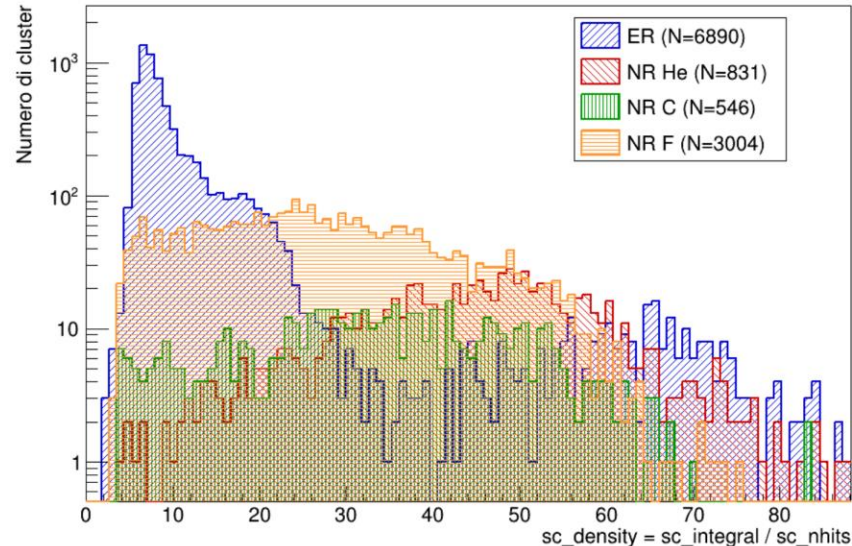
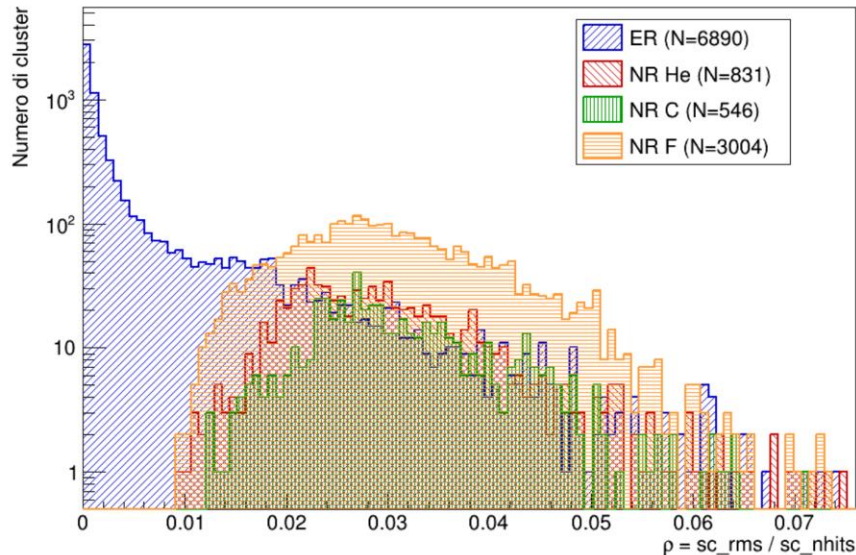


The distribution of the relevant track variables can be studied for the different particle types

For the **total light integral**:

- ERs dominate at higher values (extending beyond 250k ADU)
- NRs concentrate at lower values due to quenching effects.

# Tracks “density” variables



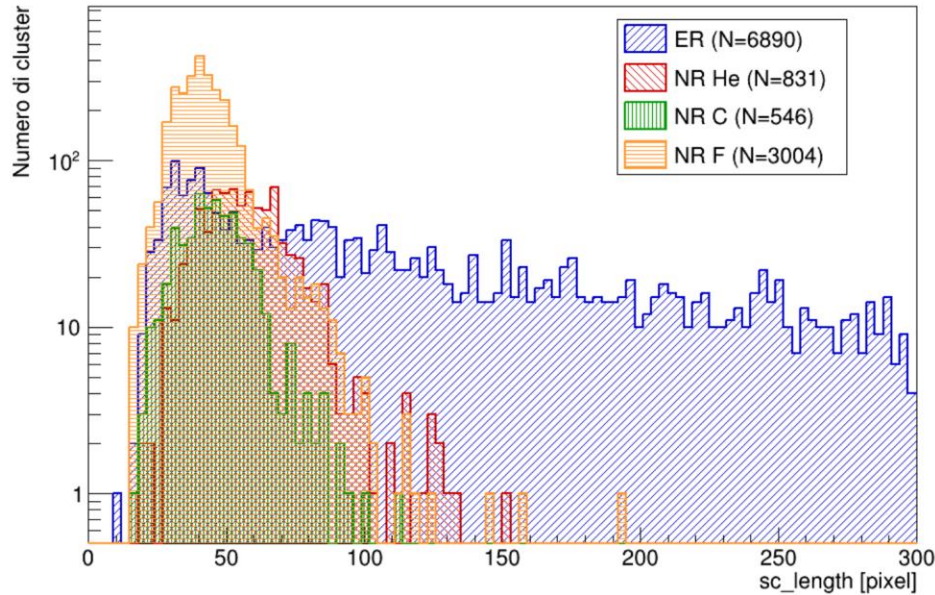
Clear separation in **rho** between ER ( $\rho < 0.01$ ) and NR ( $\rho > 0.01$ ) populations:

- Lower  $\rho$  values correspond to diffuse, discontinuous ionization (ER tracks).
- Higher  $\rho$  values indicate dense, localized charge deposition (NR tracks).

The **density** represents average charge per hit pixel:

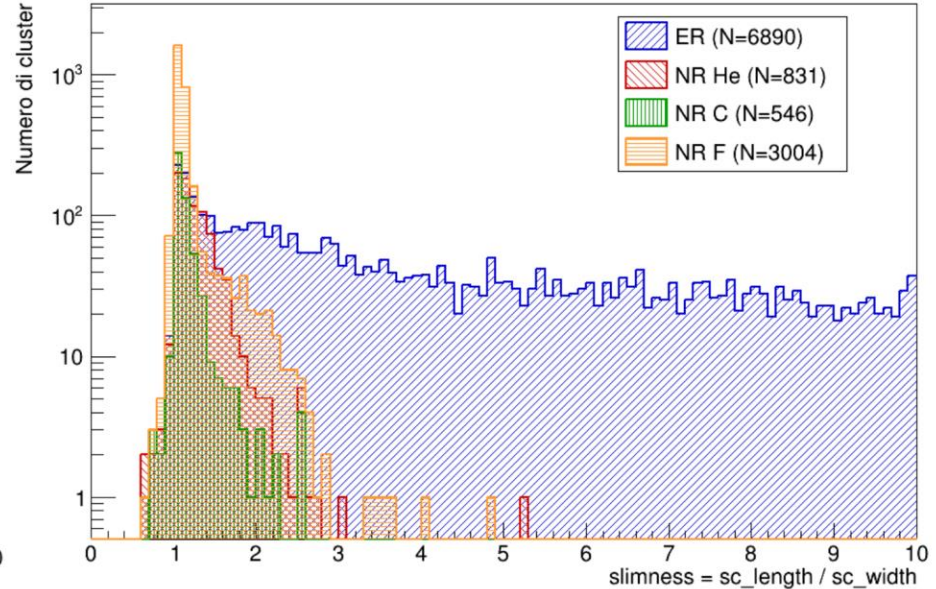
- ERs show sharp peak at low density.
- NRs exhibit broader distributions extending to higher densities

# Tracks length and slimness



The **length** reflect the characteristic ranges in CF<sub>4</sub>:He, determined by mass:

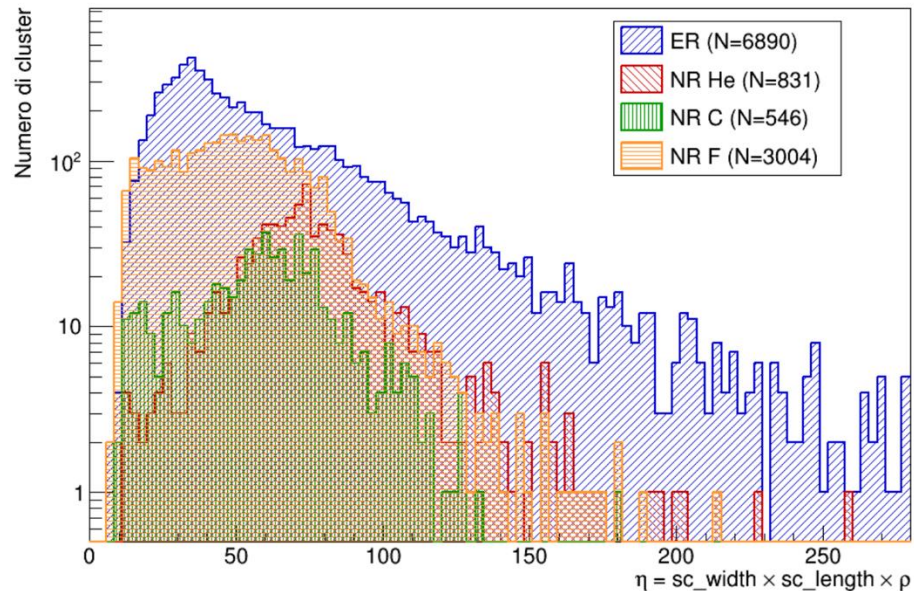
- He-NRs: longest tracks (~100 px).
- F-NRs: shortest tracks (~30 px).



The **slimness** is highly effective for discrimination:

- NRs: spot-like signatures
- Ers: elongated tracks with slimness values extending to ~20.

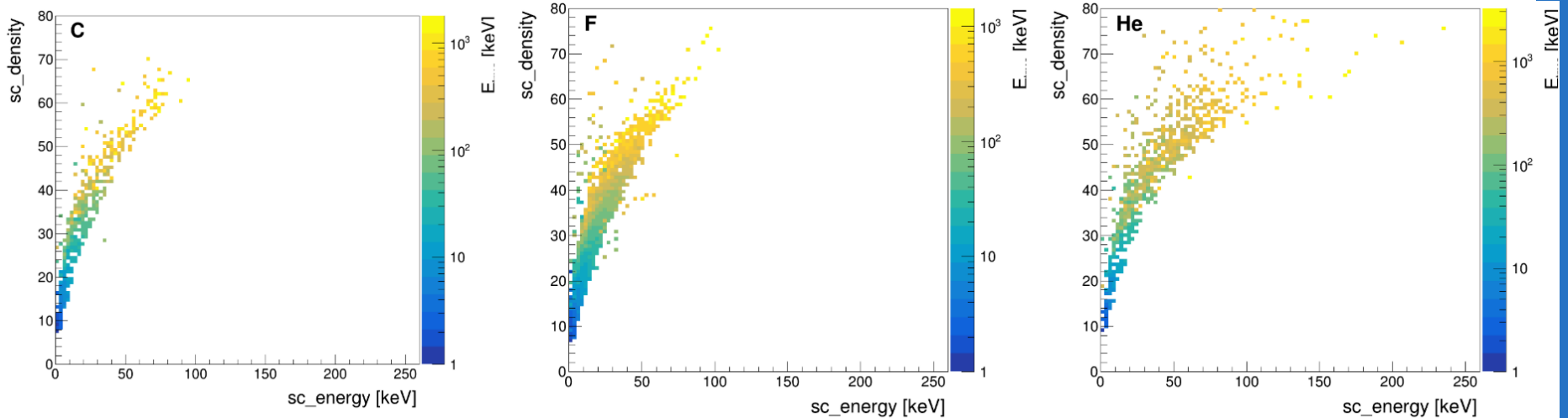
# Tracks “morphology” parameter



Combines geometric and density information:

- captures the overall "ionization volume"
- provides excellent discrimination between particle types.

# 3D track variables distribution



Distribution of sc\_density vs sc\_energy for NRs, **colored by true kinetic energy**

Low-energy NRs (~tens of keV, blue-green) cluster in low-density regime (< 20 ADU/pixel).

# Normalization to experimental exposure

## Trigger Rate Analysis (RUN 3)

- ▶ Background (no AmBe): 1.6 Hz
- ▶ With AmBe source: 1.8 Hz
- ▶ Net excess: 0.2 Hz (neutron-induced events)

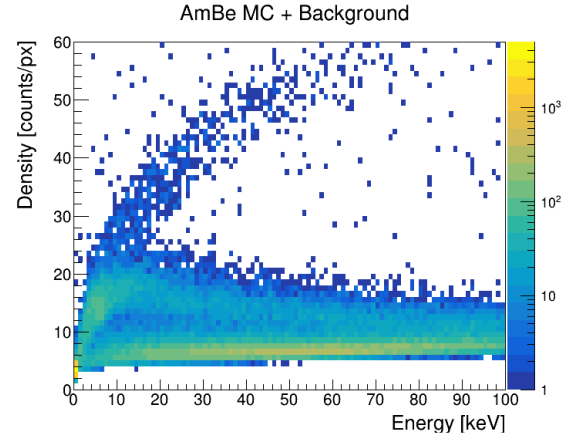
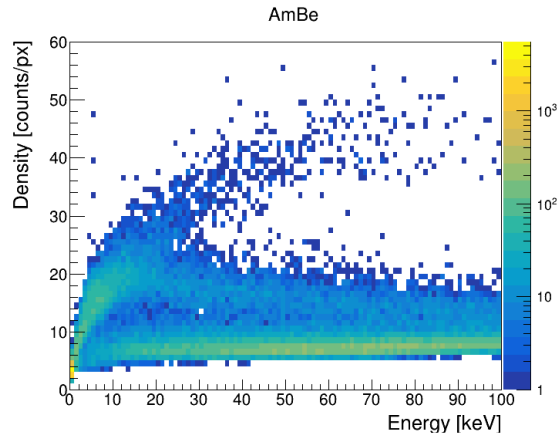
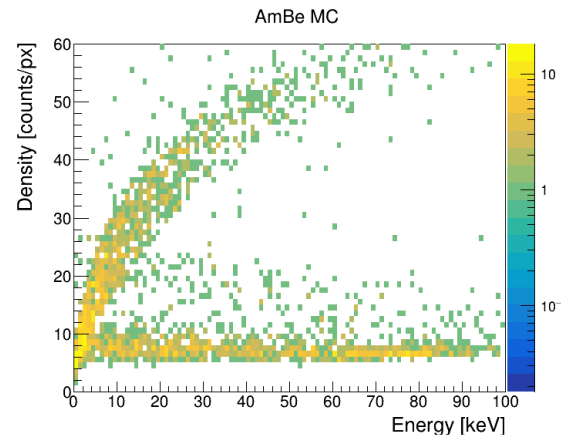
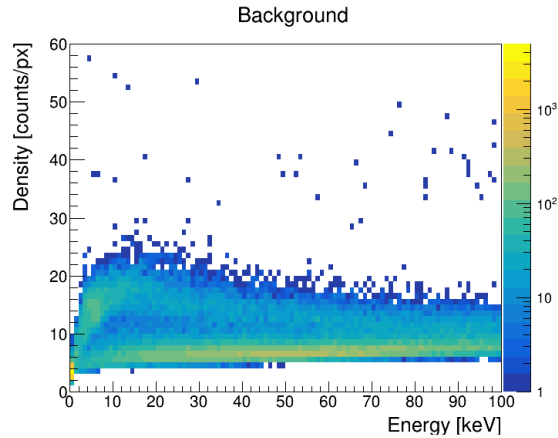
## Efficiency Calculation

- ▶ AmBe emission: 200 n/s
- ▶ Measured rate: 0.2 Hz ⇒ **Global efficiency: 0.1%**
- ▶ Geant4 interaction efficiency: 0.3% (3/1000 n interact in gas)
- ▶ DAQ efficiency: ~33% (trigger + reconstruction)

$$N_{\text{neutrons}} = 3600 \text{ s} \times 200 \text{ n/s} \times 0.33 = 237,600 \text{ neutrons.}$$

For 1h of data taking

# Data/MC qualitative comparison



# NR selection strategy

Define three selections with progressively increasing NR purity:

1. **Baseline:** Maximum inclusivity, minimal bias
2. **NR-enriched:** Enhanced NR fraction, moderate ER rejection
3. **High-purity NR:** Maximum purity, reduced efficiency

Validate MC performance under different signal conditions.

## Three Selection Strategies

### 1. Baseline Selection

Minimal requirements for genuine tracks:

- ▶  $sc\_length > 25$  pixels,  $sc\_tgausssigma > 0.5$  mm
- ▶ Radial distance  $< 800$  pixels (reject borders)

**Result:** Most inclusive sample, no particle-type bias.

### 2. NR-Enriched Selection

Additional cuts based on feature space:

- ▶  $\rho > 0.01$  (NR/ER separation in density)
- ▶  $sc\_length < 200$  pixels (removes long ER tracks)

**Performance:** Retains  $\sim 99\%$  NRs, rejects  $83\%$  ERs.

### 3. High-Purity NR Selection

$sc\_density > 25$  (almost exclusive NR region).

**Performance:** Rejects  $> 90\%$  ERs, retains  $\sim 50\%$  NRs.

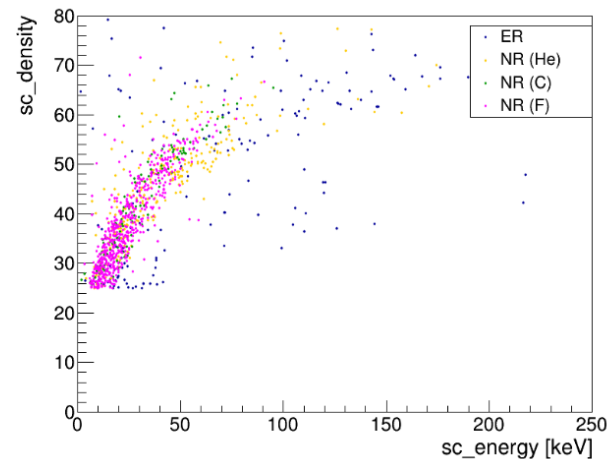
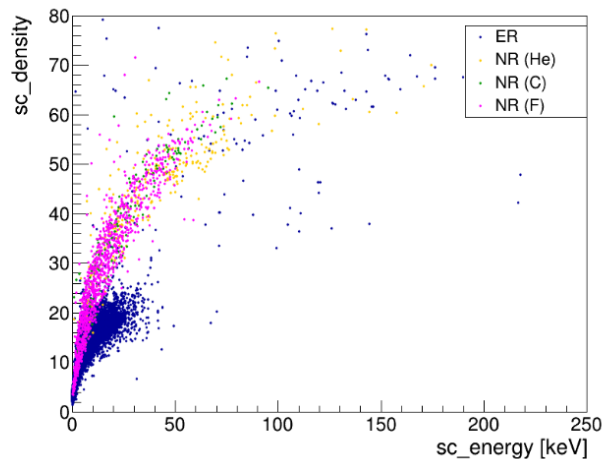
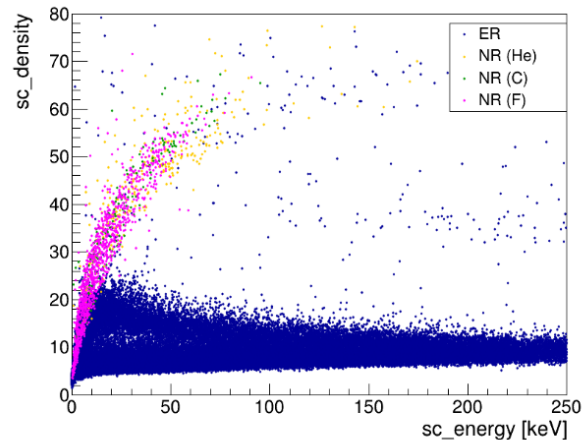
# NR selection strategy

Effects of the selection cuts on MC events

Baseline

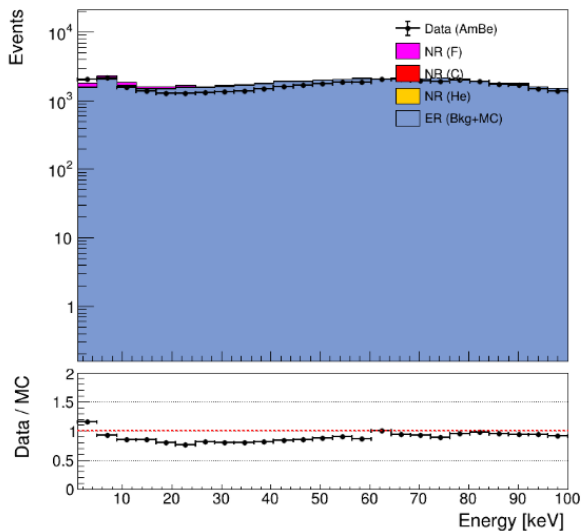
Enriched

High-Purity

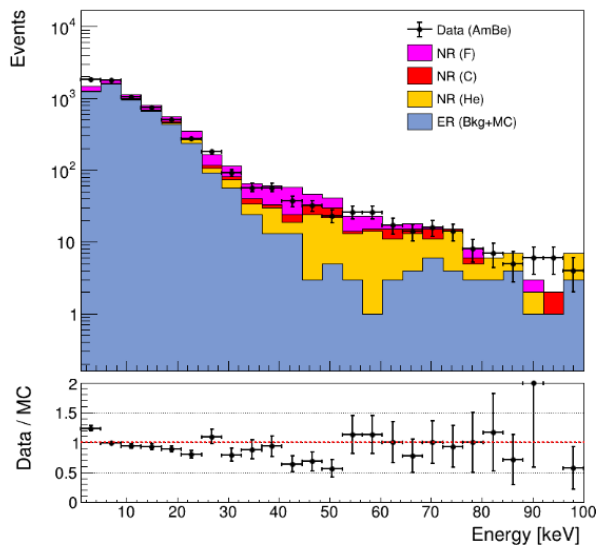


# Data/MC comparison: Energy

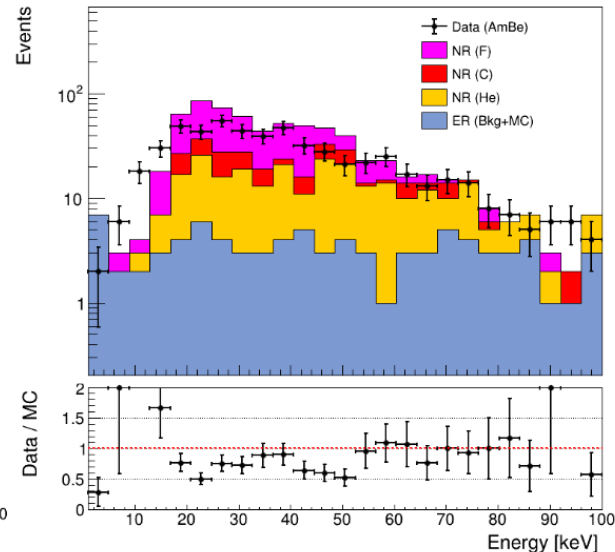
## Baseline



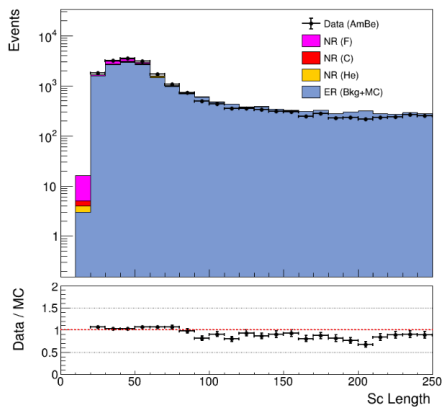
## Enriched



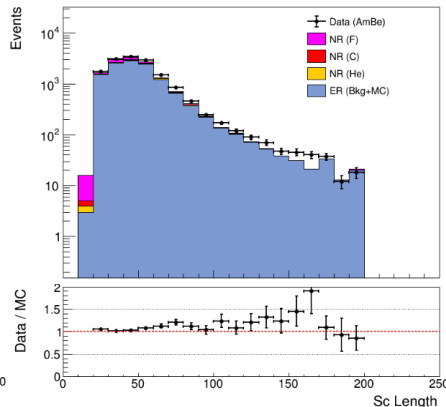
## High-Purity



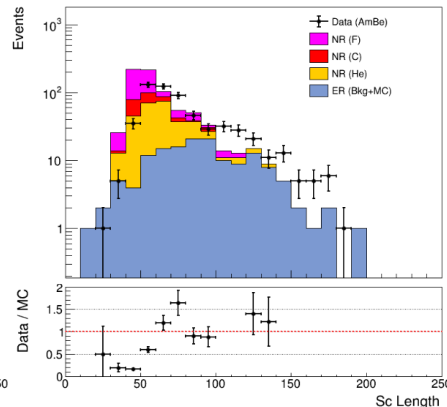
# Data/MC comparison: length and nhits



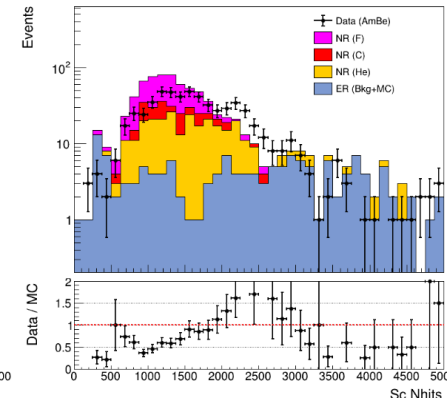
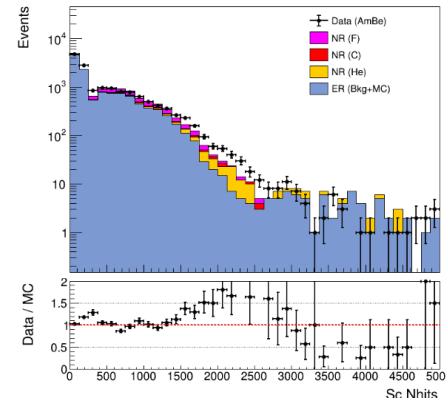
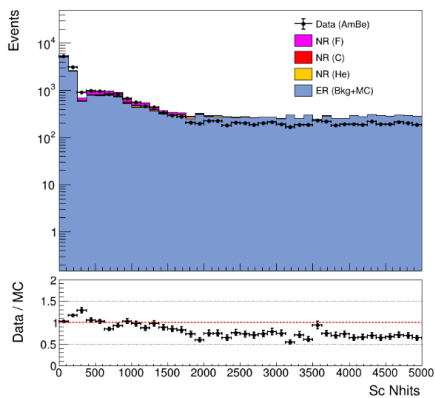
Baseline



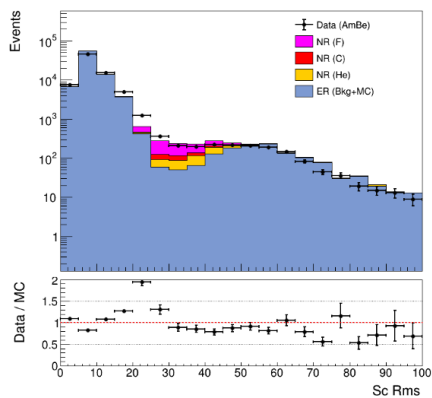
Enriched



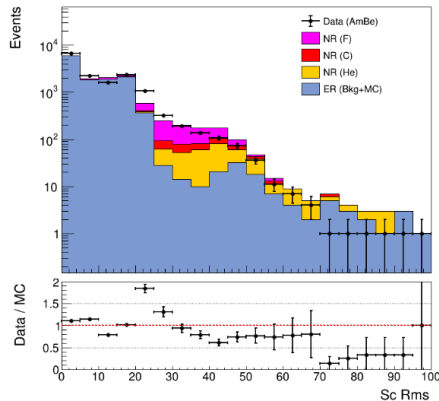
High-Purity



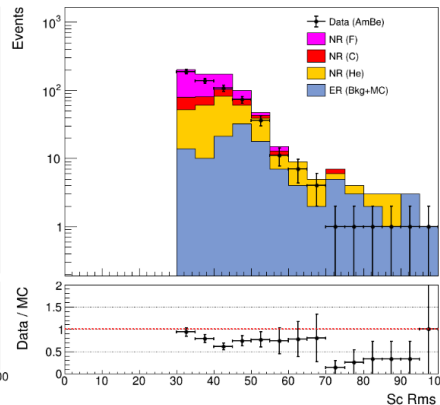
# Data/MC comparison: length and nhits



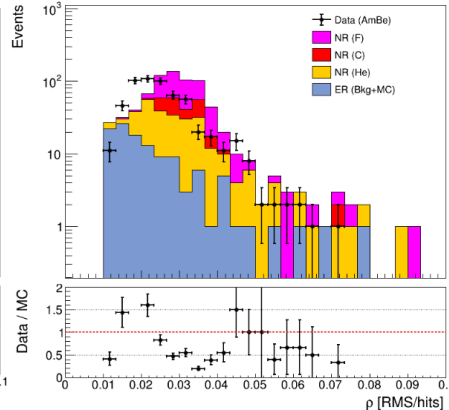
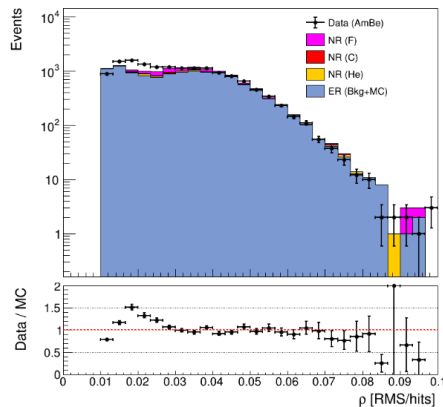
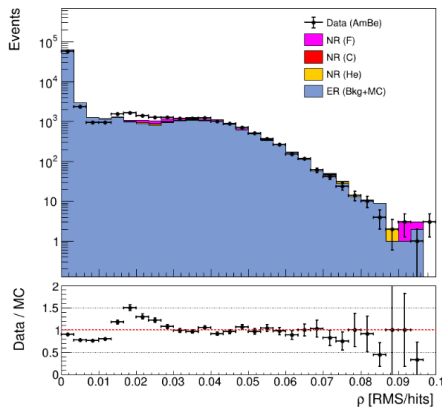
Baseline



Enriched



High-Purity



## Data/MC comparison: comments

### Remarkable Agreement Despite Complexity

- Diverse particles: ERs, He/C/F-NRs with distinct signatures
- Broad energy: Three orders of magnitude (keV to MeV)

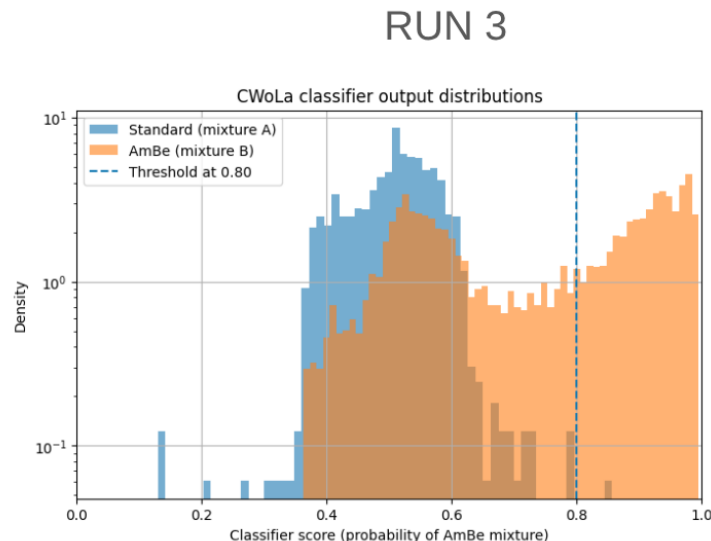
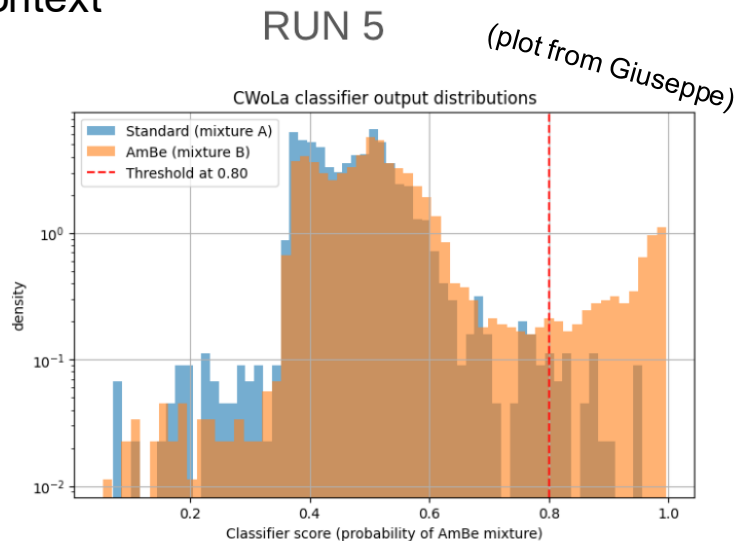
→ MC the simulation fairly reproduces main features

### Causes for the Data/MC disagreement:

- Environmental Parameter Variations: Data conditions varied (pressure, humidity, purity), MC uses fixed parameters.
- Introduce a smearing in the simulation to mimic fluctuations.
- The Digitization was optimized for X-rays spots, we are forcing
  - A better NR selection could improve Data/MC comparison

# Improving NR selection: ML approach

- The NR selection can be improved using ML techniques
- Current ER/NR separation can be used for training the Machine, providing separated samples
- An unsupervised approach was also tested (work from Giuseppe) that could be tested in this context



Number of standard images in test set: 3025  
 Number of AmBe images in test set: 5668

Number of standard images in test set: 1355  
 Number of AmBe images in test set: 3554