COSMOGENIC REJECTION STUDIES

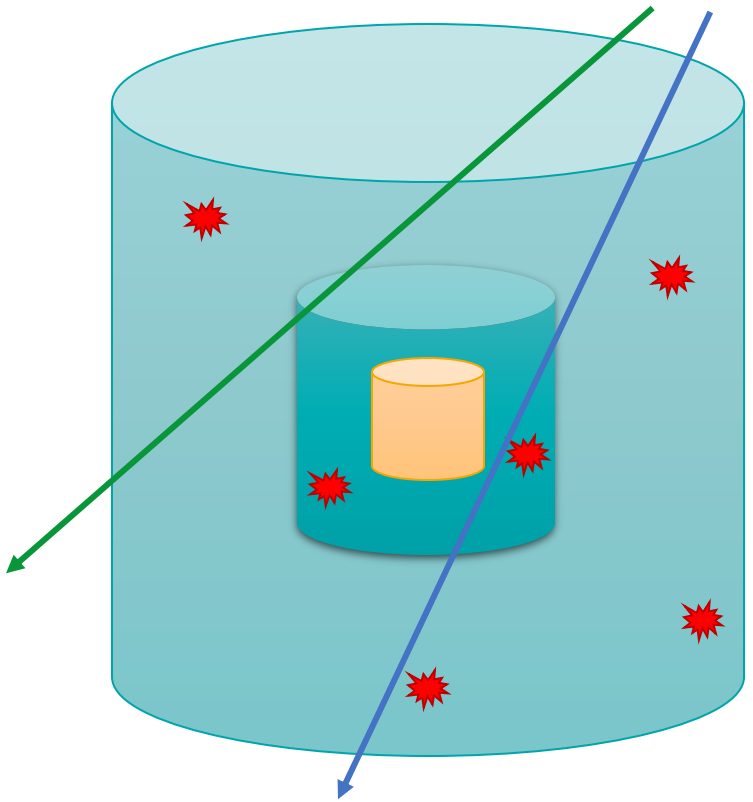
(*meeting 30.07.2020*)

Luigi Berducci, Danila Stukov, Giuseppe Salamanna

INFN Sezione Roma Tre, Università degli Studi Roma Tre

Brief recap

- **Goal**: "smart" cosmogenic rejection from LAr veto acquisitions (no active water muon veto, no Ge coincidence):
  - Target: muon-induced processes (*Ionization, Compton, Ar41 de-excitation*)
  - Other Background: Beta-decays of Ar39
- **Muon Events:**
  - High-Detection Muons: "easy" task => Cut on Number of PE Detected
  - Low-Detection Muons (*or Low-Energy Muons*):
    - Muons that **marginally** cross the **sensitive region**
    - Resulting in a small amount of energy far from the fiber shrouds
    - Are they dangerous as well?
- Why Ar39 is a problem?
  - High-activity (*1.41Bq/l*) and integration over time (*10us*)
  - would result in Pileup of Ar39 decays



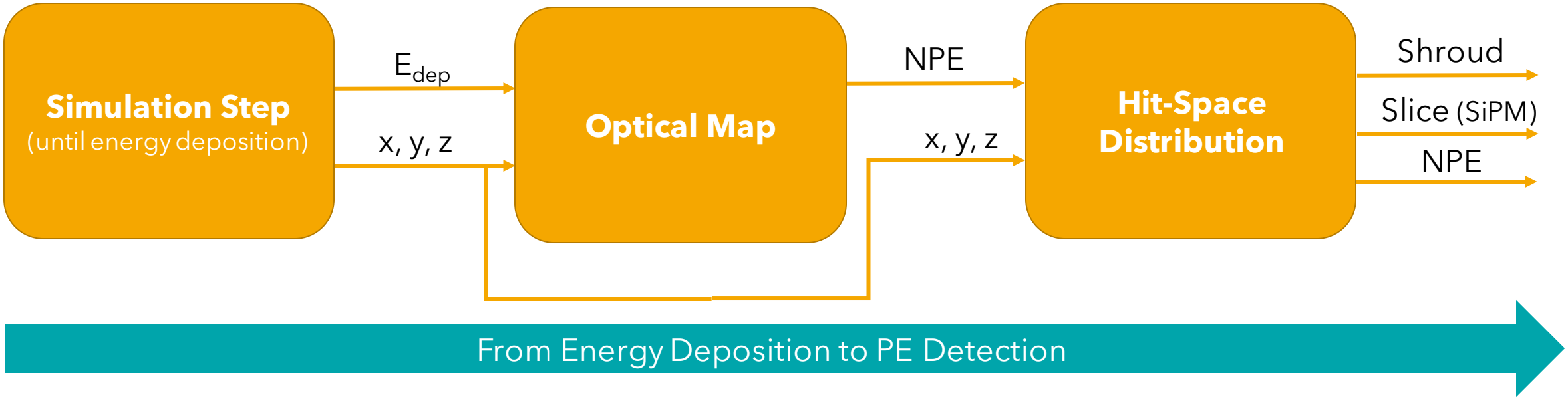
- Classification only relying on LAr:
  - **Descriptors**: amount of **energy** detected and **spatial distribution** of detections.
  - **Prompt Classification**: detections within O(10us) from trigger
  - **Delayed Classification**: correlate later detections with delayed processes in LAr (*Ar41 de-excitation*)
- **Previous presentation**: *https://indico.legend-exp.org/event/341/contributions/1800/attachments/1106/1605/ML4NP_SimGroup_20200611.pdf*

Updates since last talk (11.06.2020)

- Emulation of **SiPM readouts**:
  - Hit Space Distribution: Toy Spatial Map
- Extended **Simulation** in LGND200:
  - **3M Muons** (CJ)**:**
    - LNGS conditions: input from MUSUN + 3m rock surrounding the volume (showering)
  - **10M Ar39** (CJ):
    - Primary particles are generated randomly in LAr volume
  - **2M Neutrons** (Danila)**:**
    - Starting Energy from Muons' Simulation
    - Position of primaries sampled in a cylinder wt H=4m, R=2m (LAr volume)
    - Theta in [0, 180), Phi in [0, 360)
- **Prompt Classification** with Machine Learning

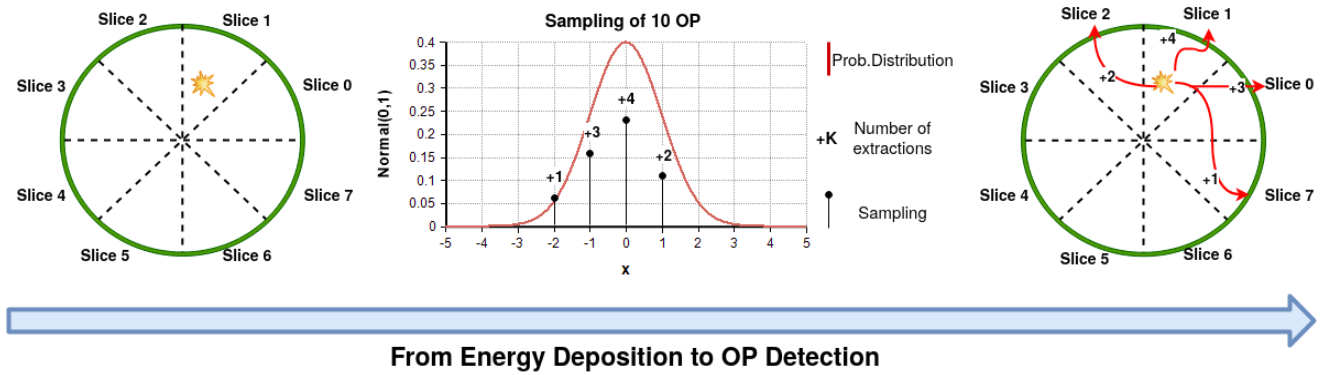
Simulation Workflow

- To simulate detections in LAr veto, without optical simulation.



Hit Space Distribution

- **Last meeting**: we use a constant Gaussian distribution to sample the spread of detection over the shroud.

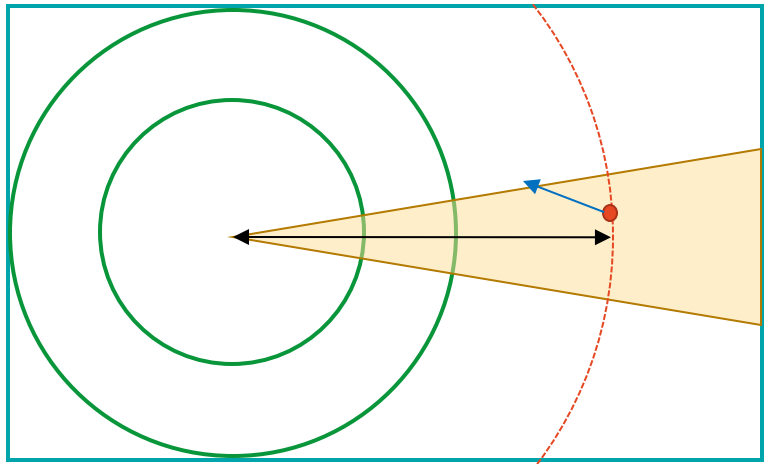


- **Problems**:
  - The **radius** of photon production point strongly affects the detection spread.
  - LGND200 has **2 shrouds** with 12 and 20 SiPM modules.
- **Idea**: use geometry!
  - Photon trajectory = straight line, Shrouds = circles
  - From the intersection of a line with circle, we can derive the hit-angle distribution
- No optical simulation (*no WLS, Radon Shroud reflections*), just a **Toy Spatial Map**

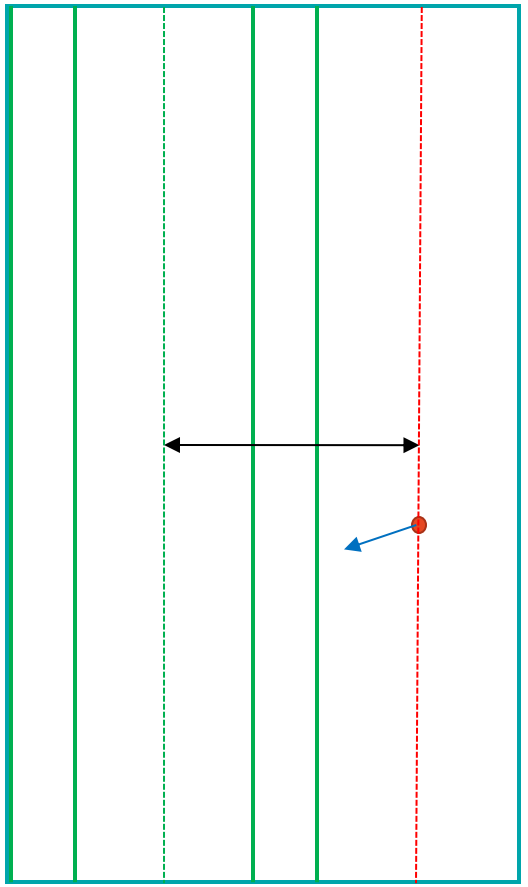
Toy Spatial Map (1)

- For each **r** in [0, 700]:
  - Sample x, y, z s.t. radius is **r** in XY plane (+- 0.22 rad)
  - Sample OP trajectory
  - See where OP trajectory intersect Ge cylinder
  - If intersection in +-425: enable 14 Ge Crystals, otherwise no Ge
  - Compute all possible intersections with inner and outer shrouds
  - Sample intersection according to Pr
  - Compute angle of hit

XY View

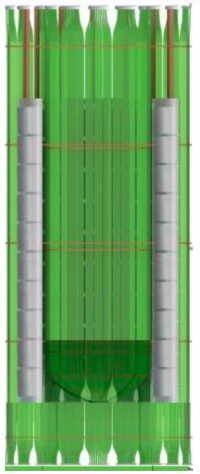


XZ View

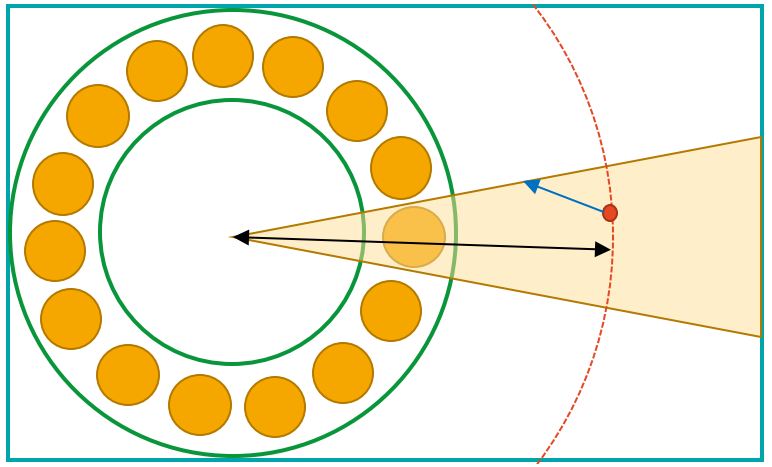


Toy Spatial Map (2)

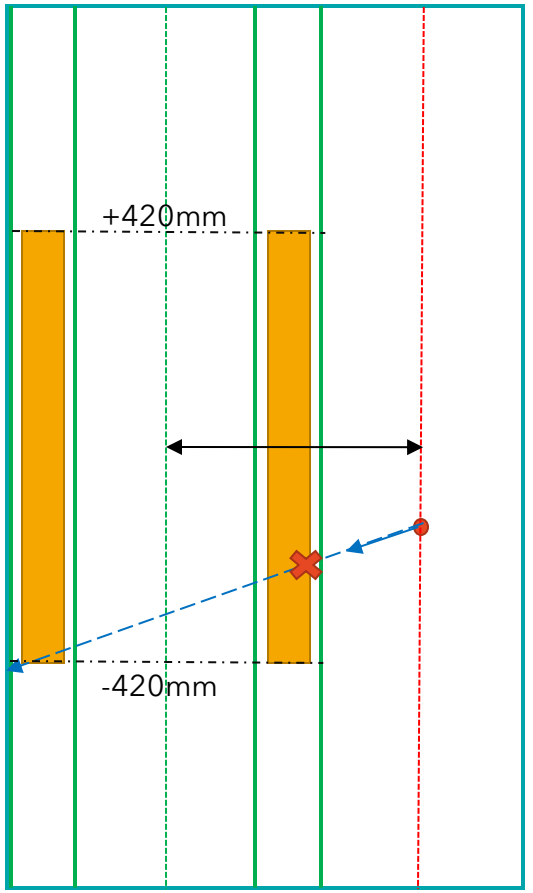
- For each **r** in [0, 700]:
  - Sample x, y, z s.t. radius is **r** in XY plane (+- 0.22 rad)
  - Sample OP trajectory
  - See where OP trajectory intersect Ge cylinders (✖)
  - If intersection in +-420: enable 14 Ge Crystals, otherwise no Ge
  - Compute all possible intersections with inner and outer shrouds
  - Sample intersection according to Pr
  - Compute angle of hit



XY View



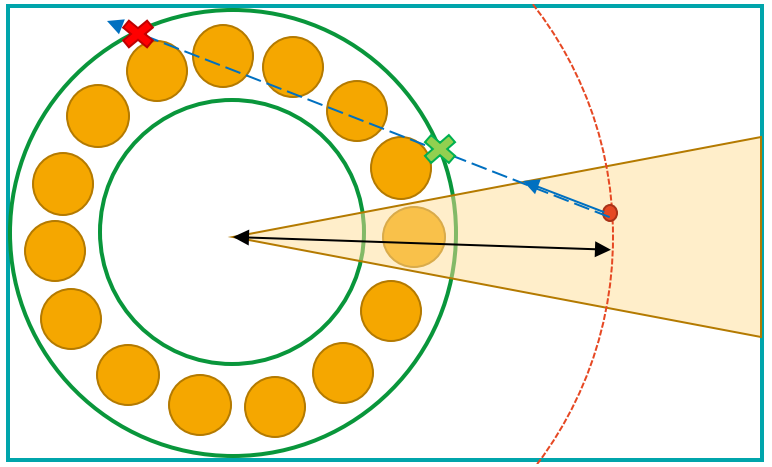
XZ View



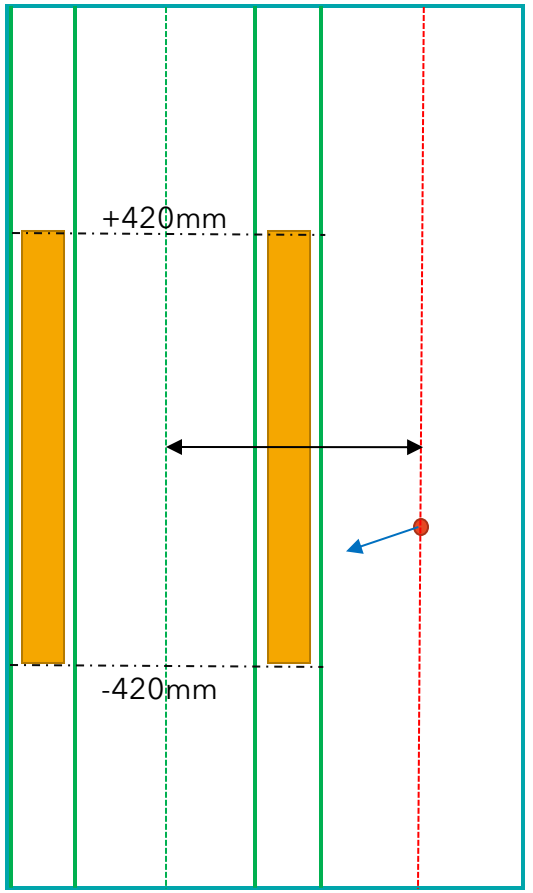
Toy Spatial Map (3)

- For each **r** in [0, 700]:
  - Sample x, y, z s.t. radius is **r** in XY plane (+- 0.22 rad)
  - Sample OP trajectory
  - See where OP trajectory intersect Ge cylinders (✖)
  - If intersection in +-420: enable 14 Ge Crystals, otherwise no Ge
  - Compute all possible intersections with inner and outer shrouds
  - Sample intersection according to Pr
  - Compute angle of hit

XY View



XZ View



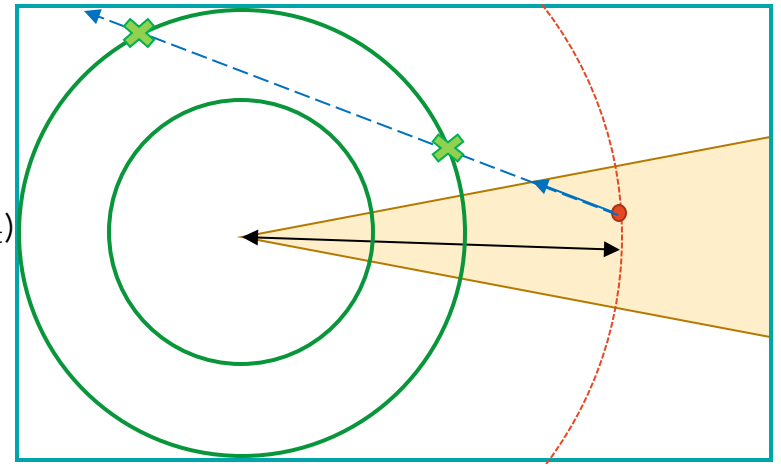
Toy Spatial Map (4)

P(geom.cov.) := shroud spatial coverage ~ 0.54

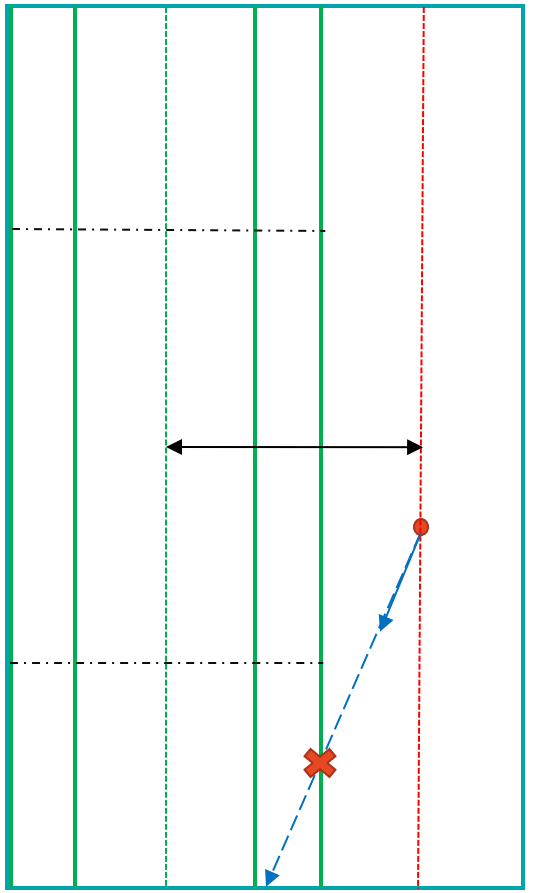
- For each **r** in [0, 700]:
  - Sample x, y, z s.t. radius is **r** in XY plane (+- 0.22 rad)
  - Sample OP trajectory
  - See where OP trajectory intersect Ge cylinders (✖)
  - If intersection in +-420: enable 14 Ge Crystals, otherwise no Ge
  - Compute <u>all possible intersections</u> with inner and outer shrouds
  - Sample intersection according to Pr
    - First hit: P(geom. cov.) * P(λ_{att})
    - Second hit: (1-P(geom.cov)) * P(geom.cov) * P(λ_{att})
    - …
  - Compute angle of hit

Another example
(Ge Disabled)

XY View



XZ View

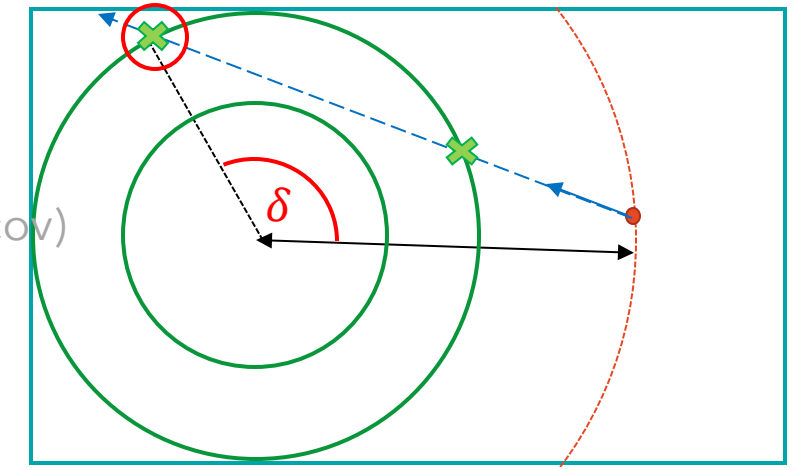


Toy Spatial Map (5)

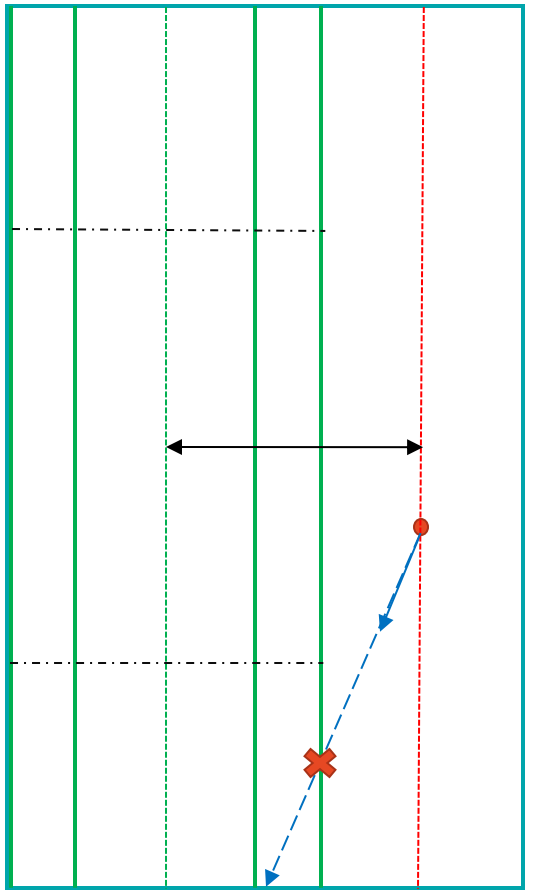
- For each **r** in [0, 700]:
  - Sample x, y, z s.t. radius is **r** in XY plane (+- 0.22 rad)
  - Sample OP trajectory
  - See where OP trajectory intersect Ge cylinders (✖)
  - If intersection in +-420: enable 14 Ge Crystals, otherwise no Ge
  - Compute all possible intersections with inner and outer shrouds
  - Sample intersection according to Pr
    - First hit: P(geom. cov.)
    - Second hit: (1-P(geom.cov))* P(geom.cov)
    - …
  - Compute angle of hit

Another example
(Ge Disabled)

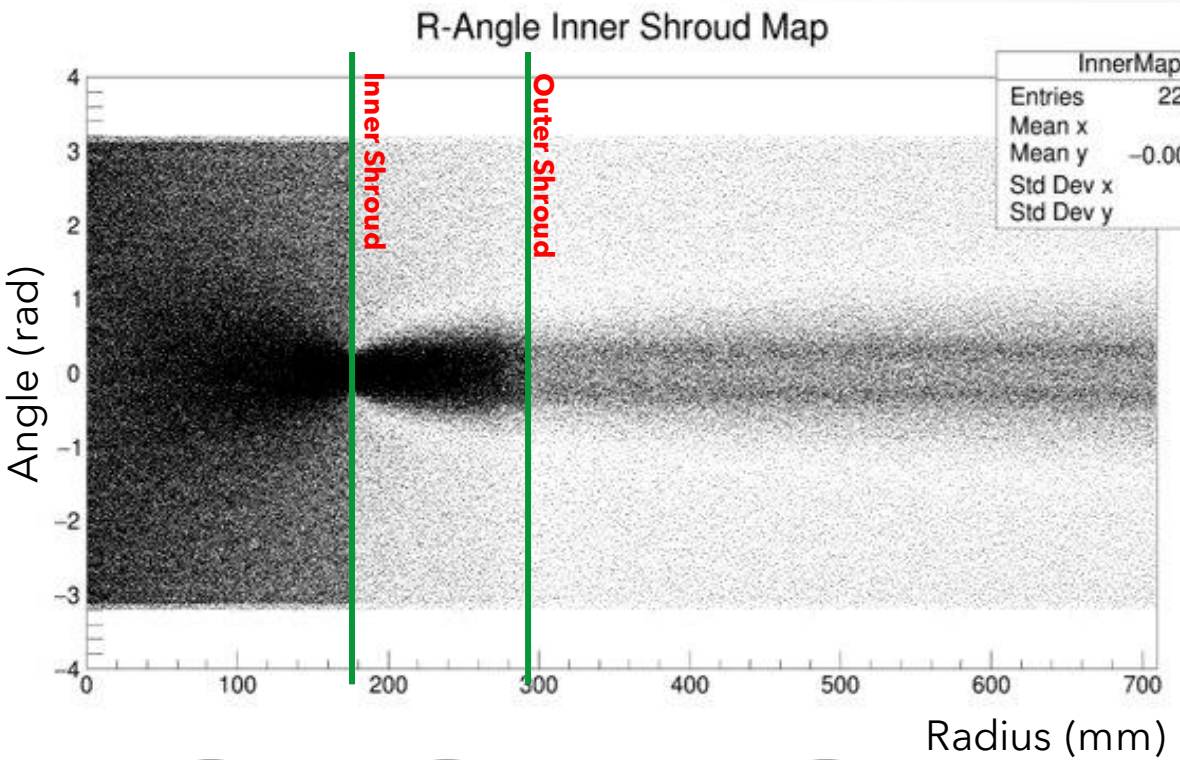
XY View

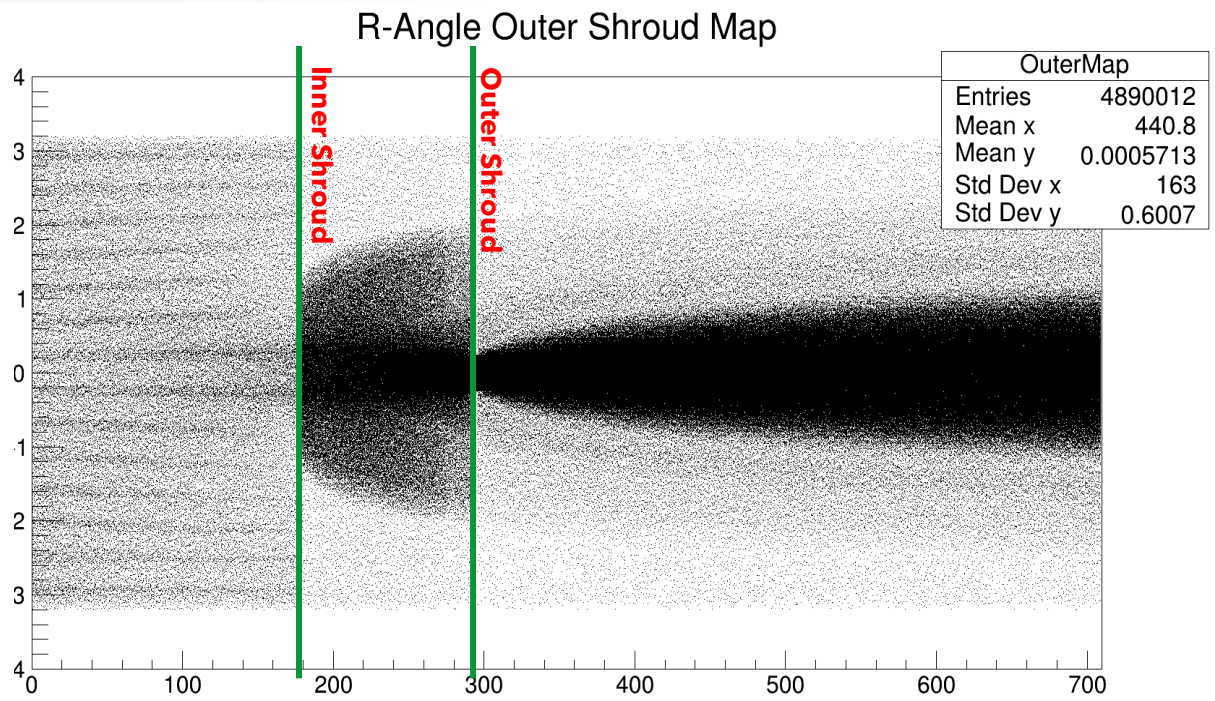


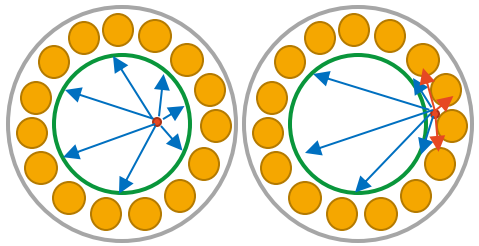
XZ View

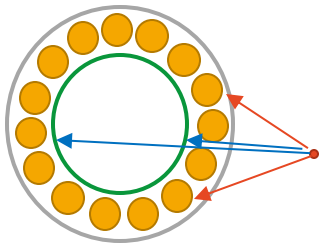


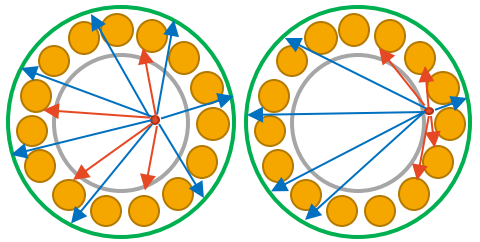
Toy Spatial Map - Result

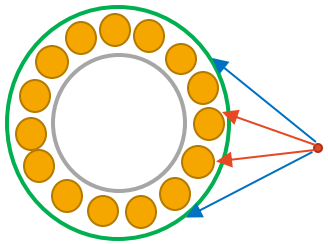






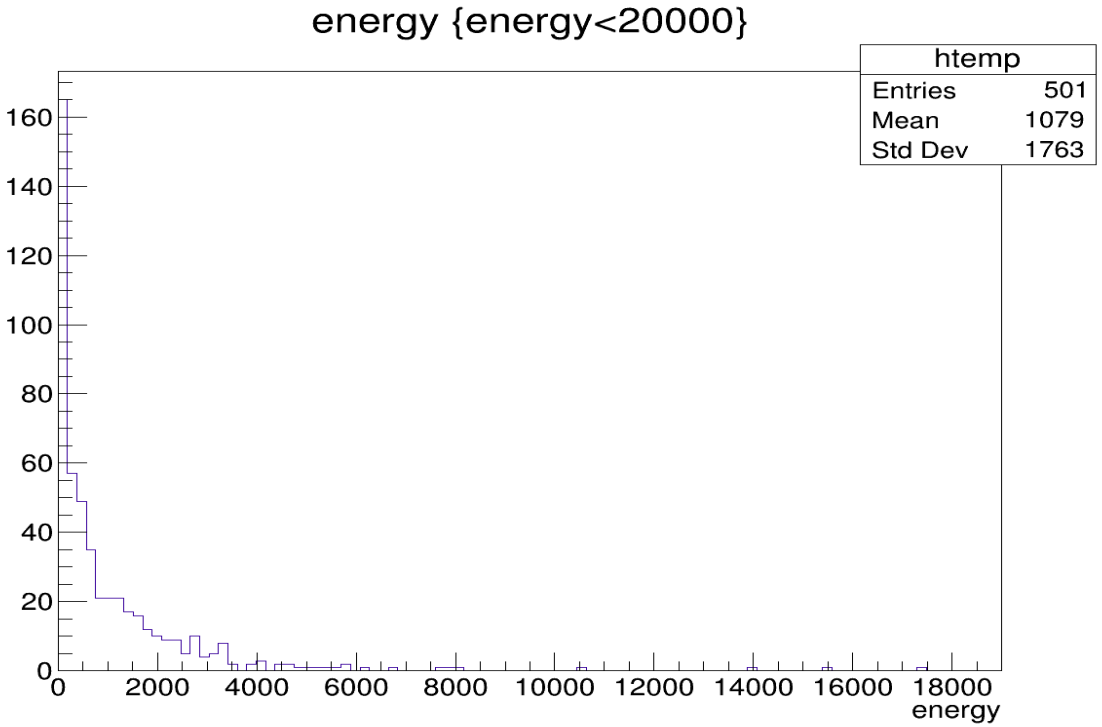




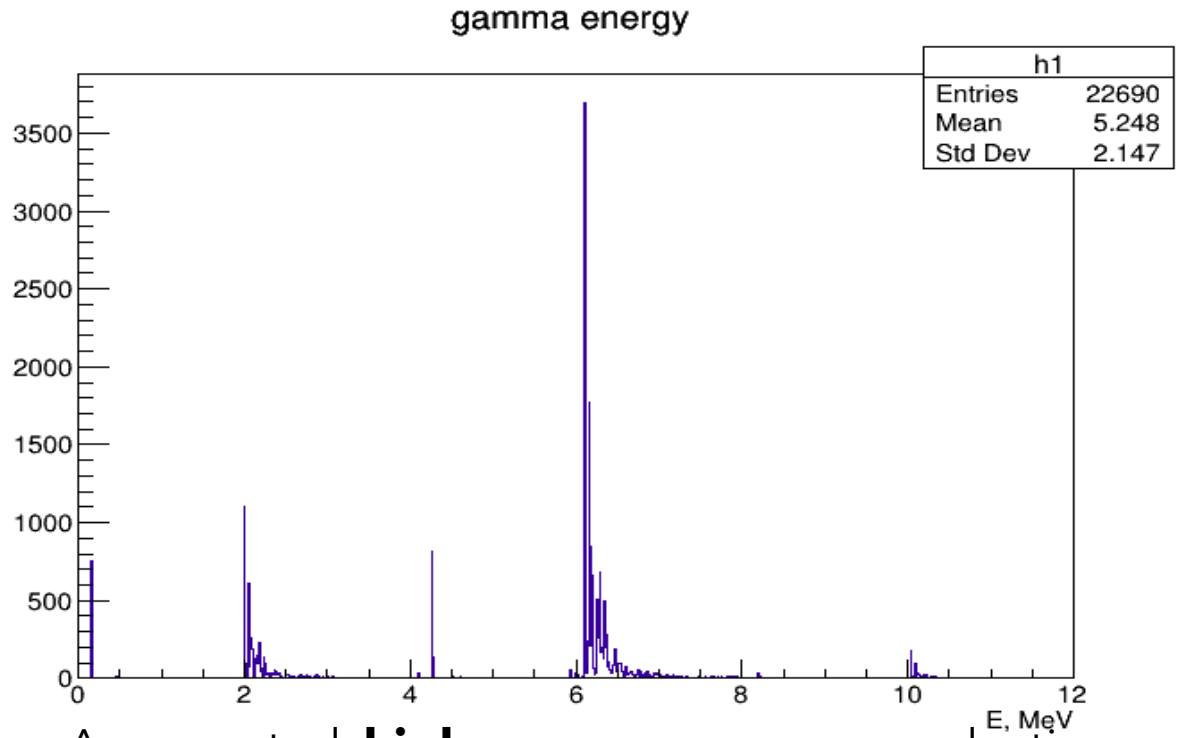


Neutron Simulation

We extract the initial energies of neutrons from Muons Simulation to create an **Energy Spectrum**
We have **~20K** events with **Ar41 de-excitations**, over **2M neutron** simulations (**1%**)



Starting Neutron Energy Spectrum (KeV)



As expected, **high-energy gammas** production by Ar41 de-excitation

Post-processing is still on going...

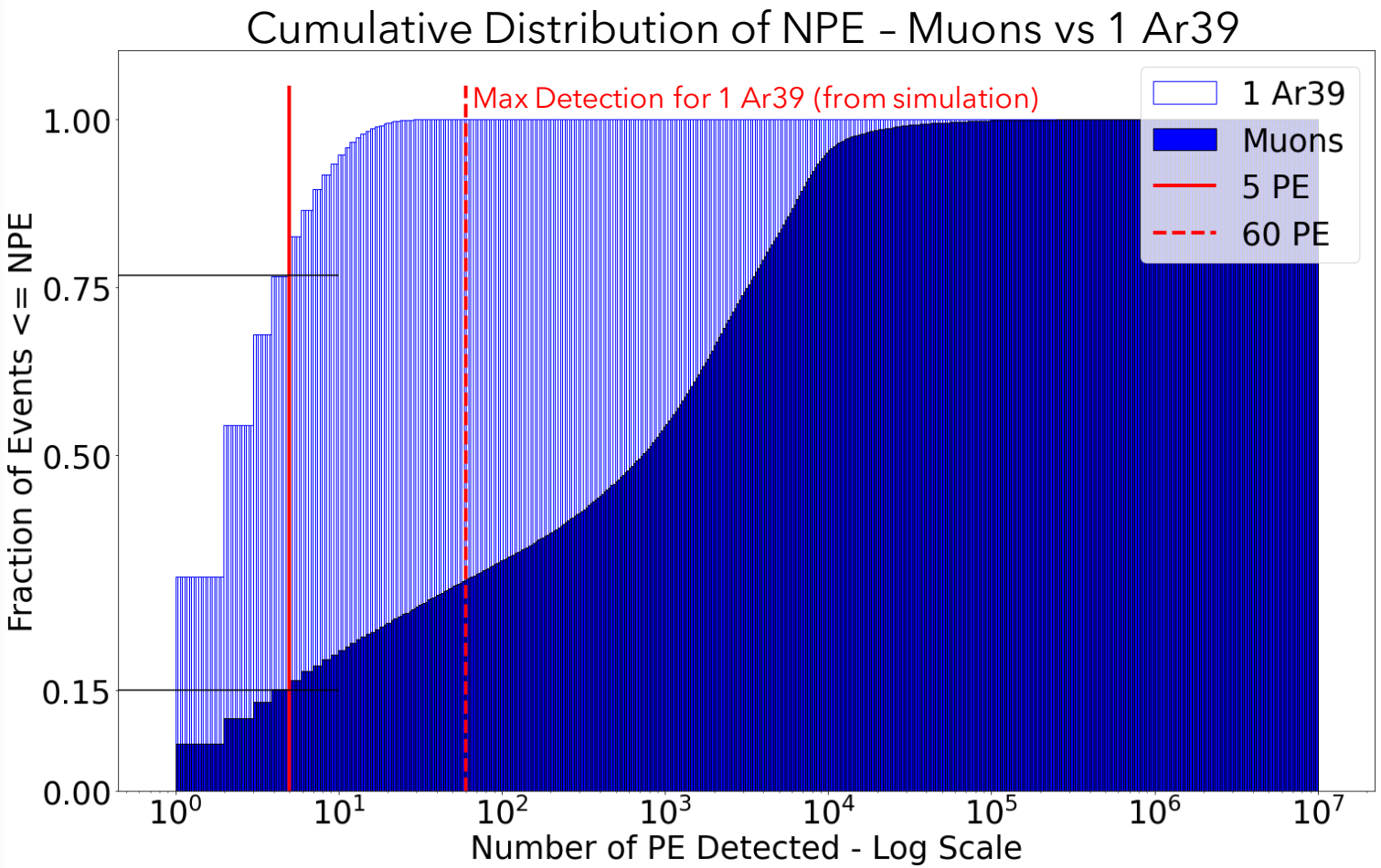
Machine Learning – Prompt Classification (1)

- Simulation Parameters: Light Yield=40 OP/KeV, Q.E.=0.40

- First Problem: Single Ar39 decays
  - E[Ar39 Rate|1.41 Bq/l, V=2.6m3] = **3666 Hz**
  - From Simulation: ~**65%** Ar39 events wt NPE>=1
  - E[Ar39 Det. Rate] = 3666 * .65 = **2353 Hz**

  - Pr[>=1 Ar39 decays] = 1-Pois(k=0, rate) = **.023255**
  - Pr[1 Ar39] = Pois(k=1, rate) = **.022983**
  - Single Ar39 decays is dominant => **98.83 %**

- Preliminary Cut: NPE>=5
  - From Simulation: ~75% of 1 Ar39 has NPE<5
  - From Simulation: ~15% of Muons has NPE<5



Including events NPE<5 PE, tough classification => Learning get worse

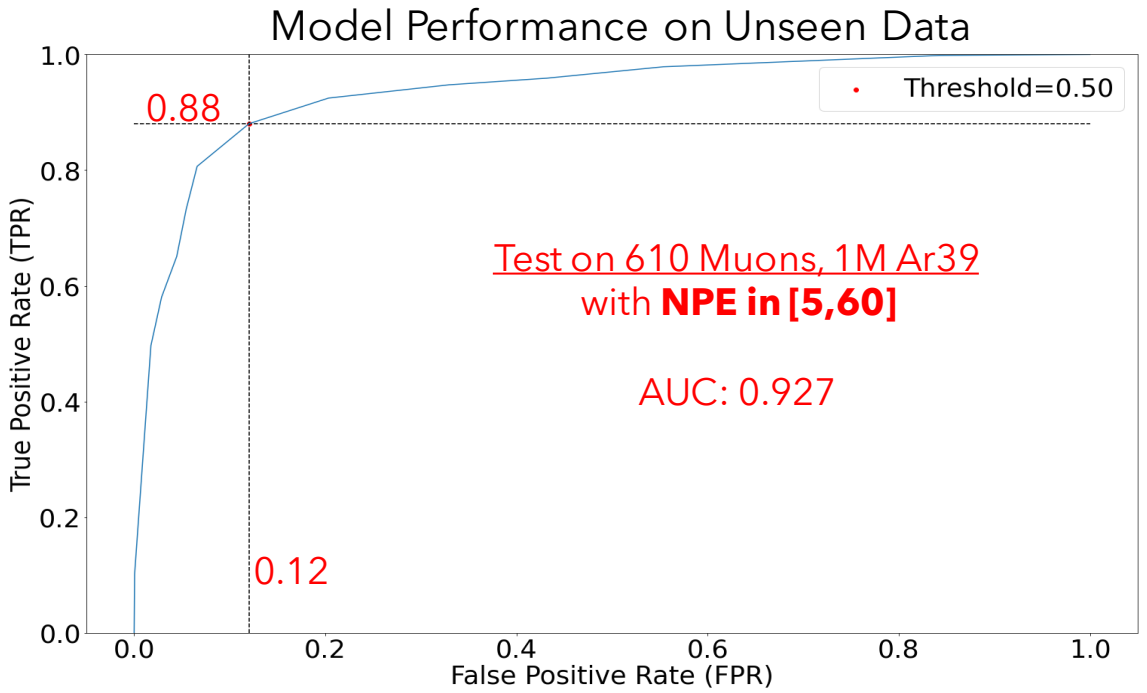
Since Ar39 has a dominant rate compared to Muons => Classify NPE<5 as Ar39

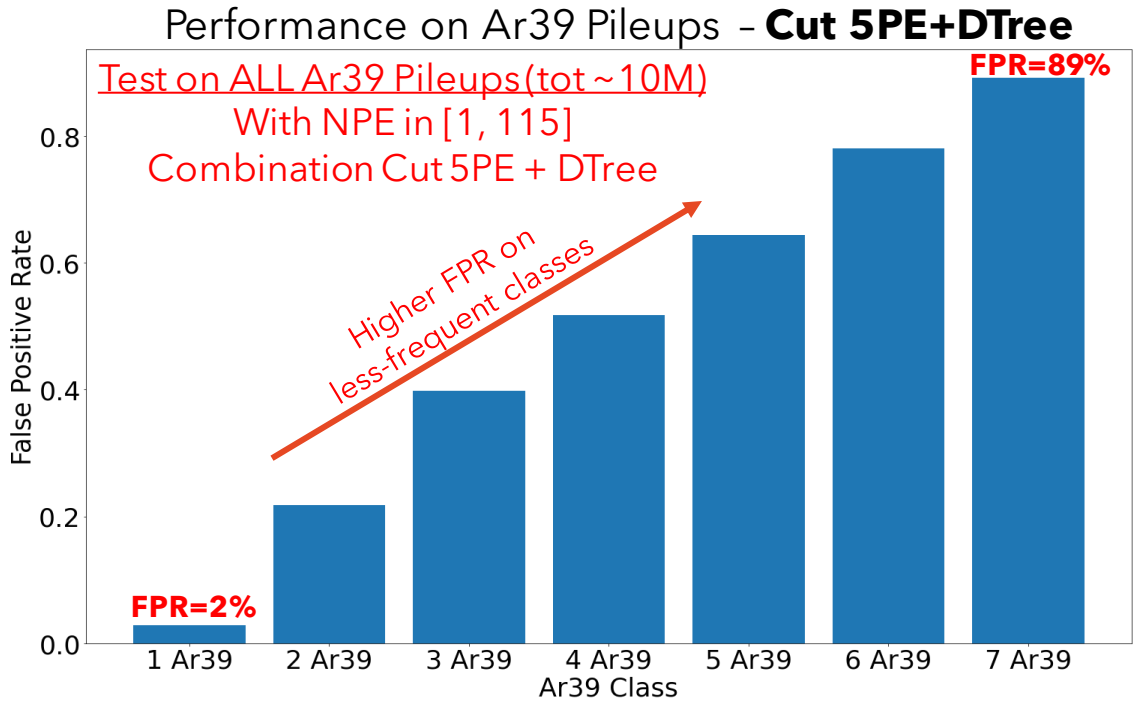
Prompt Classification (DTree)

- First Problem: Single Ar39 decays
- From Simulation: Ar39 have NPE between 0 to 60 PE
- Preliminary Cut: NPE>=5

- Specialized Model "Mu vs 1Ar39":
  - Fast inference: **Decision Tree**
  - Simple Features for each shroud: **NPE, Number of Active Slices**
  - Explainable Logic
- Training Phase:
  - Events wt NPE in [5, 60]
  - Training Data: 4500 muons, 4500 Single Ar39

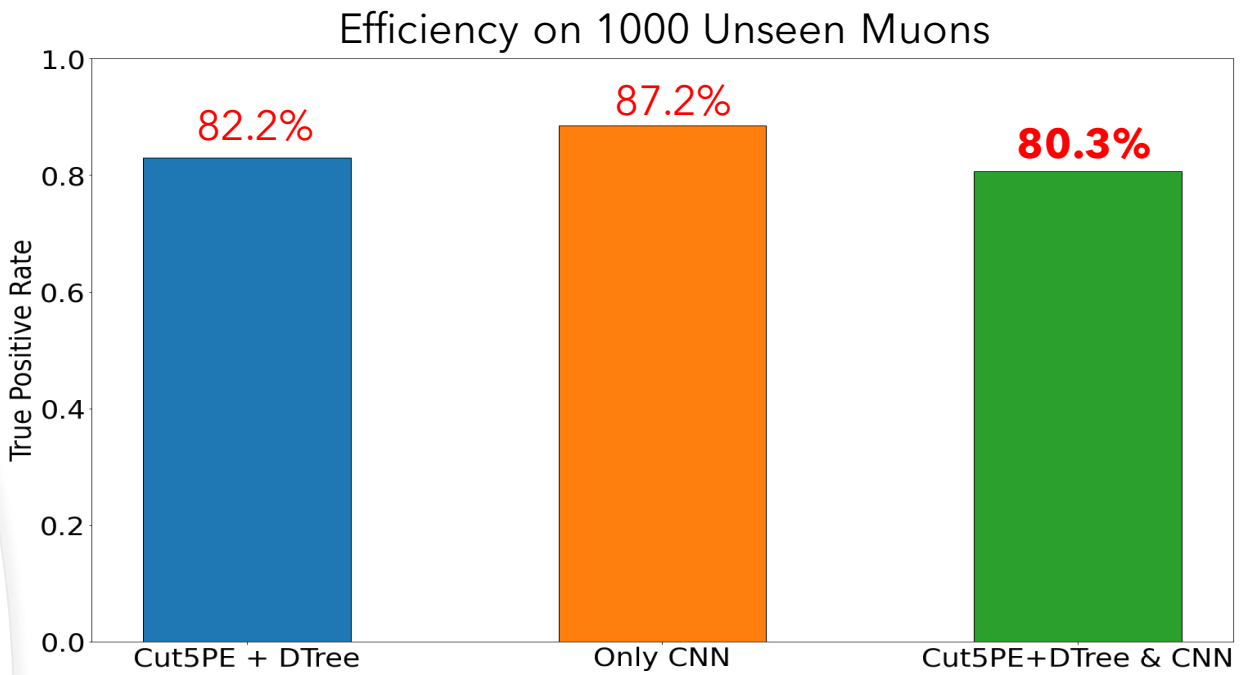
...how to deal with Ar39 pileups?

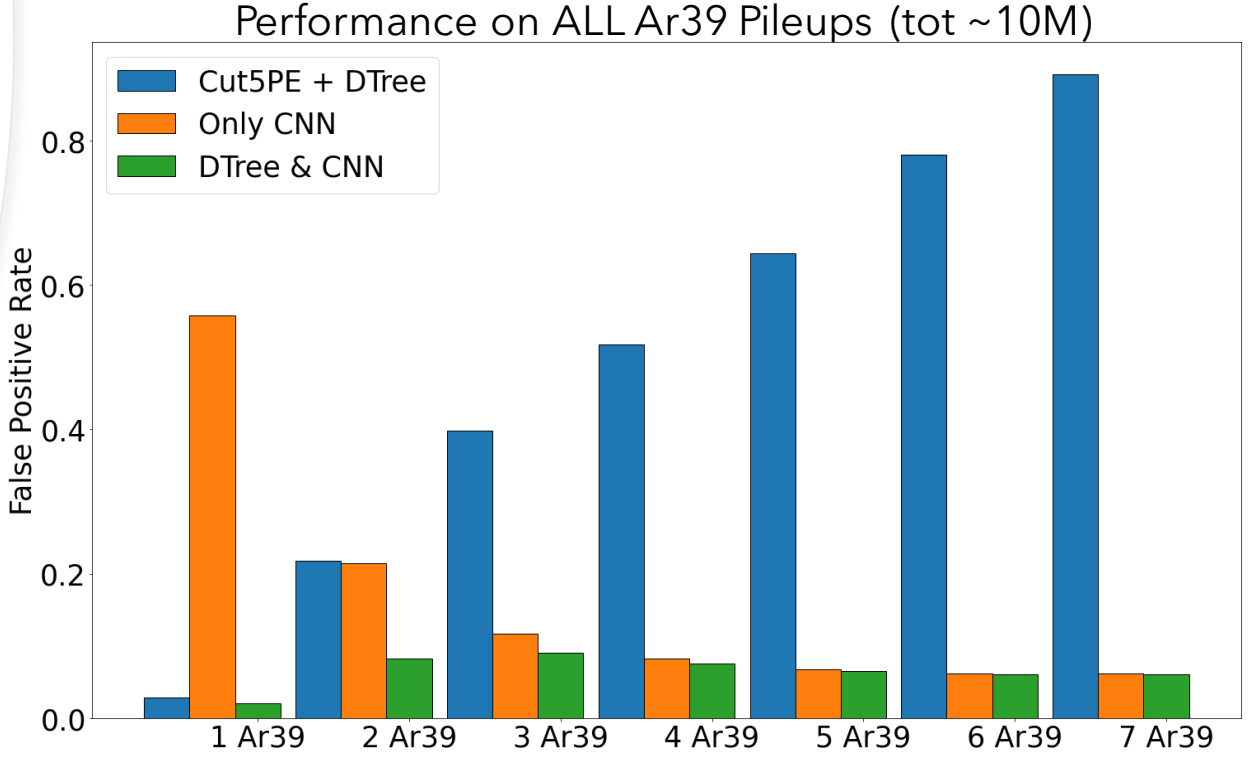




Prompt Classification (CNN)

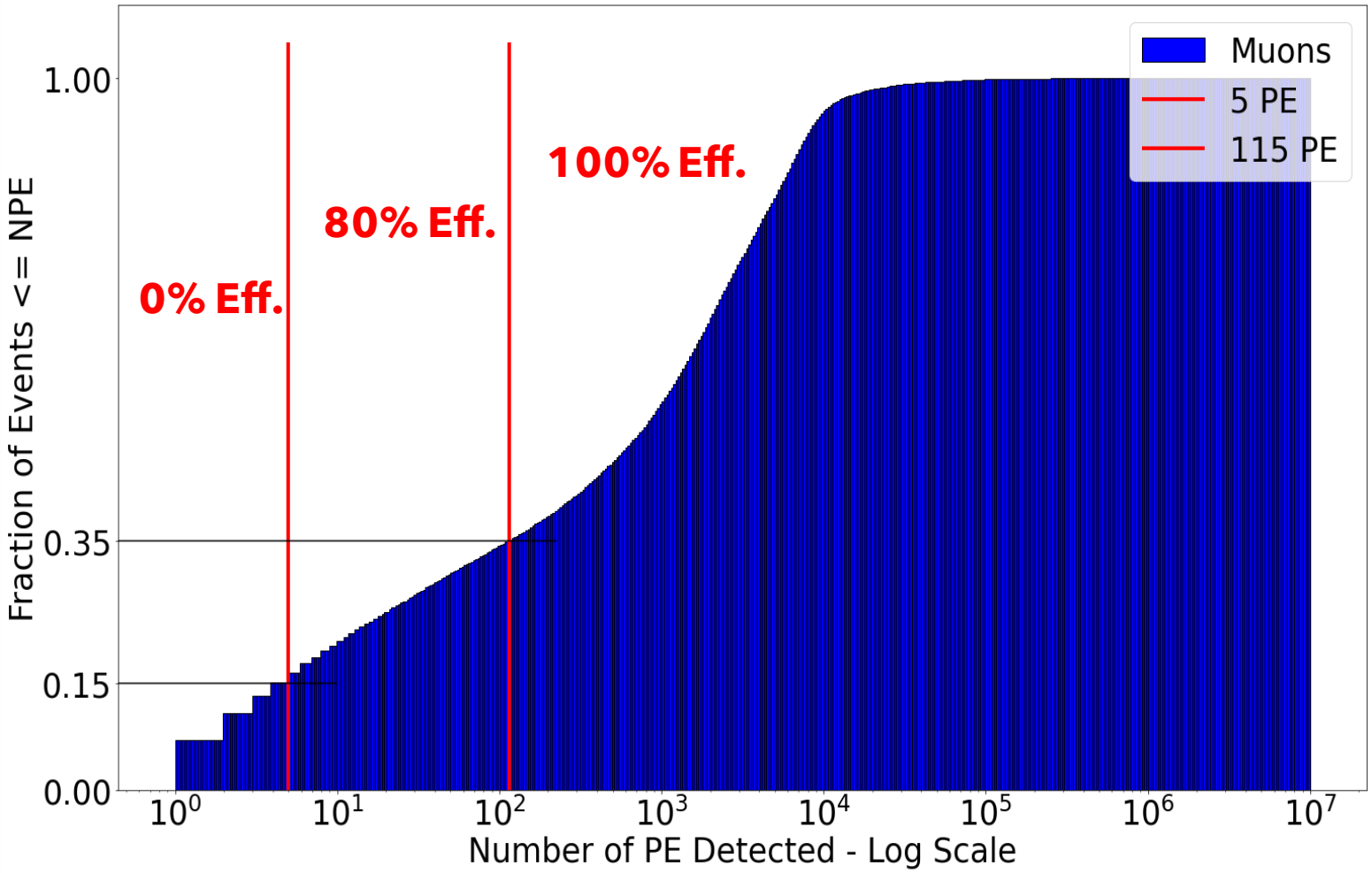
- How to deal with Ar39 pileups? **2-Steps Approach**
- General "Mu vs Ar39 Pileups":
  - CNN Model
  - Train a model on events wt 5<=NPE<=115
- Combination of DTree and CNN:
  - Still matter of investigation…
  - Example: Dtree & CNN
    - Test on Muon Sample: only **2%** Loss in Efficiency
    - Significant reduction of FPR





Prompt Classification - Recap

- In this simulated context...
  (OP Yield, Q.E., Optical Map, Toy Spatial Map)
- Classification Strategy:
  - Below 5 PE: Ar39
  - Between 5 and 115 PE: DTree & CNN
  - Above 115 PE: Muons
- Its efficiency:
  - Below 5 PE: 0%
  - Between 5 and 115 PE: ~80% *on test set*
  - Above 115 PE: 100%
- Overall expected efficiency:
  - 0.15*0 + 0.20*0.80 + 0.65*1 ~ 81%
- ...a great challenge is on purity!



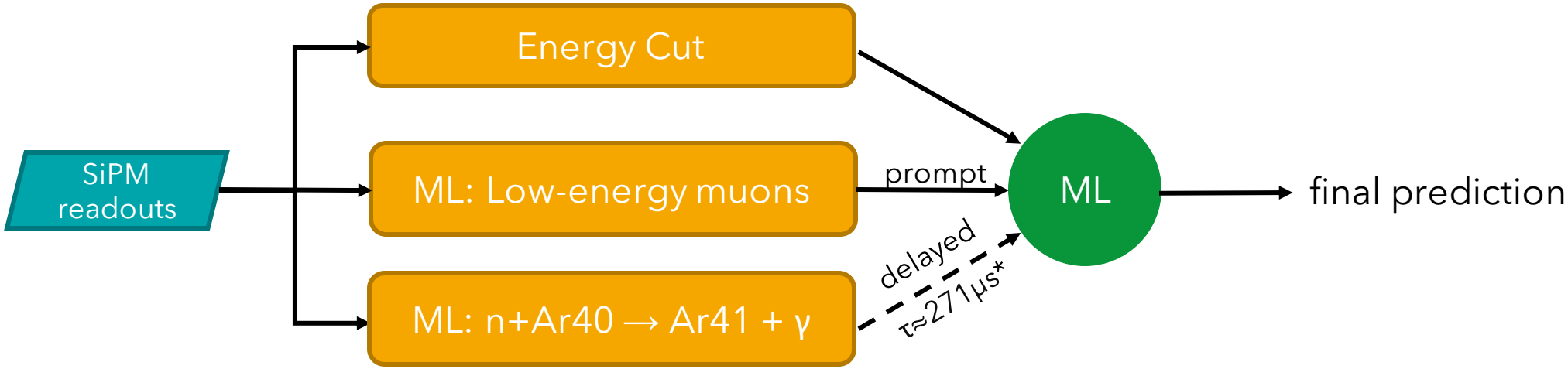
Next Steps

- Detection **Delayed Process** and correlation Prompt/Delayed Classification
- Investigation on **sensitivity** of trained models w.r.t. realistic **experimental conditions**
- Consideration on **inference time** to move to an online realm (*trigger implementation*)

BACKUP SLIDES

Proposed Strategy

- Multidimensional problem: depositions in **time and space**
- Investigate on the use of **ML** to synthesize an **automated classifier**
- Leveraging the **spatial distribution** of depositions to identify **specific topologies**
  - Low-energy muons (partially cross the sensitive region)
  - Undetected muons but delayed Ar41 de-excitation: n+Ar40 → Ar41 + γ

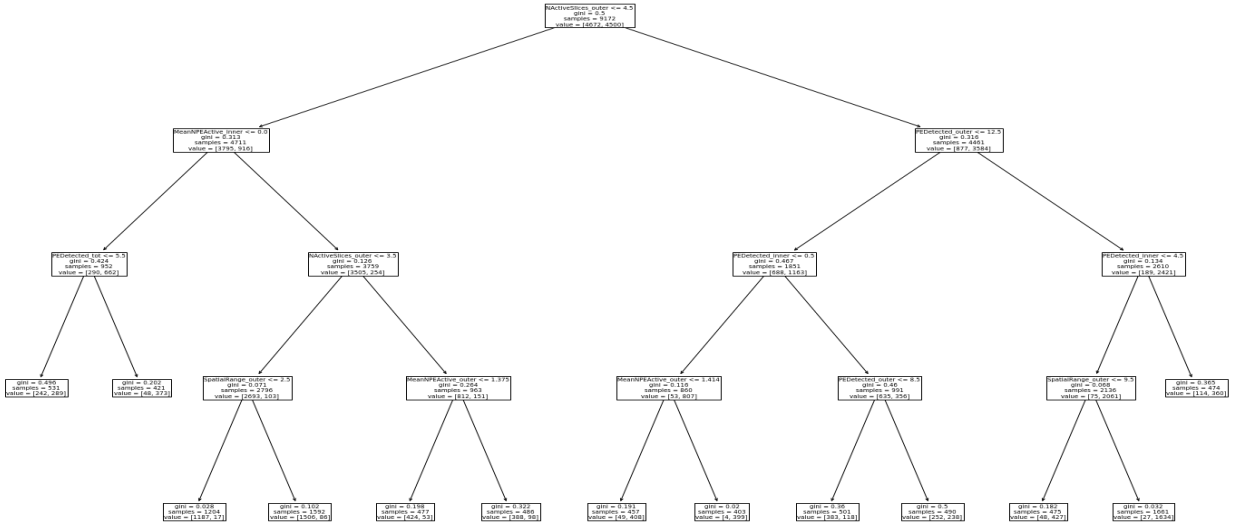


*Wiesinger, et al. "Virtual depth by active background suppression: revisiting the cosmic muon induced background of Gerda Phase II."

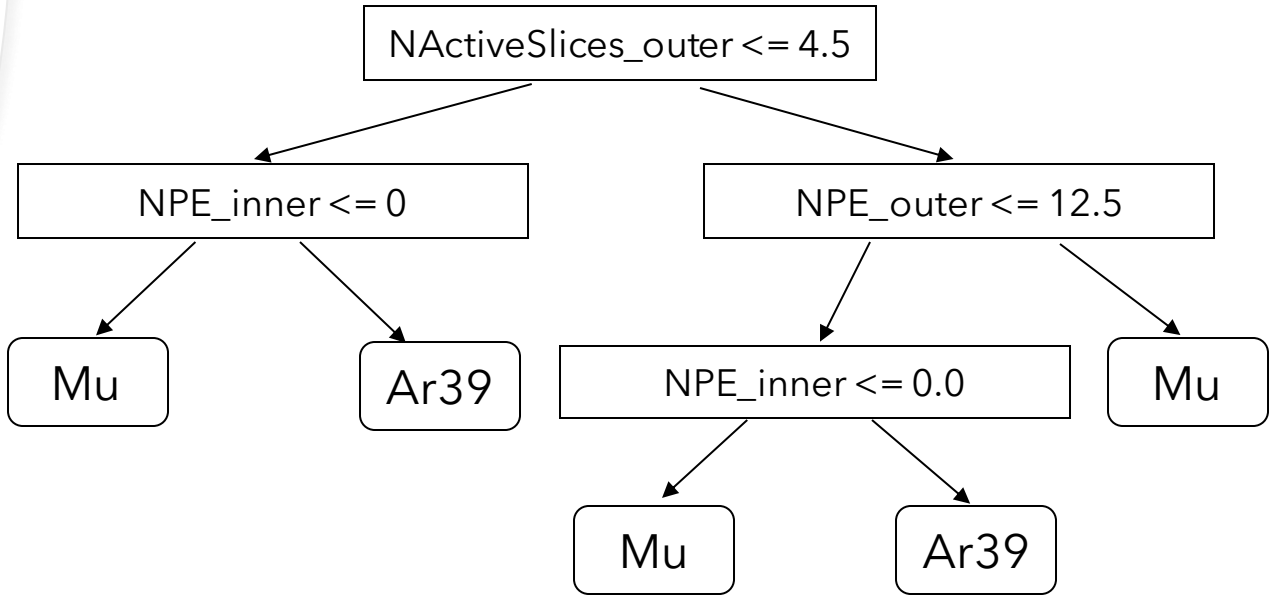
Prompt Classification – DTree

- **Feature Extraction**:
  - *NPE, MeanNPE, …*
  - *Nr Active Slices, Spatial Range, Spatial Var, …*
- **Feature Selection**: Remove too-high correlated features
- **Training**: Tune a model on selected features (*top figure*)
- **Refinement**: Remove redundant features and prune DTree (*bottom figure*)

Trained Model: 10 Features, Max Depth 4



Pruned Model: 2 Features, Max Depth 3



DTree – Zoom Trained Model