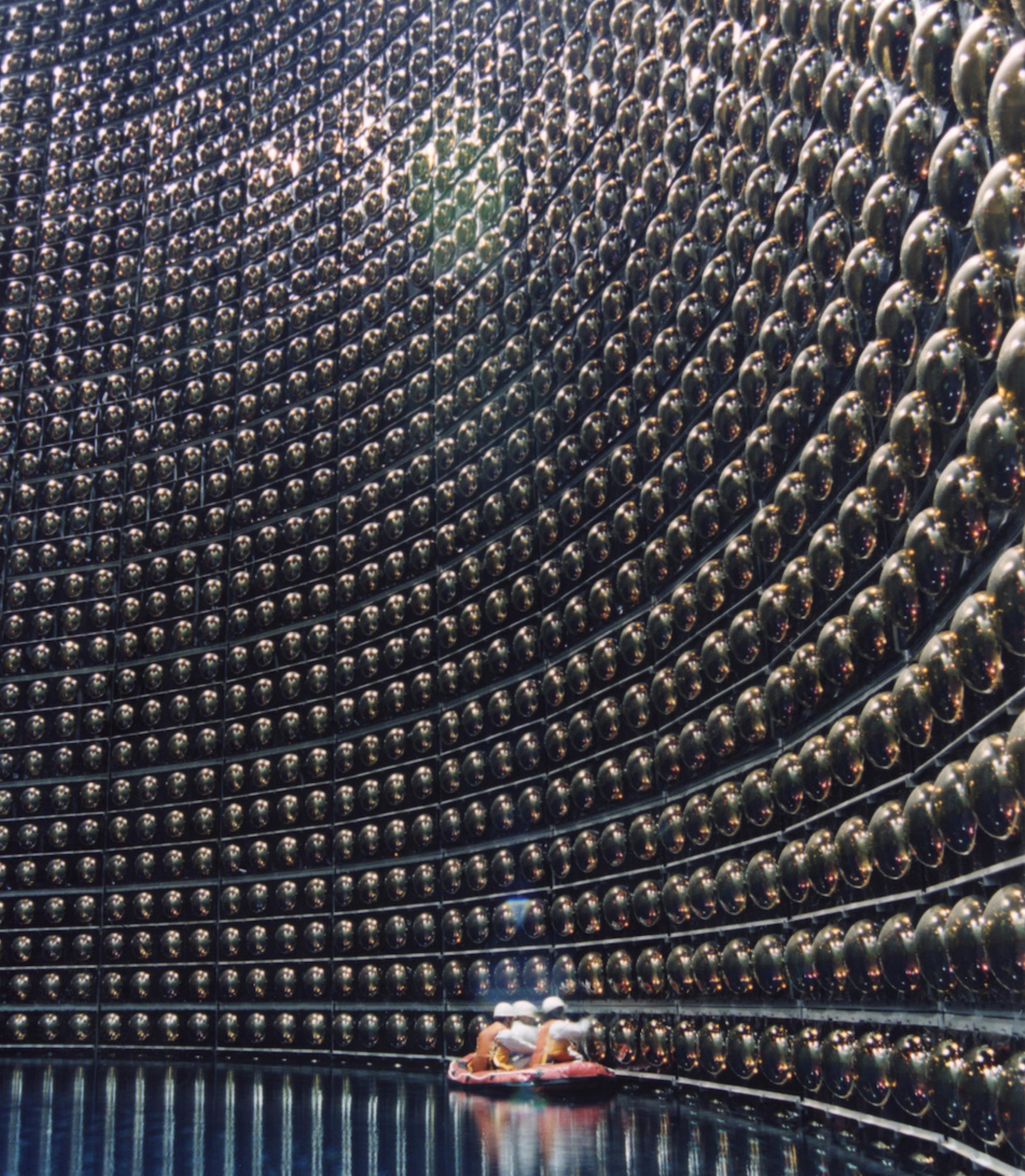


CNN4SK: Brief physical introduction to the dataset generator

Software and Computing for Nuclear and
Subnuclear Physics - Exam Project

Lorenzo Tomba
30/12/2024



Premise



This project has been *inspired* by the Super-Kamiokande (SK) neutrino observatory in Japan.

However, **the generator script (generate.py) does not perform a proper, full physical simulation** of SK events. Instead, it will generate images with some features that resemble real SK event displays.

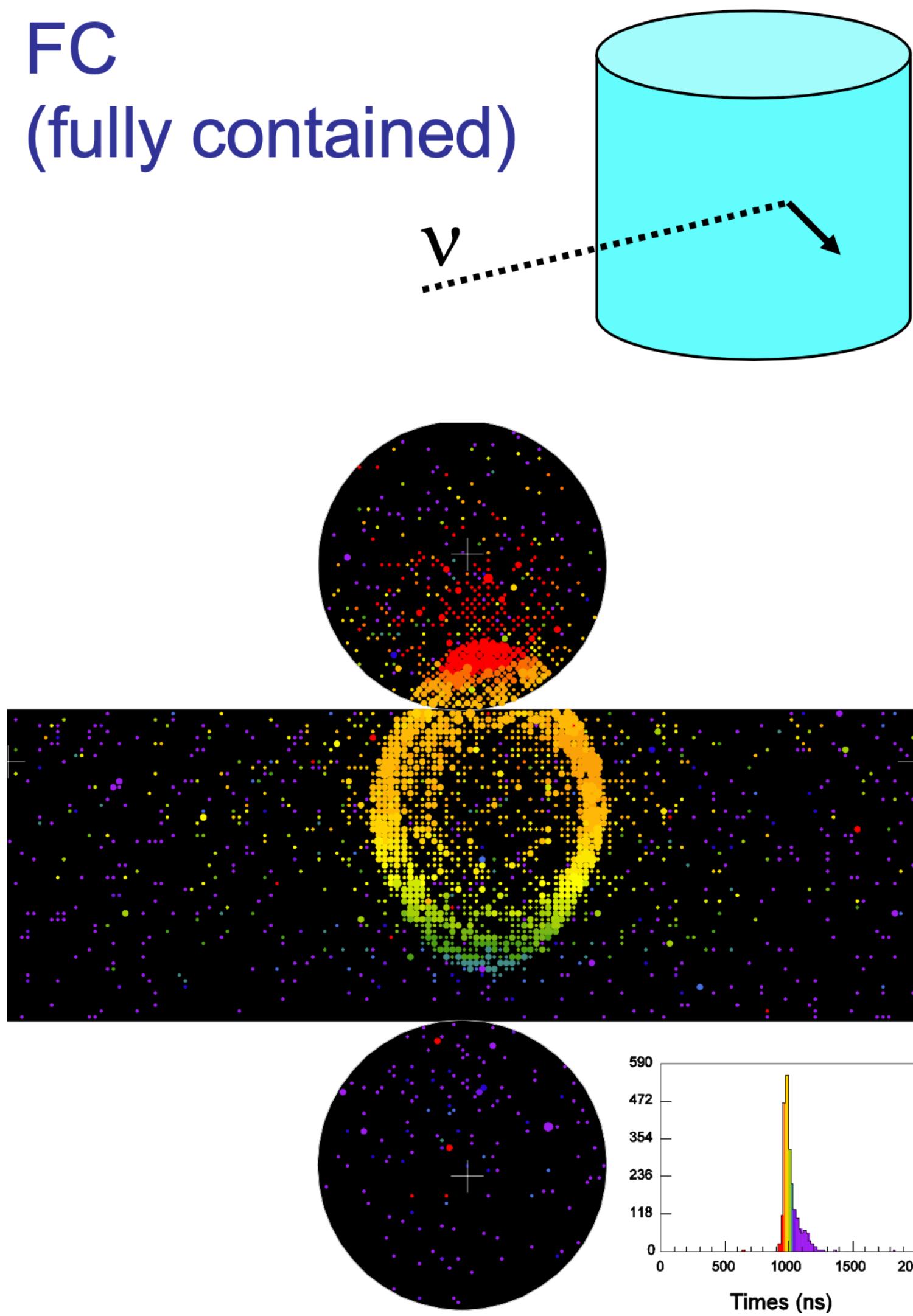
The aim of this project is to give a very preliminary answer to the question:

“Could CNN architectures such as ResNet50 be used in SK-like neutrino observatories to classify e-like, μ -like, fully contained (FC) and partially contained (PC) events?”

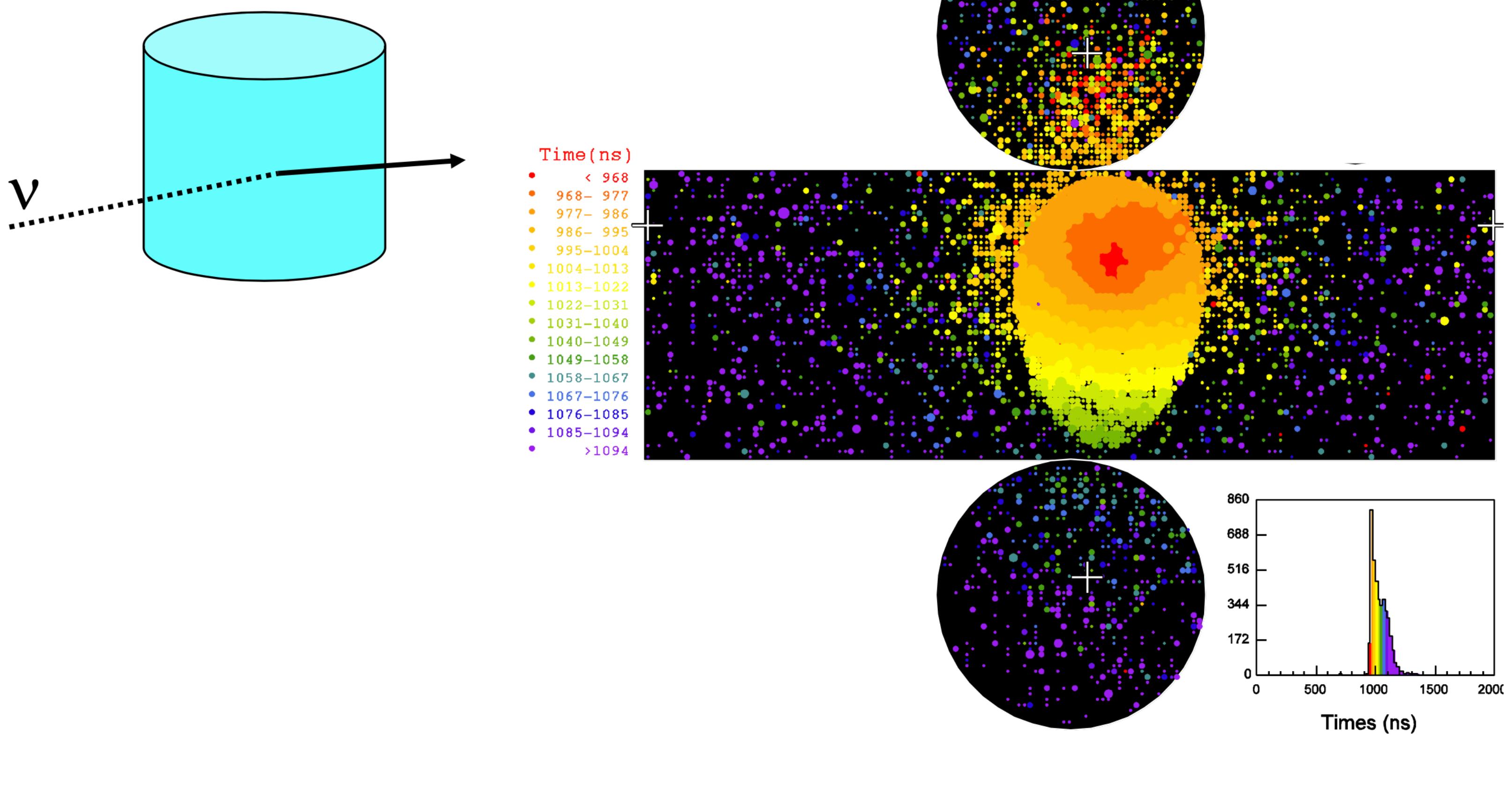
“FC” vs “PC”: real event displays

- **Fully contained (FC) events** occur when the charged particle from the neutrino interaction remains entirely within the detector. The Cherenkov light emitted by the charged particle forms a cone, and its projection on the detector walls creates a ring-like shape.
- In contrast, **partially contained (PC) events** happen when the charged particle exits the detector. In this case, the Cherenkov light produces a filled elliptical shape.

FC
(fully contained)



PC
(partially contained)



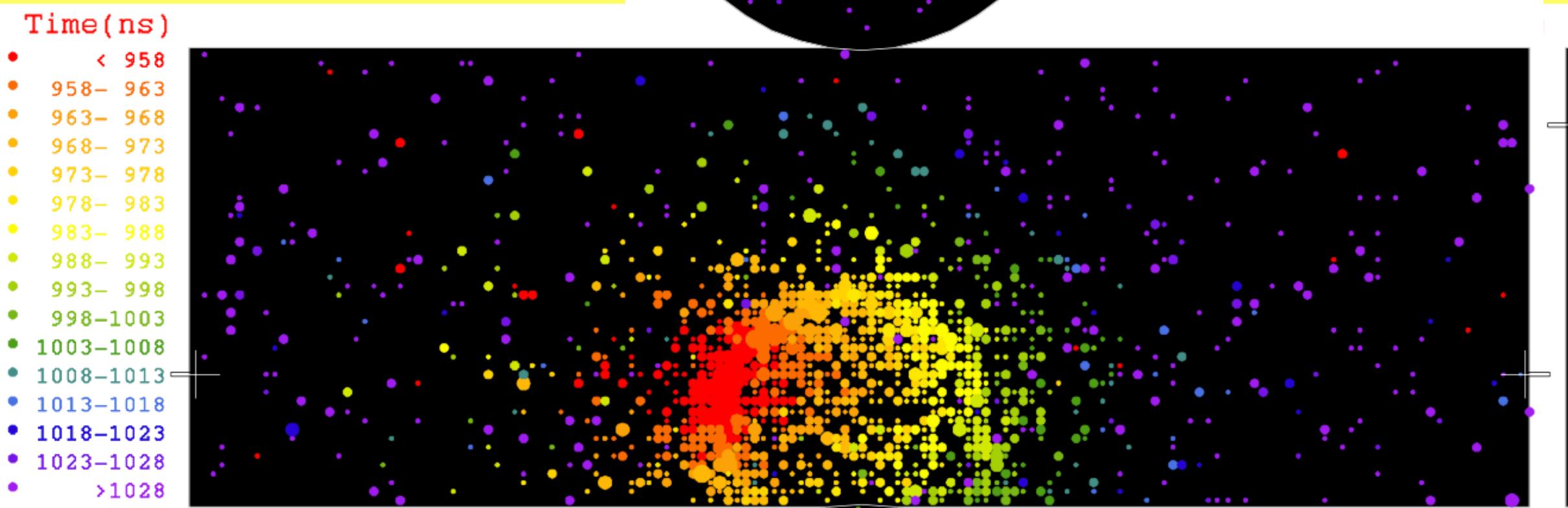
“e-like” vs “ μ -like”: real event displays

In Super-Kamiokande, e-like and μ -like events differ due to the distinct behaviors of electrons and muons when they travel in a medium and emit Cherenkov light.

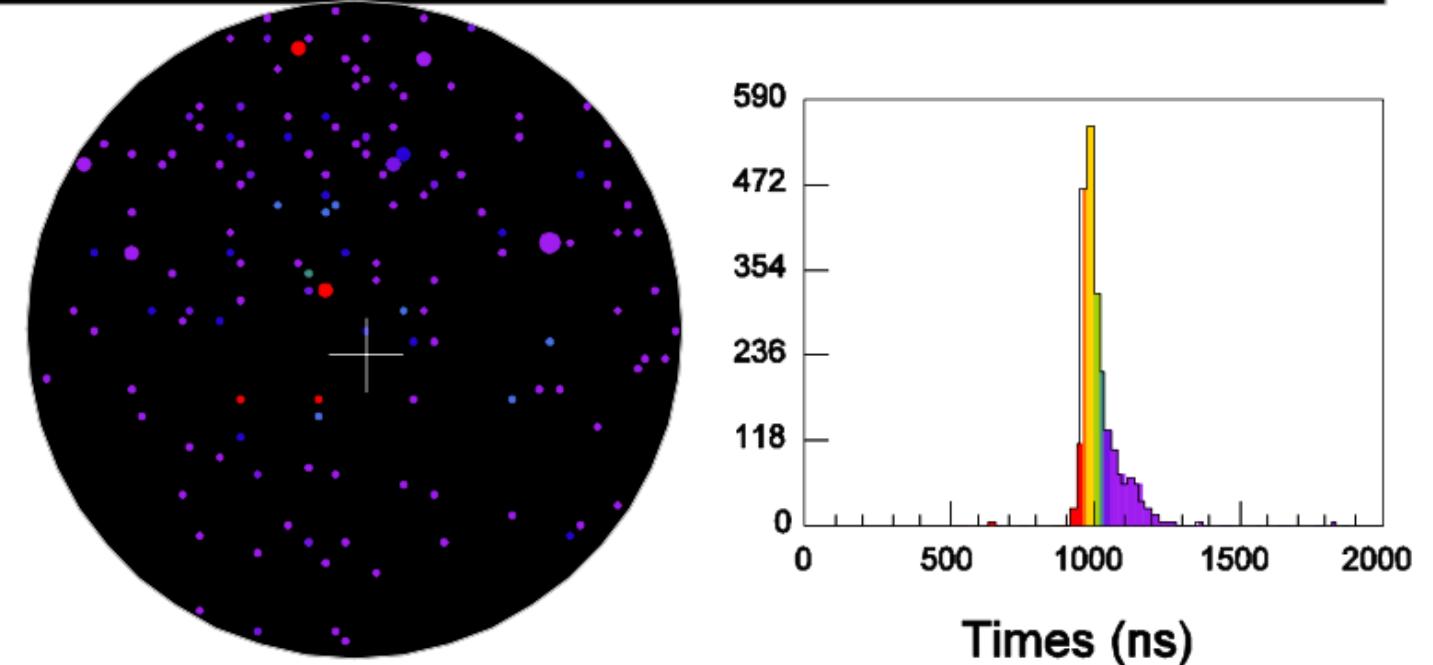
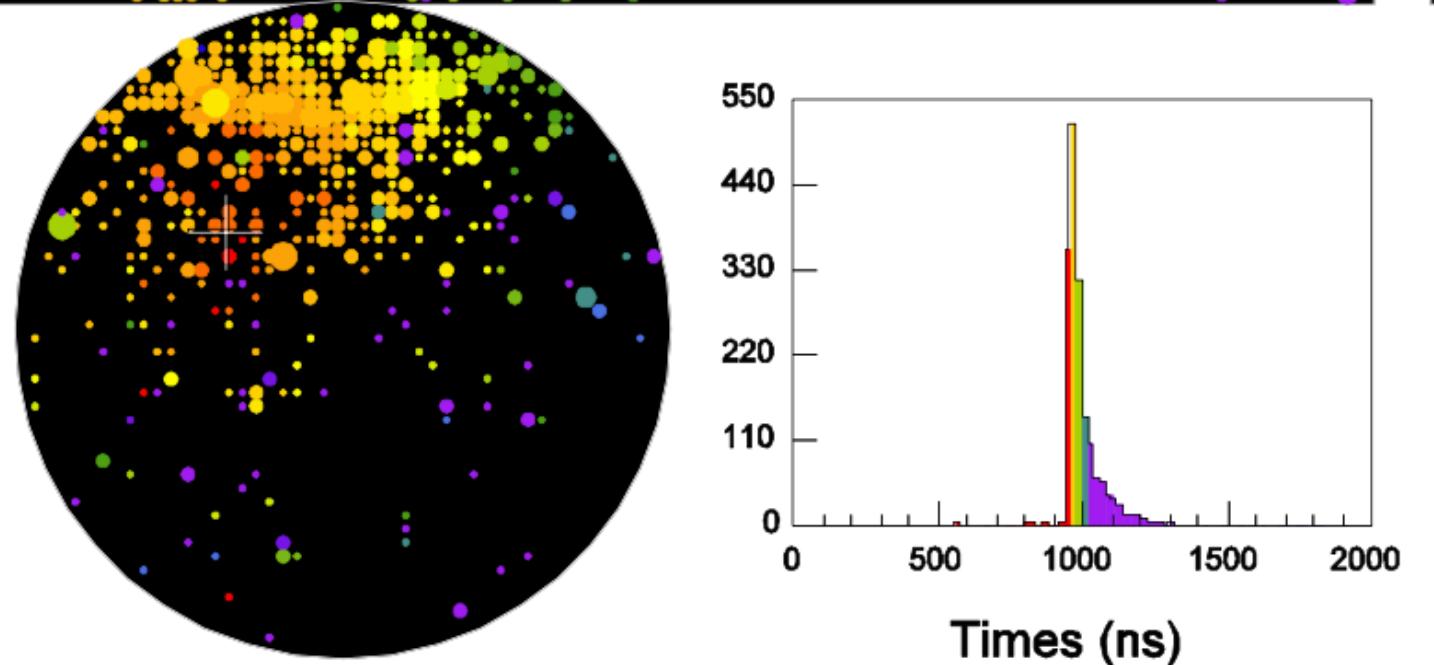
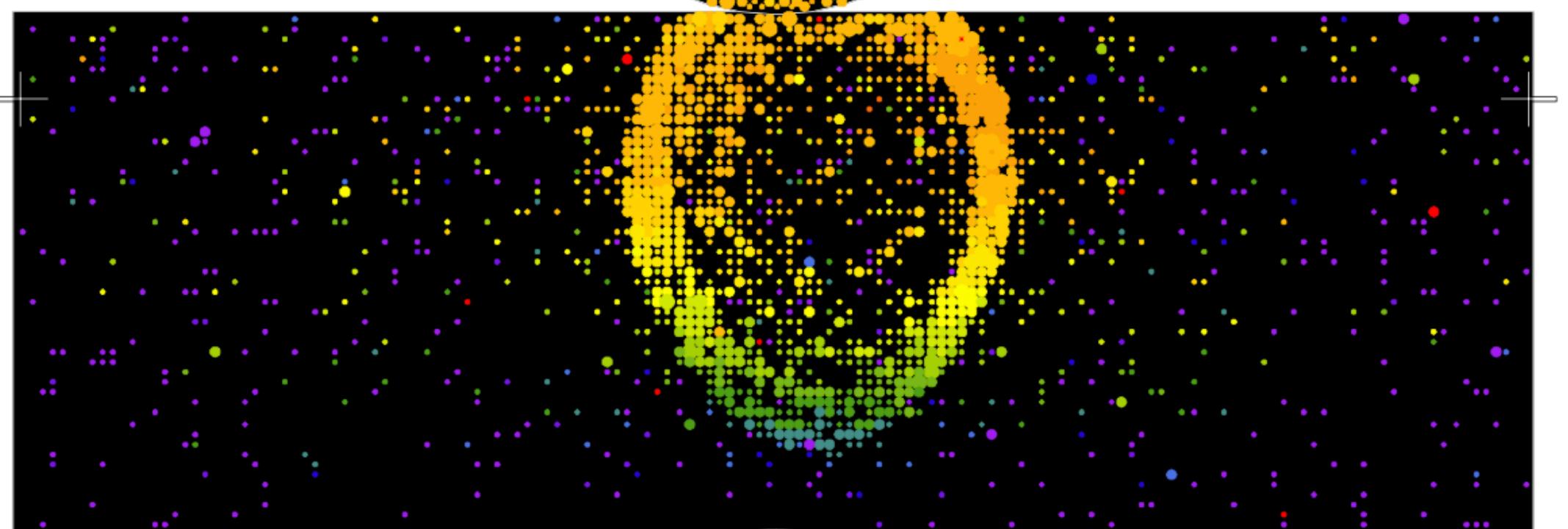
- **e-like events** Electrons interact strongly with the water via multiple scattering and electromagnetic showering. The result is a diffuse, fuzzy ring or filled pattern.
- **μ -like events** Muons, being heavier, travel in nearly straight lines and do not shower as electrons do. Their Cherenkov light is emitted along a single well-defined path, producing a sharp, clean ring or filled pattern.

This difference in the Cherenkov patterns is critical for distinguishing between interactions caused by electron neutrinos (e-like) and muon neutrinos (μ -like).

Single
Cherenkov ring
electron-like
event



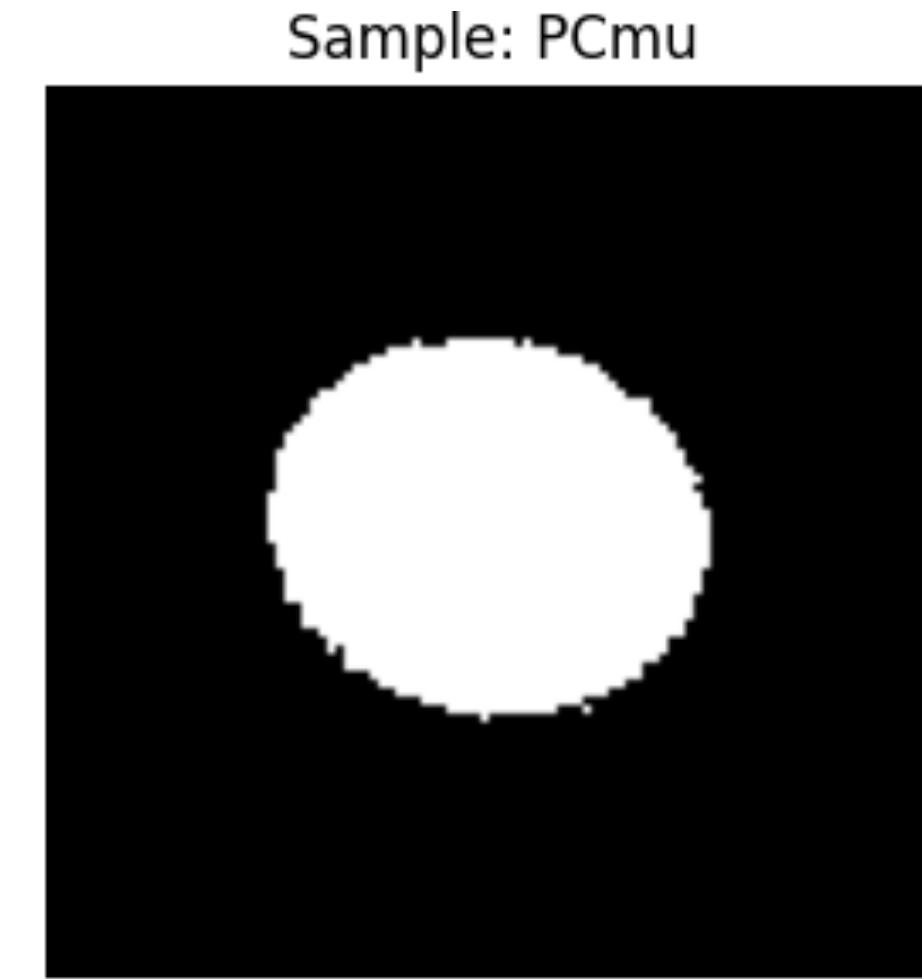
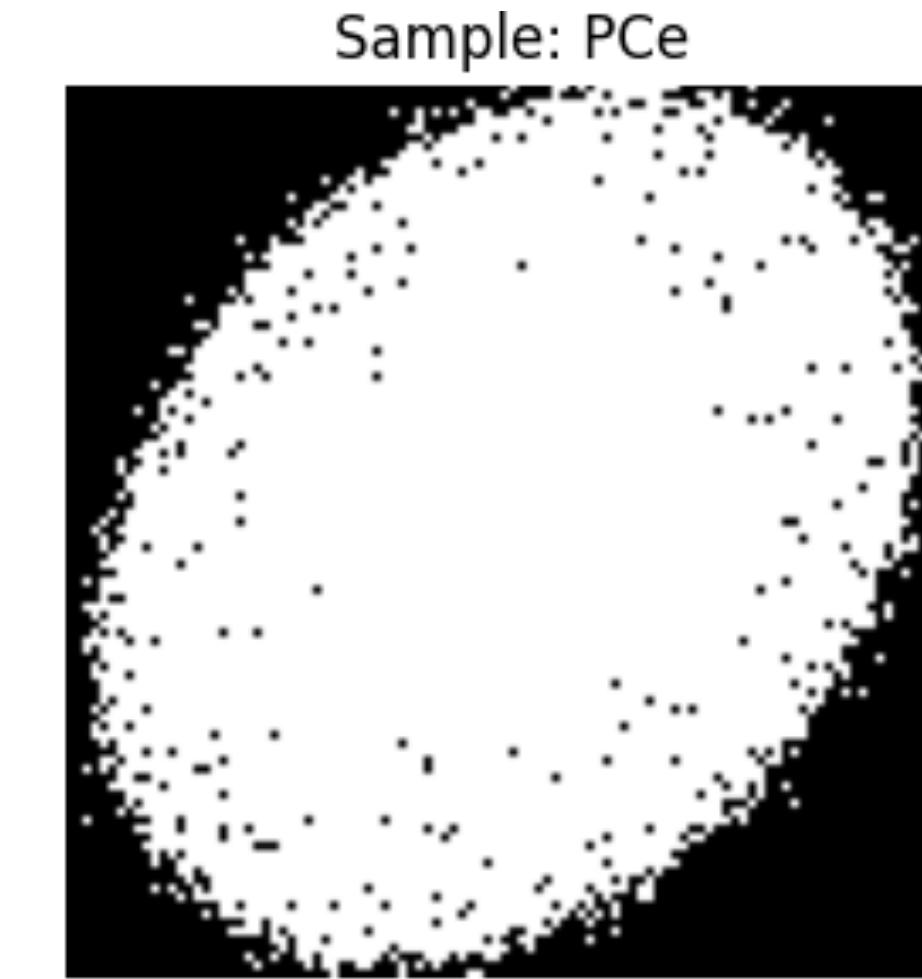
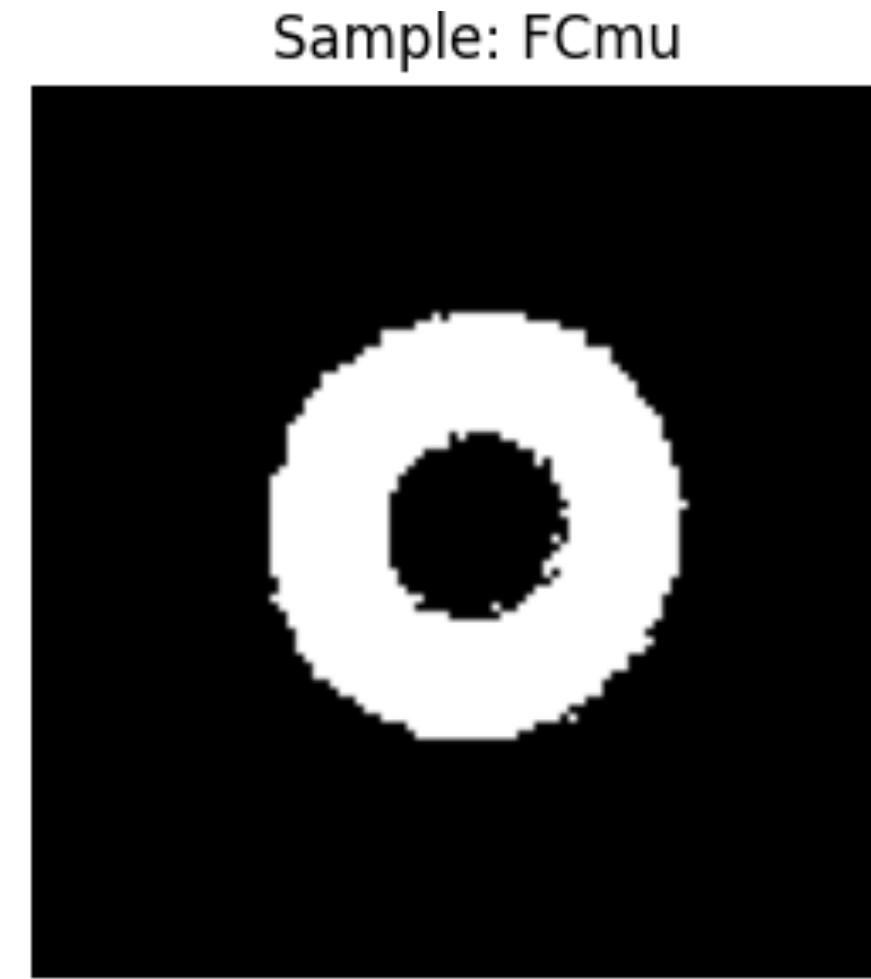
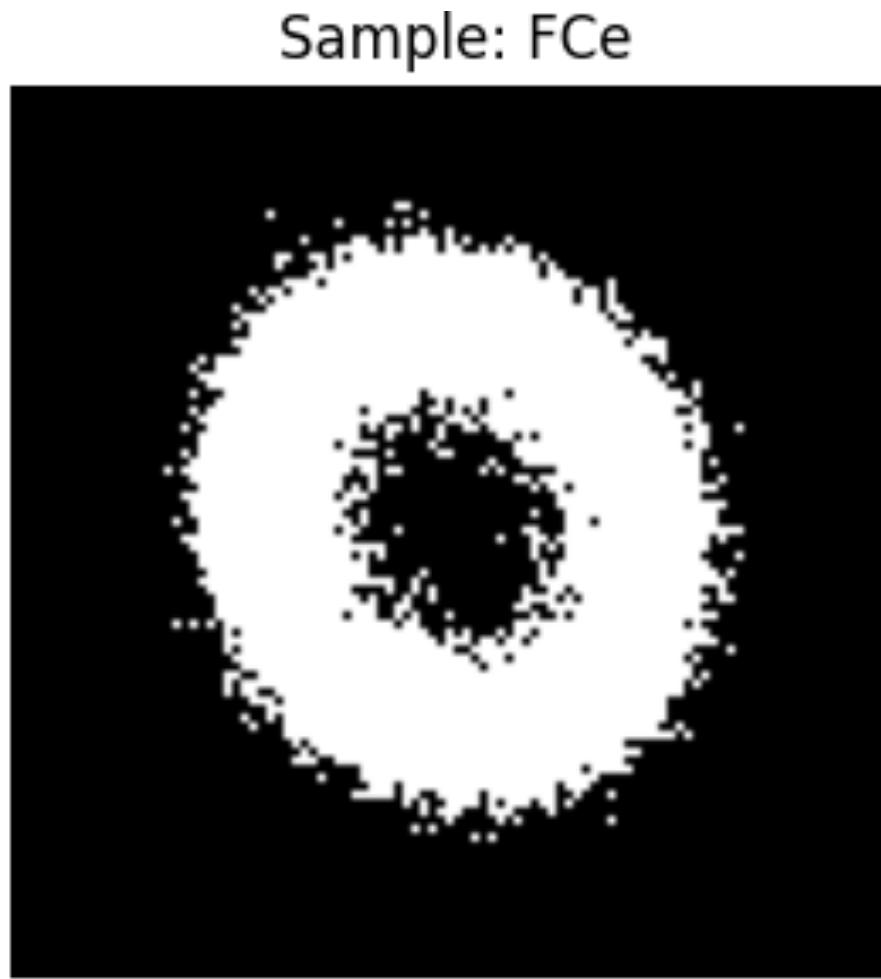
Single
Cherenkov
ring **muon-**
like event



Description of the generator script

- **Parameter Generation** The script randomly selects:
 - Particle ID: electron or muon (“e” or “mu”);
 - Topology: Fully Contained (FC) or Partially Contained (PC);
 - Angles: ϕ (azimuthal angle) within $[0, 2\pi]$ and θ (polar angle) within $[0, \pi/2]$;
 - Distances: birth distance (from the plane where Cherenkov light is being recorded) and death distance for the FC scenario (set to zero for the PC scenario).
- **Cone-Plane Intersection** Cherenkov radiation forms a cone. Its intersection with a plane is represented by an ellipse. The script computes:
 - Outer ellipse parameters based on the larger (birth) distance;
 - Inner ellipse parameters based on the smaller (death) distance.The ellipse transitions from a larger size to a smaller size as the particle moves through the detector.
- **Incremental Ellipse Drawing** A sequence of ellipses is generated at discrete intervals (`num_steps`) between the outer and inner distances. Each ellipse is scaled and oriented according to:
 - The Cherenkov angle (fixed at $\sim 41^\circ$);
 - The particle’s polar angle θ and azimuthal angle ϕ .The script calculates the major and minor semi-axes at each step and renders an ellipse on a 2D grid.

- **Noise and Multiple Coulomb Scattering** A noise model is introduced to replicate the diffuse edges observed in real Cherenkov rings:
 - Electrons typically exhibit larger scattering;
 - Muons exhibit less scattering.A Gaussian random displacement (larger for e-like events) is applied to each point of the ellipse. Noise growth over the trajectory is implemented to mimic increased scattering towards the end of the path.
- **Field-of-View Checks** If the generated ellipse lies outside the detector's 2D window, the event is discarded. In practical terms, this step ensures that the script only retains images where the Cherenkov pattern lies inside the detector plane before the application of noise.
- **Output and Labeling** Valid events are saved as 2D black and white images. The file naming encodes: topology (e.g., “FC” or “PC”), particle type (e.g., “e” or “mu”) and a sequential event index. Images are saved in subfolders named “FCe”, “FCmu”, “PCe” and “PCmu”, reflecting both the topology and the particle ID.





Main limitations

- In actual experiments, the detector geometry can be much more complex, for example a large cylindrical tank lined with thousands of photomultiplier tubes. By contrast, this dataset generator models only a single two-dimensional plane to represent light detection, rather than a curved surface. This script treats light detection as a uniform pixel grid.
- The generator assumes an unattenuated cone emission, without accounting for wavelength-dependent attenuation or other realistic optical effects. It also does not address photostatistics or dark noise, both of which can influence the actual number of photoelectrons recorded by each PMT.
- This dataset generation does not physically model the full multiple scattering process that charged particles undergo in water. Instead, it applies a simple Gaussian noise component to each ellipse point, which increases with the progression of the particle's motion. This approximation emulates a *broader Cherenkov ring edge*, but does not replicate the detailed scattering interactions that would occur inside a real detector volume.
- In the actual detector, geometric and fiducial volume cuts are applied. These have not been taken into consideration in the generator script.