ACAT 2014

- Prague -

1-5 September 2014

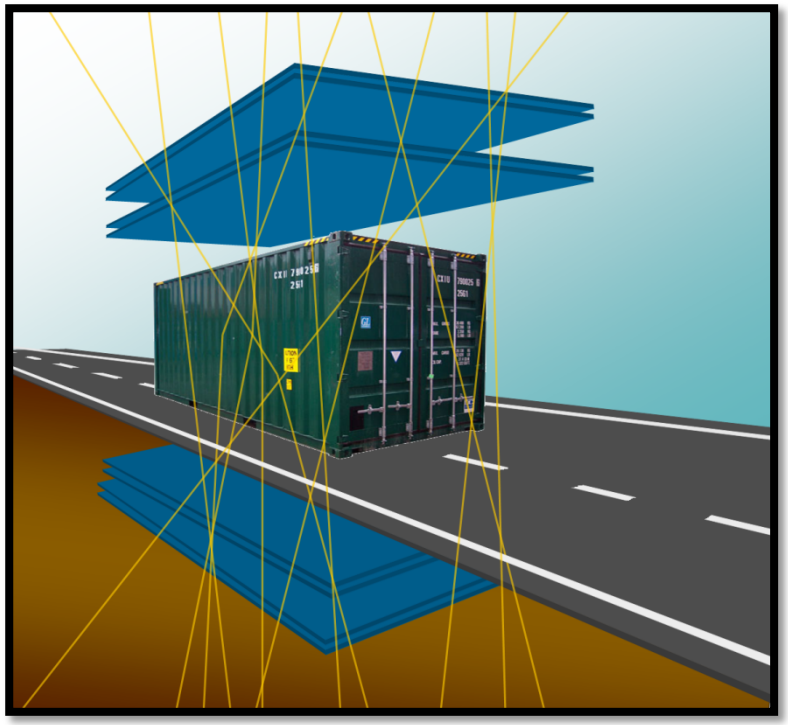
Clustering analysis on muon tomography data in the Muon Portal Project

M. Bandieramonte, V. Antonuccio, U. Becciani, A. Costa, P. La Rocca, P. Massimino, C. Petta, C. Pistagna, F. Riggi, S. Riggi, E. Sciacca & F. Vitello

INAF

Outline

- **Borders security problem**
- **Muon Tomography goals and concept**
- **The Muon Portal Project**
  - Overall Design
  - Tracking and Image reconstruction
  - Density based clustering for muon tomography
  - Simulation scenarios and rendering results
- **Conclusions and future works**

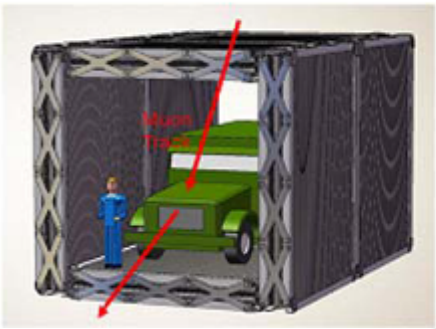


The problem

- Detection of **nuclear weapons** and **radioactive materials**
- Safety regulations are demanding a **better** and **faster** way to inspect containers, not possible by traditional techniques
- **Containers:** employed since 50 years as a standard way to transport goods by ships or trucks
  - Estimated yearly traffic: 200 M containers (in Catania ~ 15000 container /year)

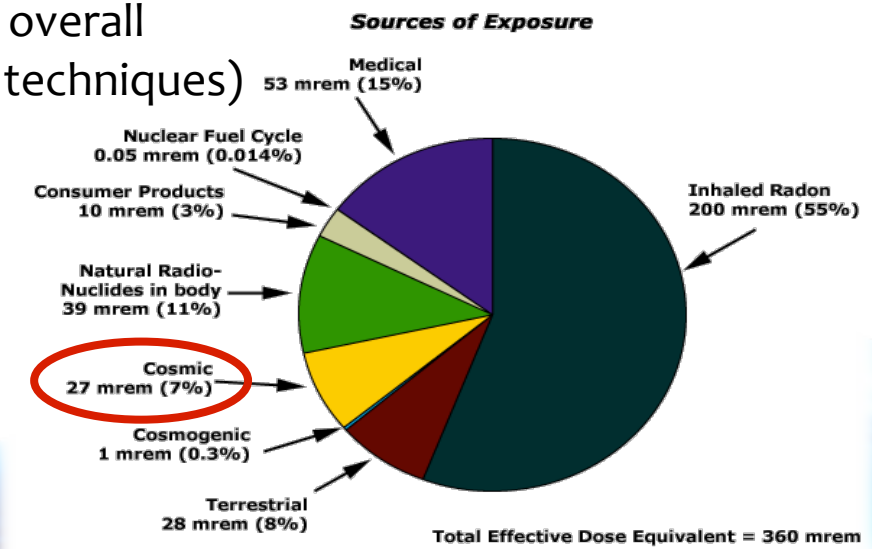
The problem

- Presently, only about 1% inspected
- **X-ray tomography**
  - Based on X-ray absorption (requires high fluxes)
  - Problems to traverse large thickness (Mean free path ~ 25 g/cm^2 ~ 2 cm Pb)
  - Mono or two-dimensional, not 3D (require multiple projections)
- Other possibilities
  - Directional imaging with **gamma rays**
  - **Neutron** radiography



Why Muon Tomography

- Secondary cosmic muons: **“Natural” radiation,** hence no additional dose delivered to users and goods
- Highly **penetrating radiations** (even more than X-rays)
- **Muon flux** at sea level relatively large: 10^{-2} cm^{-2} s^{-1}
- **Muonic interactions** well understood
- Muon scattering strongly **dependent on the Z of the material**
- **Each muon may contribute** to determine the overall imaging result (contrary to muon absorption techniques)

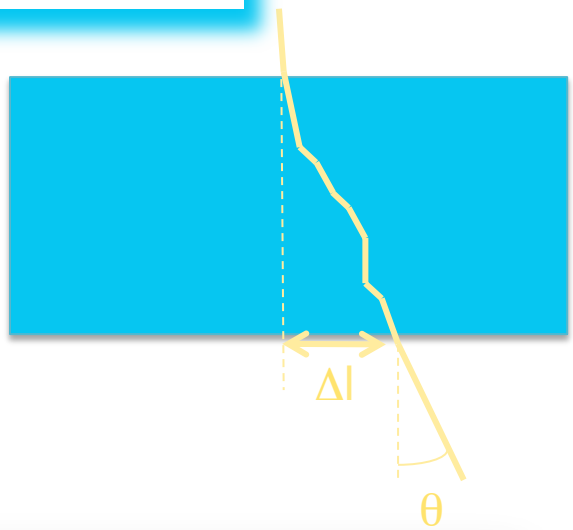


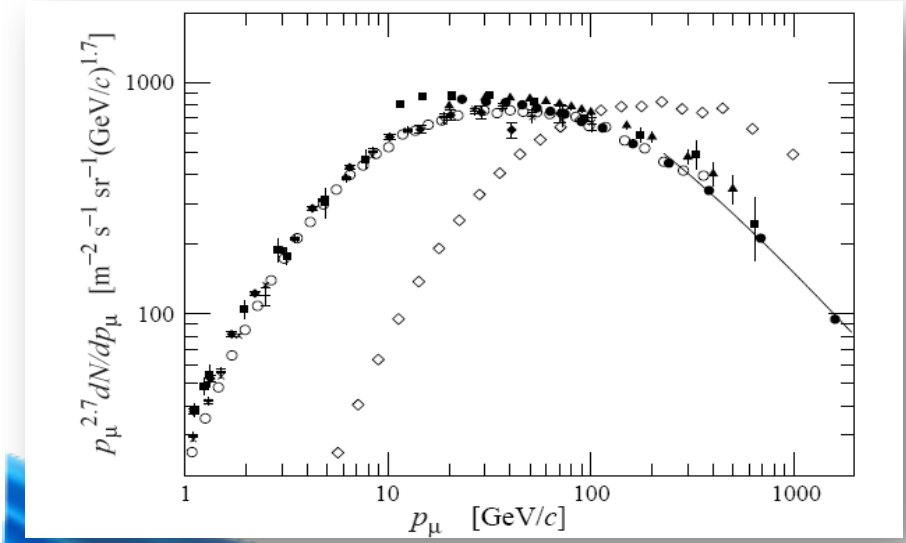
Muon Tomography

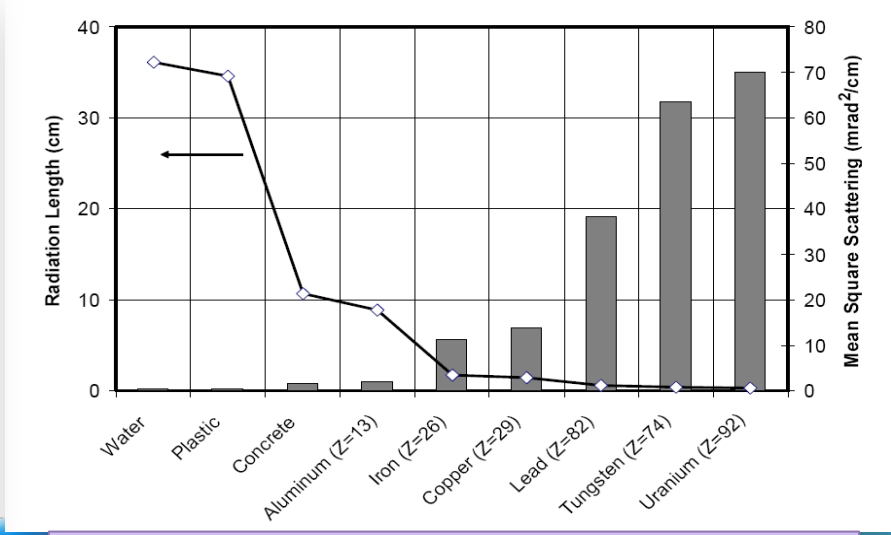
$$\theta_0 = \theta_{\text{plane}}^{\text{rms}} = \frac{1}{\sqrt{2}}\theta_{\text{space}}^{\text{rms}}$$

$$\theta_0 = \frac{13.6\text{MeV}}{\beta cp} \cdot z\sqrt{\left(\frac{x}{X_0}\right)} \cdot \left[1 + 0.038 \ln\left(\frac{x}{X_0}\right)\right]$$

$\beta c \rightarrow$ Velocity
$p \rightarrow$ Momentum
$z \rightarrow$ Charge Number
$x \rightarrow$ Width of Medium
$X_0 \rightarrow$ Radiation Length



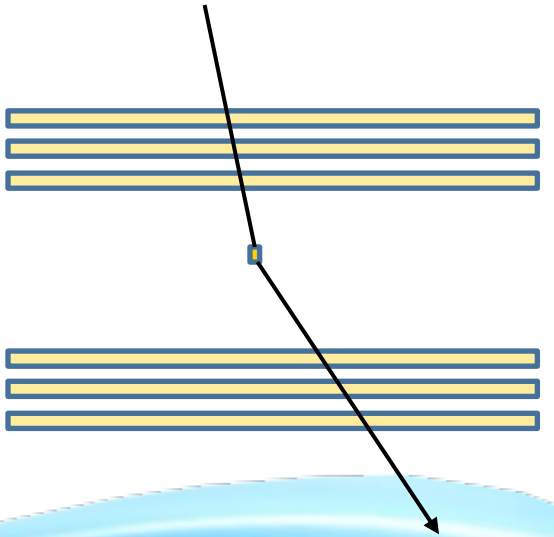




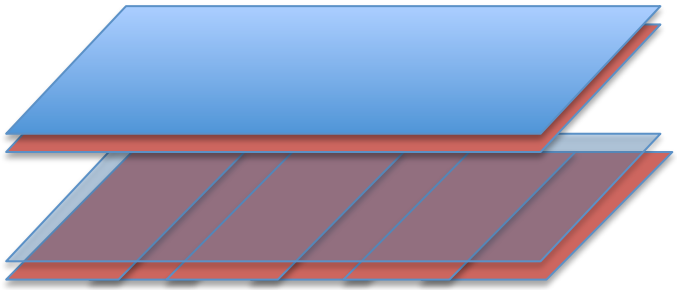
L. Schultz IEEE NSS Conference Record (2006)

Muon Tomography Detector

- A muon tomography detector is basically made by **several detection planes**, above and below the volume to be inspected
- Reconstruction of muon tracks allows, by different methods, to produce a **3D tomographic image**
- Performance of the system are given in terms of
  - **Sensitive** area / volume to be inspected
  - Spatial and angular **resolution**
  - **Time to scan** a volume
  - **Sensitivity & Efficiency** to high-Z objects
  - Sensitivity to **false positive**
  - **Discrimination** between high-Z vs low & medium-Z

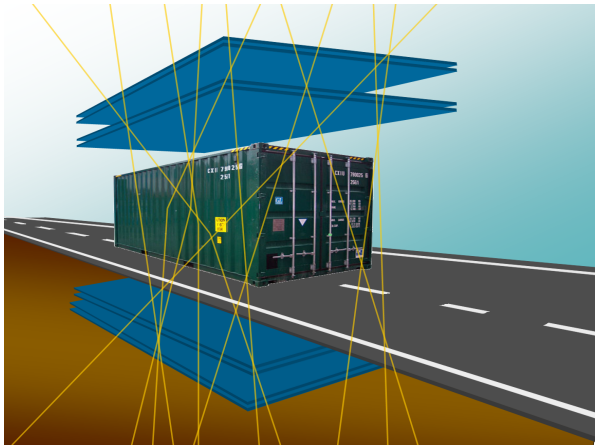


The Muon Portal project

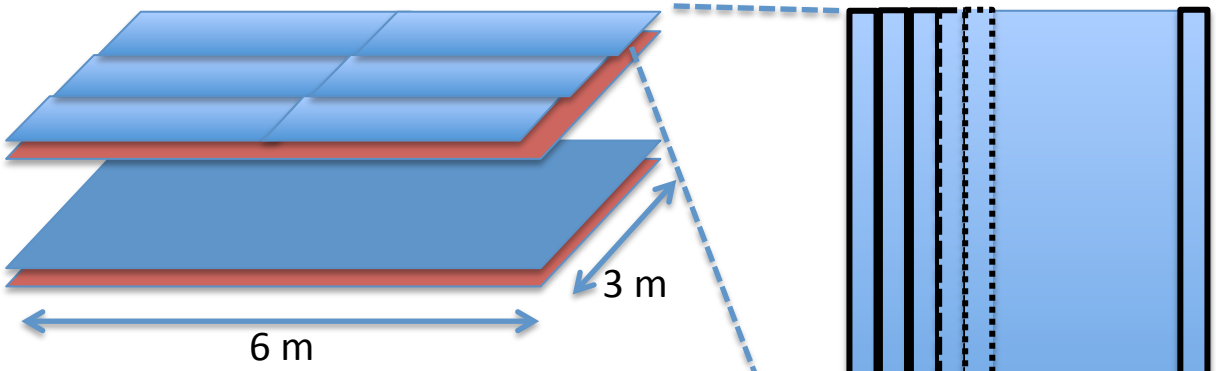


Plane X

Planes: 4 XY, 18m^2,
Separation[m]: 1.0-3.0-1.0
Modules: 2x3 in X, 6x1 in Y,
100 strips per module
Trigger: 4 planes crossed and
dE/dX > 0.5 MeV
A=10 m^{2}sr -> 20%
N_{events} ~2×10^5 for Δt=5 min
φ_μ = 1 m^{-2} s^{-1} integrated over the solid angle.

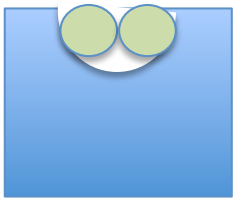


Plane Y



Module
1x3 m^2:100 extruded plastic scintillator strips 1x1 cm^2

Strip design with 2 WLS fibers



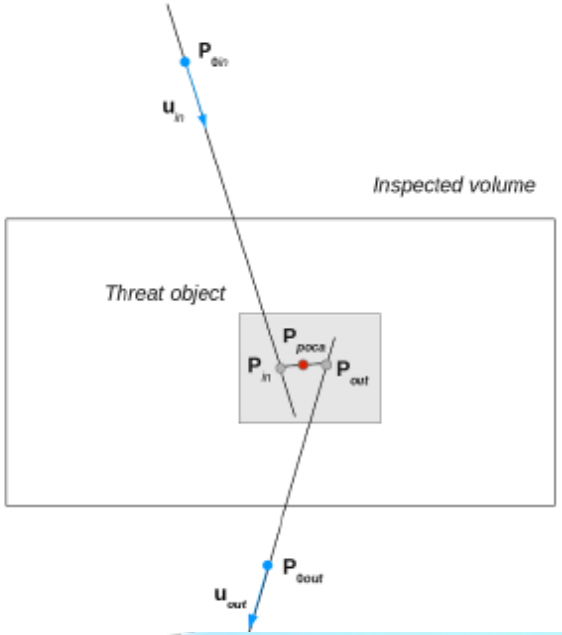
Poca Algorithm

- Simplest approach, fast and easy to implement
  - ✓ Geometrical **P**oint **O**f **C**losest **A**pproach between incoming and outcoming tracks

$$P_{poca} = \frac{1}{2} (P_{in} + P_{out})$$

  - ✓ Spatial distribution of the scattering centers, weights given by some power of the scattering angle

- However:
  - ✓ **Neglects multiple scattering** within the material
  - ✓ **Poor resolution** images
  - ✓ Critical behavior for material located close to **volume borders**

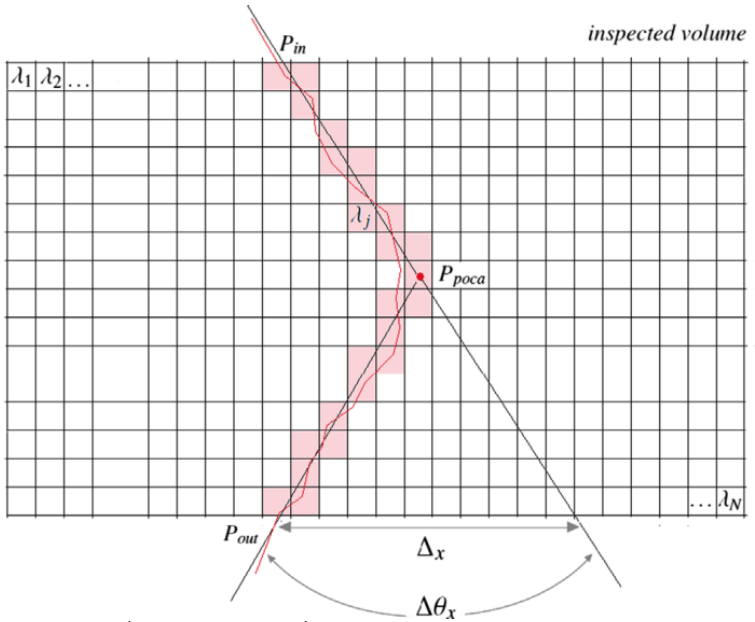


S.Riggi et al., NIM A 728 (2013) 59-68

EM-Likelihood algorithm

- Better statistical treatment of scattering processes by a log-likelihood method
  - ✓ Volume to be inspected divided into Nx ×Ny ×Nz 3D voxels
  - ✓ Scattering density defined for each voxel

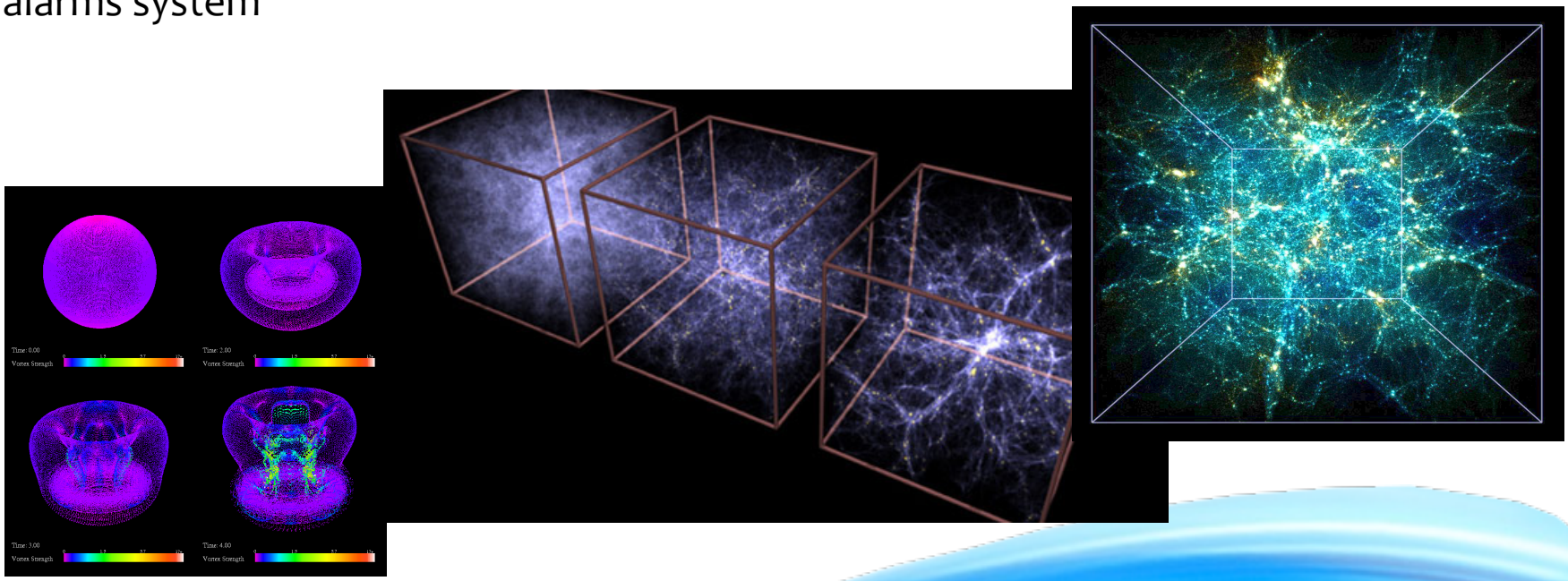
$$\lambda(X_0) = \left(\frac{15\ \text{MeV}}{p_0}\right)^2 \frac{1}{X_0}$$



  - ✓ Model the scattering and displacement data yi=(Δθi ,ΔXi) with a bivariate normal distribution
  - ✓ Iterative estimation with some stopping criteria

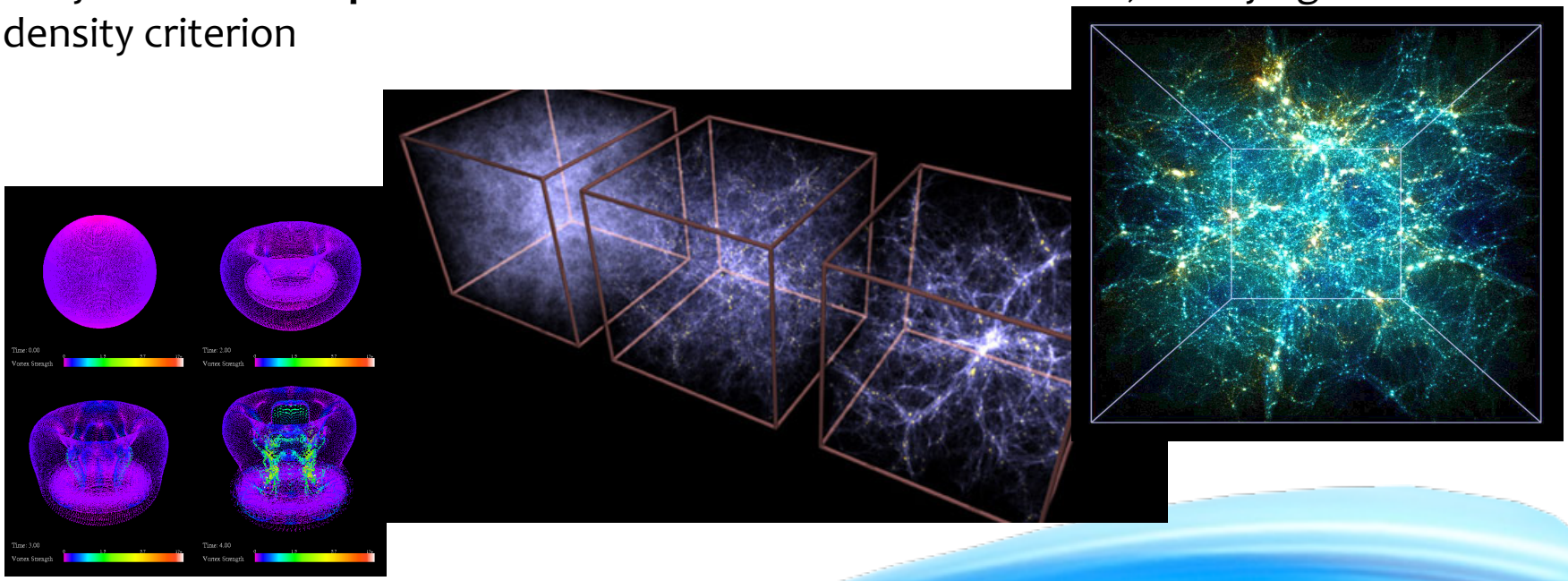
Density based Clustering: Why?

- ✓ The presence of a **three-dimensional grid** limits the automatic object identification process
- ✓ The problem is relevant in scenarios where the threat to be identified has a **comparable size** (or even smaller) to those of the single voxel and is located in a position **not aligned** with the grid.
- ✓ Risk **of losing useful statistical** and invalidate the automatic recognition and alarms system



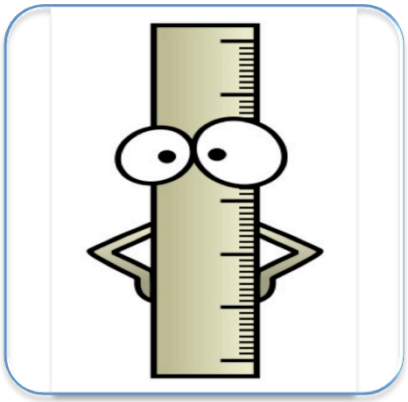
Density based Clustering: Why?

- ✓ Clustering analysis is traditionally applied to problems **like pattern recognition, image classification and color quantization.**
- ✓ It is used as a **data mining technique** to help distinguishing the structure of data and **to simplify the complexity** of huge quantities of data.
- ✓ The usage of these techniques allows **to infer the implicit information** in the data.
- ✓ They **connect data points** within a certain distance threshold ε, satisfying a density criterion



Friends of Friends Algorithm

- The **FOF** is a percolation algorithm normally used to identify dark matter halos from N-body simulations.



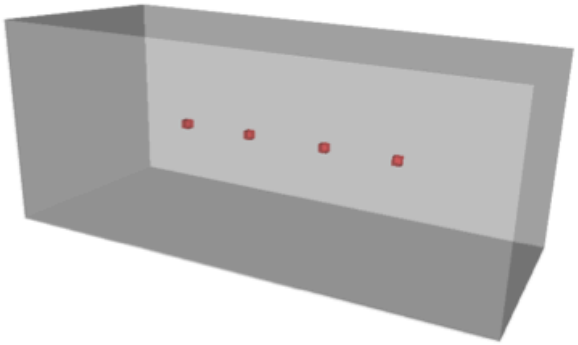
1) Defines a *Linking Length*.

2) Identifies all the pairs of "friends".

3) Identifies all the "friends of friends"

- The linking length is related to a parameter "*L*" that represents **the mean inter-particle separation** in simulations (related to the mean number density $\langle n \rangle$ as $L = \langle n \rangle^{-1/3}$
- Another parameter in FOF algorithm is **the minimum number of particles N_{min},** in a cluster. The aim is to reject spurious clusters, that is groups of friends who do not form persistent objects in the simulation

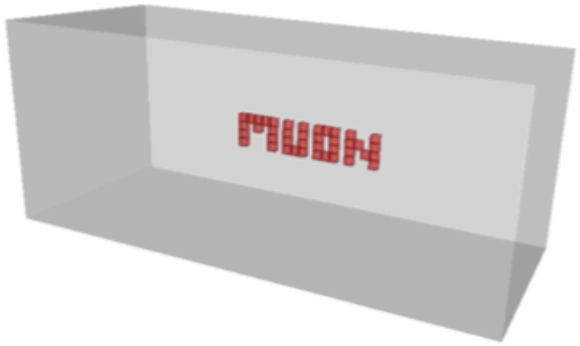
Simulated scenarios



(a) Scenario A

Four threat boxes Scenario

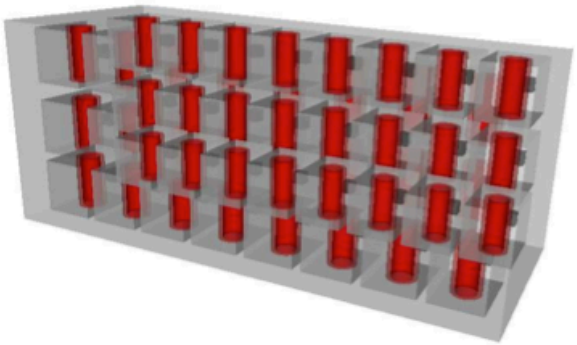
W, U, Pb, Sn of size 10 cm×10 cm×10 cm inserted at the center of a empty container.



(b) Scenario B

"Muon" Shape Scenario

'M'= Uranium (U)
'U'= Iron (Fe)
'O'= Lead (Pb)
'N'= Aluminium (Al)



(c) Scenario C

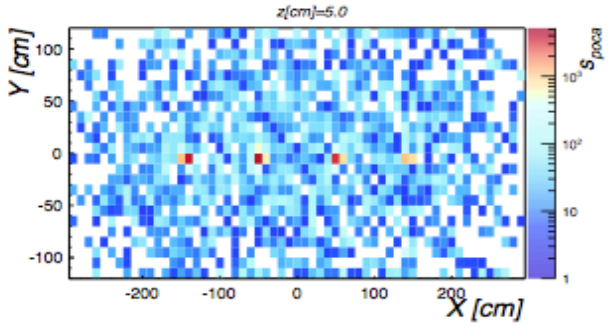
"Washing machines" Scenario

- **Fe** engine+ Concrete block+ **Al** structure

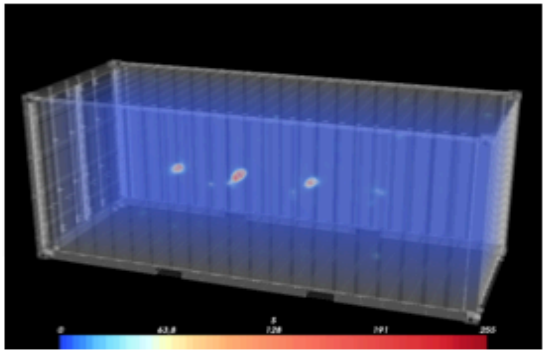
Simulations settings:

- GEANT4 simulations
- $5*10^5$ simulated events available ~ 10 min.
- realistic energy spectrum with range 0.1-100 GeV
- Only events passing the container → ~1% events rejected

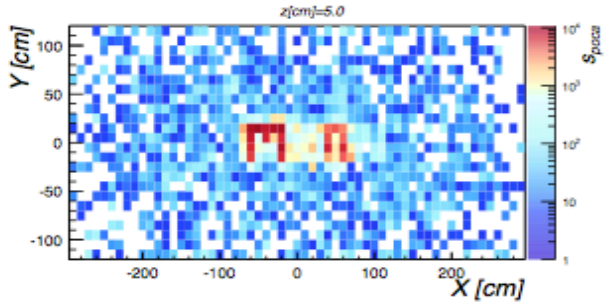
Algorithms results: POCA



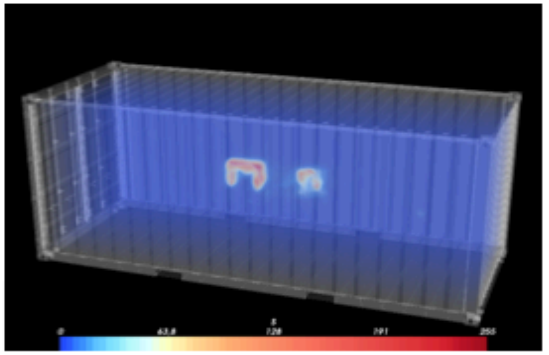
(a) POCA XY tomography - Scenario A



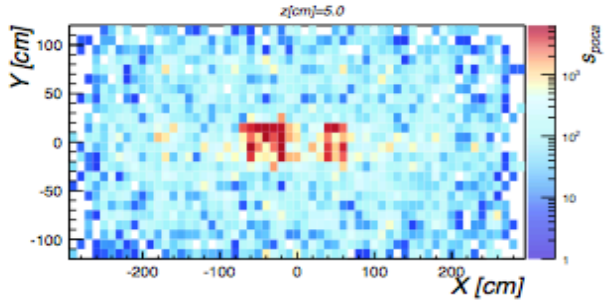
(b) POCA 3D rendering - Scenario A



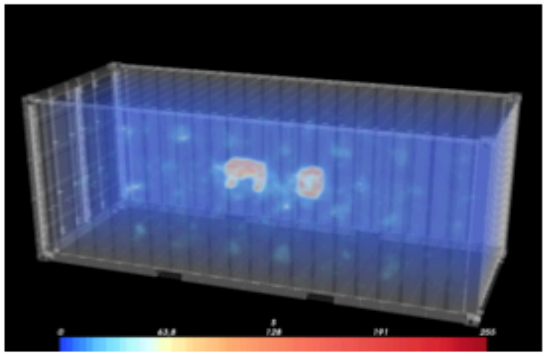
(c) POCA XY tomography - Scenario B



(d) POCA 3D rendering - Scenario B

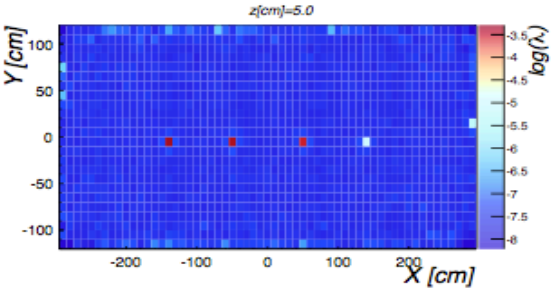


(e) POCA XY tomography - Scenario C

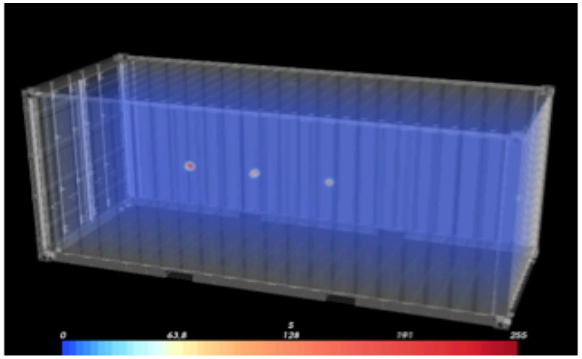


(f) POCA 3D rendering - Scenario C

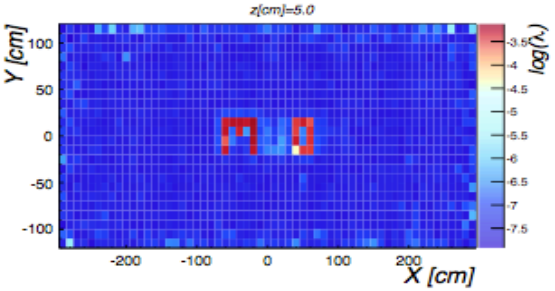
Algorithms results: EM-Likelihood



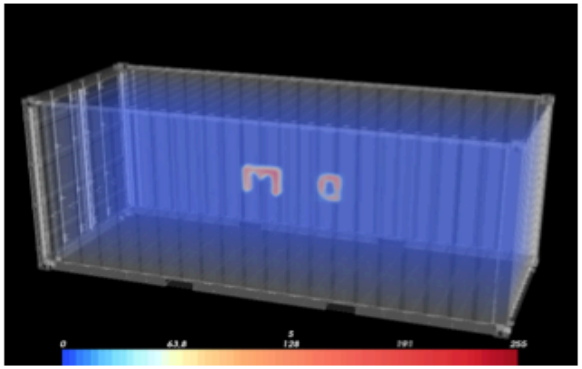
(a) EM-ML XY tomography - Scenario A



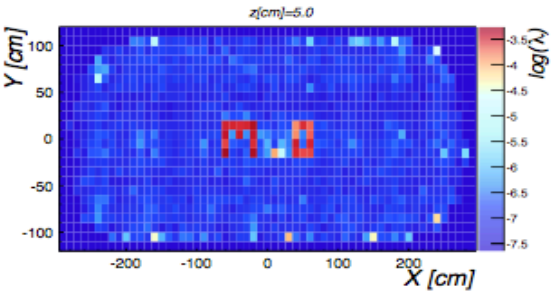
(b) EM-ML 3D rendering - Scenario A



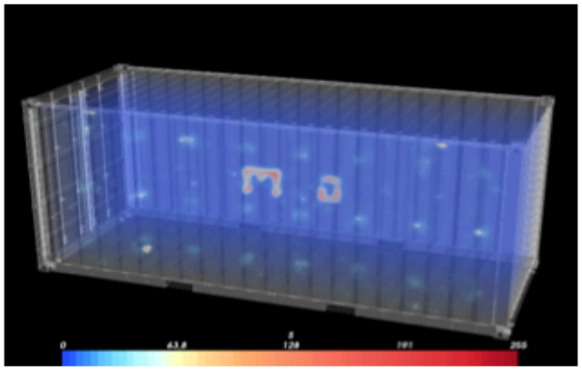
(c) EM-ML XY tomography - Scenario B



(d) EM-ML 3D rendering - Scenario B

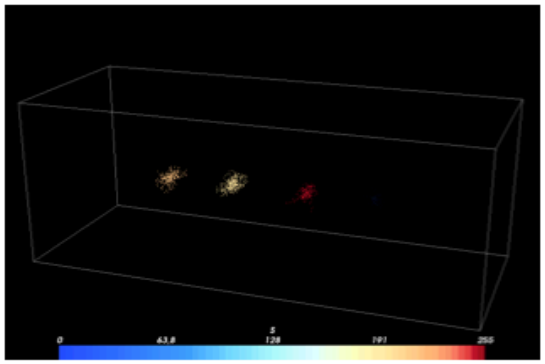


(e) EM-ML XY tomography - Scenario C

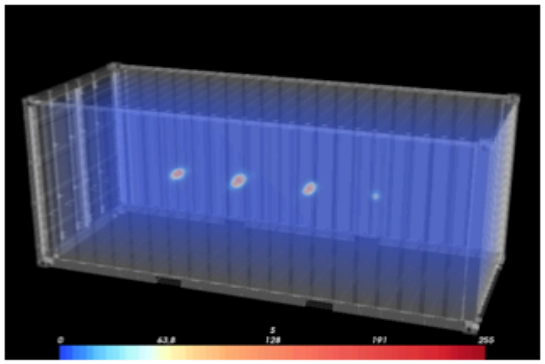


(f) EM-ML 3D rendering - Scenario C

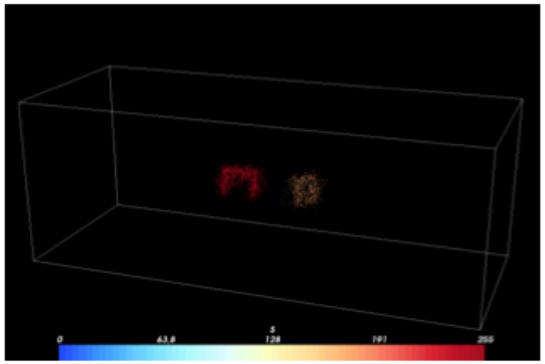
Algorithms results: Fof clustering



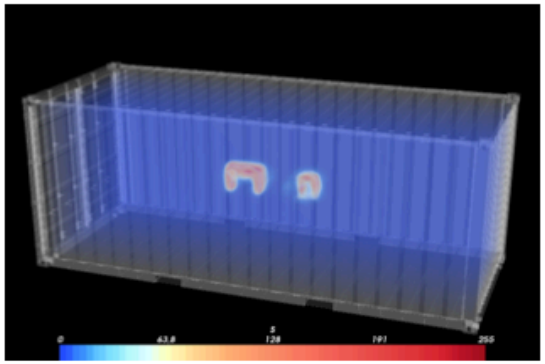
(a) Clustering - Scenario A



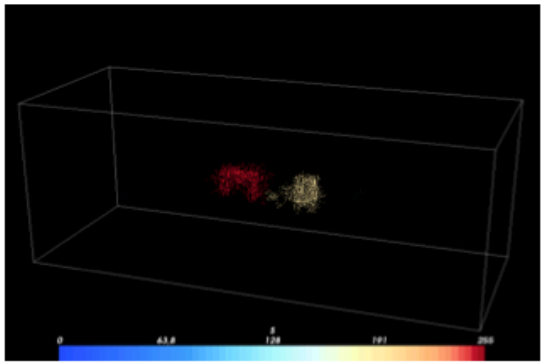
(b) Clustering 3D rendering - Scenario A



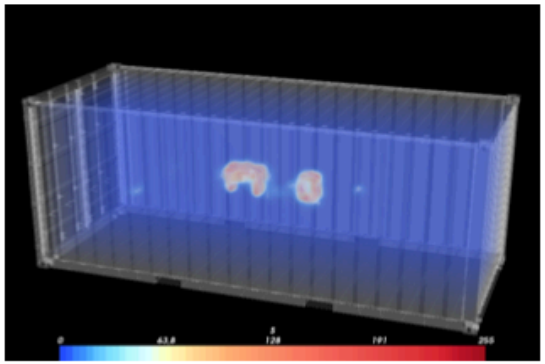
(c) Clustering - Scenario B



(d) Clustering rendering - Scenario B

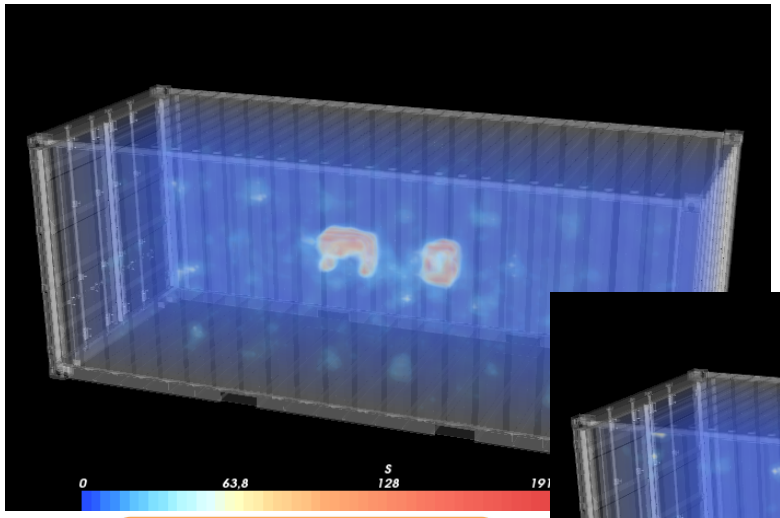


(e) Clustering - Scenario C

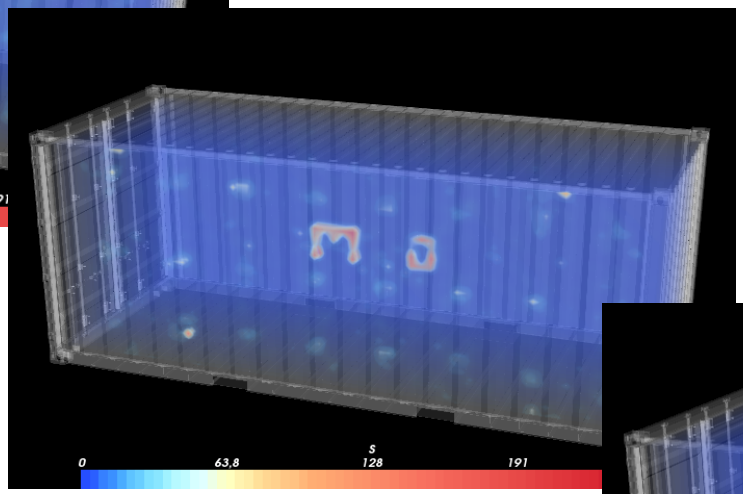


(f) Clustering 3D rendering - Scenario C

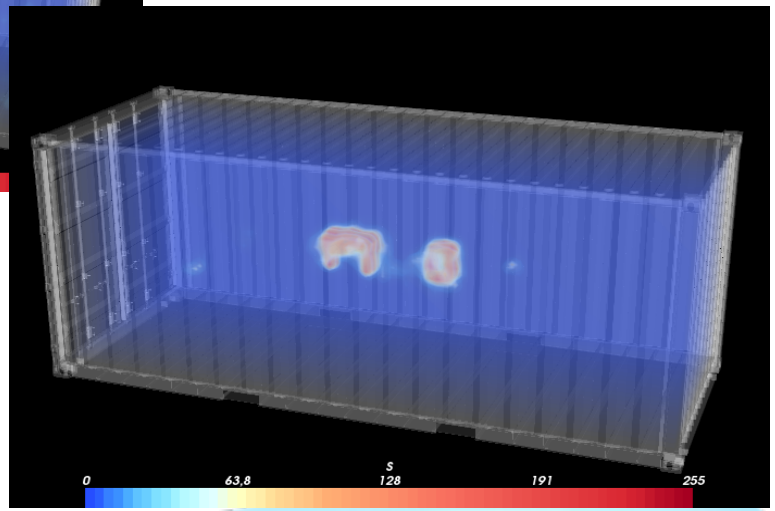
Algorithms in comparison



Poca Volume Rendering

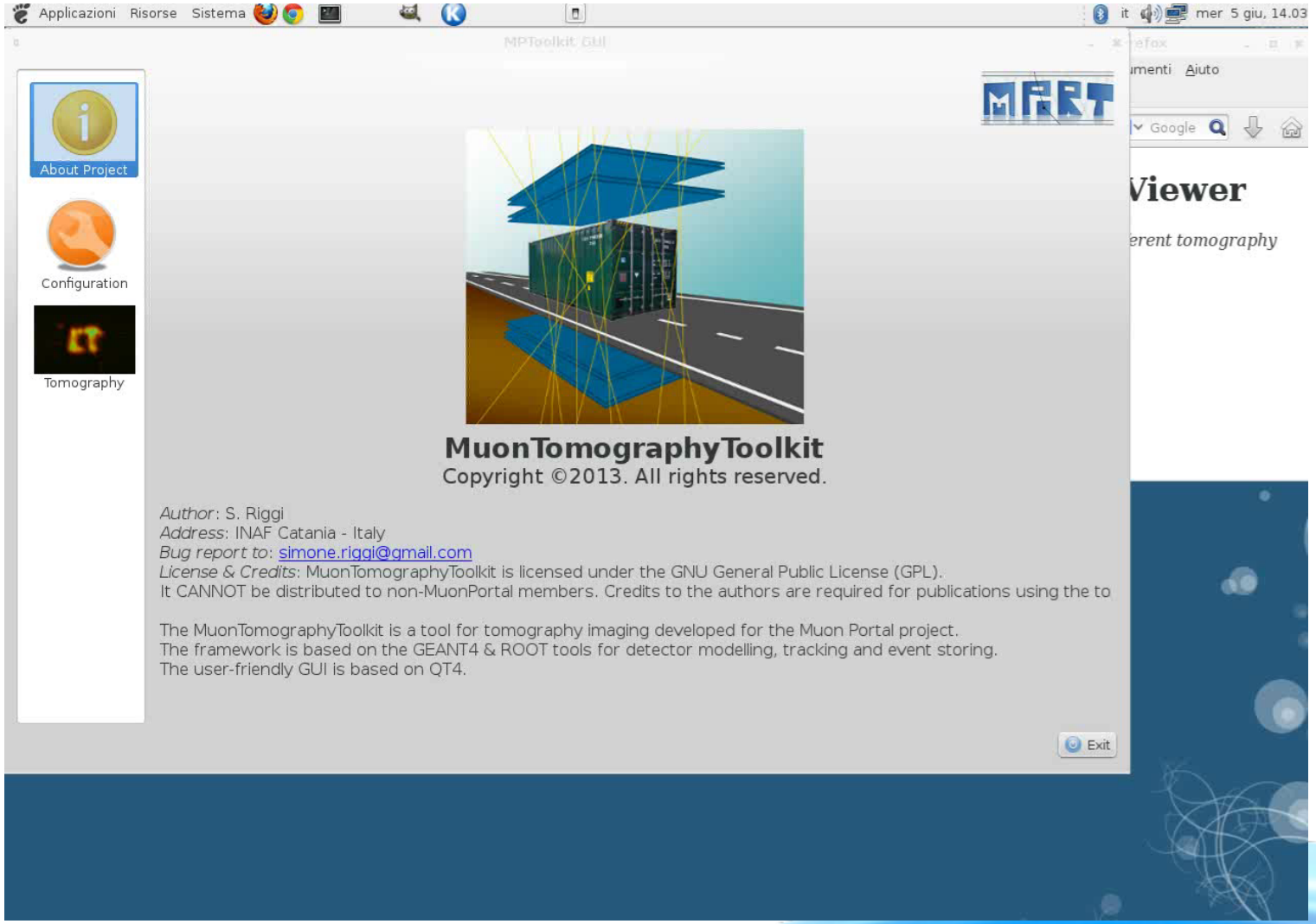


EM-LM Volume Rendering



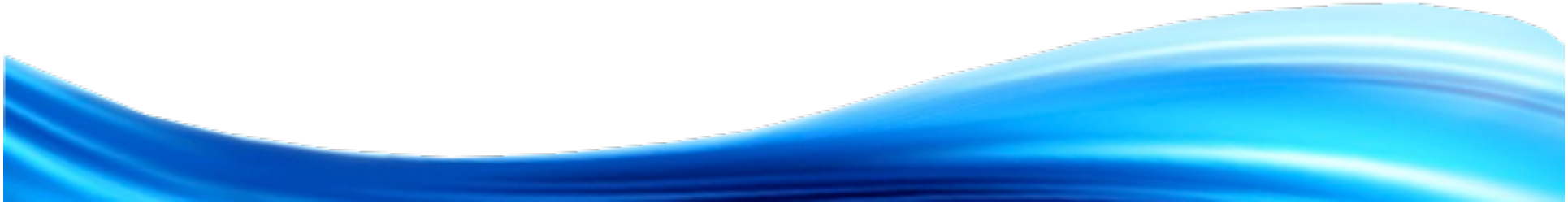
FoF Volume Rendering

Demo


Applicazioni
Risorse
Sistema
mer 5 giu, 14.03
MPToolkit GUI
About Project
Configuration
Tomography
MuonTomographyToolkit
Copyright ©2013. All rights reserved.
Author: S. Riggi
Address: INAF Catania - Italy
Bug report to: simone.riggi@gmail.com
License & Credits: MuonTomographyToolkit is licensed under the GNU General Public License (GPL).
It CANNOT be distributed to non-MuonPortal members. Credits to the authors are required for publications using the to
The MuonTomographyToolkit is a tool for tomography imaging developed for the Muon Portal project.
The framework is based on the GEANT4 & ROOT tools for detector modelling, tracking and event storing.
The user-friendly GUI is based on QT4.
Exit
Aiuto
Google
Viewer

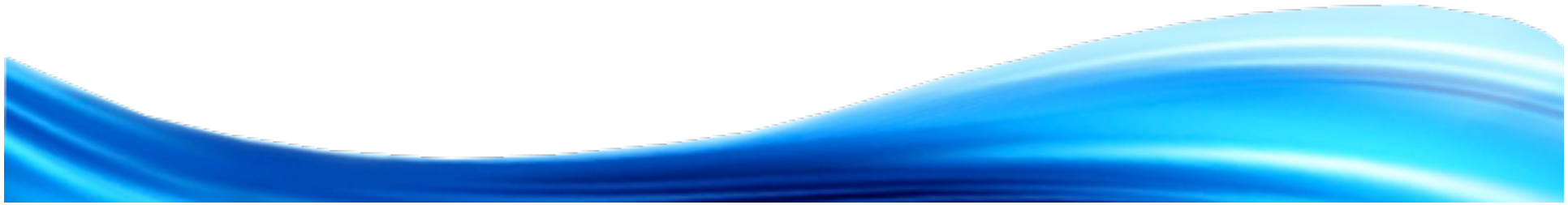
Conclusions and future works

- ✓ **Development of an integrated recostruction algorithm in order to combine the Clustering results with the E-M for better efficiency and accuracy**
- ✓ **Statistical analysis of false positive and false negative variation depending on the threshold and the scanning time**
- ✓ **Use of the particle momentum classification and borders effects study in order to improve algorithm performances**

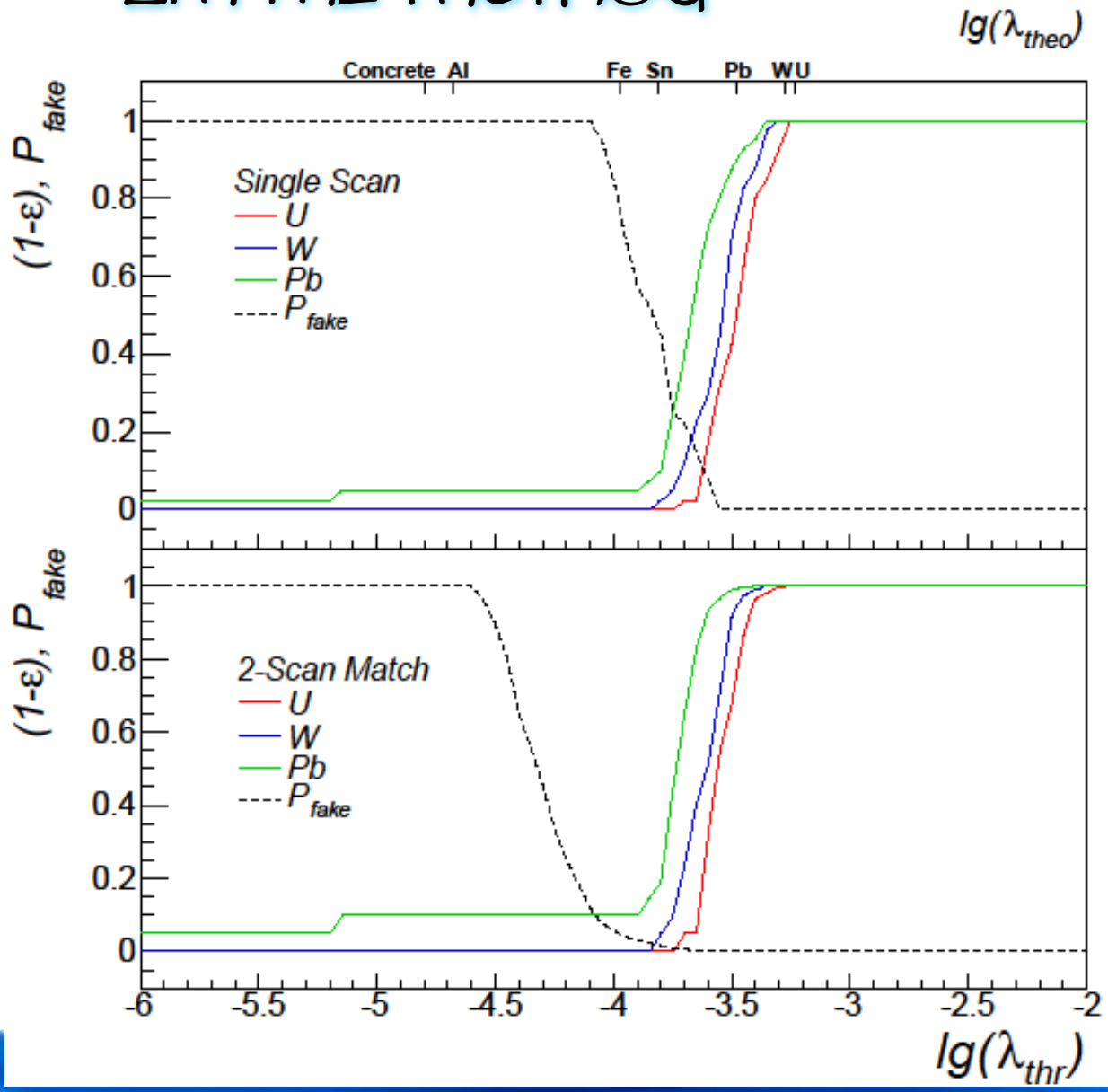


Stay human.

Backup slides



Fake probability and efficiency EM-ML method

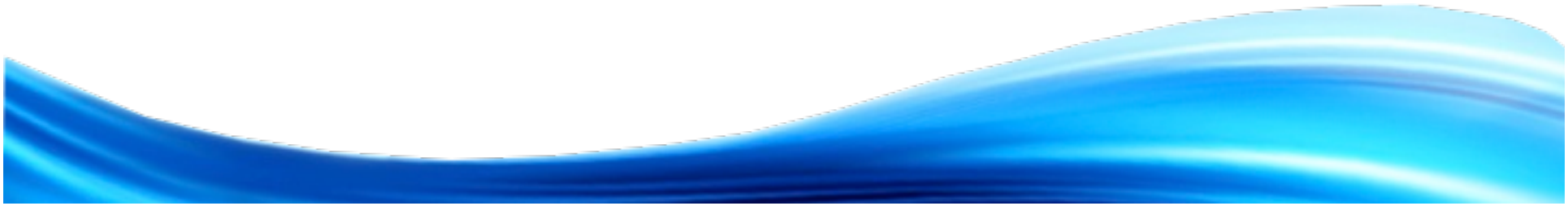


Optics algorithm

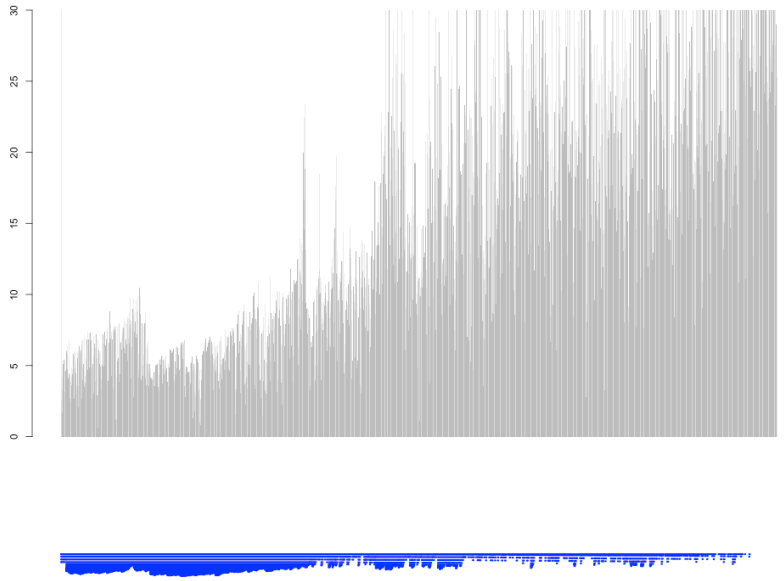
- Its basic idea is similar to DBSCAN, but it addresses one of DBSCAN's major weaknesses: the problem of detecting meaningful clusters in data of varying density.

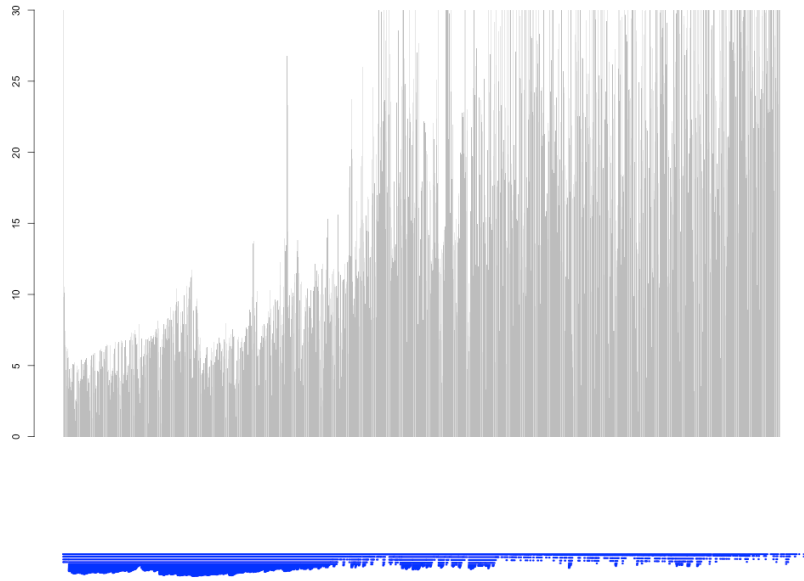
- In order to do so, the points of the database are (linearly) ordered such that points which are spatially closest become neighbors in the ordering.

- Additionally, a special distance is stored for each point that represents the density that needs to be accepted for a cluster in order to have both points belong to the same cluster. This is represented as a dendrogram.



Optics - MuonAtCenterScenario



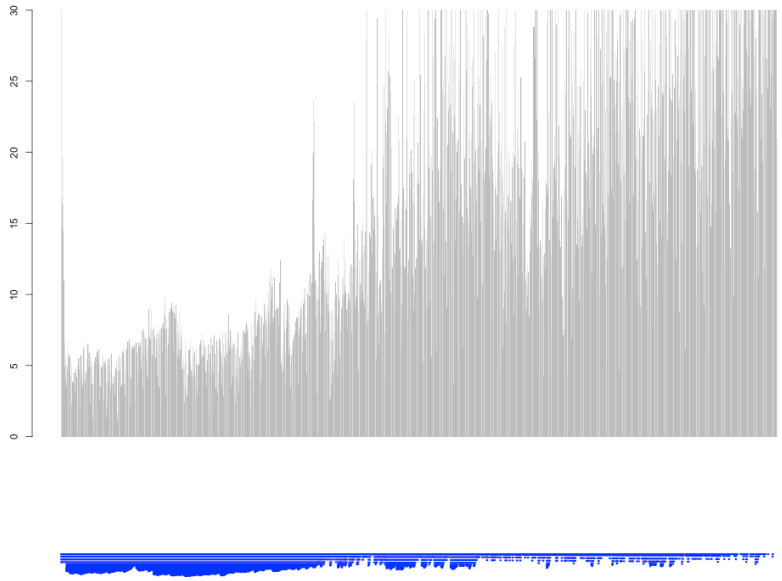


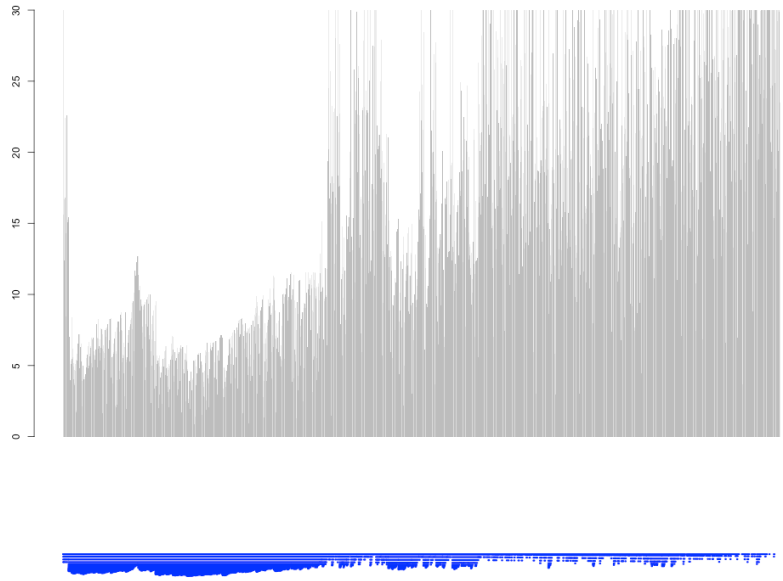
Scan1

Scan2

Cut Theta>2°

Optics - MuonAtCenterScenario



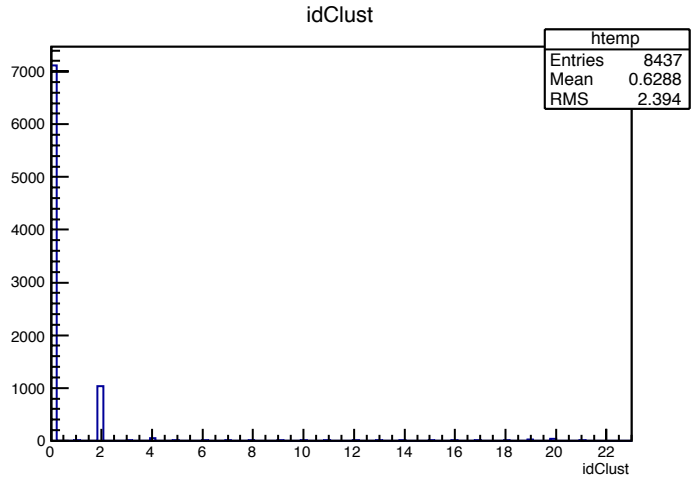


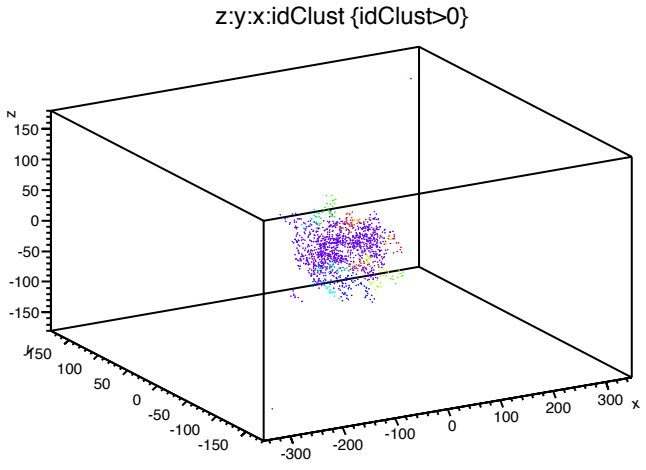
Scan3

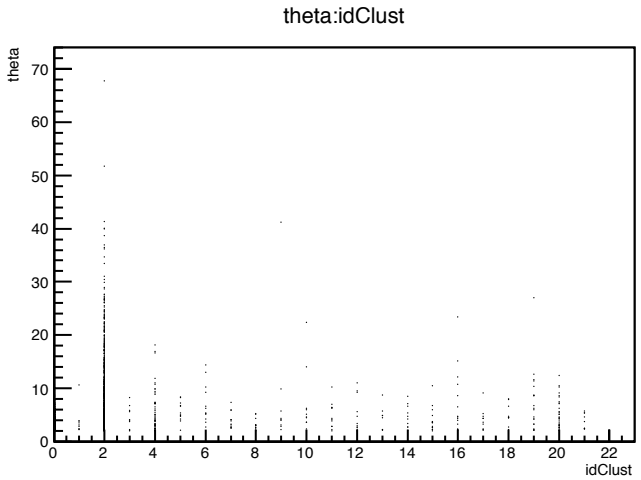
Scan4

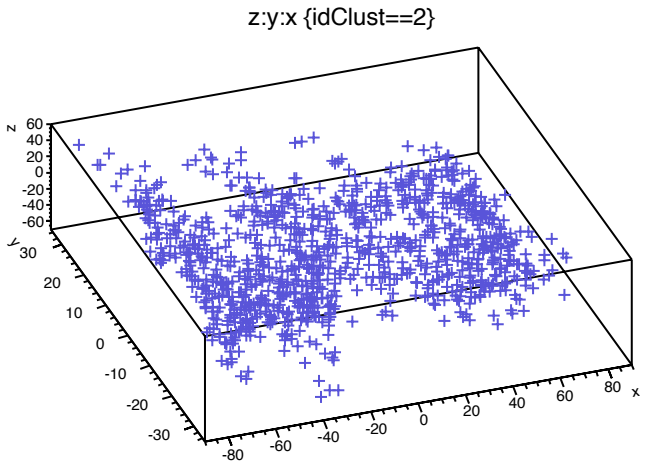
Cut Theta>2°

FOF - MuonAtCenterScenario (scan1)

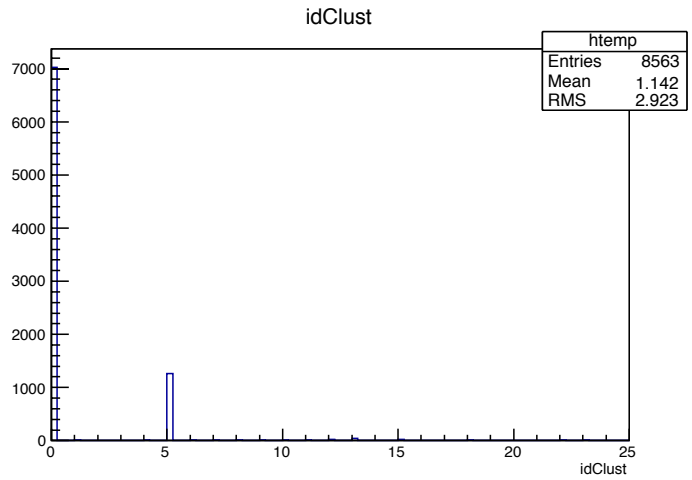


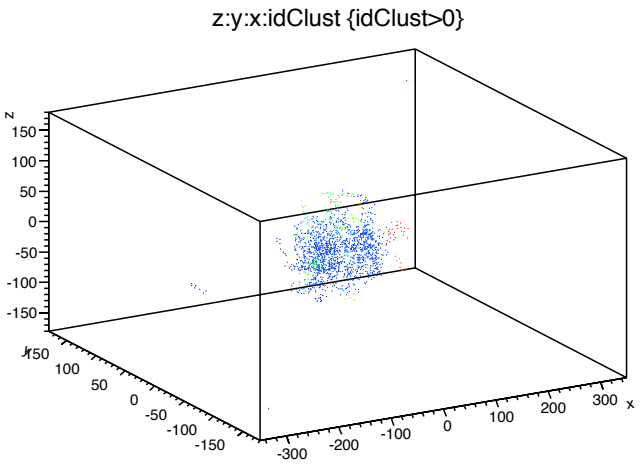


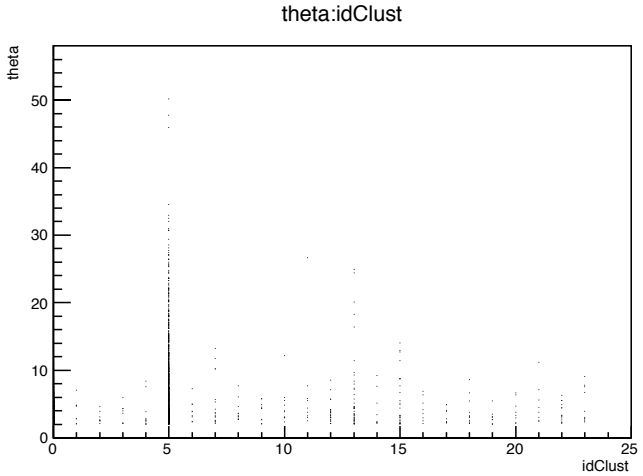


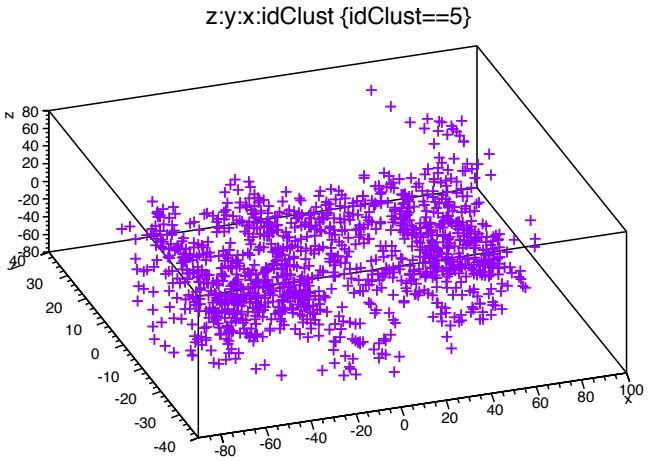


FOF - MuonAtCenterScenario (scan2)

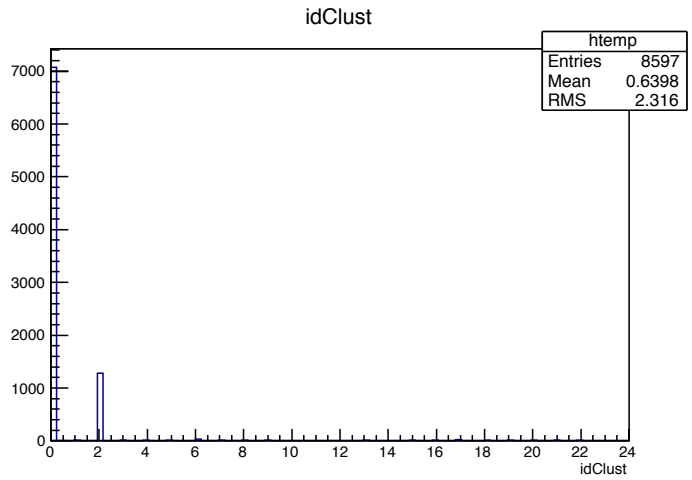


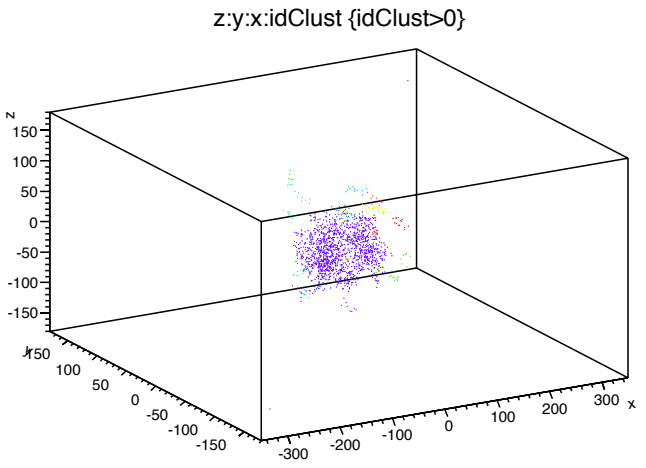


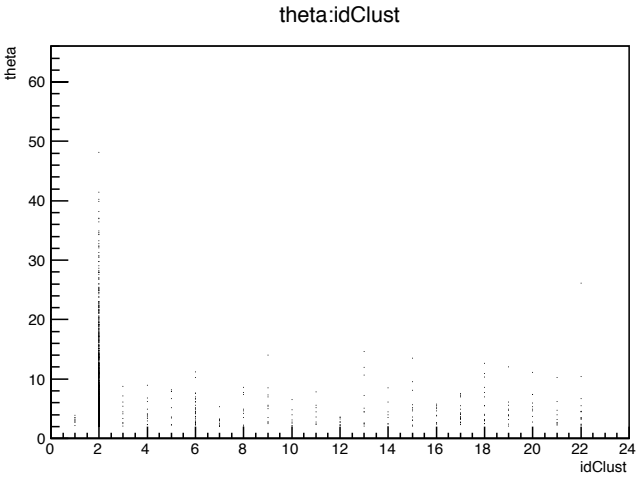


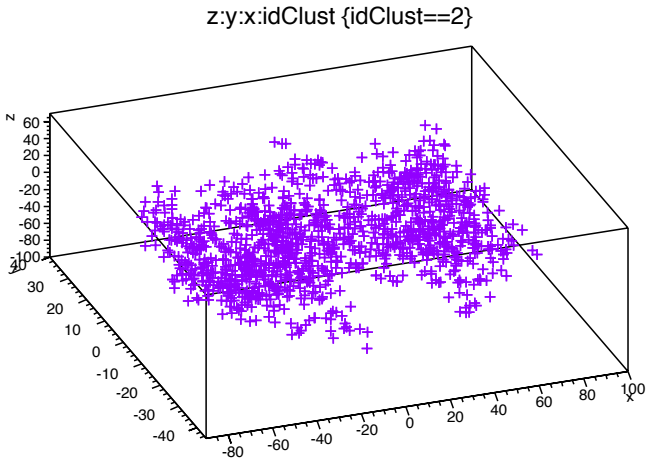


FOF - MuonAtCenterScenario (scan3)

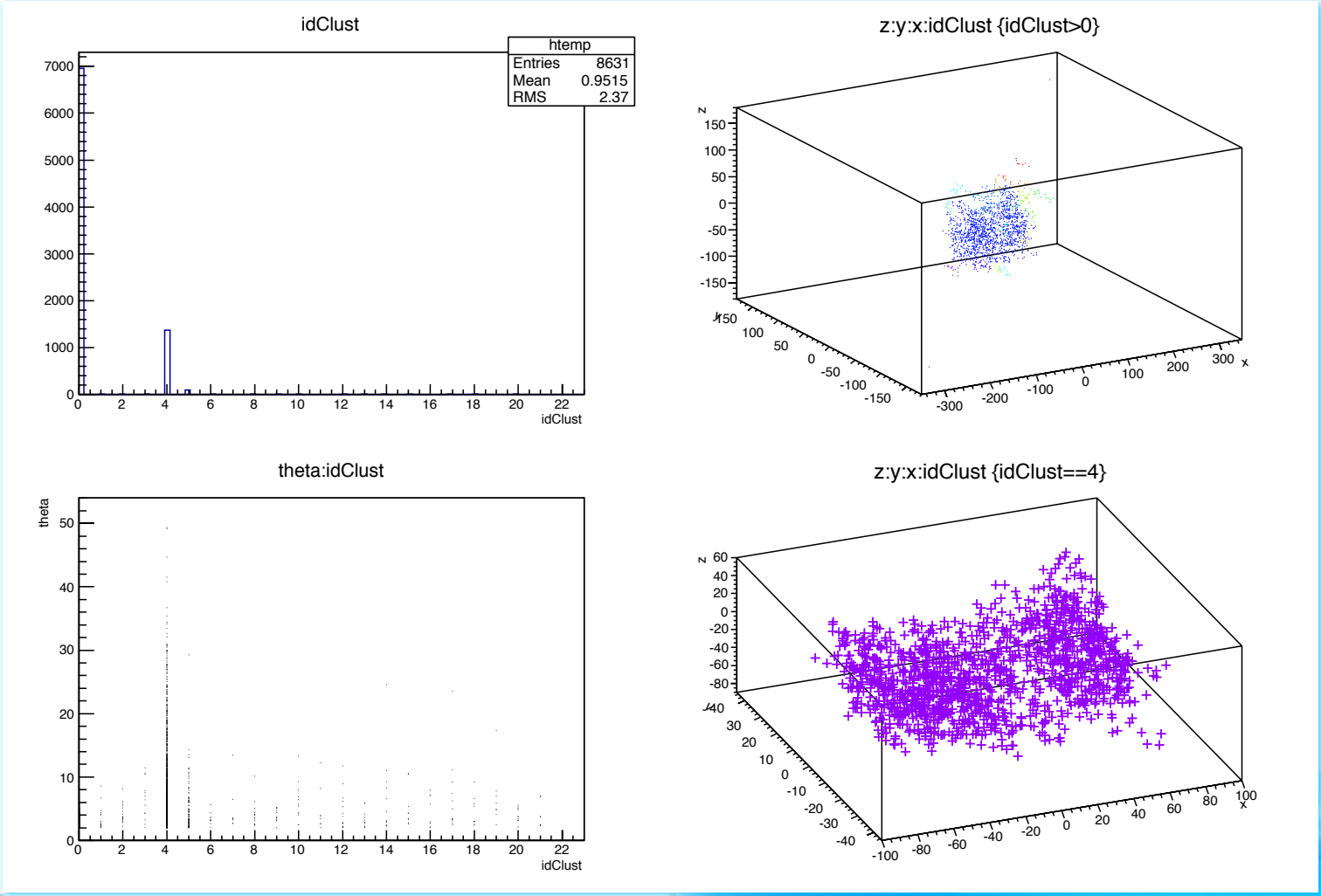




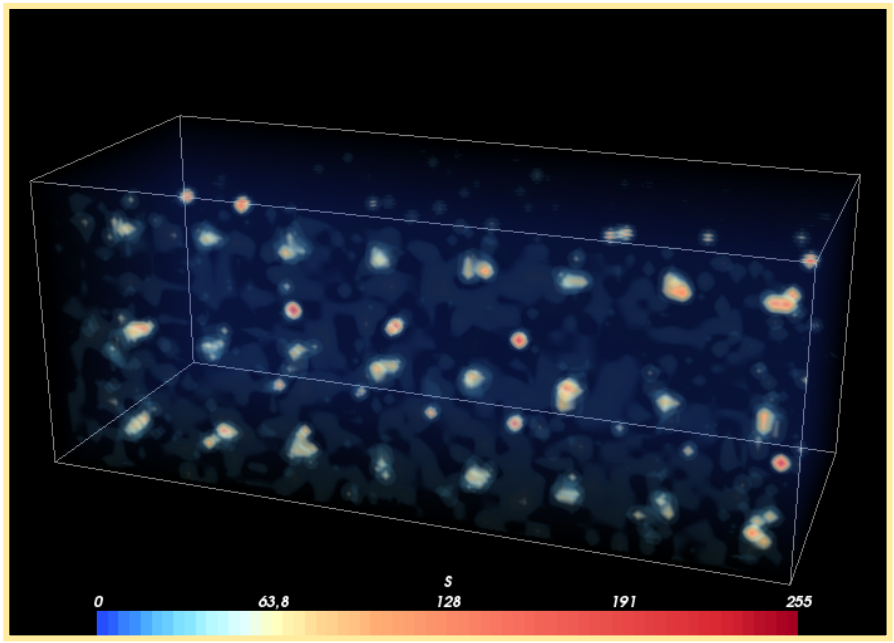




FOF - MuonAtCenterScenario (scan4)

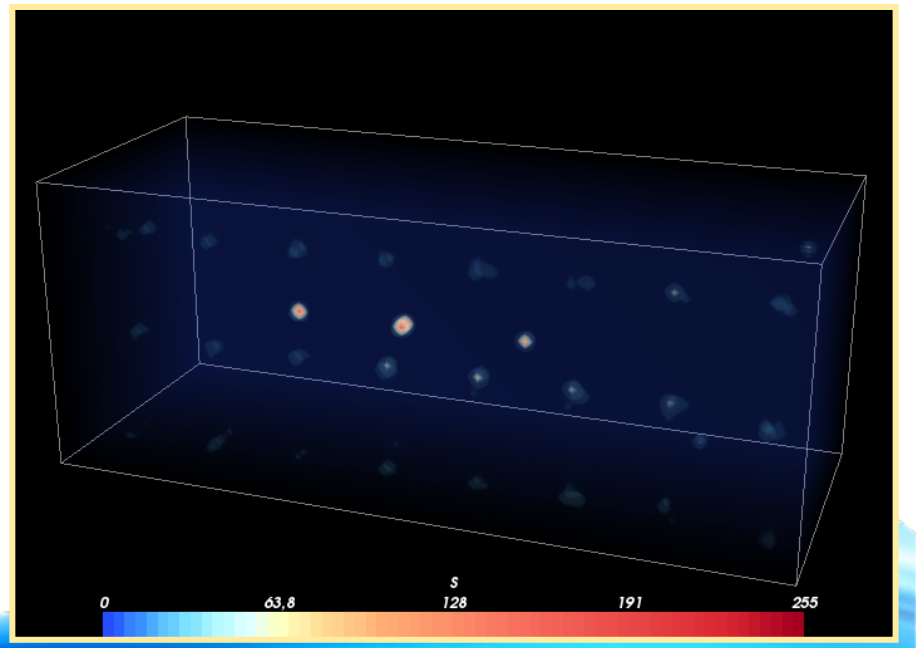


Momentum filter



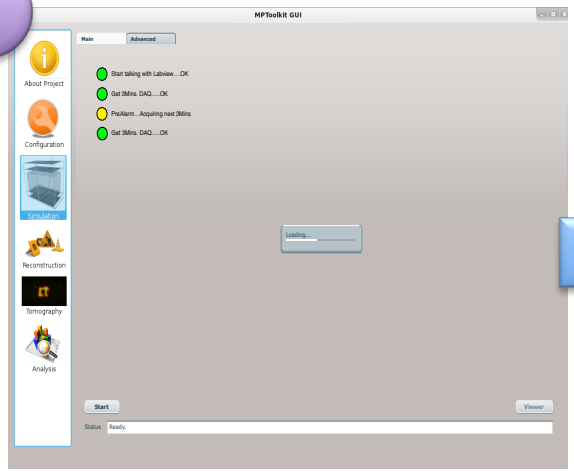
Additional prototype detector
to reject low momentum muons
being tested

After cutting low momenta
(< 1 GeV/c) muons

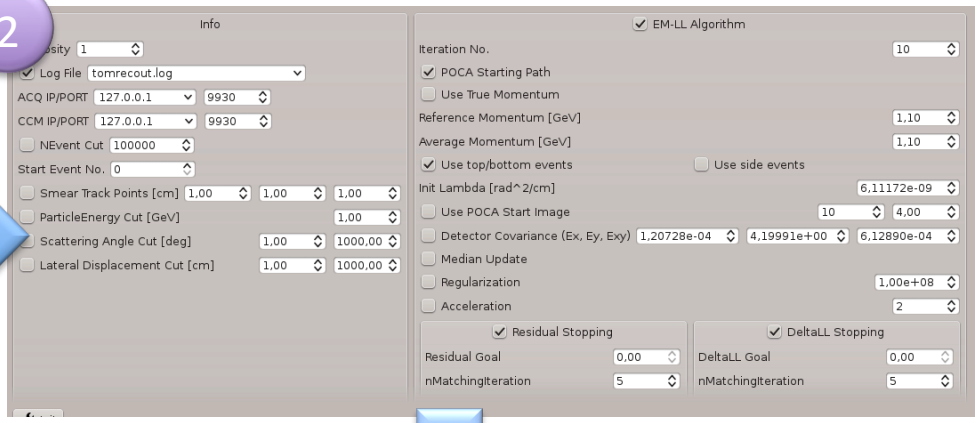


Tomographic images elaboration, visualization and monitoring UI

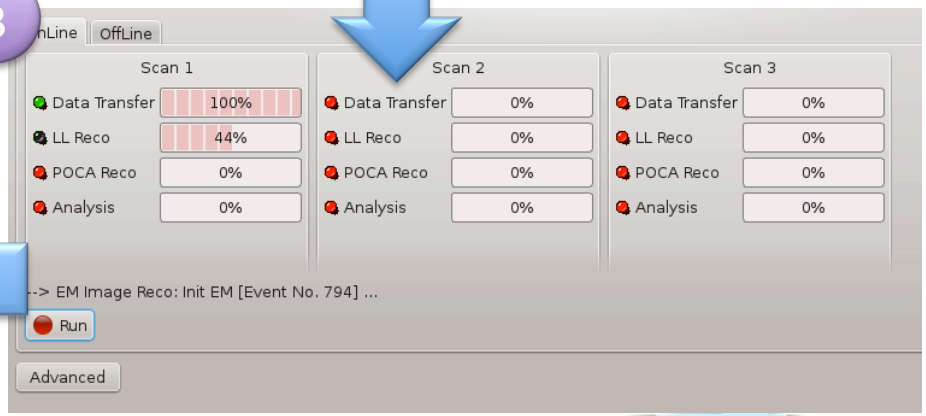
1



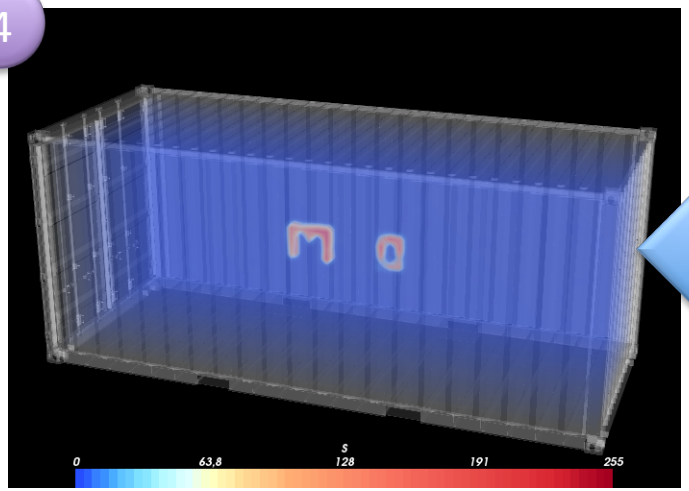
2



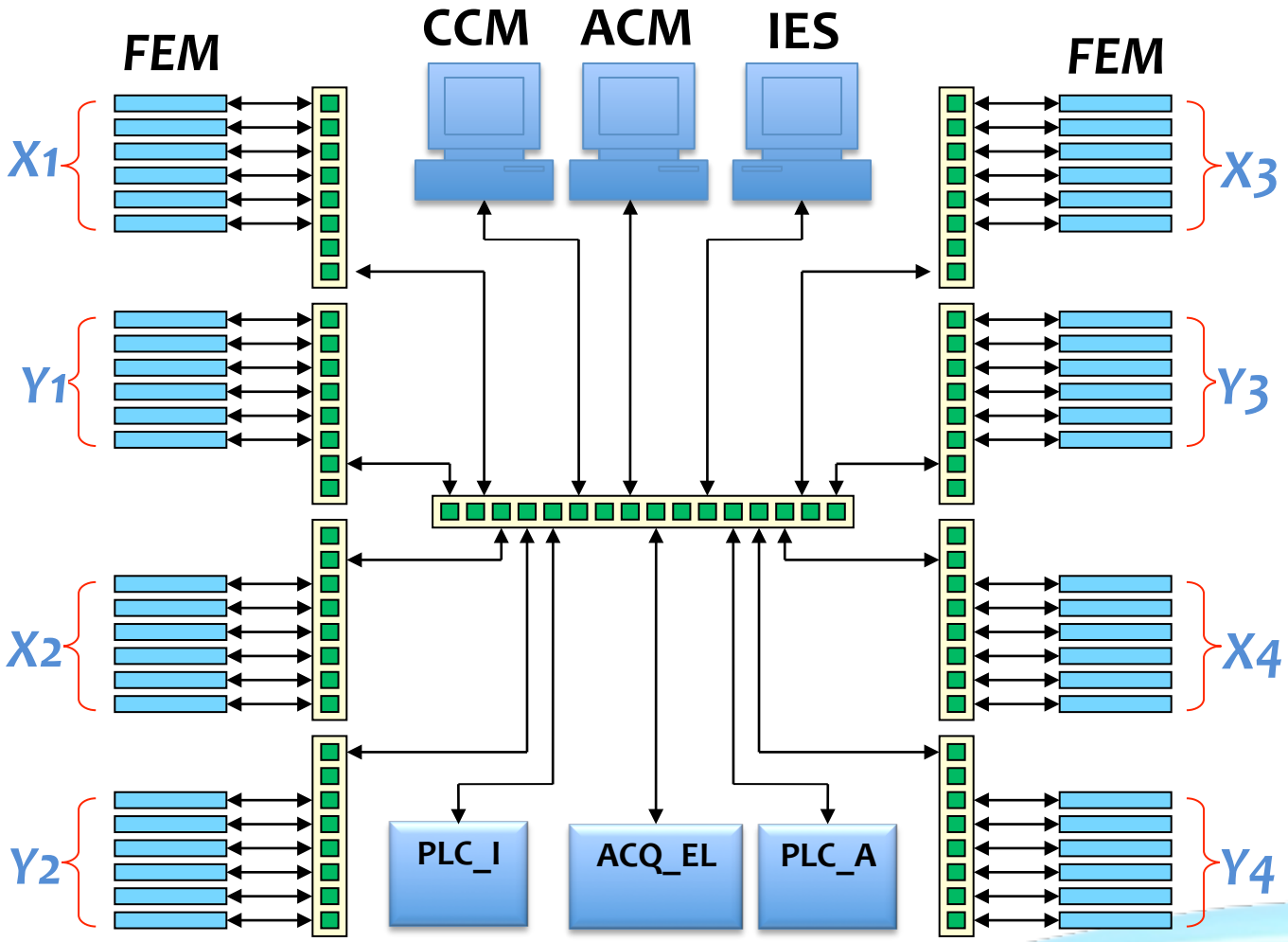
3



4



Data acquisition system



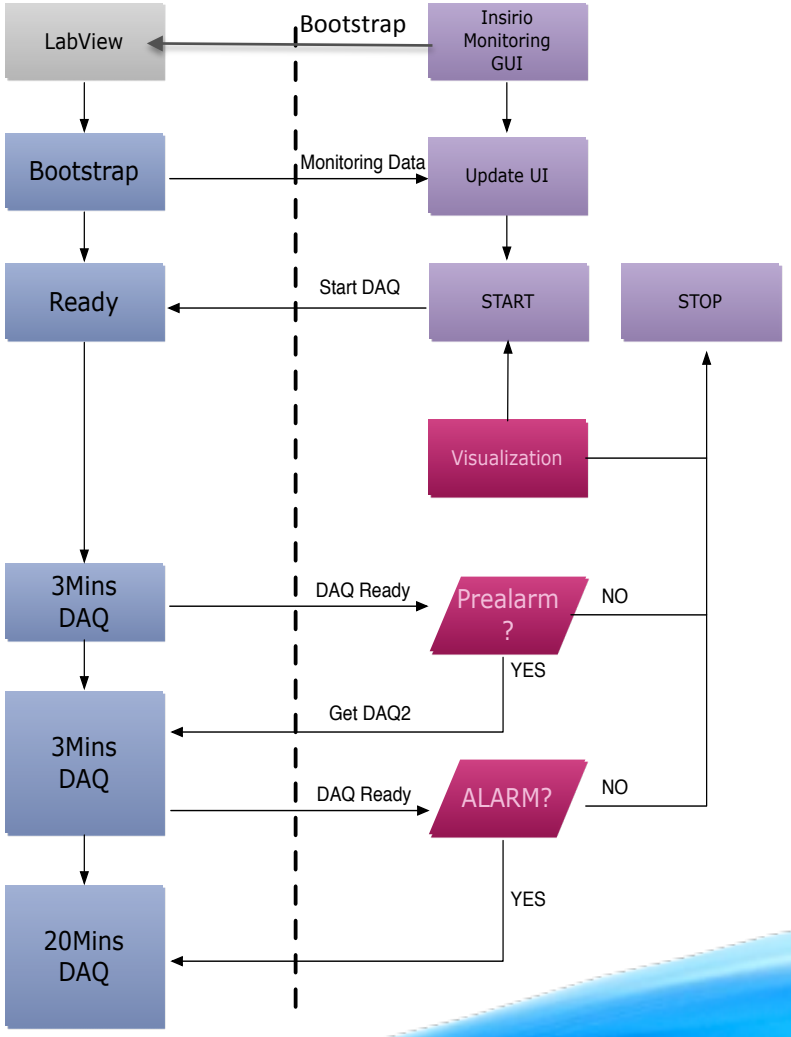
Communication workflow

Rt controller

FPGA board

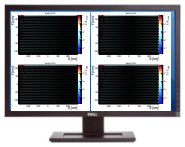
Chassis

DAQ MODULE



CCM – Monitoring

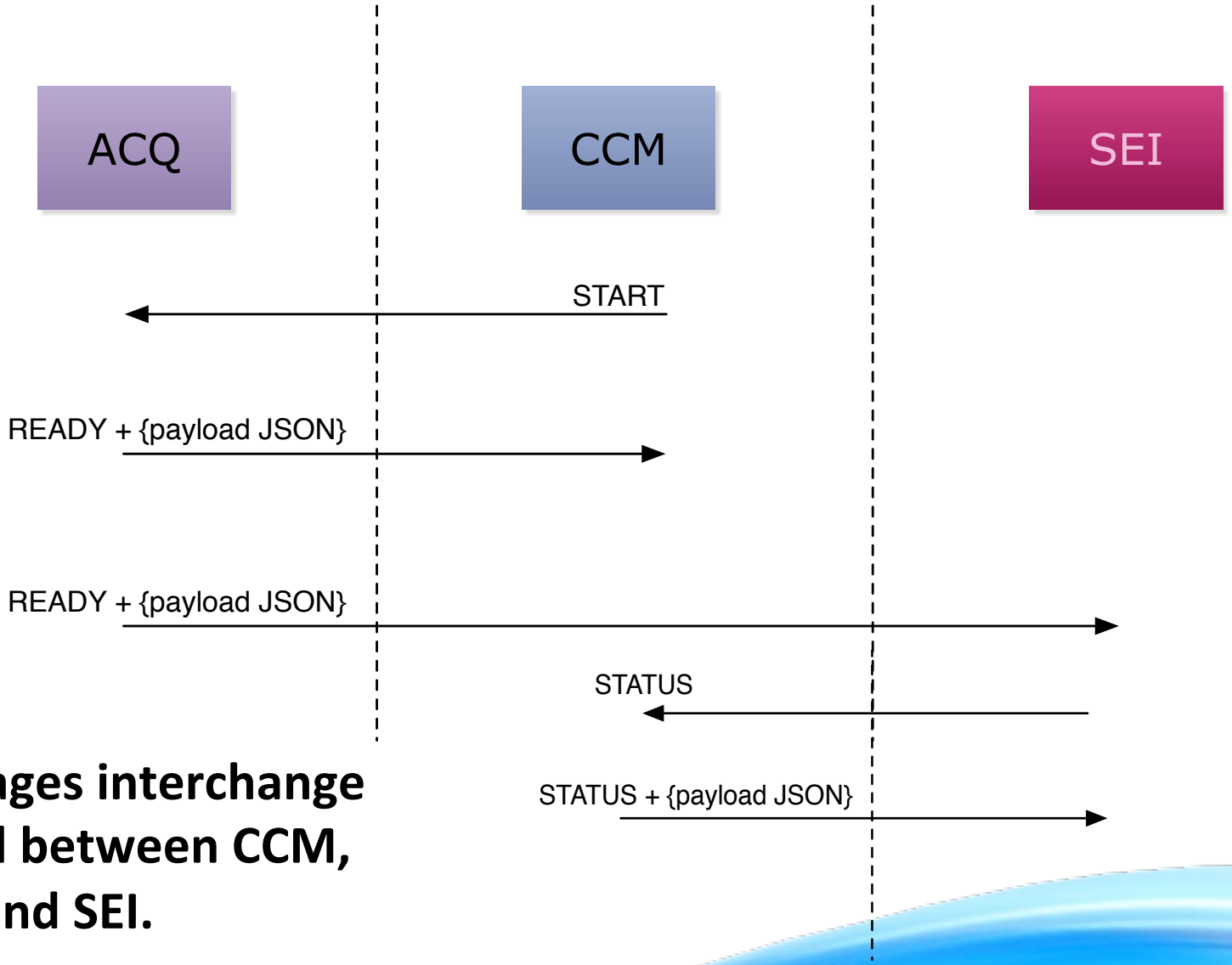
IES – Visualization



ACM - Calibration

Tomographic Images EMV SYSTEM

Interprocess communication over HTTP



Messages interchange model between CCM, ACQ and SEI.

EM-likelihood Algorithm

- Assume the container volume divided into:

  Nx ×Ny ×Nz 3D voxels

  → Each voxel has a scattering density

$$\lambda_j = \left(\frac{15}{p_0}\right)^2 \frac{1}{L_{rad}}, \; j=1,\ldots N_{voxels}, \; L_{rad}=\text{rad. length}$$

$$\rightarrow \mathcal{L}=\frac{1}{2}\sum_i^N (\log|\Sigma_i^{-1}| - y_i^T \Sigma_i^{-1} y_i)$$

- Model the scattering and displacement data yi=(Δθi ,ΔXi) with a bivariate normal distribution:

$$\text{covariance } \Sigma_i = E_i + p_{r,i}^2 \sum_{j=1}^{N_{voxels}} W_{ij}\lambda_j$$

$$\rightarrow \lambda_j^{(k+1)} = \frac{1}{M_j}\sum_i S_{ij}^{(k)} \qquad S_{ij} = 2\lambda_j^{(k)} +$$

$$p_{r,i}^2 [\lambda_j^{(k)}]^2 (y_i^T \Sigma_i^{-1} W_{ij} \Sigma_i^{-1} y_i - \text{Tr}(\Sigma_i^{-1} W_{ij}))$$

Algorithm Steps:

1. Collect data $(\Delta\Theta_{x,y}, \Delta_{x,y})_i$
2. Compute muon path lenght through the voxel (L_{ij}, T_{ij})
3. Compute the weight matrix W_{ij}
4. Inizialize $\lambda j = \lambda j^{(o)}$
5. Iterate until convergence/stopping:
   - Find median correction term $\Delta\lambda j$
   - Update λj using previous iteration estimate
   - Return vector λ

***Use** the EM algorithm to maximize Λ and find an estimate of* λ_j