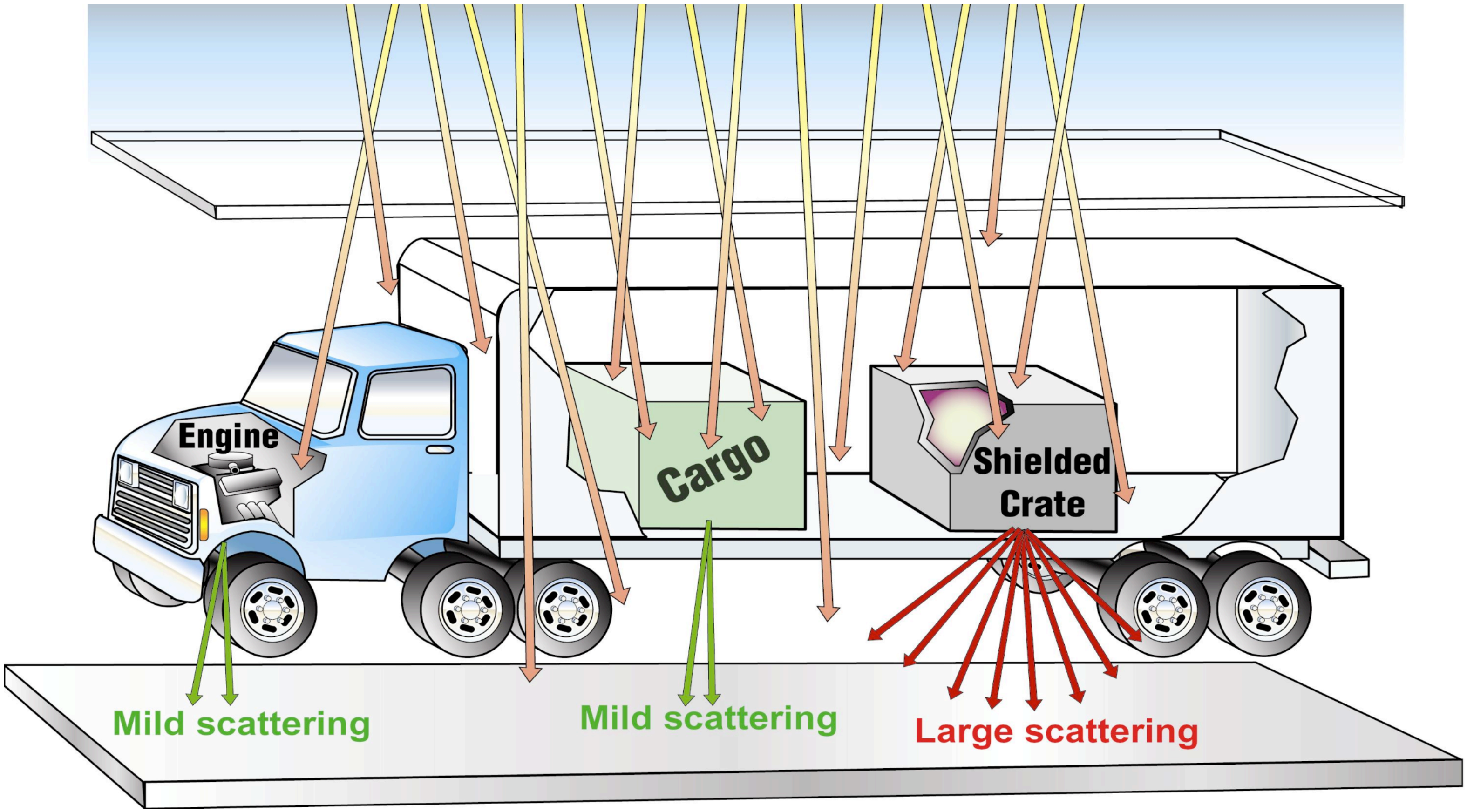
Detecting Nuclear Materials from Cosmic-Ray Muon Scattering Data

Rick Chartrand

Thomas Asaki, Gary Blanpied, Konstantin Borozdin, Mark Galassi, Andrew Green, Nicolas Hengartner, Gary Hogan, Alexei Klimenko, Christopher Morris, William Priedhorsky, Richard Schirato, Larry Schultz, Matthew Sottile, Kevin Vixie, and Brendt Wohlberg

Los Alamos National Laboratory

Muon Radiography Operational Concept

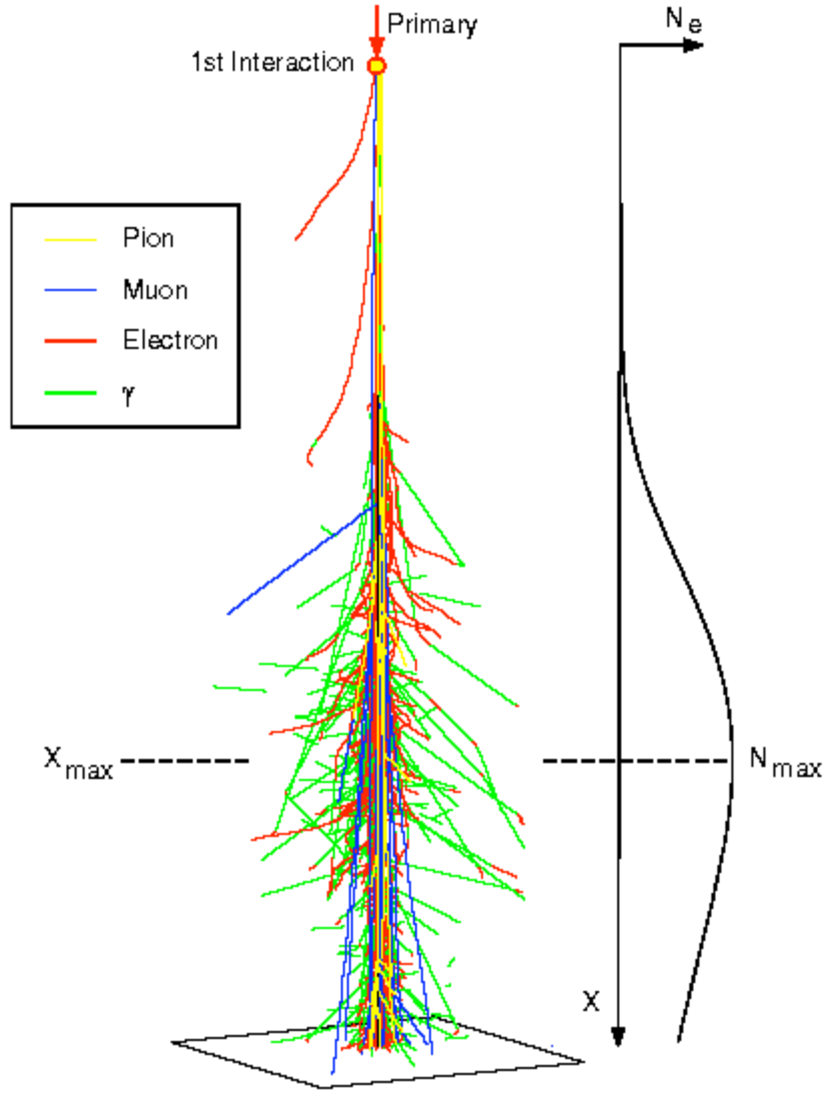


Expected muon scattering increases with increasing atomic number, due to the electric charge of muons.

Los Alamos NATIONAL LABORATORY

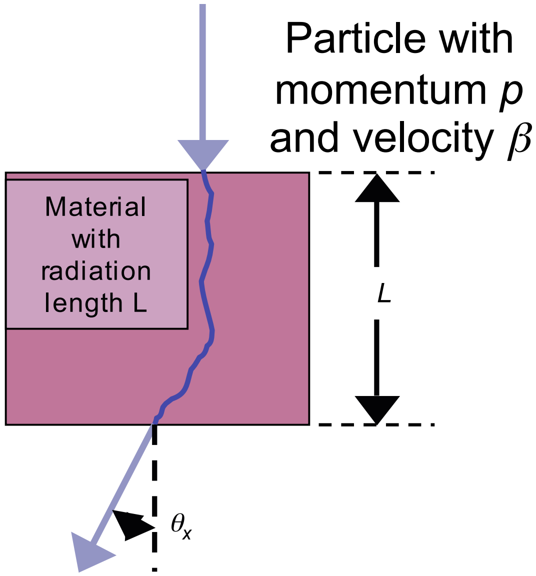
Cosmic-Ray Muons

- Primary cosmic rays strike the atmosphere and generate a cascade of secondary particles.
- Muons are the dominant particle at the Earth's surface.
- Most muons can penetrate tens of meters of rock or more.
- Muons arrive at a rate of 10,000 per square meter per minute (at sea level).
- The spectrum of cosmic ray muons (energy, arrival angle) is well documented. Peak at 3 GeV



Los Alamos
NATIONAL LABORATORY

Physics of interaction: multiple scattering



Scattering distribution is approximately Gaussian

$$\frac{dN}{d\theta_x} = \frac{1}{\sqrt{2\pi}\theta_0} e^{-\frac{\theta_x^2}{2\theta_0^2}}$$

and the width of the distribution depends on material and muon properties (λ is a radiation length)

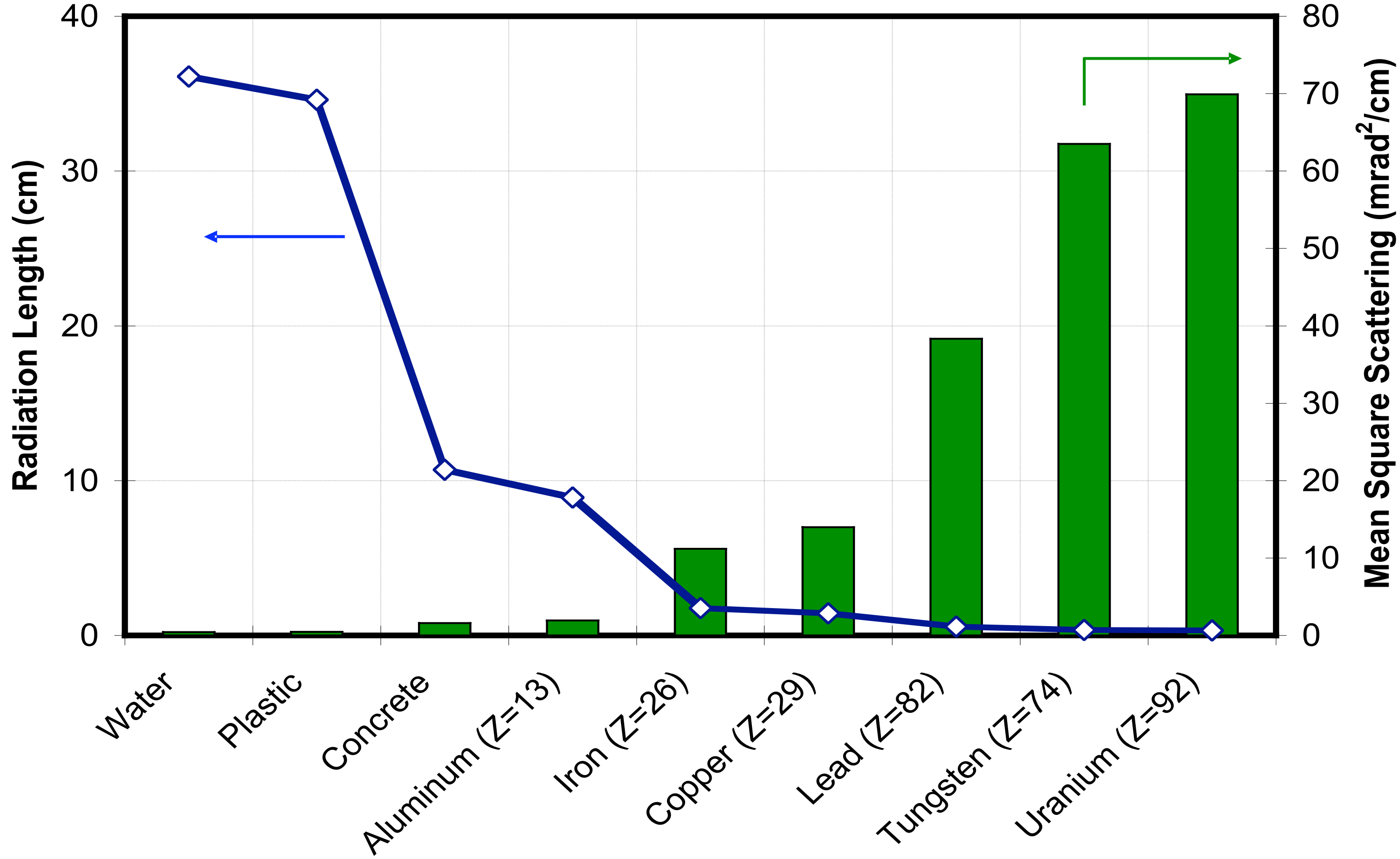
$$\theta_0 = \frac{13.6}{p\beta}\sqrt{\frac{L}{\lambda}} + H.O.T.$$

Scattered particles carry information from which material may be identified.

Material	λ , cm	θ_0, mrad*
Water	36	2.3
Iron	1.76	11.1
Lead	0.56	20.1
*10 cm of material, 3 Gev muons		

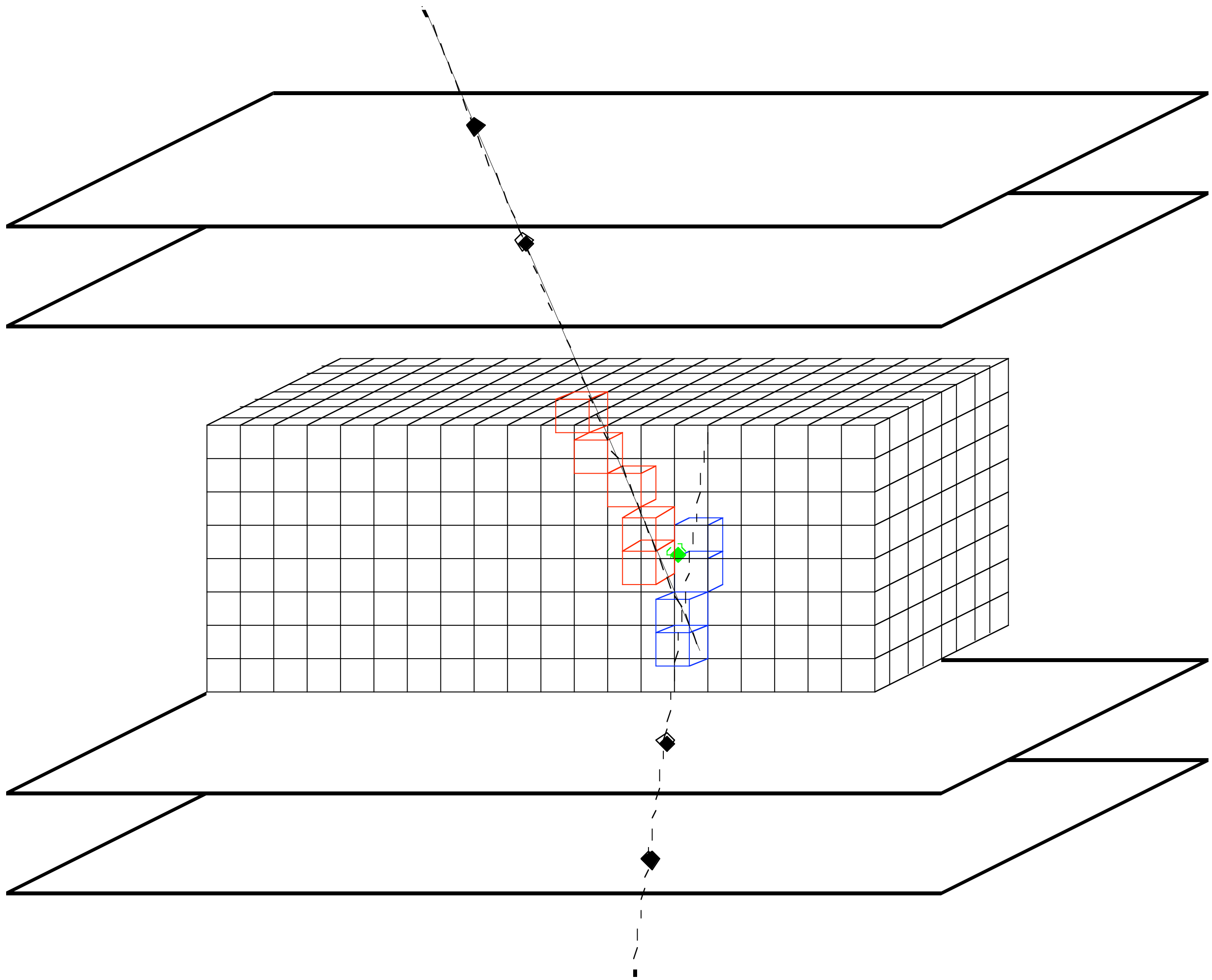
Los Alamos NATIONAL LABORATORY

Radiation Lengths and Mean Square Scattering for Example Materials



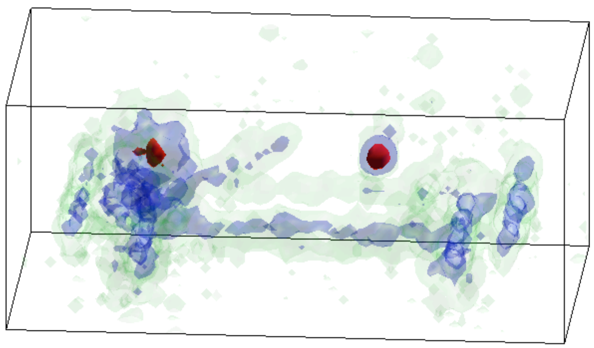
Dense, high-Z materials scatter muons more than other materials.

Voxels, Rays, and Point of Closest Approach



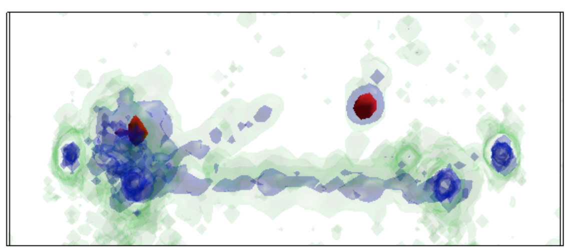
Maximum-Likelihood Reconstruction

Automobile, 60 seconds of muon exposure

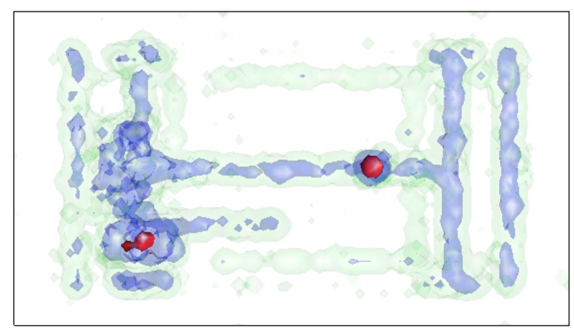


Two 20-kg U spheres

Side View



Top View



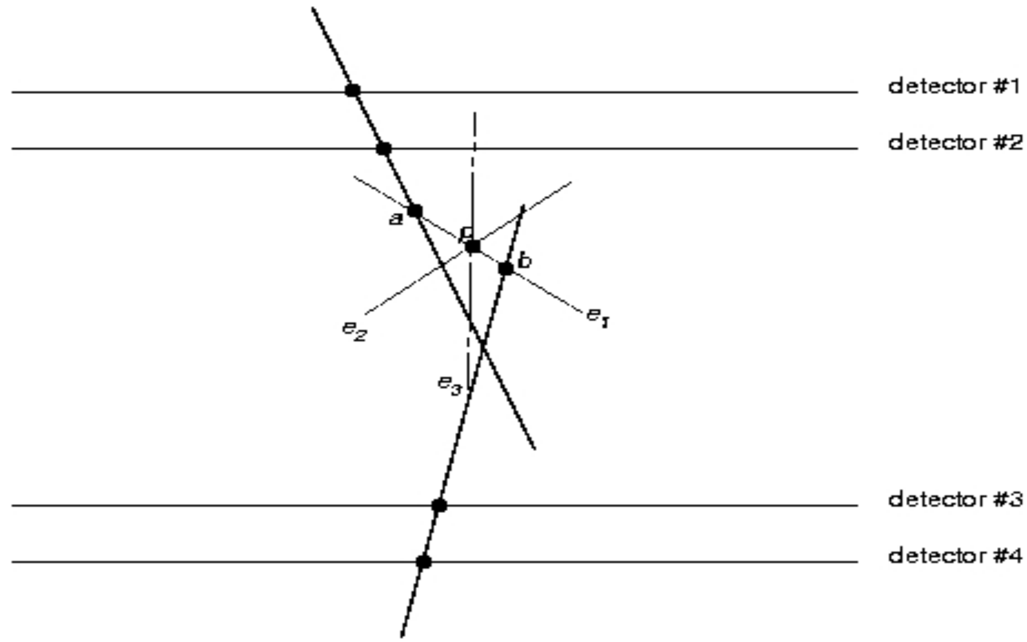
Los Alamos
NATIONAL LABORATORY

Voxel Histograms

- Several detection approaches begin with a distillation of the data into 3-D histograms of voxel scattering statistics.
- Each muon's scattering is measured, normalized by an estimate of the muon's energy, and then assigned to one or more voxel bins.
- This assignment can be done in many different ways:
  - Only to the voxel containing the PoCA:
    - over-localized, as the PoCA is a coarse estimate of scattering location
  - To each voxel containing the muon's ray tracks:
    - under-localized; doesn't take into account location information from PoCA
  - Weighted among voxels according to a measure of PoCA uncertainty

Muon Metric

- A distance function is constructed to quantify and orient uncertainty in the PoCA.
- Orthonormal coordinate vectors for each muon:
  - e_1: orthogonal to both paths
  - e_2: in direction of the deflection
  - e_3: follows tracks most closely
- Choice of weights quantifies the uncertainty in appropriate directions.

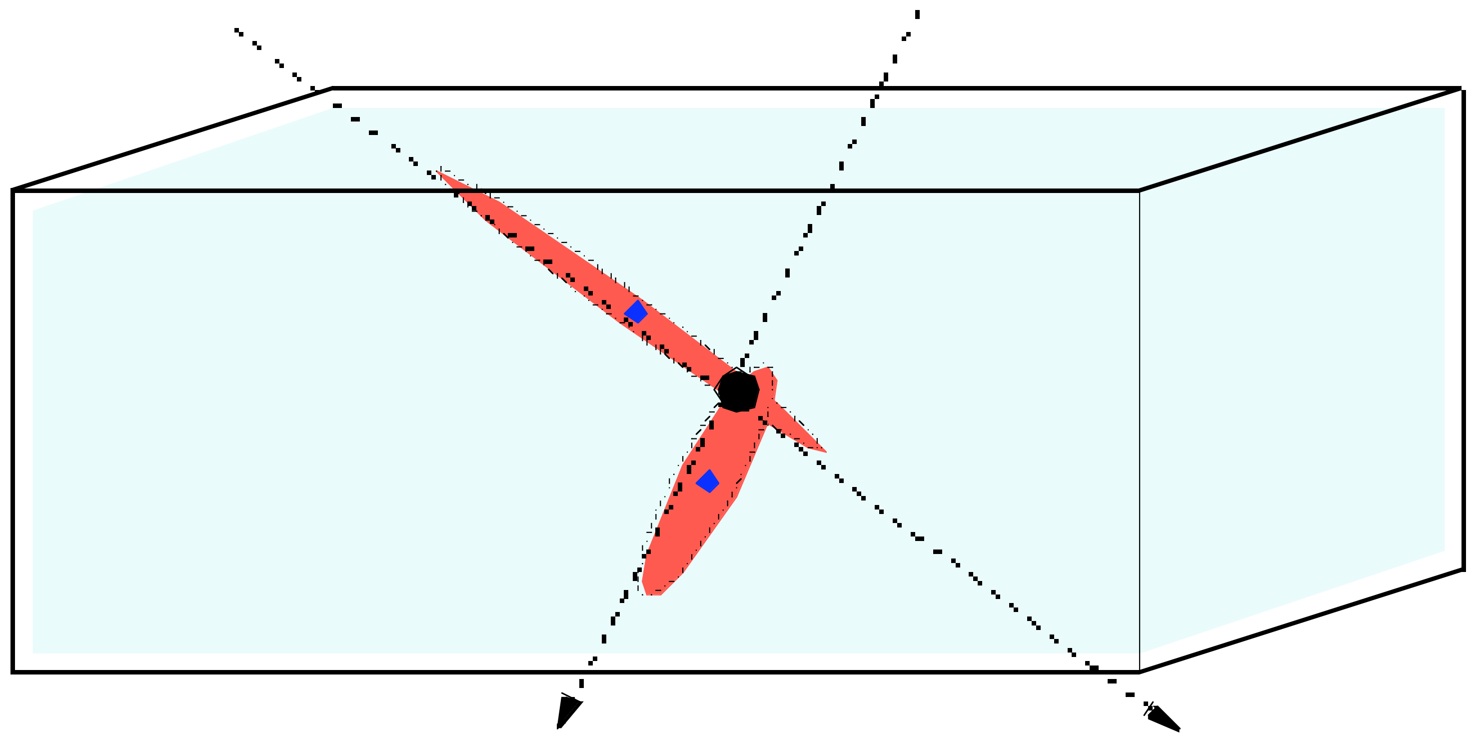


$$d(x) = \sqrt{\sum_{i=1}^{3} \left(\alpha_i (x - p) \cdot e_i\right)^2}$$

$$\alpha_1 = \alpha_2 \propto \left(|a - b| \cdot \Delta\theta\right)^{-1},$$

$$\alpha_3 \propto |a - b|^{-1}$$

Muon Metric Neighborhoods



Muons passing through a threat object in a cluttered container.

PoCAs in “wrong” place, but metric neighborhoods overlap at object.

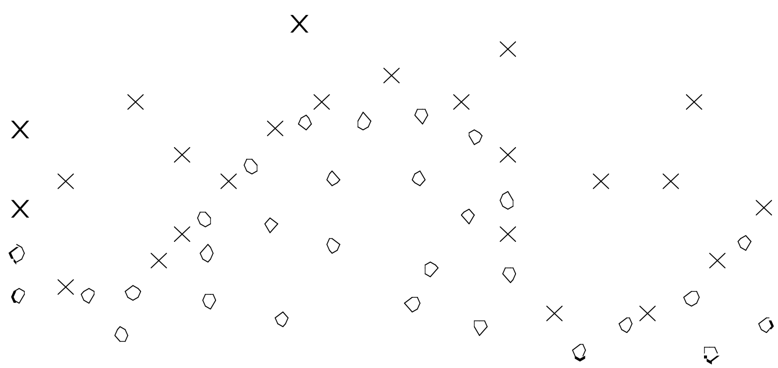
Classification

- Interpreting images can be difficult, and requires human involvement.
- Instead of using the data to create images, we sought to determine directly from the data whether a threat object is present.
- As an alternative to trying to determine how to get an answer from the voxel histograms, we used *machine learning* techniques.
- A computer algorithm is “trained” using many examples of datasets where the answer is known and provided to the algorithm.
- The algorithm “learns” how to discern the answer from the data.

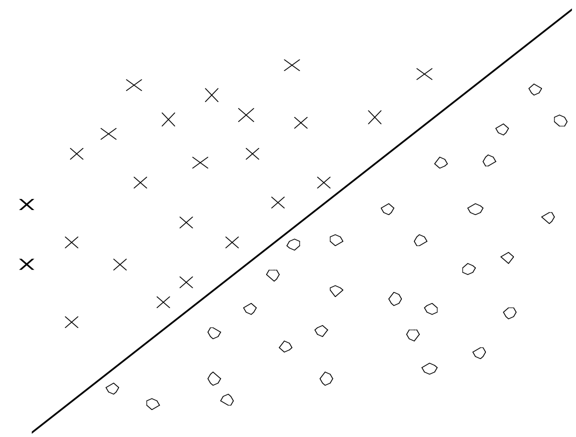
Machine Learning Implementation

- The classifier is trained using muon-scattering data generated by computer simulations of cargo containers or vehicles containing a variety of materials, some with high-Z material and some without.
- The features for each sample are the mean squared scattering (angle times energy) for each 10-cm voxel in a 3-voxel cube surrounding the sample location.
- We chose the Support-Vector Machine method for the classifier, as it is easy to implement and computationally efficient.

Support-Vector Machines



Thousands of points in 27-dimensional space



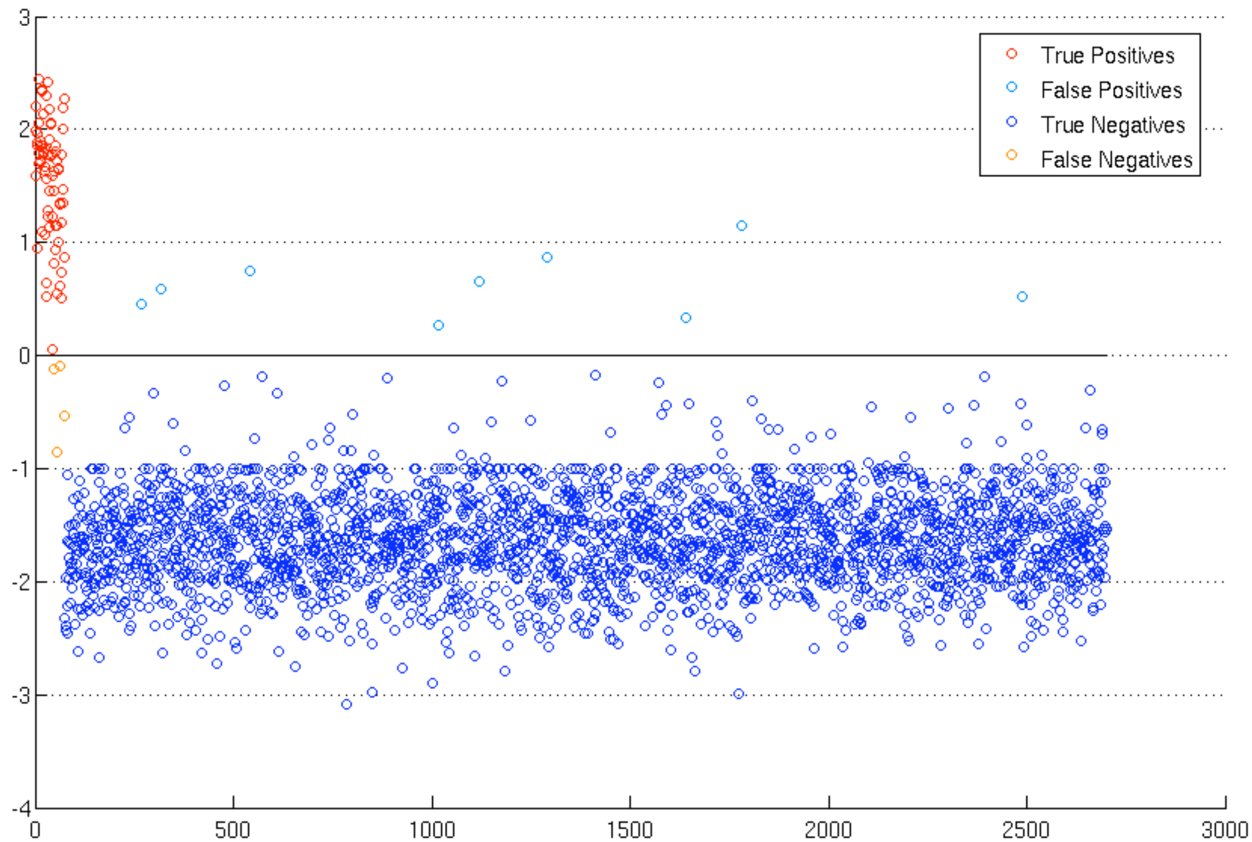
Thousands of points in infinite-dimensional space

- Linear classification is made possible by nonlinearly embedding the data into an infinite-dimensional function space.
- Given a kernel $k(x,y)$, the embedding maps y to $x \mapsto k(x,y)$.
- For $k(x,y) = e^{-\Gamma\|x-y\|^2}$, the data are mapped to a simplex; any two subsets can be separated by a hyperplane.

Los Alamos NATIONAL LABORATORY

Results---Test Data

15 seconds of muon exposure

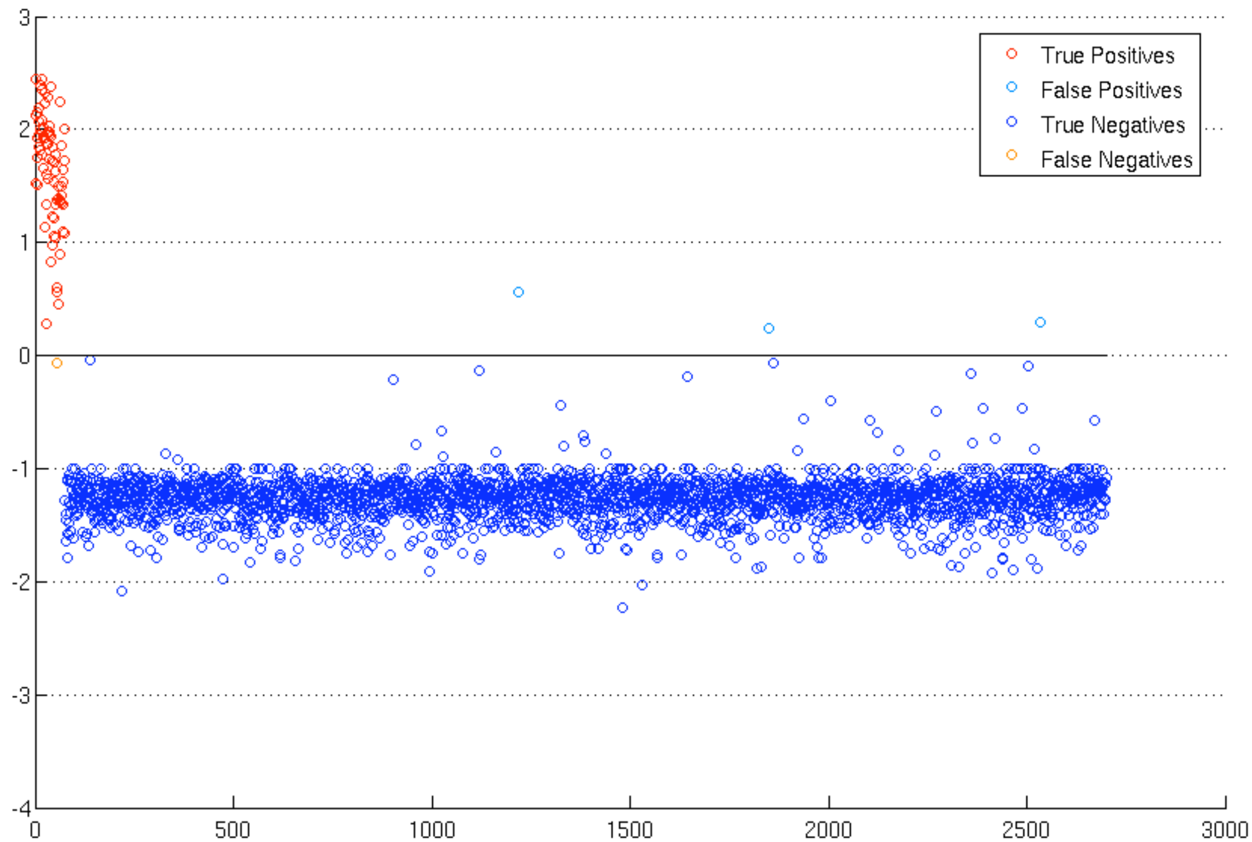


Positive samples have 20-kg U sphere centered on the central voxel.

Los Alamos
NATIONAL LABORATORY

Results---Test Data

30 seconds of muon exposure

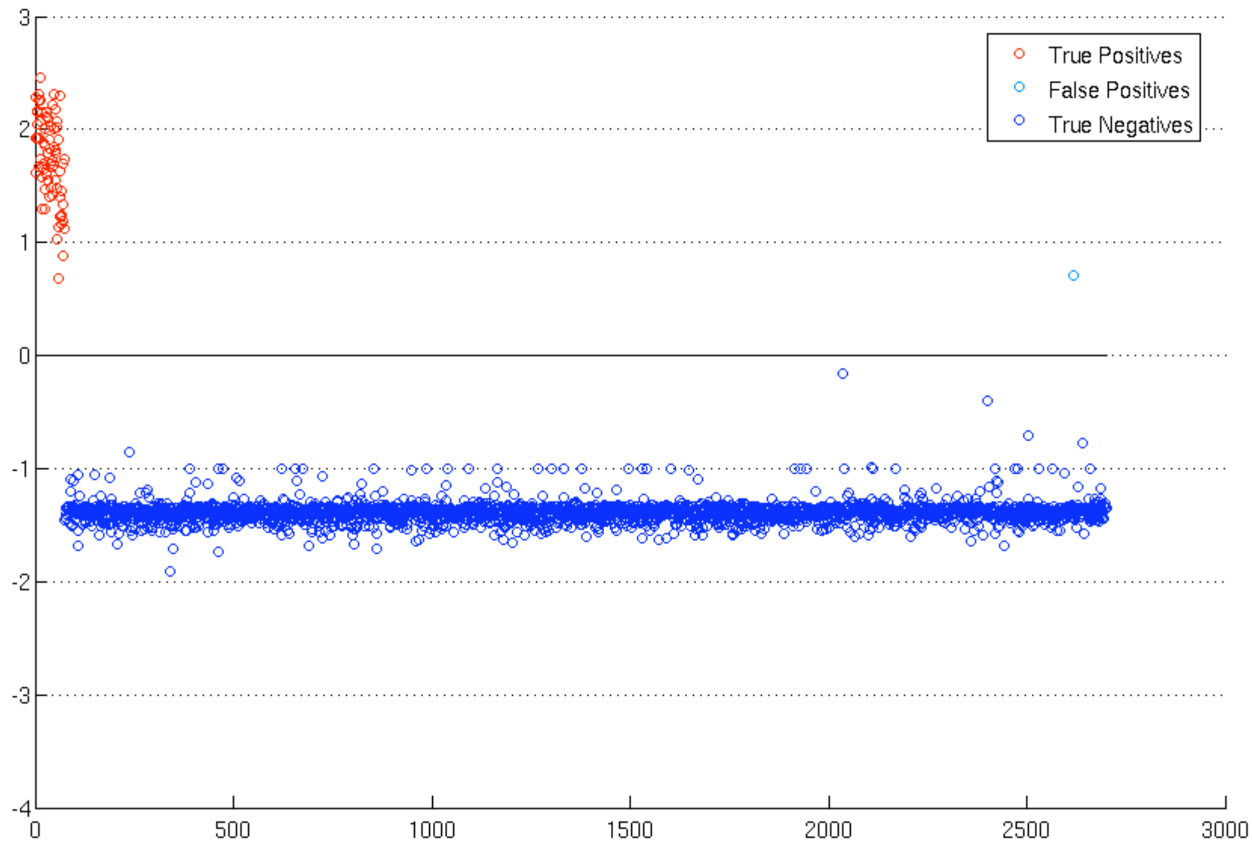


Positive samples have 20-kg U sphere centered on the central voxel.

Los Alamos
NATIONAL LABORATORY

Results---Test Data

45 seconds of muon exposure

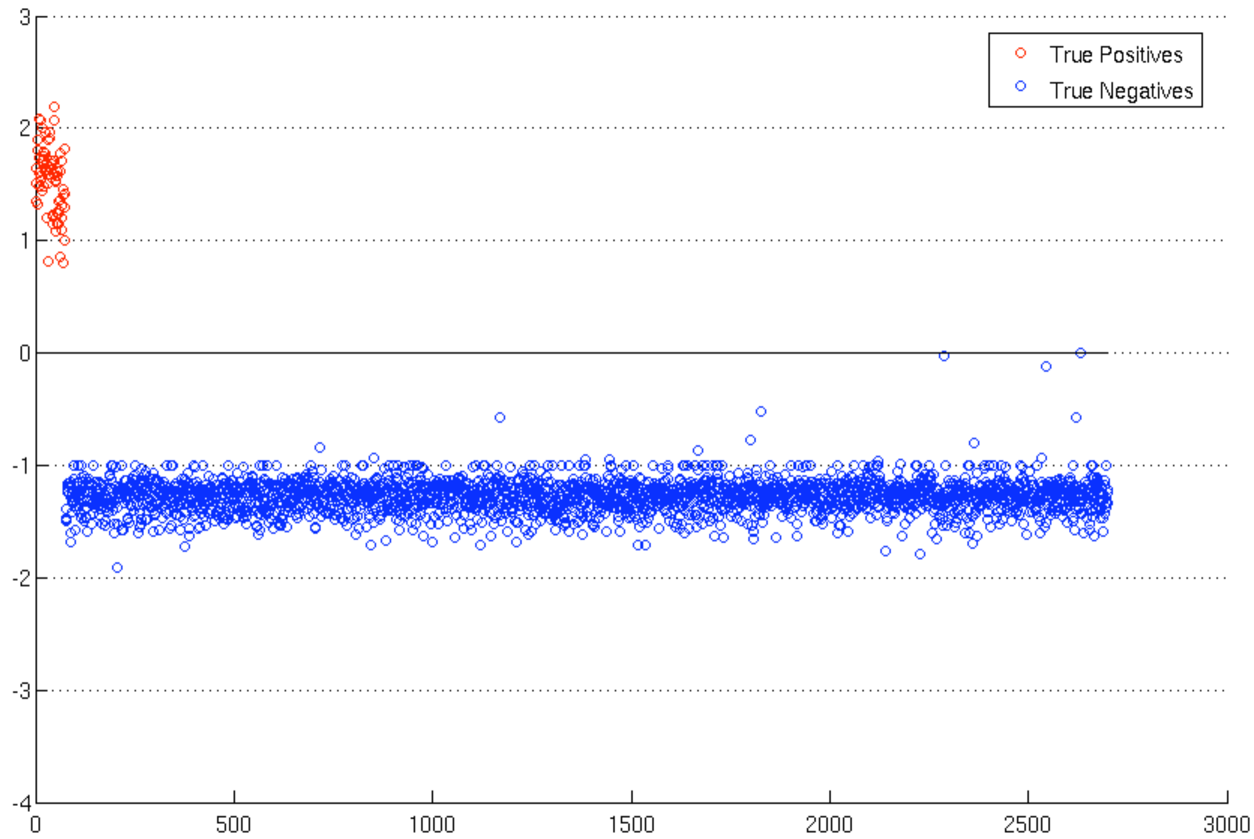


Positive samples have 20-kg U sphere centered on the central voxel.

Los Alamos NATIONAL LABORATORY

Results---Test Data

60 seconds of muon exposure



Positive samples have 20-kg U sphere centered on the central voxel.

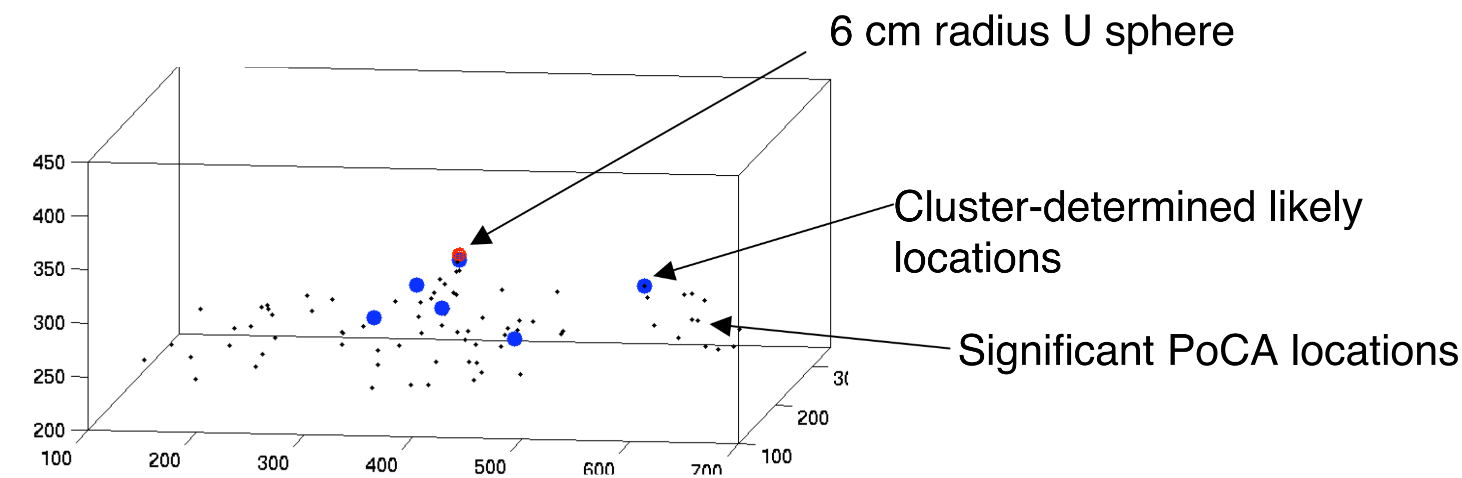
Los Alamos
NATIONAL LABORATORY

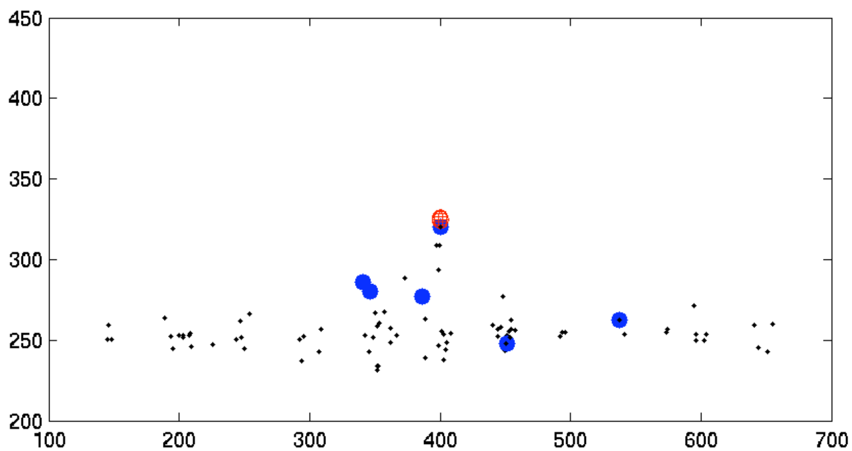
Clustering for Location Selection

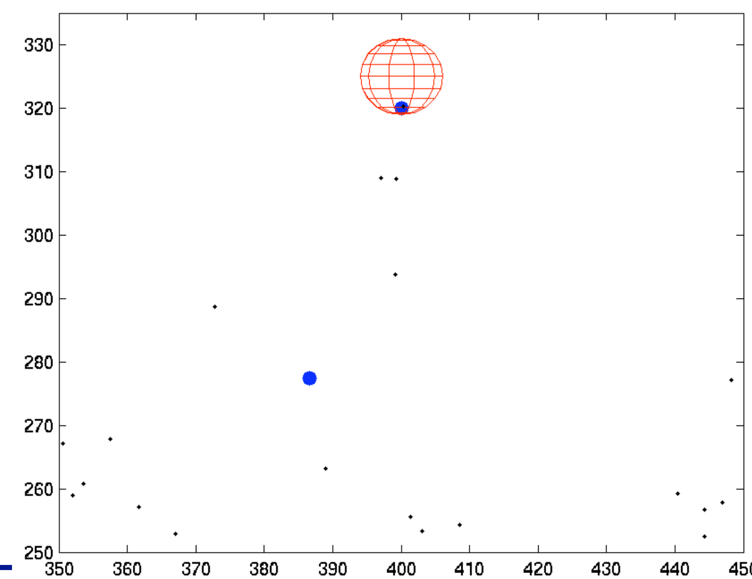
- The classifier works well for a given location; what about an entire vehicle or container?
  - Testing each voxel would lead to accumulation of test errors.
- A selection of the most scattered muons is used for a clustering algorithm.
  - The number of muons selected, ~ 100, can vary with the cargo weight, the duration of muon exposure, and the minimum size of threat object to be detected.
- A *k*-means clustering algorithm is used to find the clusters.
  - The largest cluster is iteratively checked for possible division, according to a criterion of the form $r_{p+} + r_{p-} < \gamma\, r_p$.
  - Stopping criteria involve the size of γ, the number of clusters, and the size of the smallest cluster.
- The cluster centroids are determined and passed to a classifier.
  - The centroid of each cluster is the point having the least total distance to each point of the cluster, as measured by the muon metric.

Clustering Results

20-kg U sphere, 14000 kg of Fe spheres







Los Alamos
NATIONAL LABORATORY

Likelihood Function for Location Selection

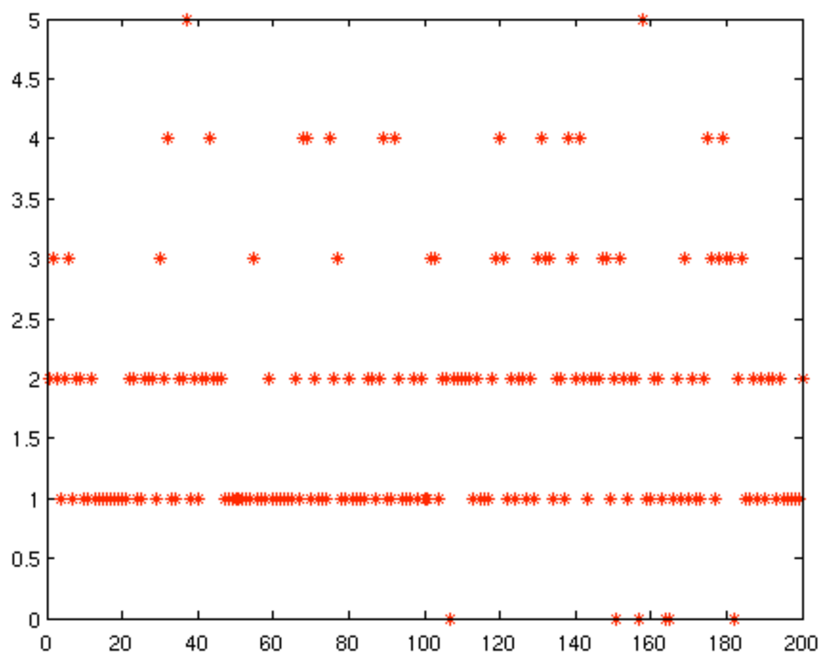
- We identify a small number of locations to test with the classifier by finding local maxima of a likelihood function.
- The function is a sum over a selection of the most scattered muons of scattering times a function of the physics-based distance.
- More principled functions of distance will become feasible with optimization improvements (with John Dennis, Rice U.)

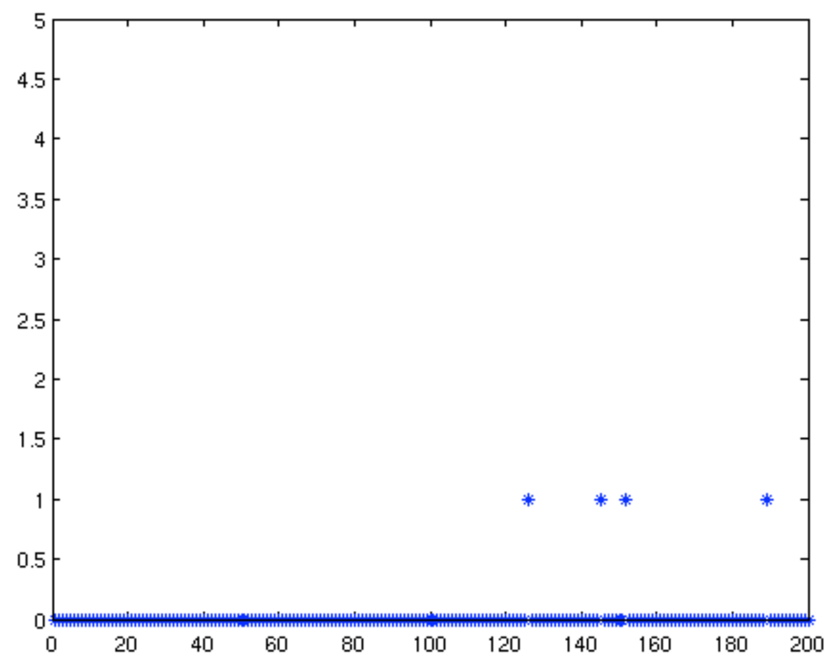
$$U(s) = \sum_j \frac{\Delta\theta_j E_j}{\max(d_j(s), a)}$$

Los Alamos NATIONAL LABORATORY

Preliminary Results---Classifying Candidate Locations

Number of positively-classified locations per container



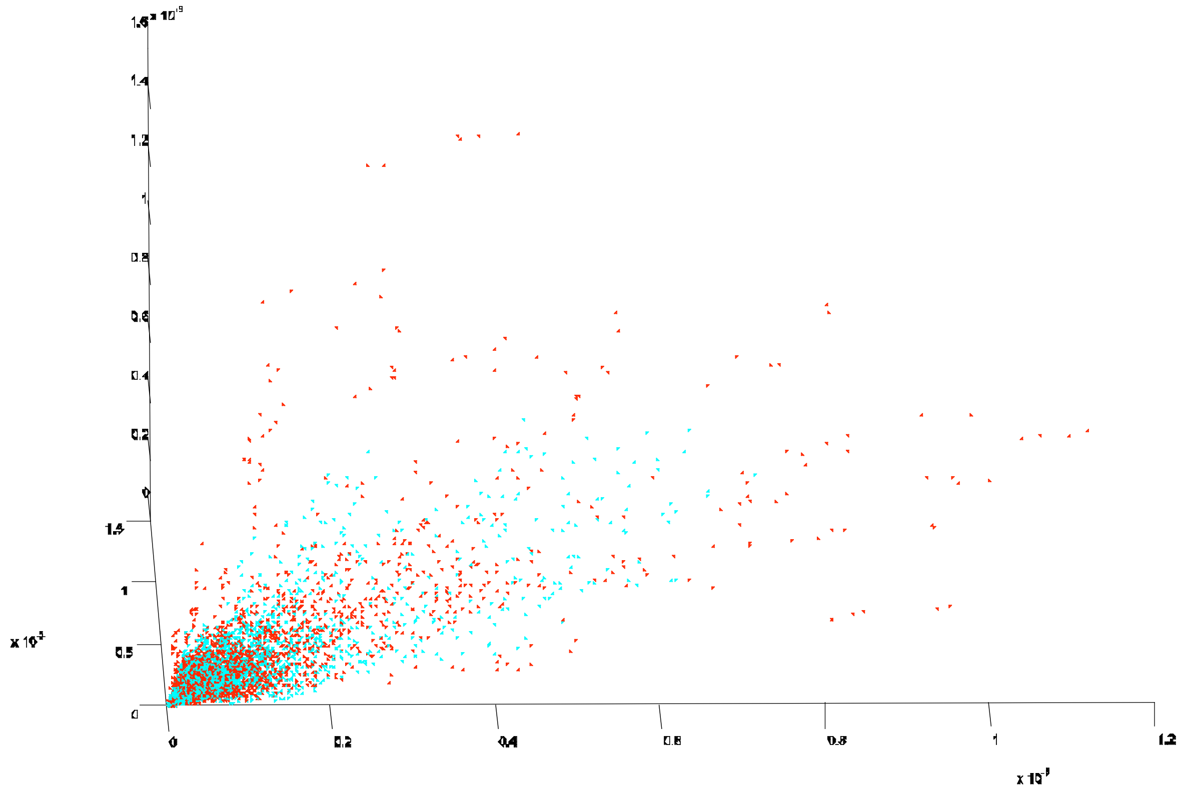


97% of containers with a 20-kg U sphere have at least one positive location

98% of containers with no high-Z material have no positive locations

Detection via Multiscale Geometry of the Data

Values of all 27-voxel cubes for each of two automobiles:
one with a threat object and one with no high-Z material.



Multiscale geometric analysis can identify and quantify outliers of the 27-dimensional data (with Gilad Lerman, UMN).

Los Alamos
NATIONAL LABORATORY