TPC digitization and track reconstruction: efficiency dependence on noise

Daniel Peterson, Cornell University, Vienna, November-2005

A study of track reconstruction efficiency in a TPC
using simulation of the FADC output and the CLEO reconstruction program.

This talk:
- description of the response simulation and hit clustering
- (old) results on reconstruction efficiency w.r.t. readout segmentation
- (old) results on reconstruction efficiency w.r.t. detector size
- discussion on noise
- description noise generation
- results on noise tolerance

Refer to the previous talks for:
- more description of the track reconstruction algorithm

Previous presentations:

* LCWS meeting in Paris,	19-Apr-2004
* ALCPG Victoria	29-July-2004

Funded by the U.S. National Science Foundation (Cornell LEPP cooperative agreement).

Goals

The Goal of this study is to measure the **reconstruction efficiency**
for complicated events simulating Linear Collider processes,
w.r.t. design parameters of a TPC based tracking system:

pad size,
charge spreading,
detector radius
noise occupancy.

Real detector effects are (will be) simulated as much as required:

track overlap,
ionization noise,
electronic noise, and inefficient pads.

This study includes **pattern recognition based on digitized pad-level signals**.
The pattern recognition does not start with 3-D space points.
It includes a simulation of the FADC output as is necessary to provide sensitivity to hit overlap.

Simulation of charge spreading is sufficient to study the effects of signal overlap.
Signals are not generated from microscopic effects.

So far, I have concentrated on track recognition that requires only modest resolution requirements.
This study does not provide details on the reconstruction resolution.

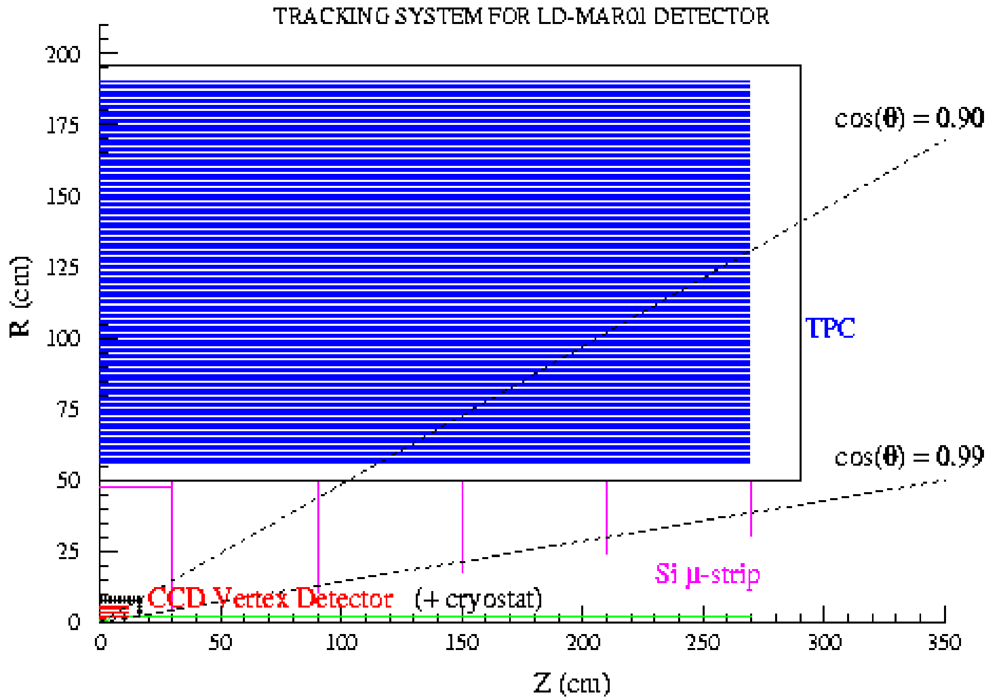
Which TPC ?

This study starts with the "NA Large Detector" design, which includes a TPC with 1.9 meter outer radius.

NA Large Detector:
- TPC outer radius: 1.9m
- magnetic field: 3T

LDC Concept
- TPC outer radius: 1.6 m
- magnetic field: 4 T
- cell size: 2mm x 6mm

In the near future, the study will be updated to use LCIO and simulation files for the LDC and GLD concepts.



However, results on readout segmentation and noise tolerance should apply to a different design with similar **track curvature separation**.

Compare BR^2. "NA large" 1.08 "Tesla" 1.02

Complicated event simulating a Linear Collider Process

A sample event, $e^+ e^- \rightarrow ZH$, from the LCD simulation illustrates the complication due to overlapping tracks.

(All hits are are projected onto one endplate.)

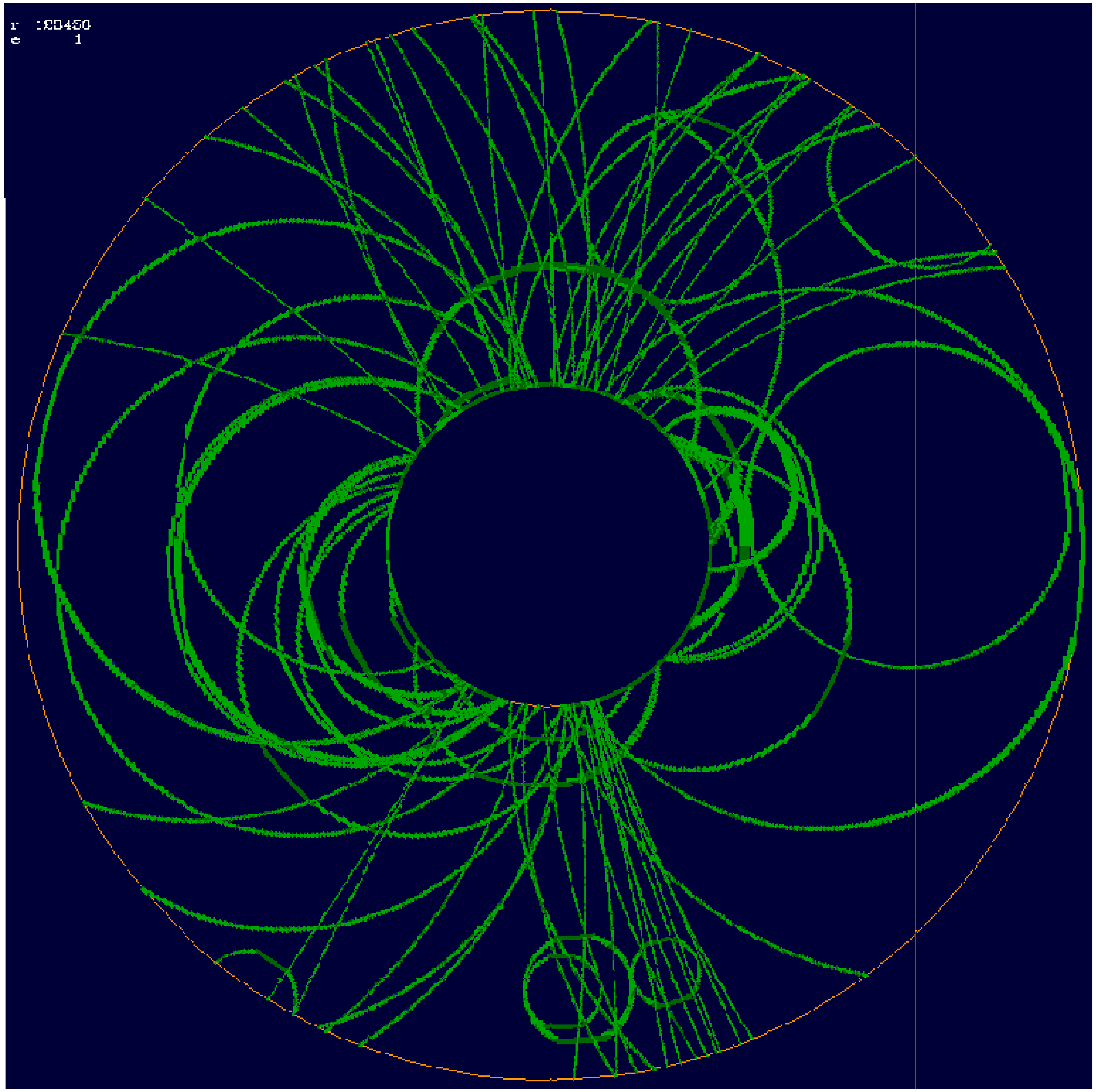
143 layers from 56cm to 190 cm

2 mm wide pads, 1cm radial “height”

- **charge spread is minimal**
- **no noise**

It is not clear from this simple picture if the separation would be sufficient.

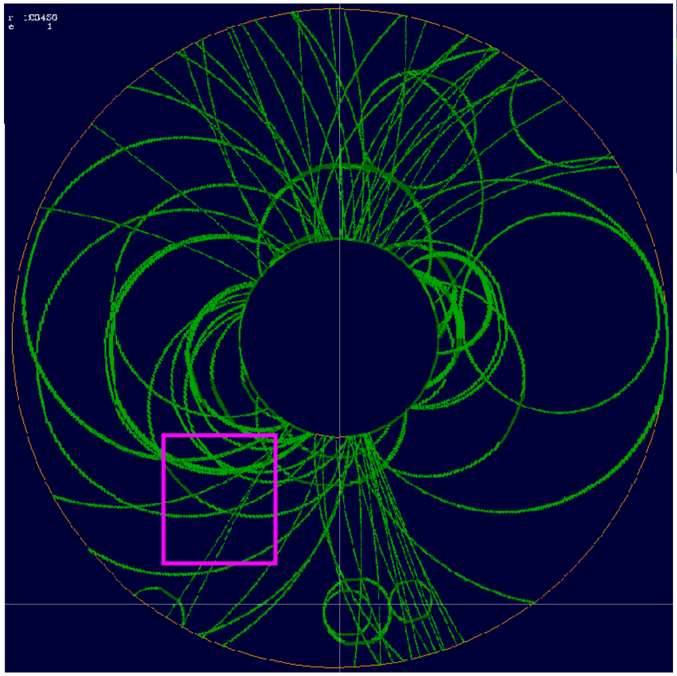
However, we expect that the overlap can be reduced by taking advantage of z separation, making pattern-recognition straight-forward.

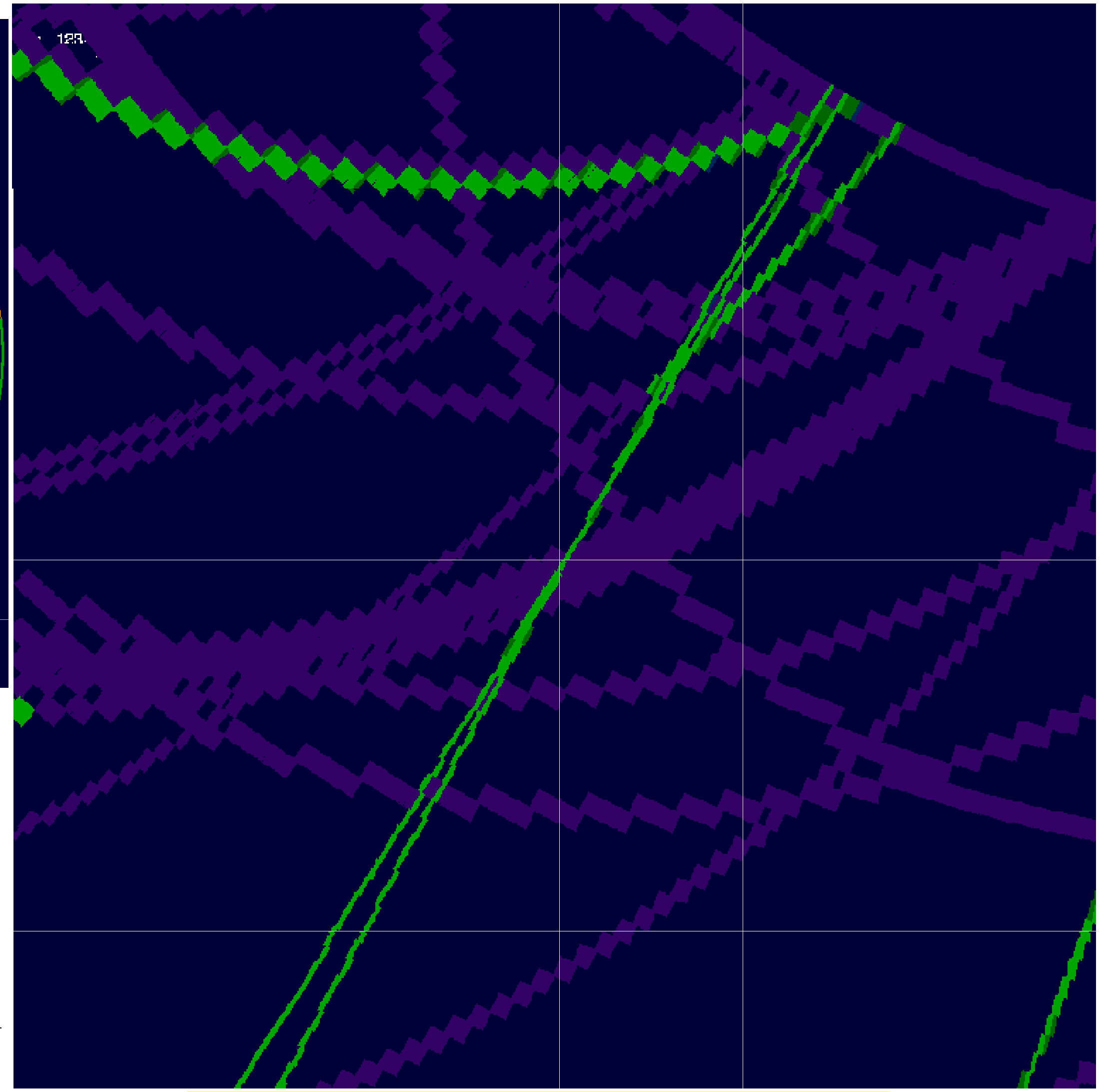


CORNELL UNIVERSITY LEPP LABORATORY FOR ELEMENTARY-PARTICLE PHYSICS

Remaining track overlap when taking advantage of Z separation

(Same event, same pad response)





Active cone: Z=[r * (-6 / 40)] +/- 4.7 cm

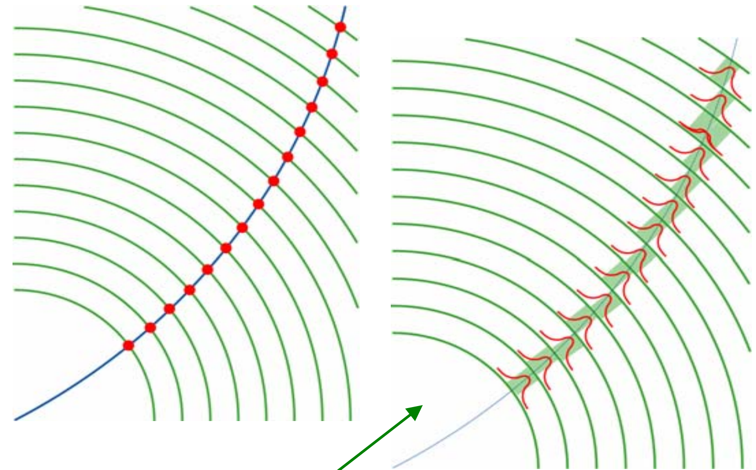
The z separation is often too small to provide straightforward track separation.

crossing tracks in r-f, and
z-separation = **1 mm** .

Track reconstruction can be efficient for very close tracks by selecting information from those regions where the tracks are isolated, as in the CLEO reconstruction program.

Detector Simulation: pad response

The LCD simulation provides only crossing points;
This study uses extensions to the simulation to achieve the goals.

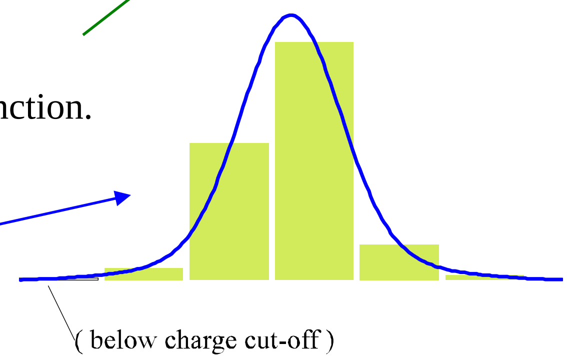


Geometry:

144 crossing points are treated as entries & exits for 143 layers.
143 layers are segmented into pads.

Create ionization using the entry and exit positions for the cell.

Create a ionization centered on the average track position in each cell.
The pulse height is distributed to neighboring cells using a pad response function.

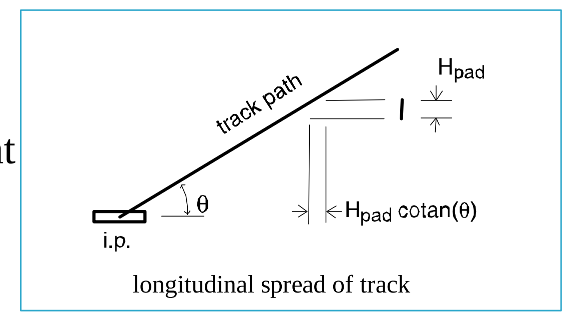


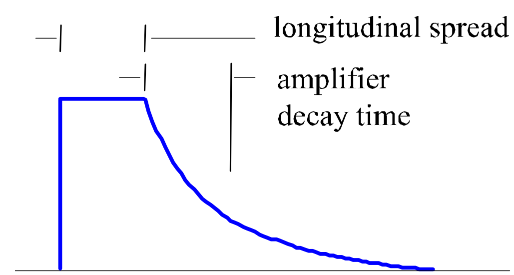
Charge spreading on the pads:

Gaussian width (70% of pad), **cut-off** (~ .002 of min.ion.),
charge is renormalized to provide a **total of min. ion**.
Cells typically have contributions from several ionization centers.

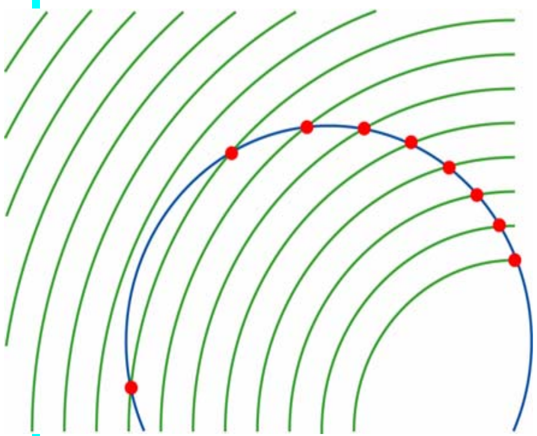
Create a time response

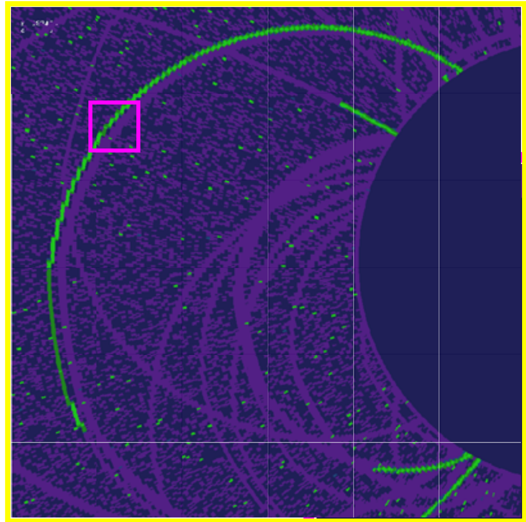
for each contribution to pulse height based the Z projection of the track and the amplifier recovery time.





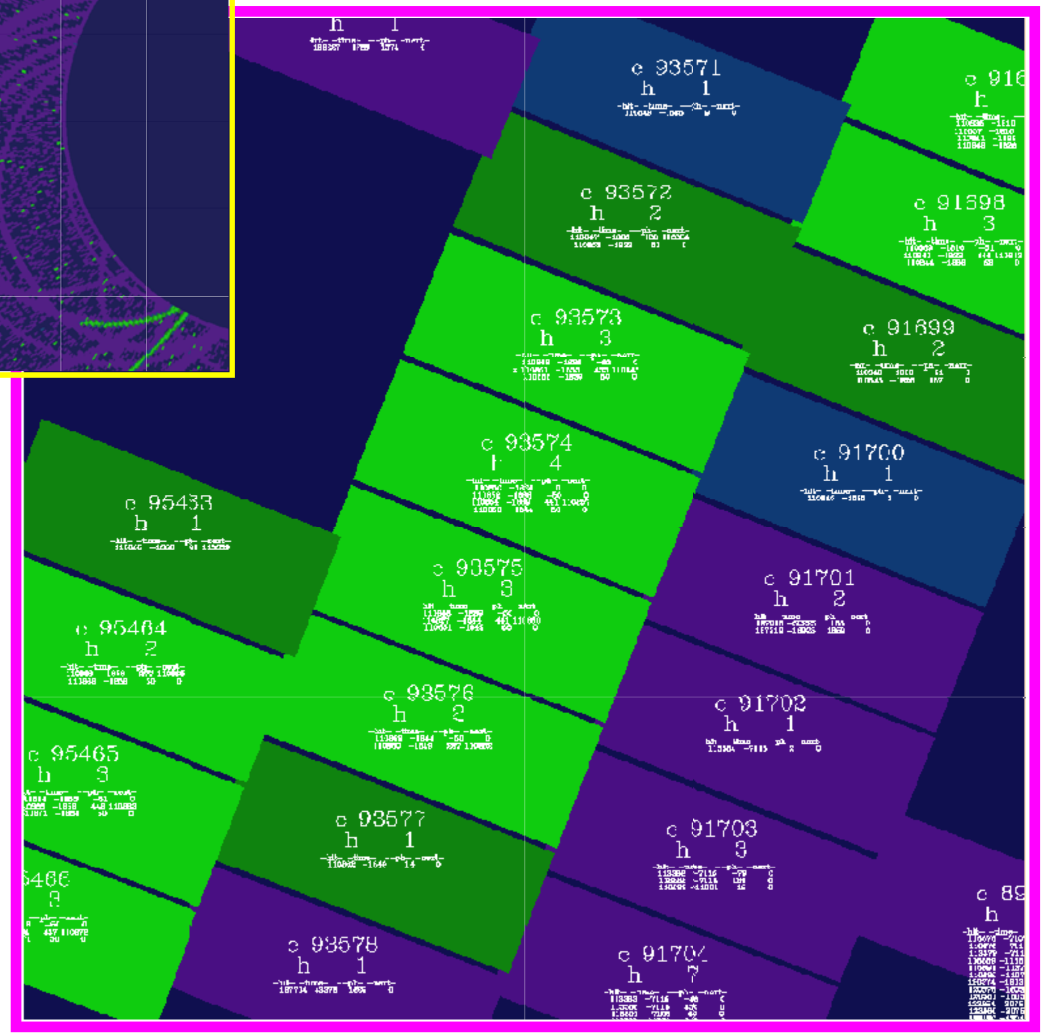
Ionization distribution for large entrance angle





Cell width: 4mm

Active cone: Z=[r * (-3 / 40)] +/- 4.7 cm



While treating the cylinder crossings as layer entry and exit positions, it is easier to identify and properly treat multiple cell crossings.

Again, ionization is deposited in the cells depending on the path length in each cell.

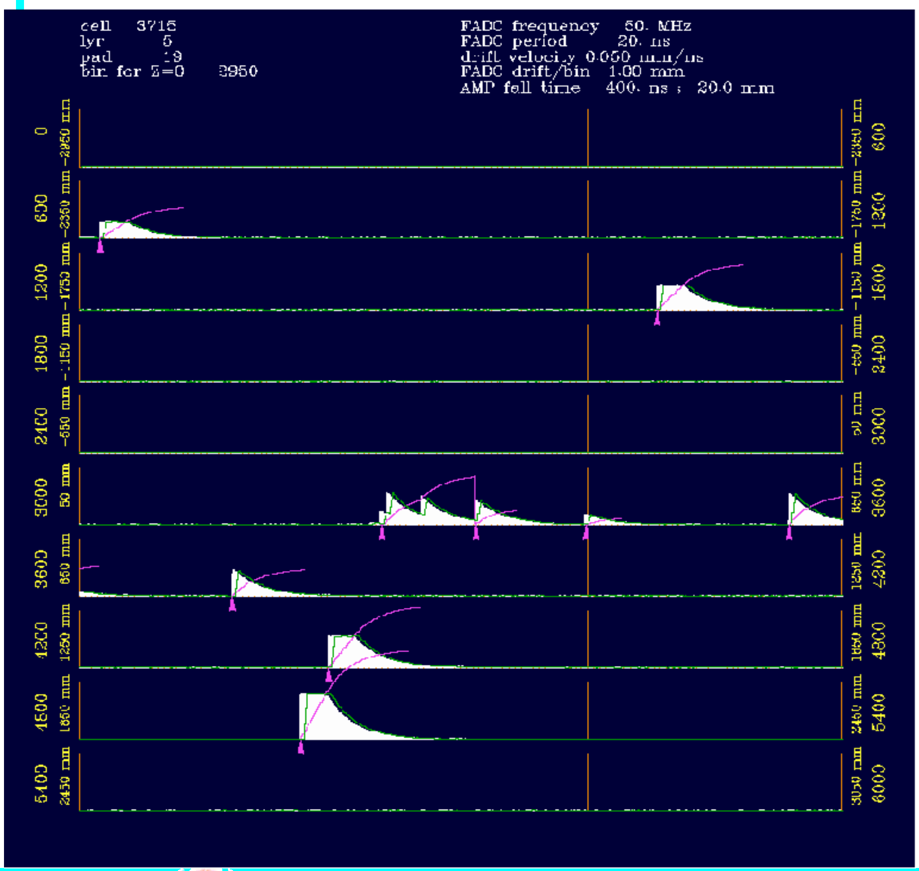
Diagnostic information on the cells provide the contributions from pulse height centers assigned to that cell.

Detector Response: merging overlapping hits and time pattern recognition

After signals are generated on pads as described in the previous 2 slides,
pads may have overlapping signals that would be merged in the hardware readout.

Each signal, **including noise hits**, is described by a pulse height, time, and duration at max. pulse height.

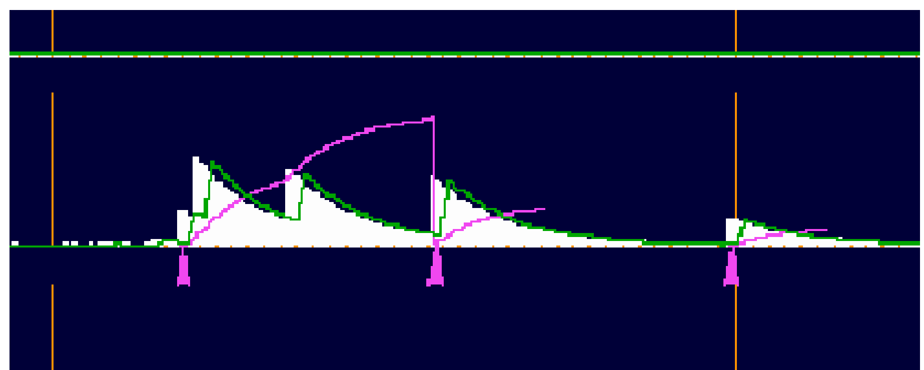
This information is used to simulate a FADC response in which overlapping signals add.



The FADC response is analyzed to determine the *unambiguous threshold crossings* indicated by (▲).

The example shows merged and separated hits.
(Also, note low level noise.)

Threshold crossings found in this procedure replace the original pad signals.

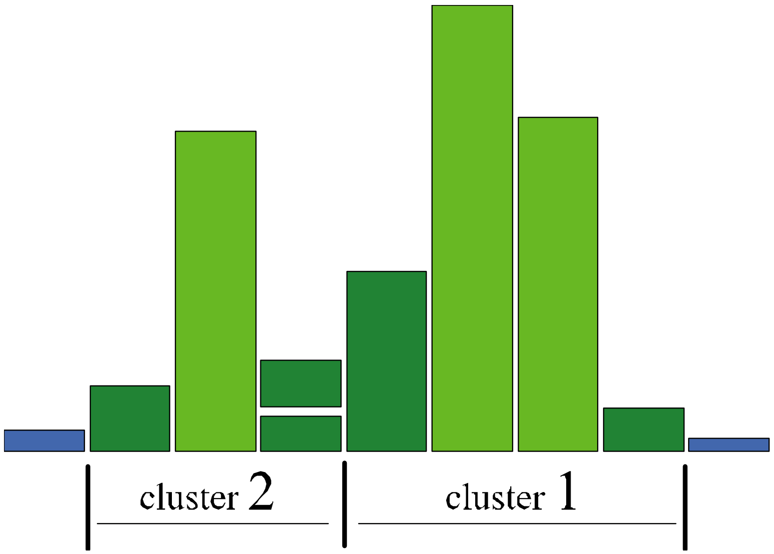


Event reconstruction: pad clustering

Previous slides have described how the generator track crossing of ideal concentric cylinders are converted to FADC signals.

The FADC signals were then processed to recognize "unambiguous" threshold crossings – single pad hits.

Now these single pad hits are clustered in ϕ to locate the significant centers of ionization that can be used by the pattern recognition.



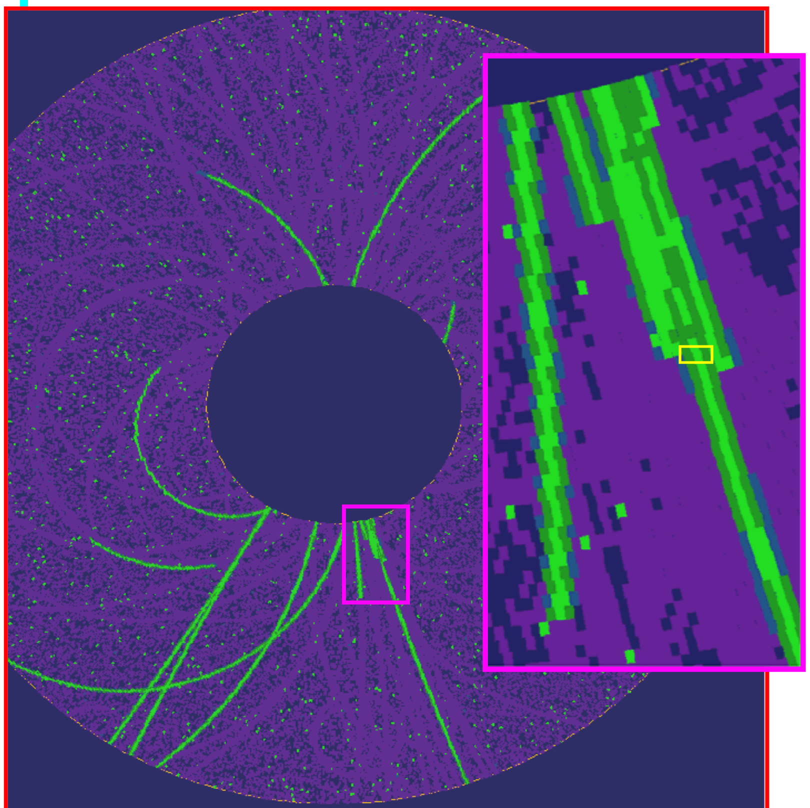
Pads with > 0.51 of the maximum are treated as "core pads".
(a detail of the primary pattern recognition)

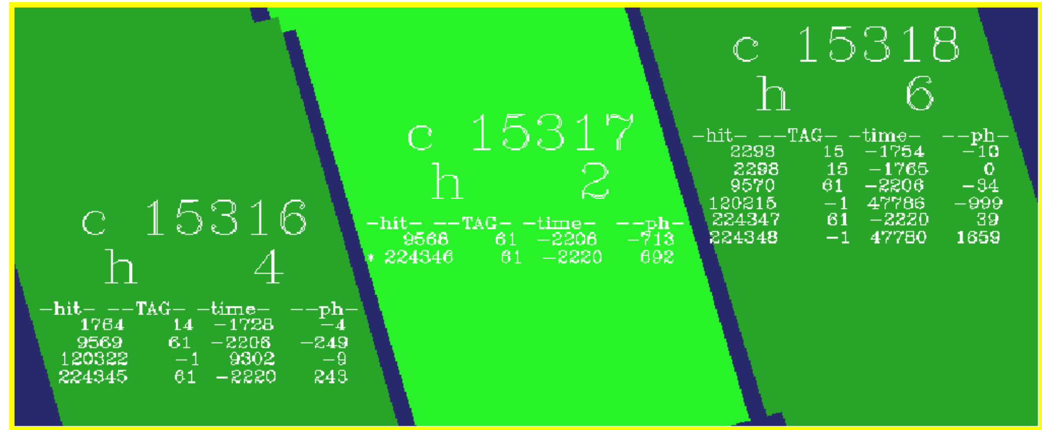
Clustering in r-ϕ

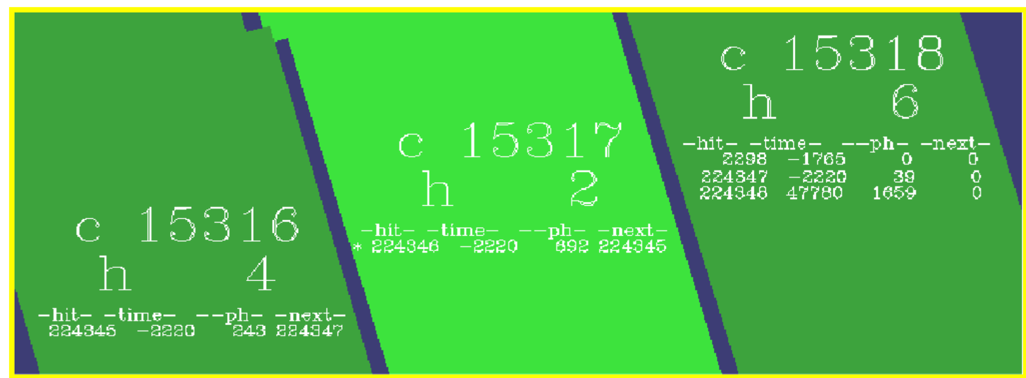
A local maximum, above a threshold, defines a **central pad**.
Adjacent pads, above a lesser threshold are added to the cluster.
Difference in Z of adjacent pads is required to be less than a threshold.
Clusters are **Split** at local minima, less than a fraction of the lesser peak.

Splitting of overlapping cluster **is not precise**.
A pad, which may have contributions from 2 (or more) sources, is assigned to the larger neighbor as shown.
This may lead to non-gaussian smearing of the central position.

Projected hits for event, after detector response simulation and clustering







Active cone: Z=[r * (-7 / 40)] +/- 4.7 cm

Active hits in green

Ignored hits in purple

This is the information input to the pattern recognition.

The pad response includes merged hits with time and pulse height information.

Simple, pre-merged, hits have been “hidden”.

Clustering has been completed for the initial pattern recognition.

MC tracks selected for efficiency studies; the denominator

Not Used

MC generated track list
1) $curv_1, \phi_1$, $impact_1$, $Z0_1$, $COS(\theta)_1$
2) $curv_2, \phi_2$, $impact_2$, $Z0_2$, $COS(\theta)_2$
3) $curv_3, \phi_3$, $impact_3$, $Z0_3$, $COS(\theta)_3$
…
…
N) $curv_N, \phi_N$, $impact_N$, $Z0_N$, $COS(\theta)_N$

Plausible tracks recognized in the hit list

MC generated hit list
1) gen. $track_1$, $layer_1$, X_1, Y_1, Z_1
2) gen. $track_1$, $layer_2$, X_2, Y_2, Z_2
3) gen. $track_1$, $layer_3$, X_3, Y_3, Z_3
…

i) gen. $track_i$, $layer_1$, X_i, Y_i, Z_i
…
j) gen. $track_i$, $layer_n$, X_j, Y_j, Z_j

…
M) gen. $track_M$, $layer_M$, X_M, Y_M, Z_M

Sub-list of contiguous generated hits satisfying…
a) same generated track number
b) starts at layer 1
c) increasing layer number
d) truncated if layer number decreases (top of curler)
e) continues through at least 30 layers

TRACK FIT

"Plausible Track" List
1) $curv_1, \phi_1$, $COT(\theta)_1$, $impact_1$, $Z0_1$
…
n) $curv_n, \phi_n$, $COT(\theta)_n$, $impact_n$, $Z0_n$

Match $\chi^2 = (\Delta C/.002)^2 + (\Delta\phi/.003)^2 + (\Delta COT/.002)^2$

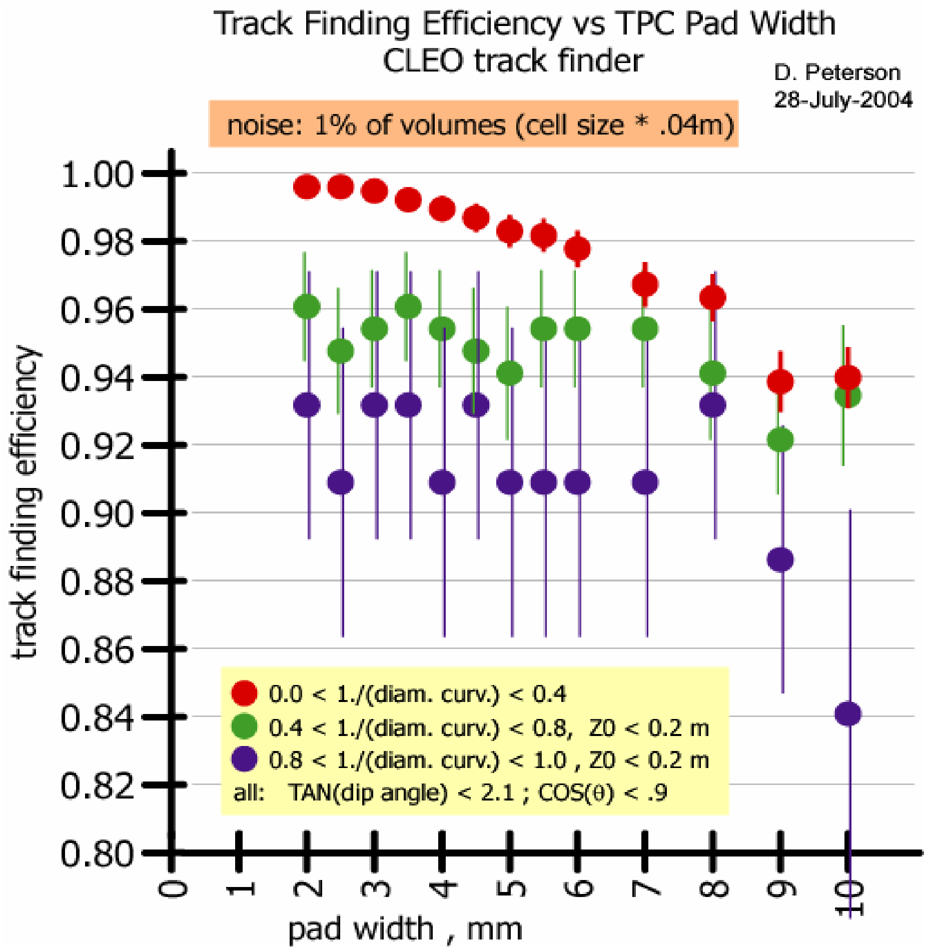
Track finding efficiency dependence on pad width and track "curvature".

Require χ < 25 (defined on previous slide.)

Low curvature tracks: defined to NOT curl within the TPC volume.

Medium curvature tracks:
curl-over radius: 1.2 to 2.5 meters, $Z_0 < 0.2$m.

High Curvature tracks:
curl-over radius: 1.0 to 1.2 meters, $Z_0 < 0.2$m.
(The inner radius of the chamber is 0.56 m.)



Within error, the efficiency for MED/HIGH curvature tracks is largely independent of pad size; these tracks are spread outside the jets.

Statistical errors larger than the fluctuations indicate that, mostly, the same tracks are lost regardless of pad size.

Efficiencies for MEDIUM and HIGH curvature are consistent: average ~ 94%.

The efficiency for LOW curvature tracks is >99% for pad size <4mm.

Pathologies: low curvature

The loss seen at 4mm pad width is dominated by track overlap.

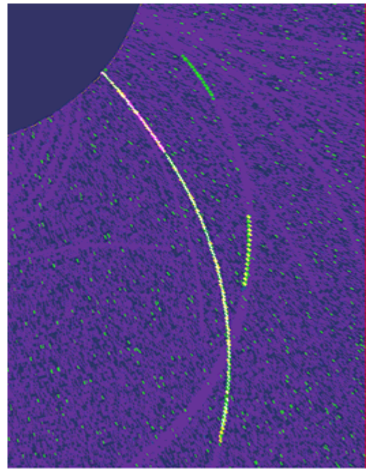
Two tracks that are usually NOT found with smaller pad width are identified as decays-in-flight.

(There are only 766 tracks in the low curvature (straight tracks) sample; this is a large contribution.)

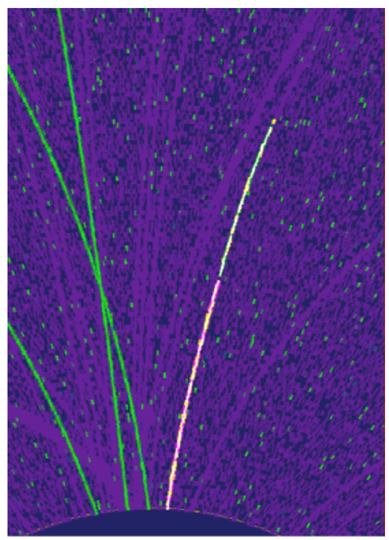
(A change of generator ID number indicates that this is decay rather than hard scatter.)

These tracks can be recovered.

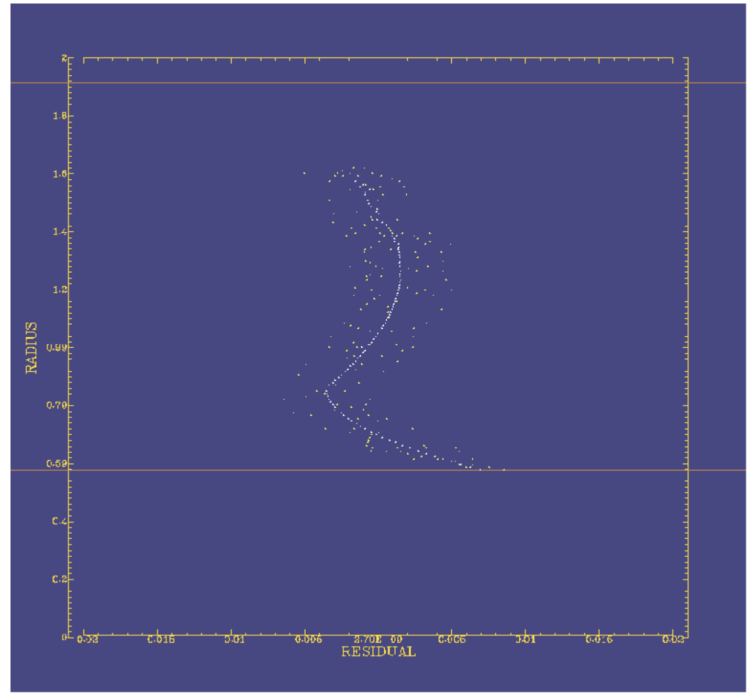
The CLEO reconstruction includes a procedure for recognizing decay-in-flight. This is implemented for CLEOc *where many processes involve low momentum K's. This is not yet implemented for the TPC.*

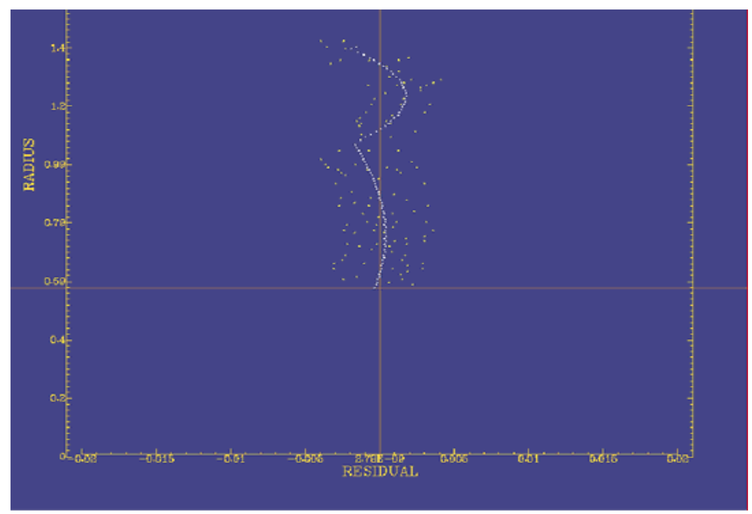


Active cone: Z=[r * (+31 / 40)] +/- 4.7 cm



Active cone: Z=[r * (+30 / 40)] +/- 4.7 cm





CORNELL UNIVERSITY LEPP
LABORATORY FOR ELEMENTARY-PARTICLE PHYSICS

Efficiency dependency on Chamber radius

In a full detector design, chamber radius may be compromised by the calorimetry.

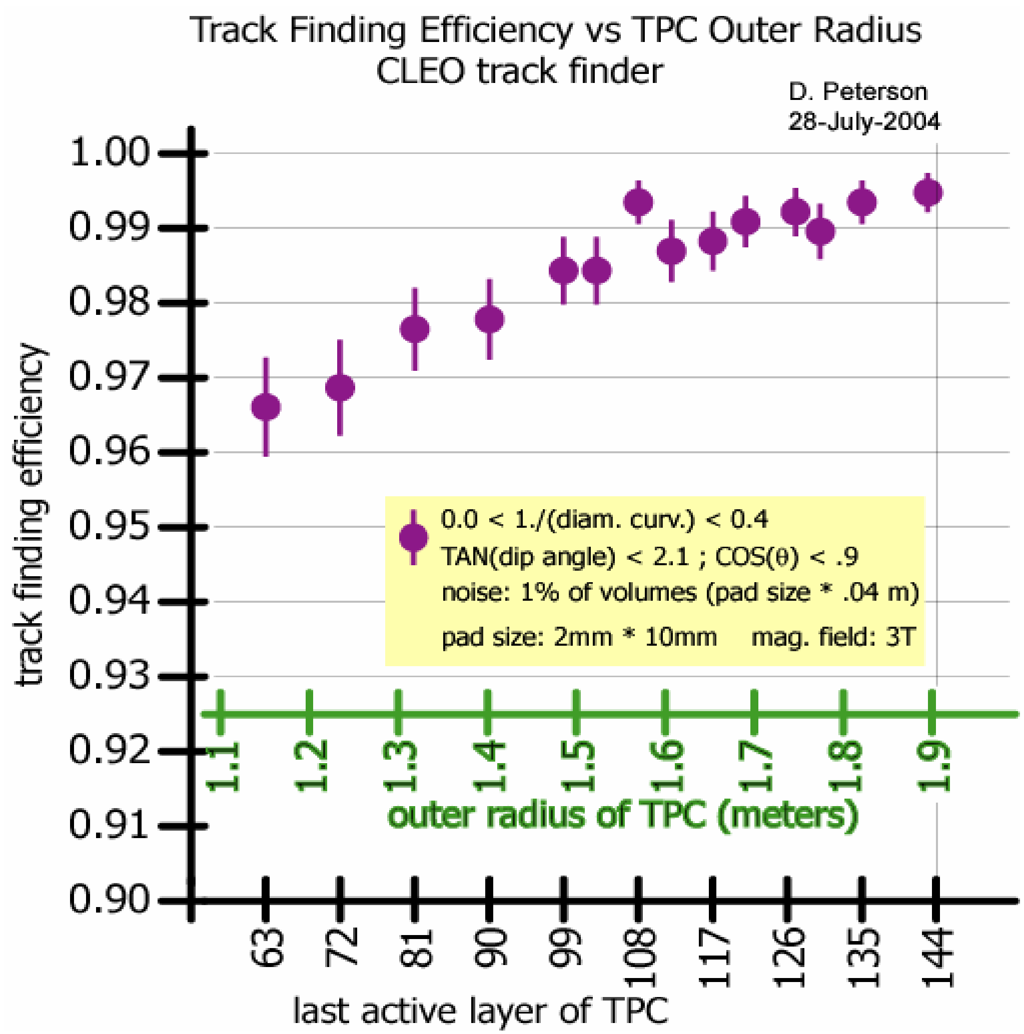
This is the simplest study possible;
cell "height", inner radius, and B field
are not adjusted for smaller outer radius.
Hits are not created beyond the selected "last active layer".

The efficiency for low curvature tracks
is above 99% for
chamber radius above 1.7 m (maybe 1.6 m).

High redundancy in the detector provides efficient reconstruction with greatly reduced information.

Small cell height or smaller inner radius (more cells) may improve the efficiency.

Higher B field (more track separation) is expected to improve the efficiency.



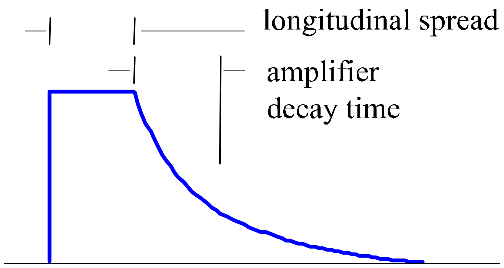
Efficiency dependency on Noise

We would like to claim that the detector can tolerate 1% noise ;   1% of what?
The voxel might be defined by the (pad size) x (FADC bucket).

But a track measurement signal in one layer occupies many more that 1 voxel.

The signal is spread in Z.
Noise signals within (+/- 1 characteristic signal length) will affect the signal. (factor of 2)

The signal is spread in ϕ.
Noise signals in the central cell and in (+/- 1 cell) will affect the signal. (factor of 3)



For data hits,
longitudinal spread
depends on the longitudinal angle,
decay time = 2cm

For noise hits, define
longitudinal spread = 2cm ,
decay time =2cm,
total=4cm

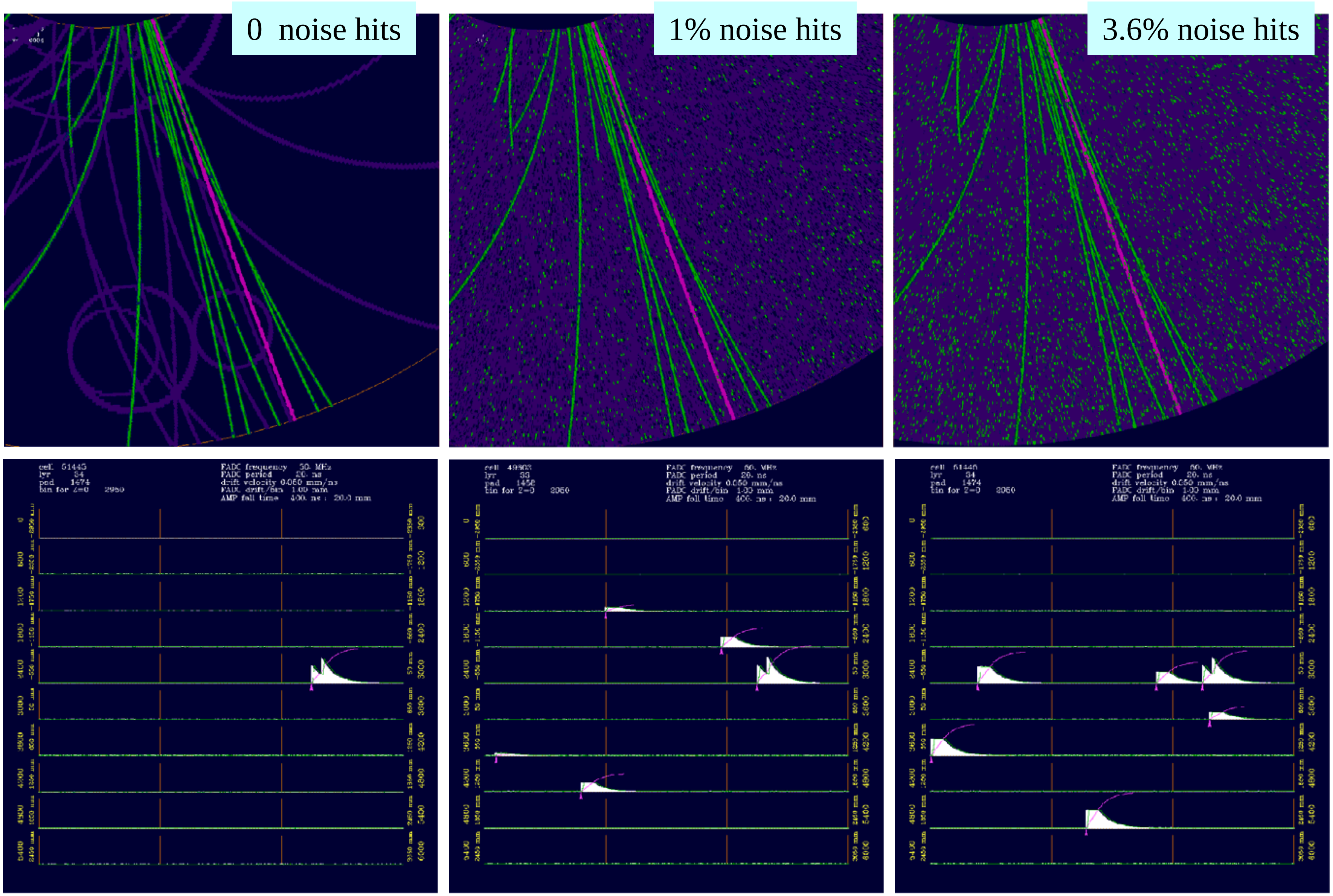
The amplitude of the noise signals is 0:2 x min-ionizing; ave=1 min-ionizing.

The probability of affecting a track measurement signal is
2 * 3 * (occupancy for voxels of characteristic signal length (4cm))

Efficiency dependency on Noise

3 mm cells

Active cone: Z=[r * (-4 / 40)] +/- 4.7 cm



Efficiency dependency on Noise

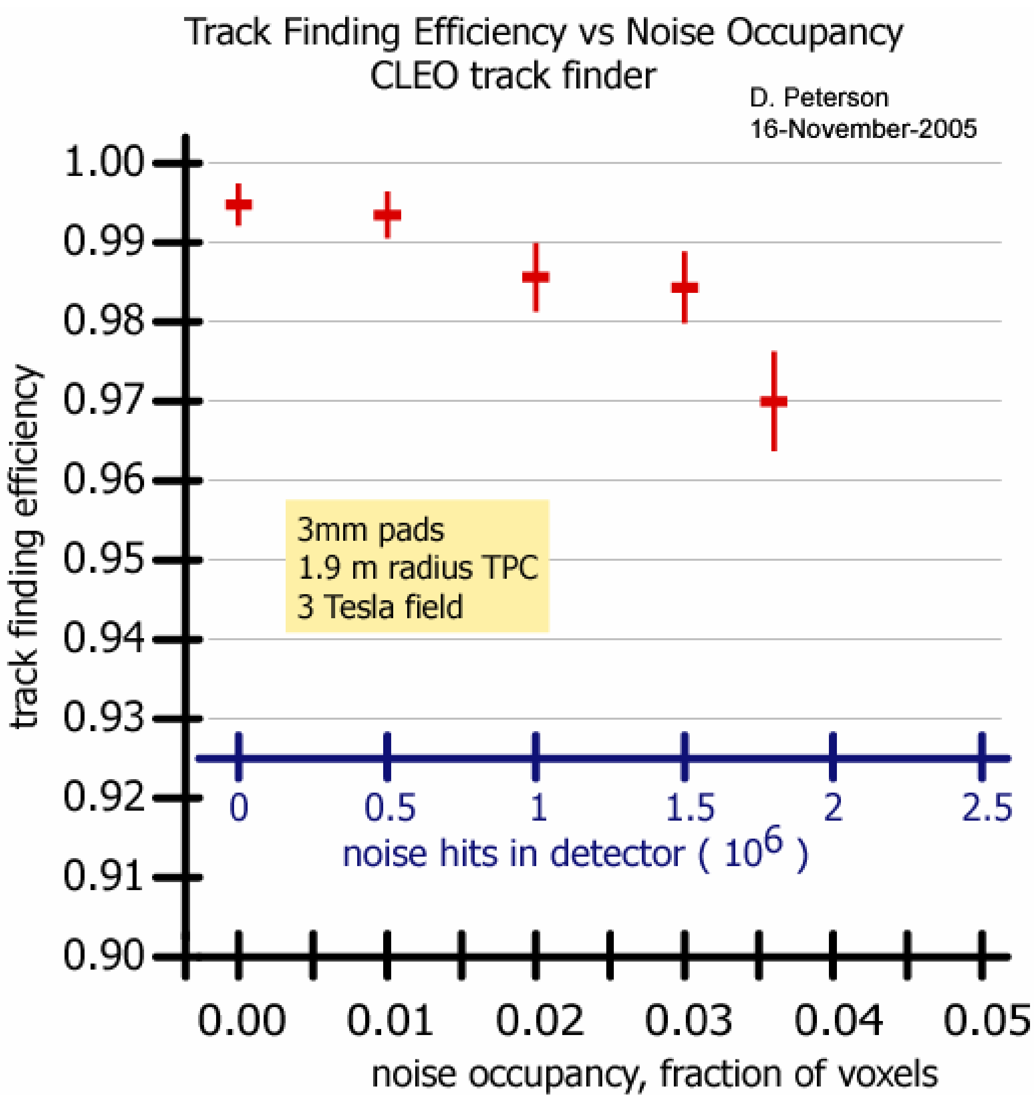
For 3.6% occupancy
1.8 x 10^6 noise hits
(21 % of hits affected by noise)

The efficiency decreases 2.5% .

And, the running time increases 50x .

In this case, "noise hits" are single-pad, not spread by pad-response-function. However, "noise hits" have the full time distribution.

The result is for a particular pattern recognition package. Another pattern recognition may may do better.



Summary, Outlook

The state of the simulation and reconstruction at the ECFA Vienna meeting:

Simulation of TPC signals and adaptation of the CLEO reconstruction for a TPC are largely complete.

Time pattern recognition: investigated, but did not install, change in the "new hit" criterion.

Tuning for high-noise environment: installed tighter hit selection criteria.

Results: (for non curling tracks, within the search volume)

Victoria: TPC reconstruction efficiency: > 99% for pad size < 4mm (1.9m radius chamber, B=3T).

Victoria: TPC radius, can be reduced to 1.7 meters with little change in efficiency (with 2mm pads, B=3T).

Vienna: Efficiency loss is 2.5% (incremental) with 3.6% volume noise occupancy.

(MDI: Simulations of beam related backgrounds must fill the volume related to the signal at the readout;
it is not realistic to fill only one FADC time-bucket voxel.)

Possible improvements to the study:

Clustered ionization noise: simulate the true readout signal width and the correlation due to tracks.

Higher electronic noise, Track background.

Investigate dependence on a parameterization of track isolation.

Quantify the rate of non-removable spurious "found" tracks; this is equally important to energy flow.

Investigate dependence on signal spreading. This could be relevant to, *e.g.*, resistive spreading.

Comment: The study used spreading σ=0.7(pad width), FWHM for 3mm pads is 5mm.

Currently using "DOIT-FAST" (sophistication equivalent to the CLEO on-line processor).

There are many extensions available in "DOIT-RIGHT": track extension, decay-in-flight pattern recognition.

Future:

I am still using the SIO Java interface provided by Mike Ronan. (many thanks)

The framework (soon) will be upgraded to use LCIO and simulated LDC and GLD concept events.