MIPP Software

Mark Messier
Indiana University

Fermilab
9 December 2006

What's in this talk

Overview of the MIPP software

- Framework and basic services
- Quick guide for how to get started
- Reconstruction status

Thoughts on software and analysis needs for upgrade will be sprinkled throughout

Overview: Data format

MIPP uses two data formats

Raw format

Used by DAQ

Optimized for fast instantaneous write speeds

Defined in MippIo package

```
Block
•   Fixed size prefix: ID/Version/Block size
•   Data block
    .
    .
```

New detector systems will need to define format, packing, and unpacking code in MippIo package

ROOT format

Used by offline

More flexible format. Allows streaming of any user-defined objects

Defined in EDMEventDataModel package

I think all this code is ready for reused w/o modification

Online/Offline interface: Online monitoring

<u>OnlineMonitoring</u>

- The "onmon" program runs in the control room and displays detector status and data quality to shift workers. (package: OnlineMonitoring)
- Will require some rethinking and work for upgrade
  - Data passed from DAQ to onmon. Currently done through NFS mounted disk. Need something better?
  - Analysis and GUI merged into single program. Will want to separate these onto their own threads
  - Analysis performed will need to be streamlined to keep up
  - Will want to create more "spill summary"-type plots

Will need to identify a person to take this on

Online/Offline interface: Online monitoring

- It should be possible to reconstruct and display data in roughly real time in the control room.
- This “ExpressLine” processing worked at only a very low level in MIPP-I
- Now we have a pretty reliable reconstruction which could run in the control room and produce summary output

This will need a name assigned for the upgrade

Getting started: Finding information

All software related information posted here:

http://ppd.fnal.gov/experiments/e907/OfflineSoftware/

Offline installation guide:

http://ppd.fnal.gov/experiments/e907/OfflineSoftware/InstallationGuide/

Offline user's manual:

http://enrico1.physics.indiana.edu/mipp/OfflineUsersManual/

To get access to software you will need an FNAL kerberos principle or an SSH key

Getting started: External packages

Detailed in the Installation guide and users' manual. MIPP uses:

- C++ almost exclusively (MC is GEANT3 based)
- Software Release Tools (SRT) for code release management and building http://runiicomputing.fnal.gov/cmgt/SoftRelTools-Manual/
- CVS for version control http://ximbiot.com/cvs/manual/
- ROOT http://root.cern.ch/
- Xerces C++ (XML parser) http://xml.apache.org/xerces-c/
- Postgresql (Database) http://www.postgresql.org/
- CLHEP http://proj-clhep.web.cern.ch/proj-clhep/
- CERNLIB/GEANT3 http://cernlib.web.cern.ch/cernlib/

Throughout development process goal has been to balance minimization of number of external packages while avoiding reinvention

Most useful utilities

Several useful utilities to help get started

psql: The database server. To find runs taken by MIPP-I

% psql runs

runs=> select * from runs where momentum>40 and momentum<60

dumpRunConfig: Prints summary of run

% dumpRunConfig 12340

edm_dump: Prints event data structure for an event

% edm_dump -e 10 file.root

Using the code: JobControl

Users interface to the event data through “anamipp” program defined in JobControl package

anamipp is extended and configured at run time using XML files

Example:

Configuration files

Additional libraries

Analysis modules to run

```
<jobdoc>
 <xmlfile>
      TPCReco.xml   TrkRBase.xml   SPSegAssn.xml   SPFit.xml   SPTrkBuilder.xml   VtxDAFit.xml
 </xmlfile>
 <link>
      Minuit   TPCResCor   NumericalMethods   Swimmer   SPFit   VertexReco
 </link>

 <job name="Tracking">
  <node sequence="TPCReco"  filter="off"/>
  <node sequence="TrkRBase" filter="off"/>
  <node module="SPTrkBuilder" config="default" reco="1" ana="0" filter="off"/>
  <node module="VtxDAFit"     config="default" reco="1" ana="0" filter="off"/>
 </job>

</jobdoc>
```

Analysis unit is the "JobCModule"

Well described in user's manual

User implements "Reco" and/or "Ana" methods.

Reco: read/write access to event

Ana: read only access to event

```
JobCResult VtxDAFit::Reco(EDMEventHandle& evt)
{
 unsigned int i;
 int run = evt.Header().Run();
 int evn = evt.Header().Event();
 // Pull the beam track out...
 try {
  evt.Reco().Get(fBeamTrkDir.c_str(), gsBeamTrack);
 }
...
 // Finally, store the vertices in the event
 const char* dir = "./vtxdafit";
 if (evt.Reco().GetFolder(dir)==0) evt.Reco().MakeFolder(dir);
 for (i=0; i<vtxList.size(); ++i) {
  evt.Reco().Put(vtxList[i],dir);
 }
 return JobCModule::kPassed;
}
```

Event data object

Pulling out previous reconstruction results

Storing new reconstruction results

Status of the MIPP Reconstruction Software

Since completing our running in February of this year, we have been focused on completing the reconstruction algorithms, alignment, and calibration of the detector elements. This task is divided into two essential elements: tracking and particle ID.

The tracking is performed by first forming track segments in the TPC and in the wire chambers. These segments are then joined to form global tracks and fit using detailed maps of the JGG and Rosie fields. The track lists are searched for vertex candidates. We are working on the code to refit tracks using a common vertex constraint.

The tracks are then associated with particle ID information from the TPC (dE/dx), TOF, CKOV, and RICH counters. Tracks are also matched to the calorimeters

TPC Tracking

Track finding and fitting in the TPC follows this basic outline:

[1] Cluster finding: In each slice of pads along the beam direction, clusters of hit voxels are formed

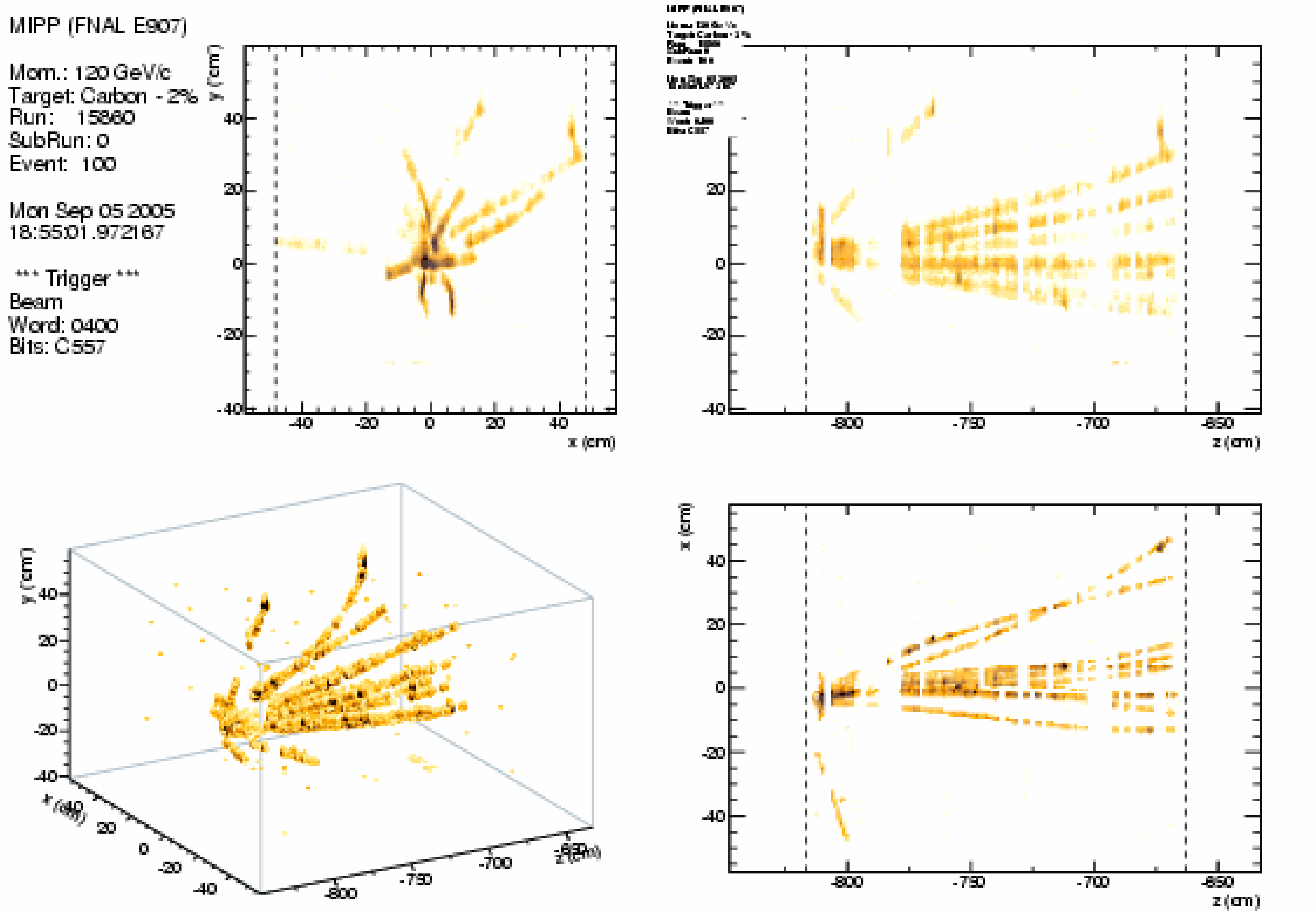
[2] Cluster fitting: These clusters are searched for significant peaks and fit to models of a Gaussians in the x direction and Gamma functions in the drift direction.

[3] Distortion corrections: The position of the resulting hits are calculated by drifting electrons through a map of the Jolly Green Giant B field backward in time from the pad plane up to the measured drift time. These positions differ by as much as 10 cm from a straight vertical drift

[4] Residual distortion corrections: After correcting for all known effects, we still observe biases in the track fit residuals as large as 1 cm in some regions of the TPC volume. These are mapped as a function of volume in the TPC and removed “by hand”. The data used for these corrections requires another iteration as our track reconstruction and vertexing algorithms have improved.

[5] Helix fit: The hits in the TPC are collected into tracks and fit to a helix model

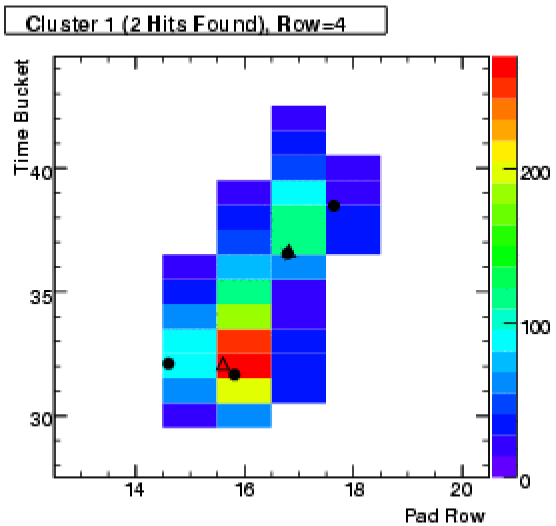
The following slides document the progress on each of these steps.

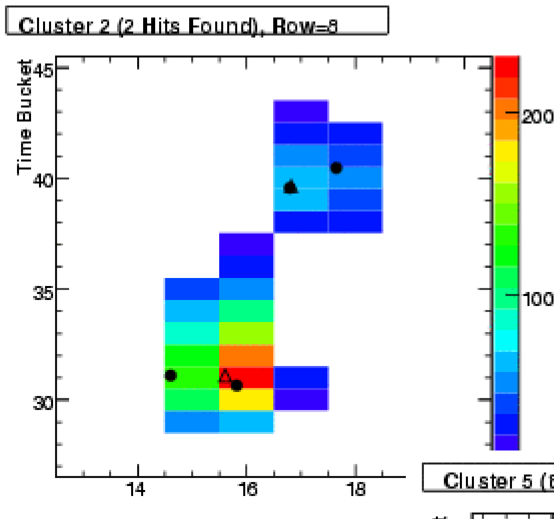


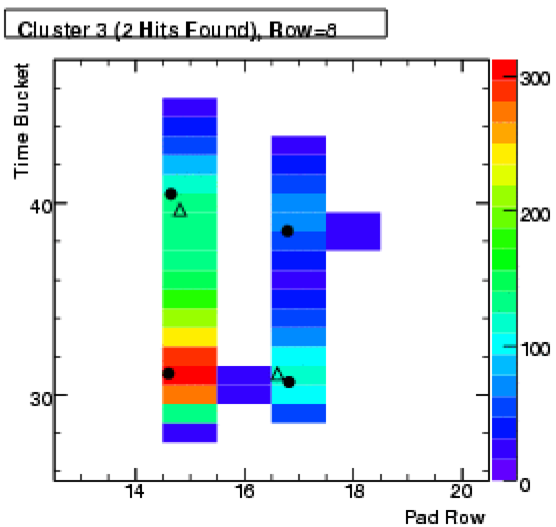
Raw TPC data in the MIPP event display

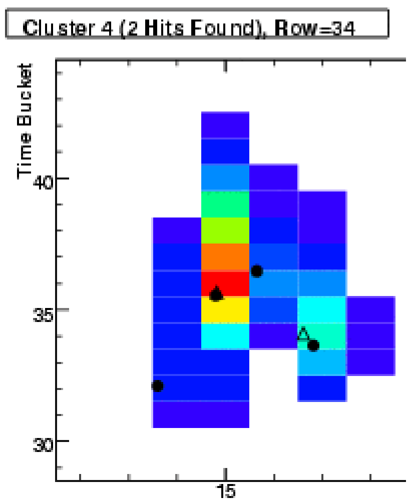
TPC Clusters

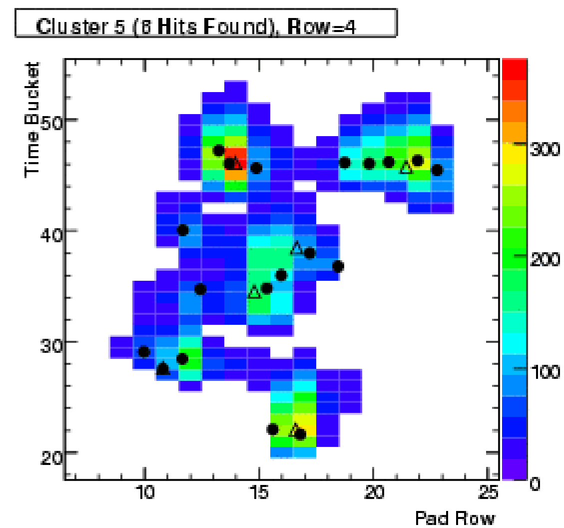
The left panels show some two-hit clusters. Below are example of clusters where more than 3 hits were found.

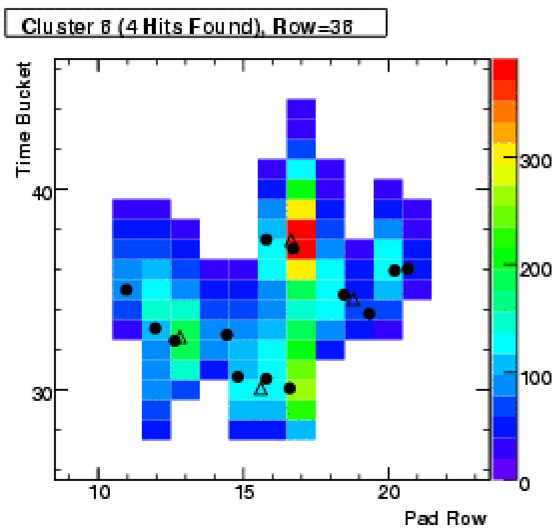


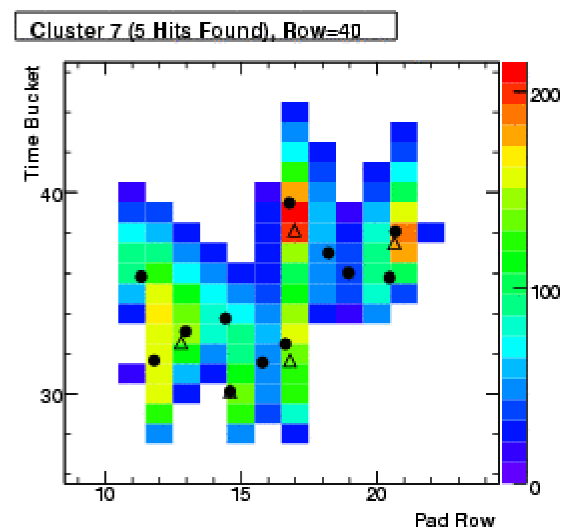


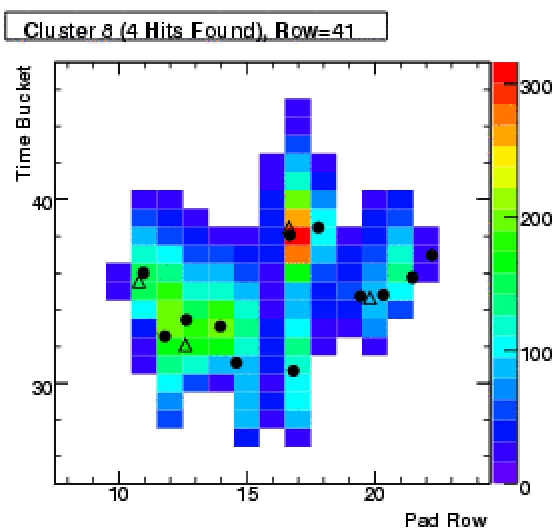






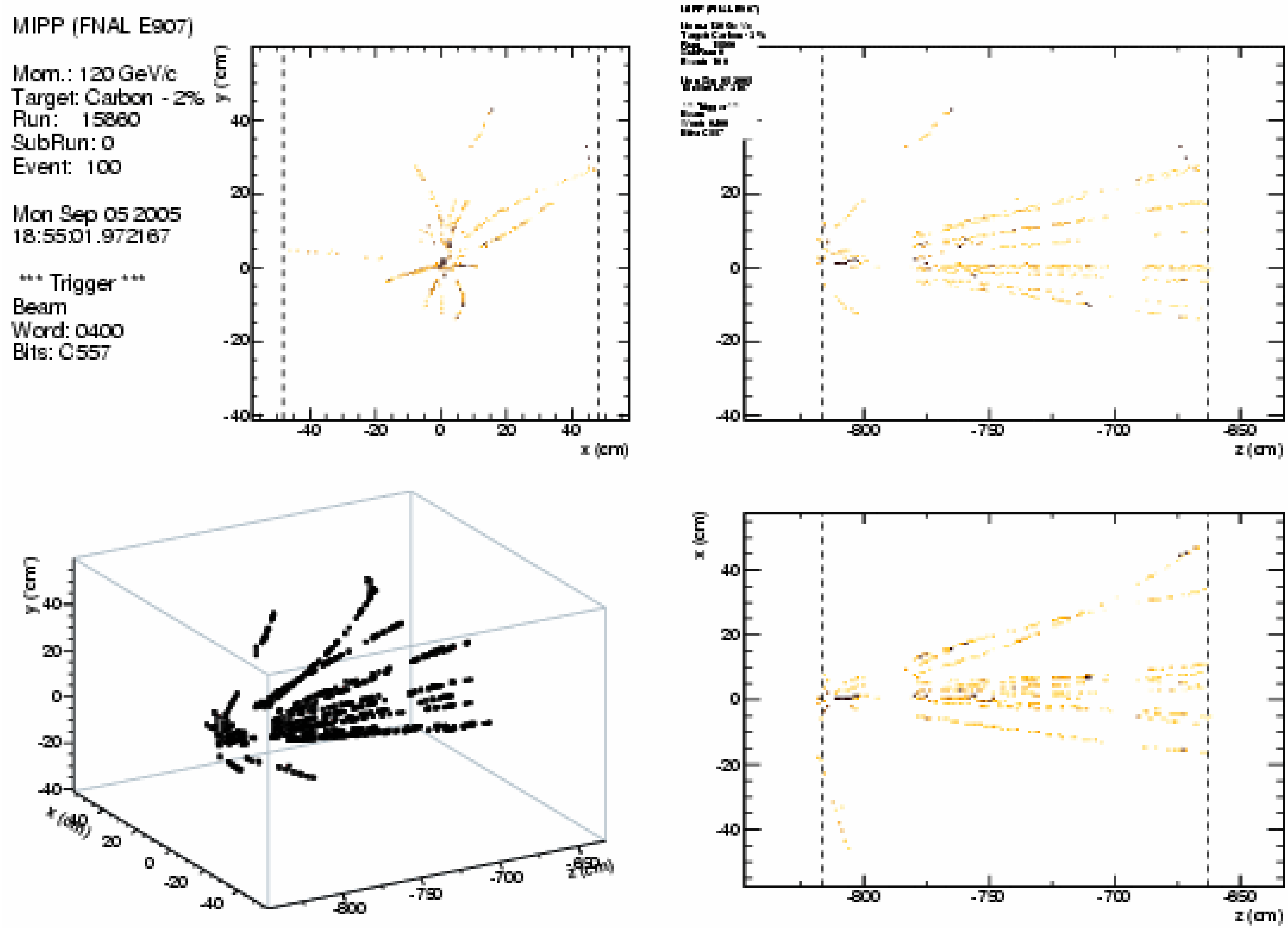




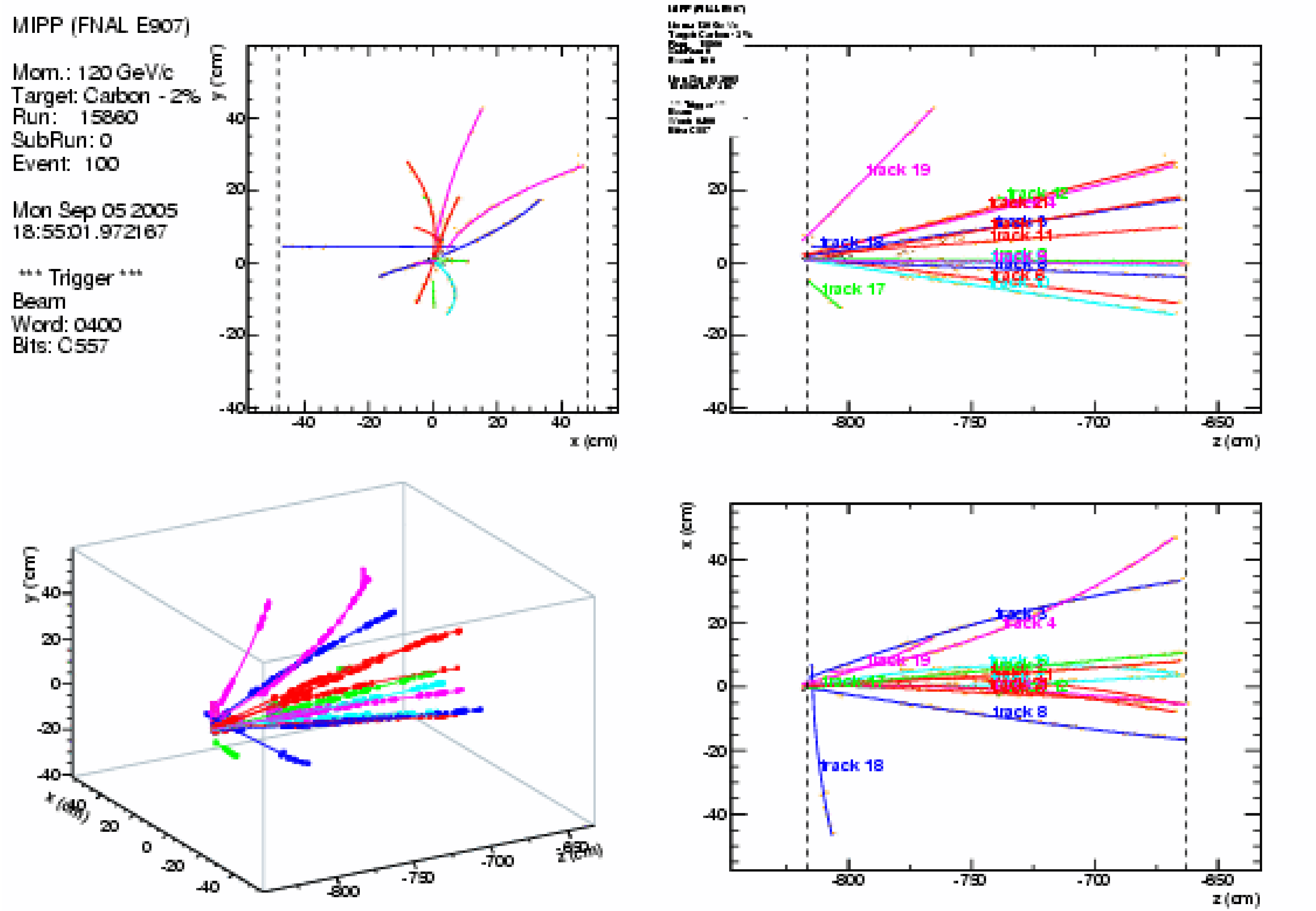


These plots show clusters in a single slice of the TPC along Z. The columns are searched for local peaks (dots) and hits are fit assuming a Gaussian shape in x and a Gamma function along y (drift) direction (triangles).

Color indicates the ADC information for the voxel.



Reconstructed TPC hits, after distortion correction, in the MIPP event display

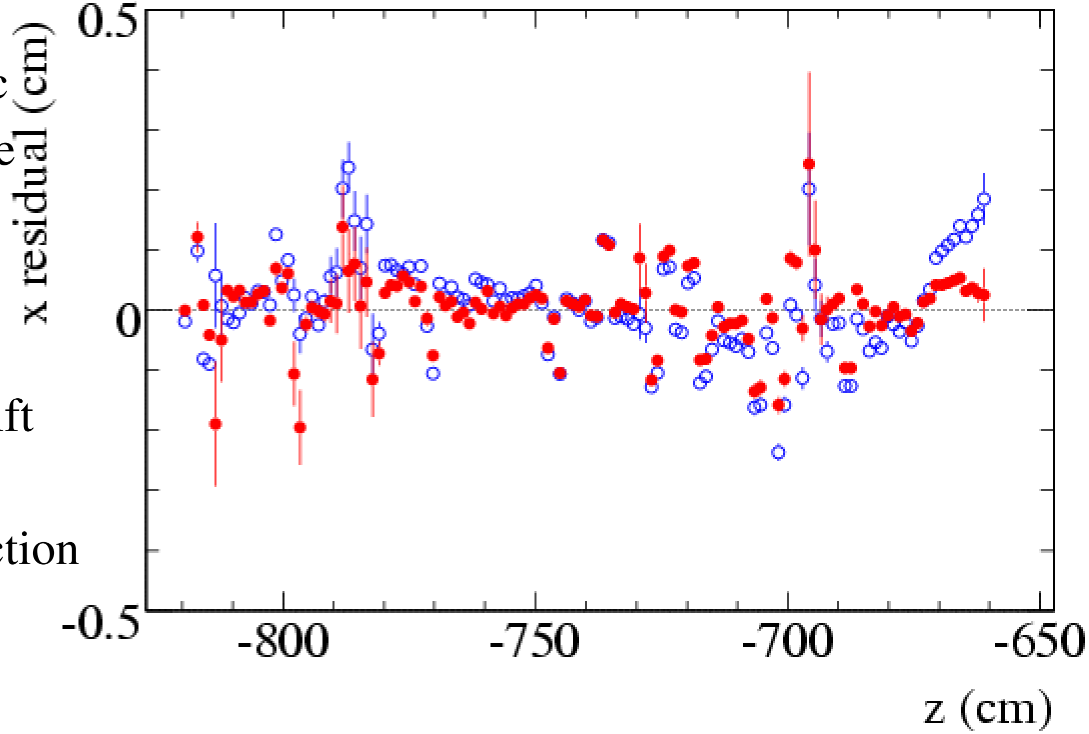


Reconstructed TPC tracks in the MIPP event display

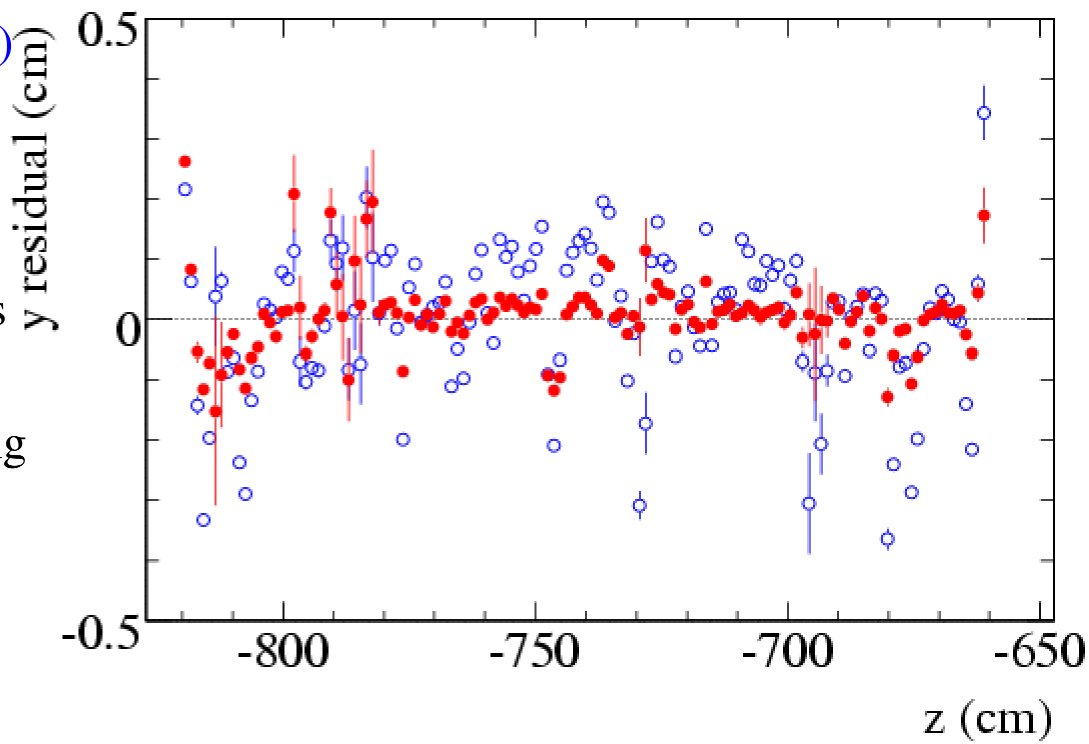
TPC Distortion Corrections II

After correcting for all known effects, systematic biases in the track fit residuals are still seen at the few mm level

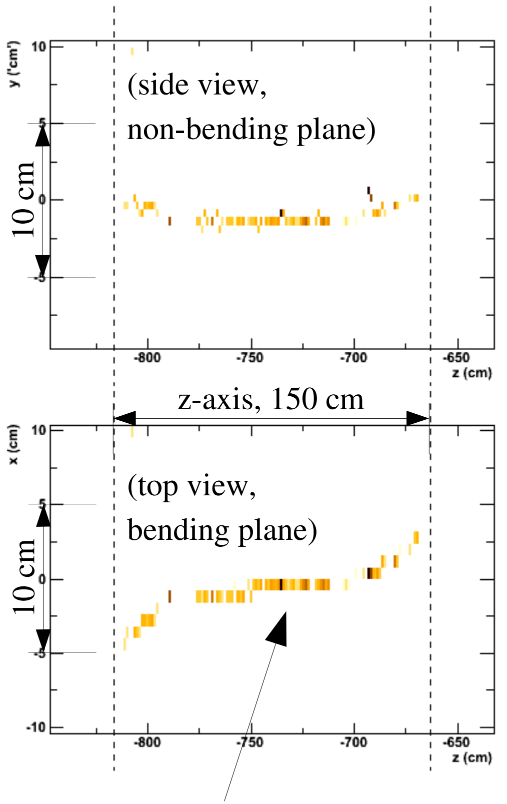
These are removed by mapping out the track fit residuals (using a global track fit to the TPC+Drift chambers and our best estimate of the magnetic fields) and storing them in our database as a function of position throughout the TPC



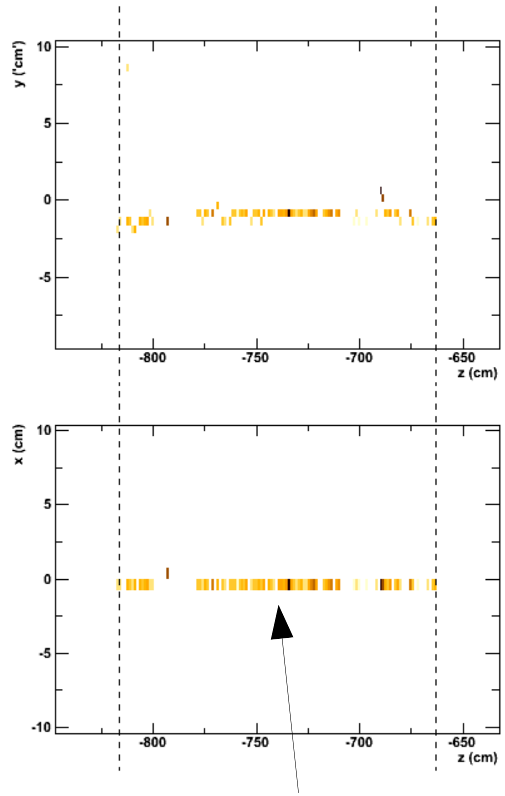
The plots at the right show the average track fit residuals prior to the final fine adjustments (blue) and after the fine adjustments (red). The top plot is for the x residuals where the remaining differences are most likely due to our limited knowledge of the Bfield lines. The bottom shows residuals in the y direction. Variations here are most likely due to differences in the pulse shaping capacitors. After all corrections, the biases are <1 mm



TPC Distortion Effects

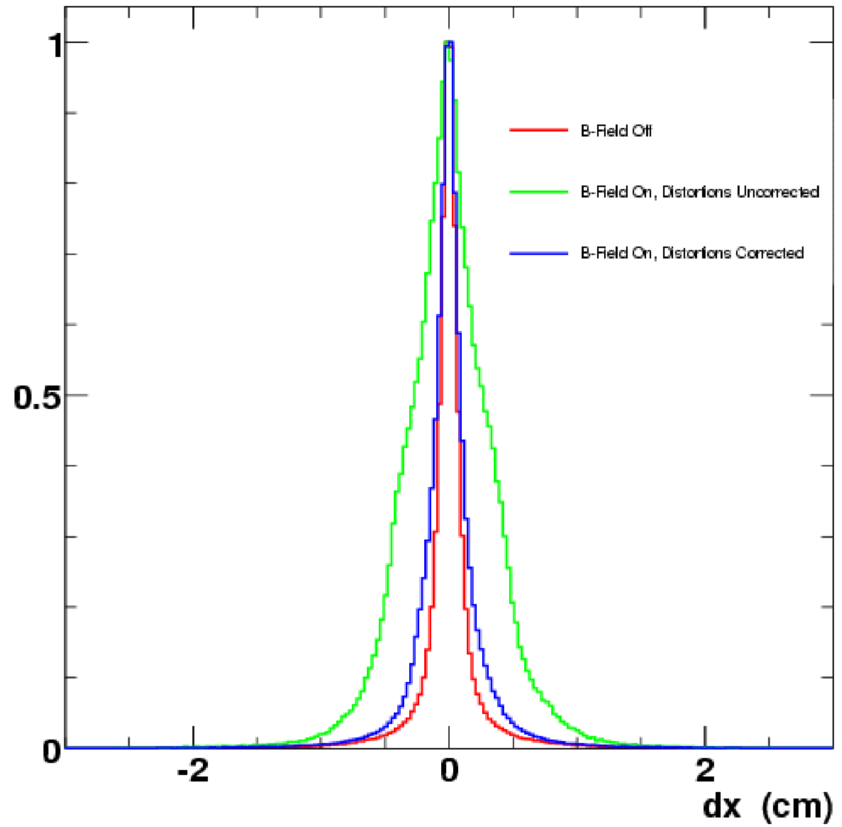


Before correction



After correction

TPC Helix Fit x-Residuals



- Inhomogeneous magnetic field causes drift electrons to deviate from straight-line path to pad plane on bottom of TPC. Deviation is ~5 cm at edges.
- Corrections are applied using a measured map of the field; distortion effects are now < 3 mm.

Alignment Status

- We have gone through a lot of effort to make sure that
  - Chamber wire planes are aligned
  - Chamber z-locations in geometry are consistent with data
  - Relative B-field strengths of JGG and Rosie are matched
  - Z-locations of B-field maps are consistent with data

MIPP

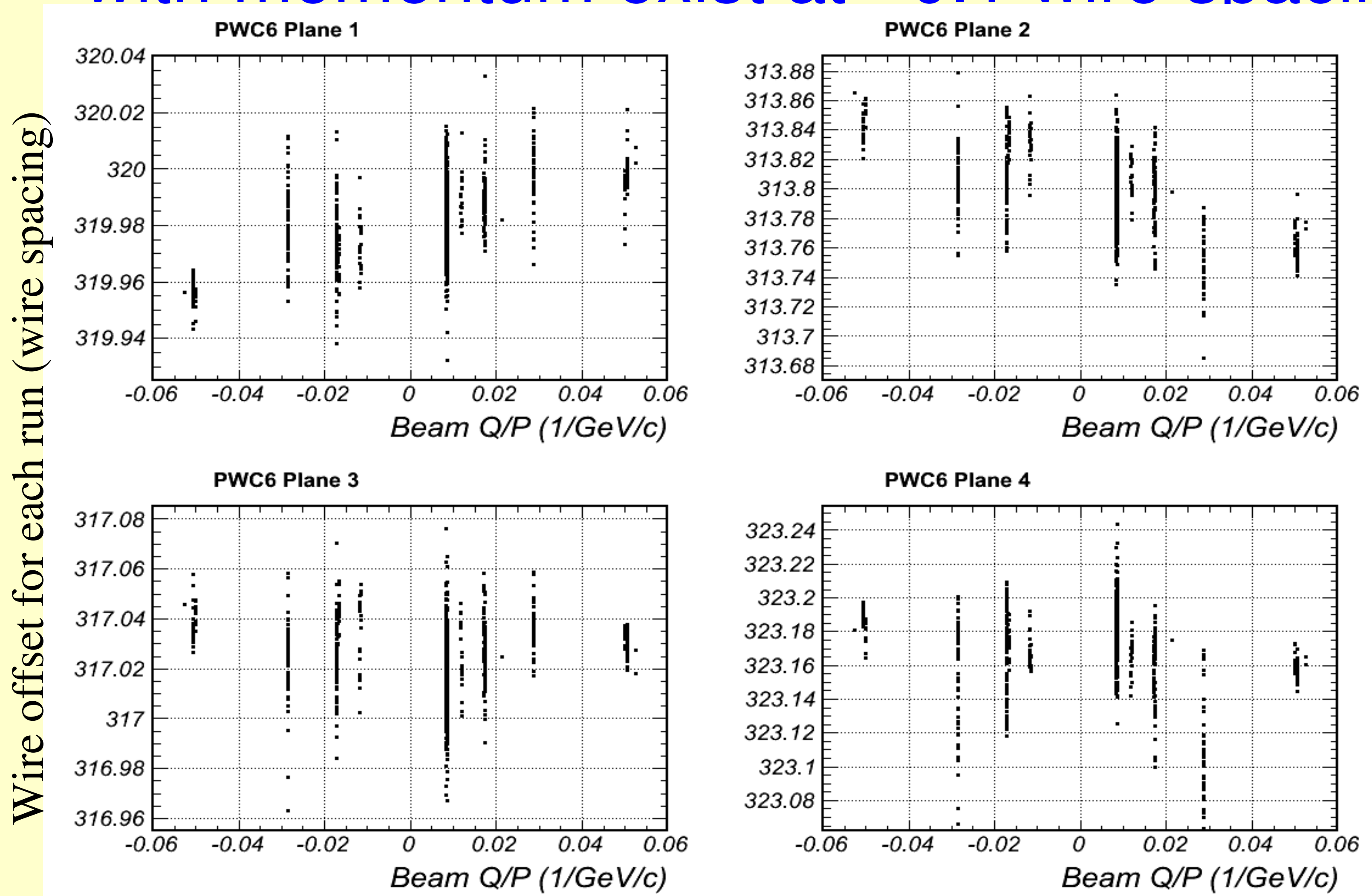
Alignment Status (cont)

- We have aligned EMCal and RICH to chambers
- TPC alignment is work in progress
  - Field on studies are complicated by large distortions
  - Field off data is scarce

MIPP

Wire Chamber Alignment

- Chambers are not fully understood: correlations with momentum exist at ~0.1 wire spacing



Wire angles wrt vertical:

Plane 1: 0^{o}

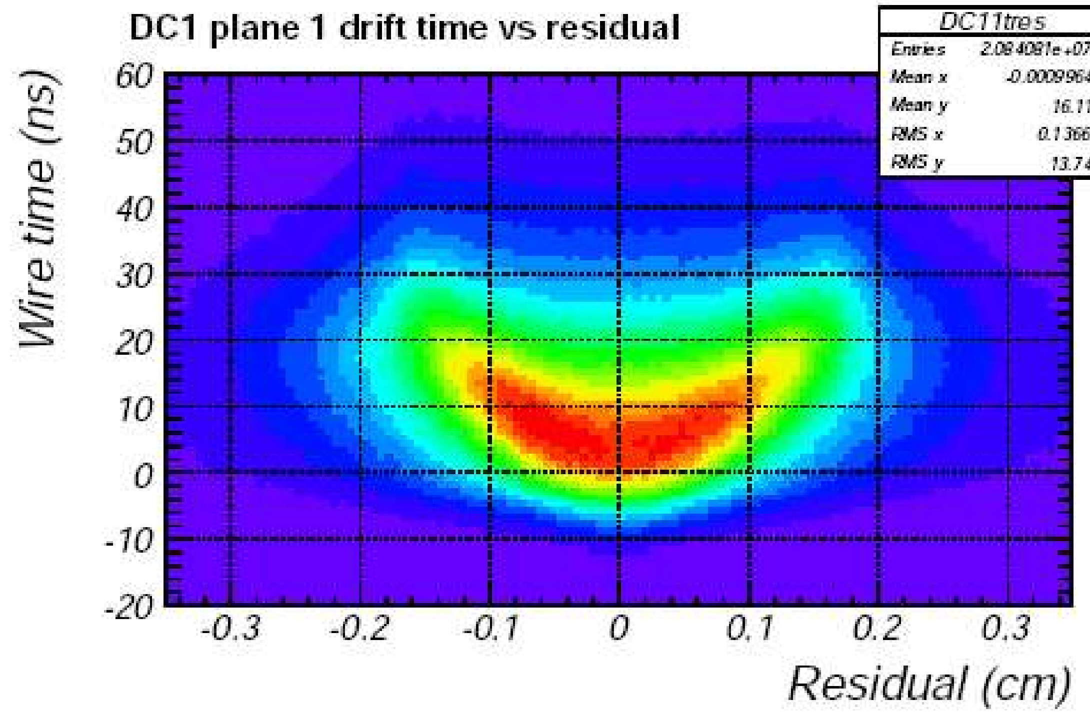
Plane 2: 90^{o}

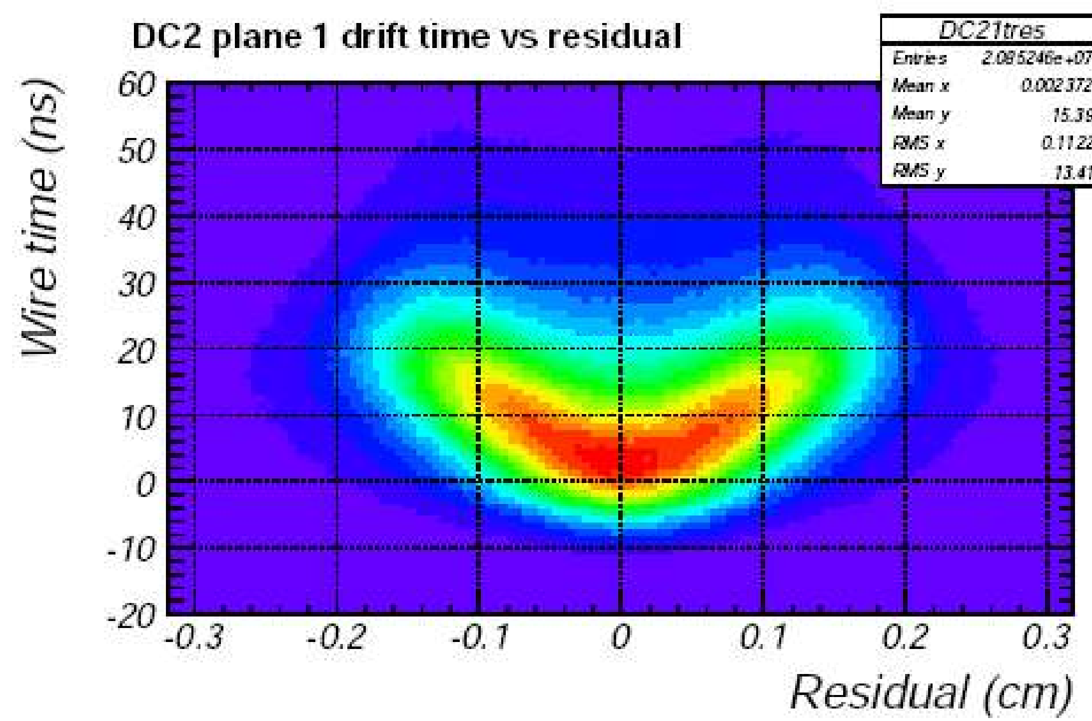
Plane 3: $+28^{o}$

Plane 4: -28^{o}

Wire chamber alignment and calibration

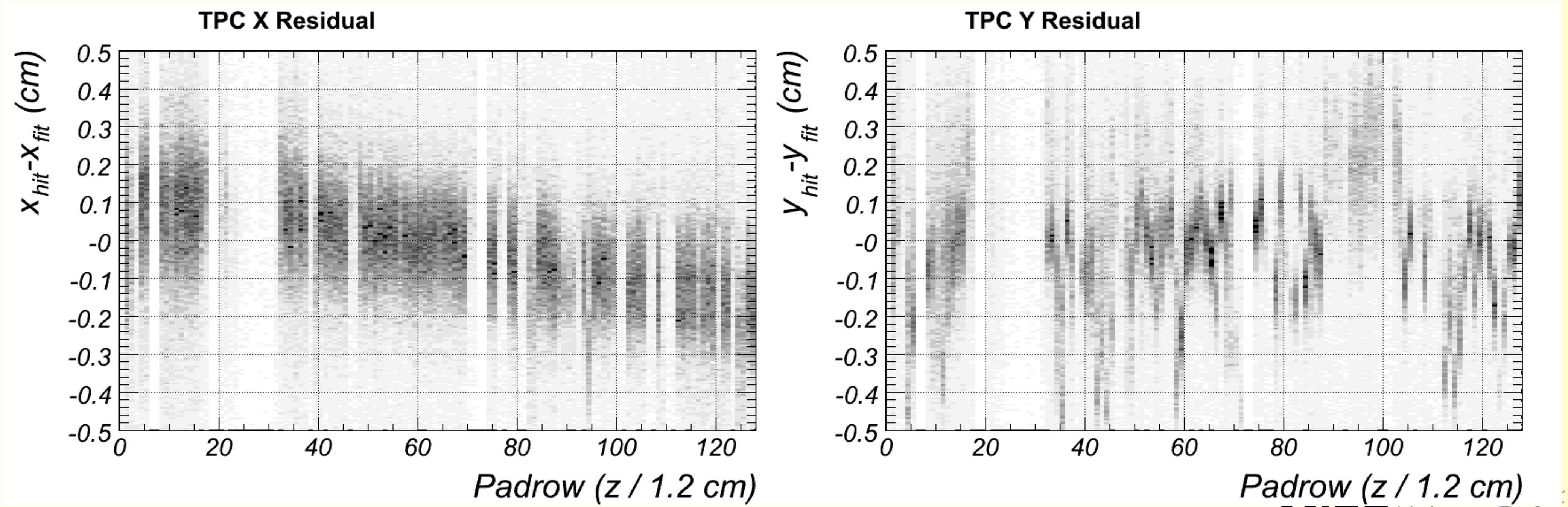
- Completed T0 and wire-by-wire timing calibration of drift chambers
- Completed alignment of wire chambers over the summer.
- Difficult job as there were several discrepancies with survey numbers
- With these alignments and calibrations complete, ready to make use of timing information during tracking
- Plots at right show residual vs. time for two of the 28 wire planes in the experiment. Once the time-to-position conversion is parameterized it can be used in fits
- At the moment, fit TPC+wire chamber information using just wire spacing as resolution





TPC Alignment

- TPC appears to be rotated by ~2 mrad in xz-plane
- Y fits depend on modeling shaped time with Gamma function
- Plots below are from global fits of field off data



MIPP

Global Tracking Status

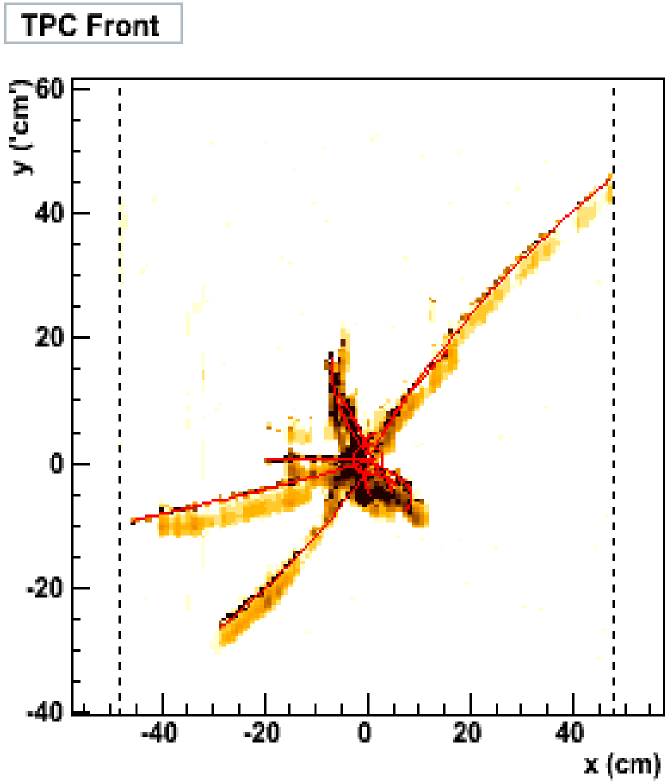
- Track fits are currently done through template method
  - Very fast, requires few swims through B-field
  - Not reliable for large-angle tracks – use TMinuit for those (work in progress)
- Current implementation of global track finding algorithm is done
  - Fit TPC tracks
  - Match TPC tracks to chamber tracks
  - Save unused chamber-only tracks

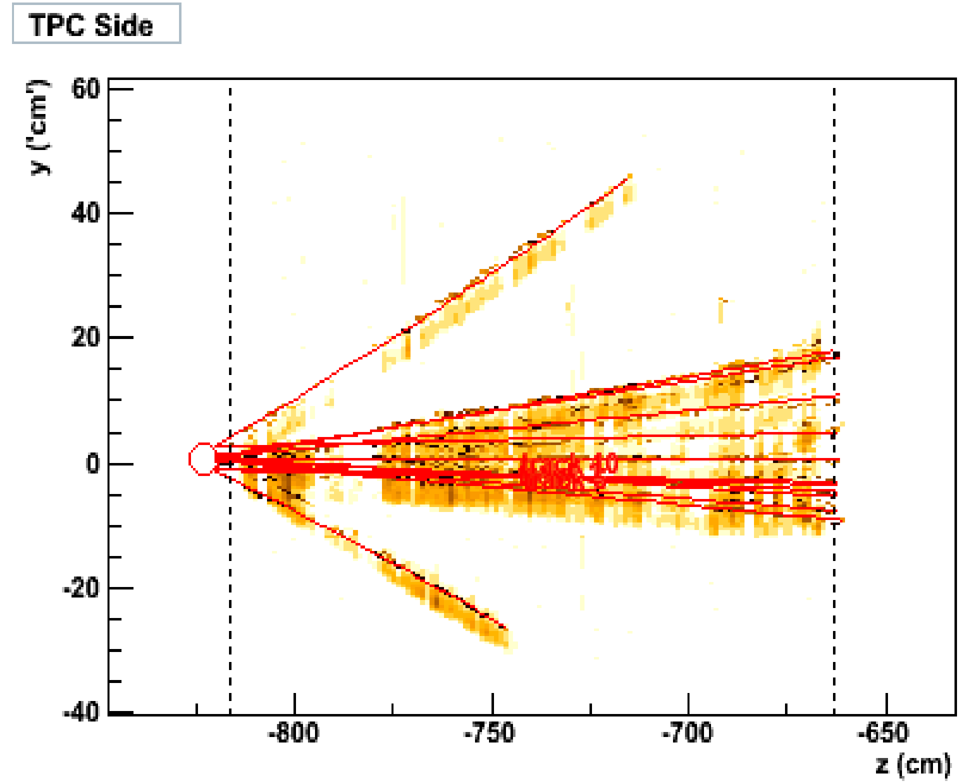
MIPP

Vertex Reconstruction

- Vertex finding is done through deterministic annealing filter (DAF)
- Vertex-constrained fit is then done on identified vertices
- We incorrectly reconstruct target z-position
  - Possible reasons are mis-alignment, B-field maps, TPC distortion corrections
  - The most important problem to be solved ASAP

MIPP





MIPP (FNAL E907)

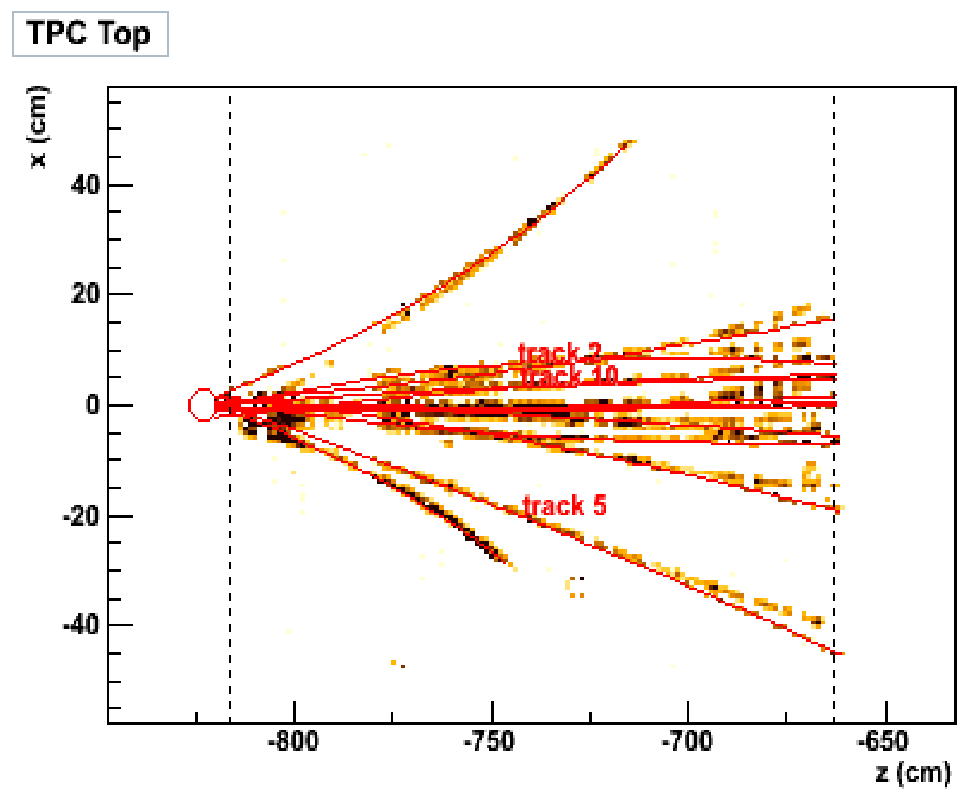
Mom.: 120 GeV/c
Target: Empty
Run: 15634
SubRun: 0
Event: 2

Fri Aug 19 2005
18:28:21.744479

*** Trigger ***
Beam
Word: 0400
Bits: C557

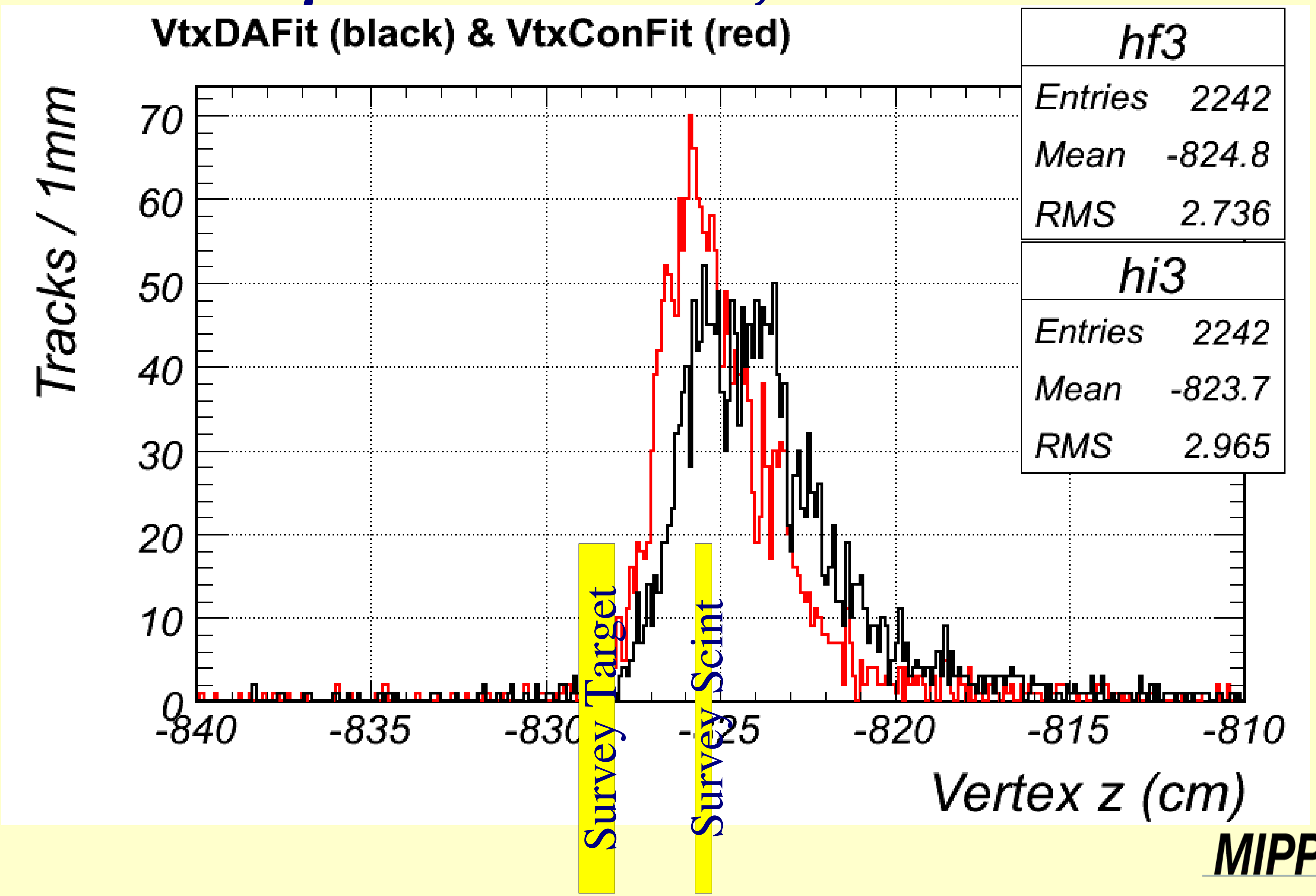
Vertexing

Sample event. All tracks in red are associated to common vertex marked by circle



Target Z with Vertex Fits

pC at 120GeV/c, >3 track vertex



MIPP

Kalman Track and Vertex Fitting

- Advantage of Kalman Filter: take into account multiple scatter and dE/dx in track fits.
- We have been using the RecPack toolkit (used by Nomad, HARP, MICE, etc.)
  - Pros:
    - Convenience of using off-the-shelf package, don't have to reinvent the wheel.
    - We had relatively rapid start up with this package.
  - Cons:
    - Still under active development.
    - Lack of documentation.

Kalman Status

- At this point, we can:
  - Refit TPC helix tracks using the field map
  - Refit beam tracks
  - Form vertices near the target
- Need to:
  - Move to standard RecPack release
  - Figure out how to swim TPC tracks downstream and pick up DC hits
  - Find secondary vertices
  - Add more geometry (eg, ToF wall)

Particle Identification

As tracking algorithms converging, it is possible to calibrate the responses of the particle ID detectors and the ECAL and HCAL.

PAC has seen the dE/dx plot in the TPC. Gain calibration there is done, major gains to be made there are in momentum resolution (fine tuning of TPC distortion corrections, vertex constrained fits, and wire times)

Here I add some plots from the TOF and DCKOV, and RICH counters

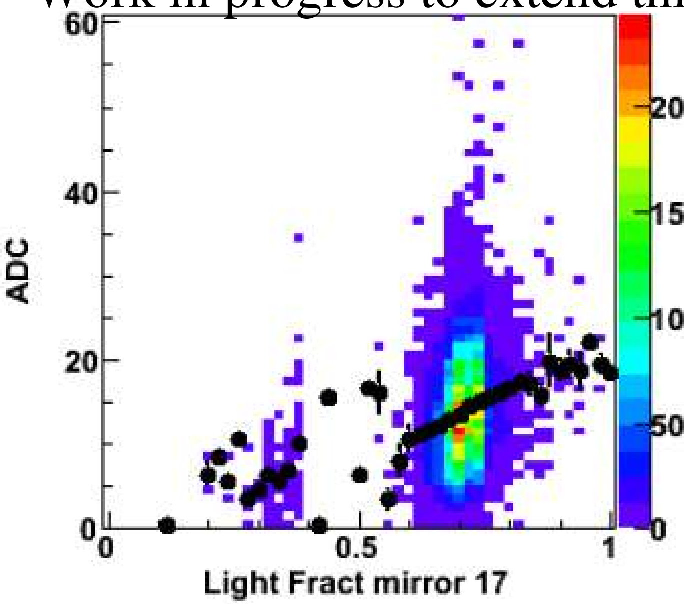
The calorimeter calibration was presented

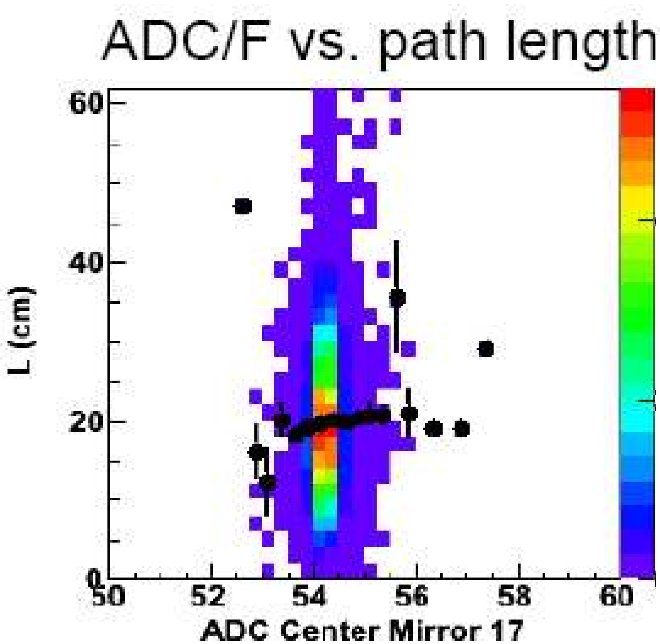
DCKOV

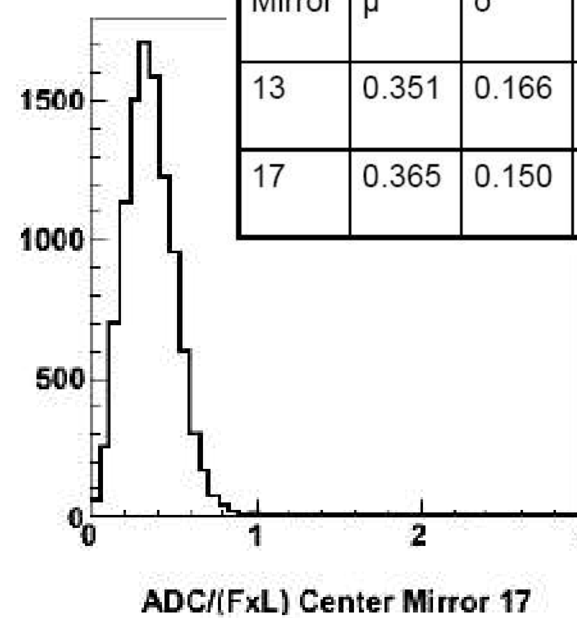
Calibration of inner two mirrors complete. These have the highest statistics. Extending calibration from these two central mirrors to the side mirrors will require much higher statistics available after our next reconstruction pass at the data

Plots below show the correlation of the mirror ADC value with the predicted light fraction predicted from the track fit. Center plot shows the (slight) correlation with ADC on the predicted path length through the mirror. Right most plot shows the ADC spectrum corrected for light-sharing and path length. #pe's found from (mean/sigma)^2

Work in progress to extend thi







NPE~$(\mu/\sigma)^2$

Mirror	μ	σ	NPE
13	0.351	0.166	4.47
17	0.365	0.150	5.93

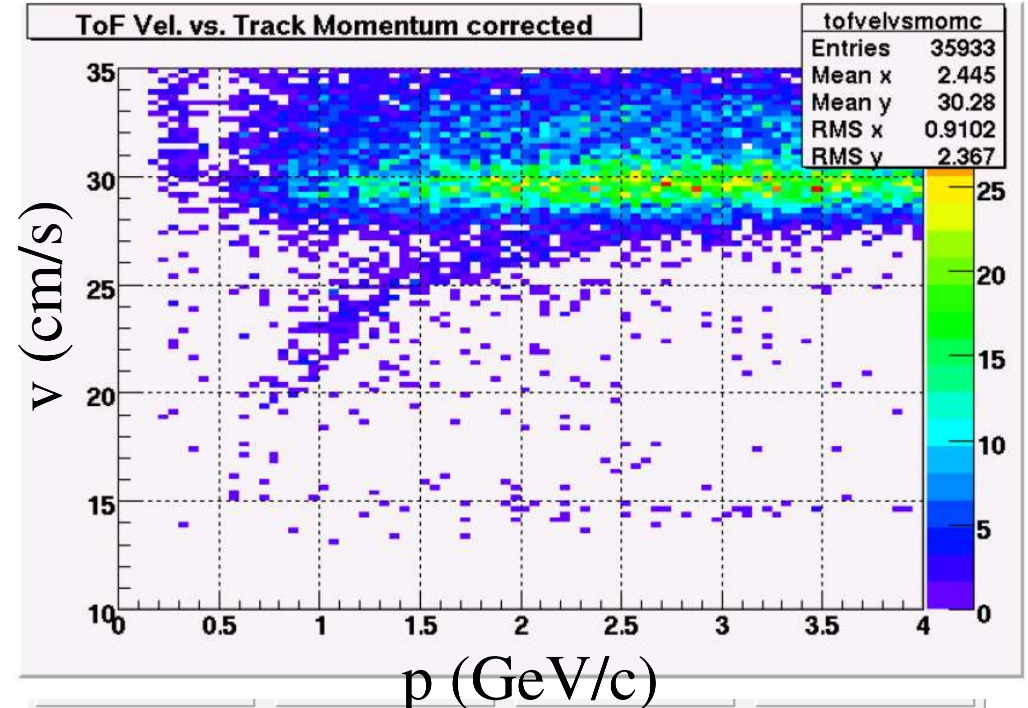
Time of flight

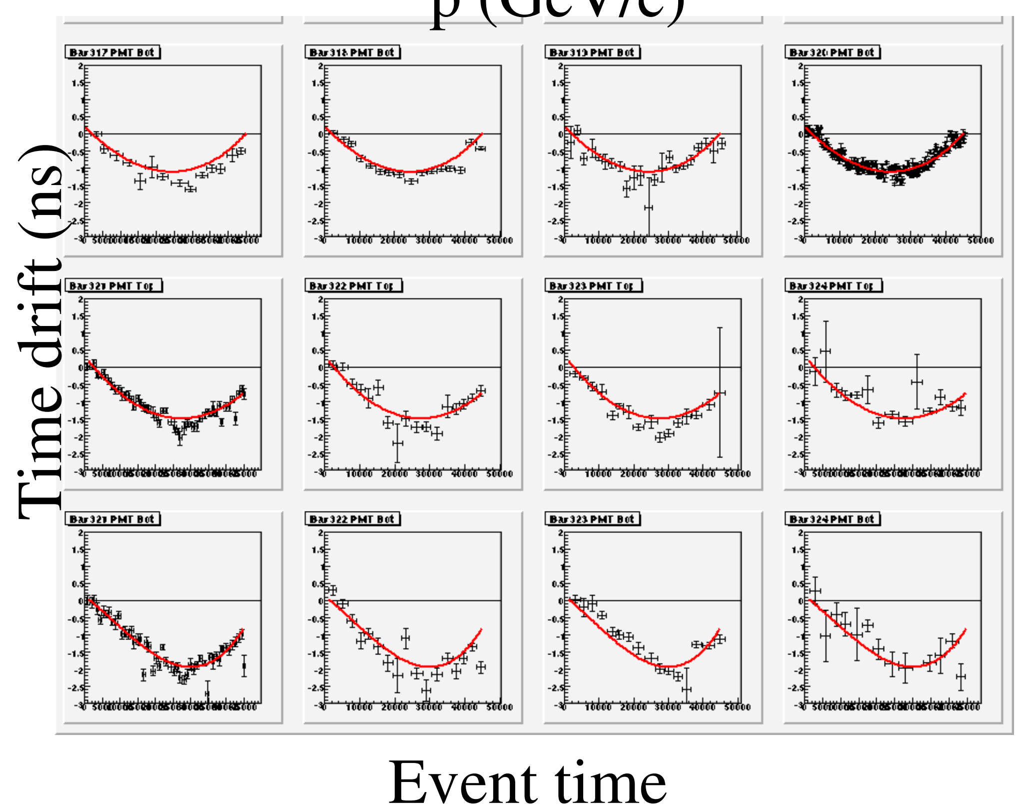
Able to associate tracks to hit ToF bars.

Top right plot shows proton band separate from v=c tracks.

Calibration very sensitive to temperature variations. Calibration procedure to remove these drifts has been worked out. (See bottom plot) and work is continuing on the remaining corrections (trigger time offsets, time walk, speed of light in bar, etc.). For bars on wings will require large statistics resulting

Results





RICH Reconstruction

Sin Man (Sharon) Seun

Dec 9, 2006

- Reconstruction Algorithm
- RICH Reconstruction Performance for p/K only

  MC p-C run: 10,000 events

Reconstruction Algorithm

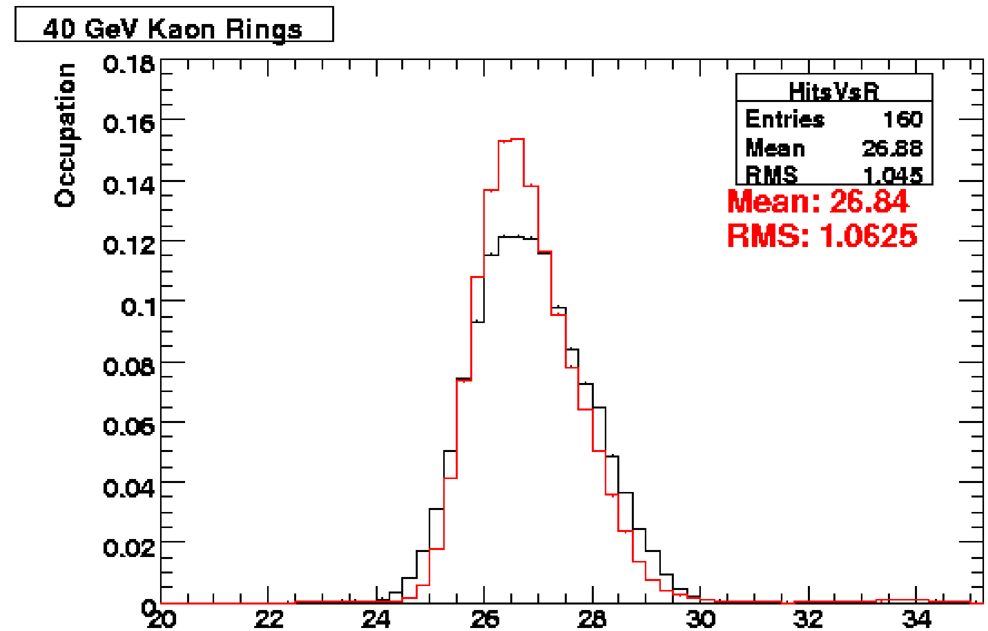
- Get global track Info
  - momentum & ring center
- Analysis
  - Given a momentum, define signal region: smallest r_p to largest r_e. If proton is below threshold, $r_p = 0$
  - For each particle hypothesis j, compute the expected number of photoelectrons n_i for every PMT i in the signal region and calculate Log(likelihood) L_j
  - Fit ring with reconstructed particle ID radius
    - For all good fits, eliminate overlapping digits
  - Redo likelihood calculation

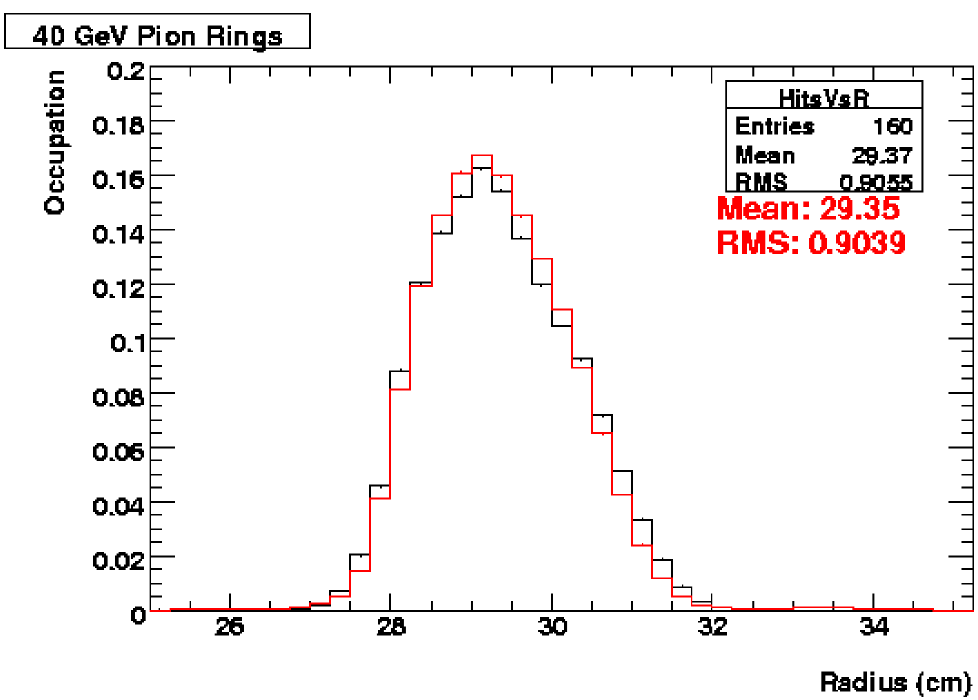
Tuning expected light in RICH (Nick Graf)

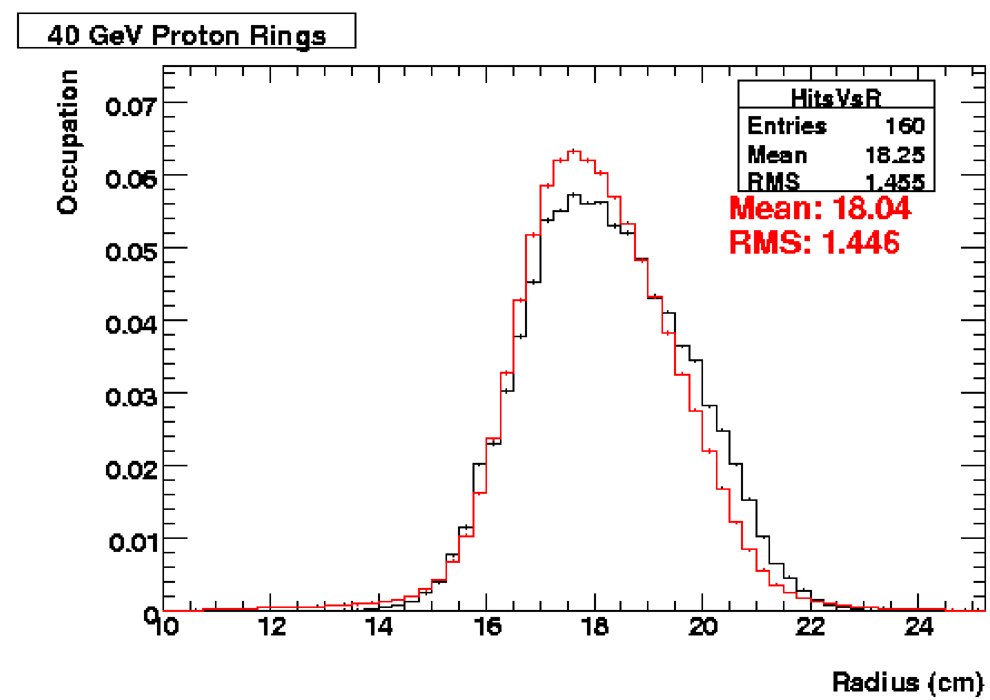
PMT occupancy as function of distance from ring center

Red: Calculation

Black: Data

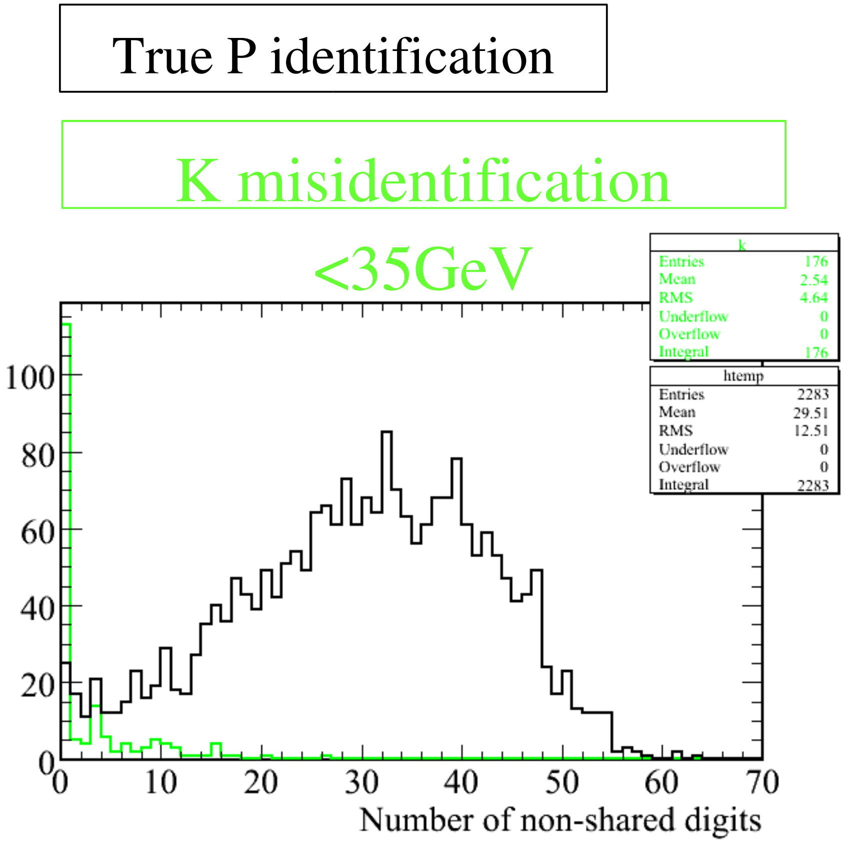


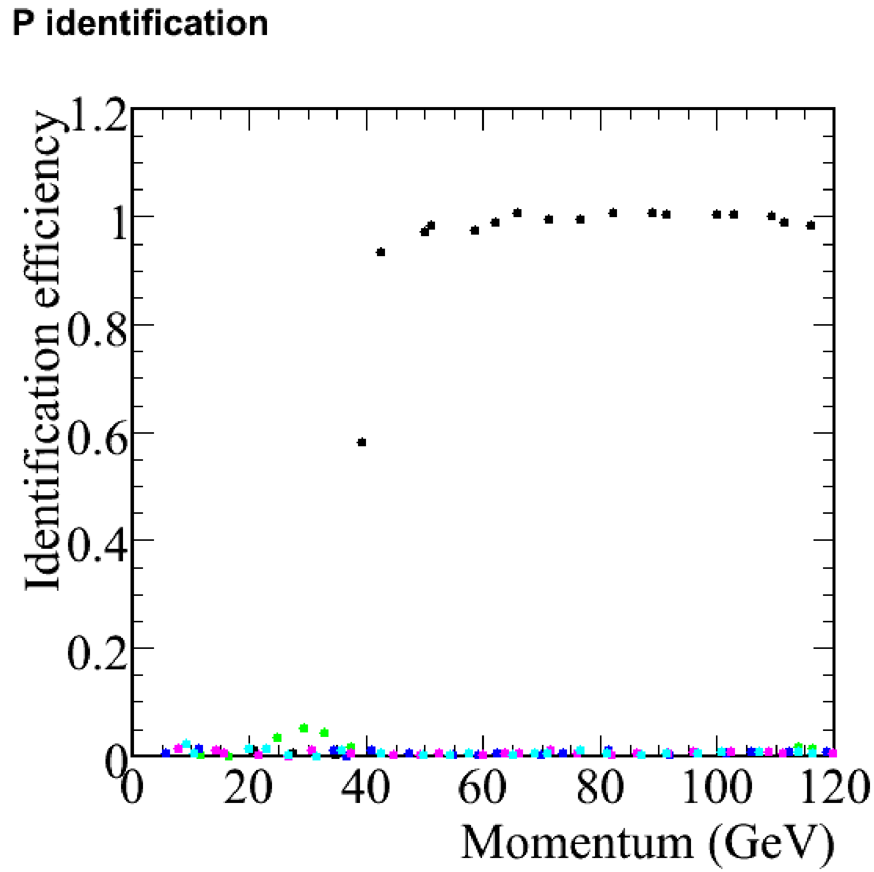




True p Identification

Efficiency of p identification vs momentum for (number of non-shared



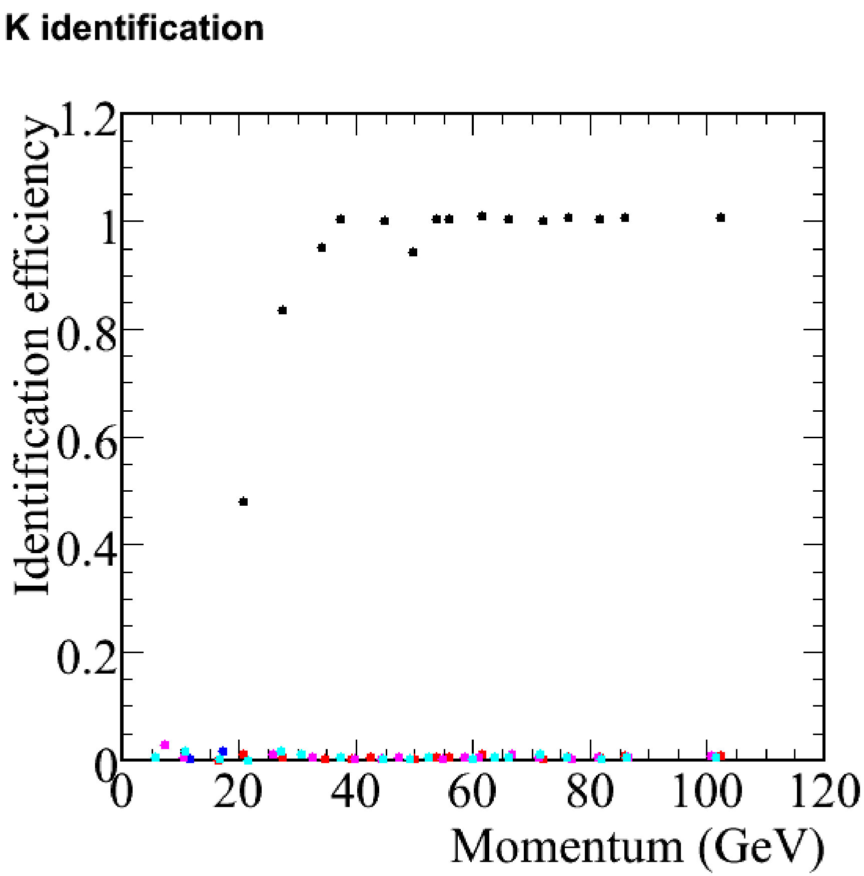


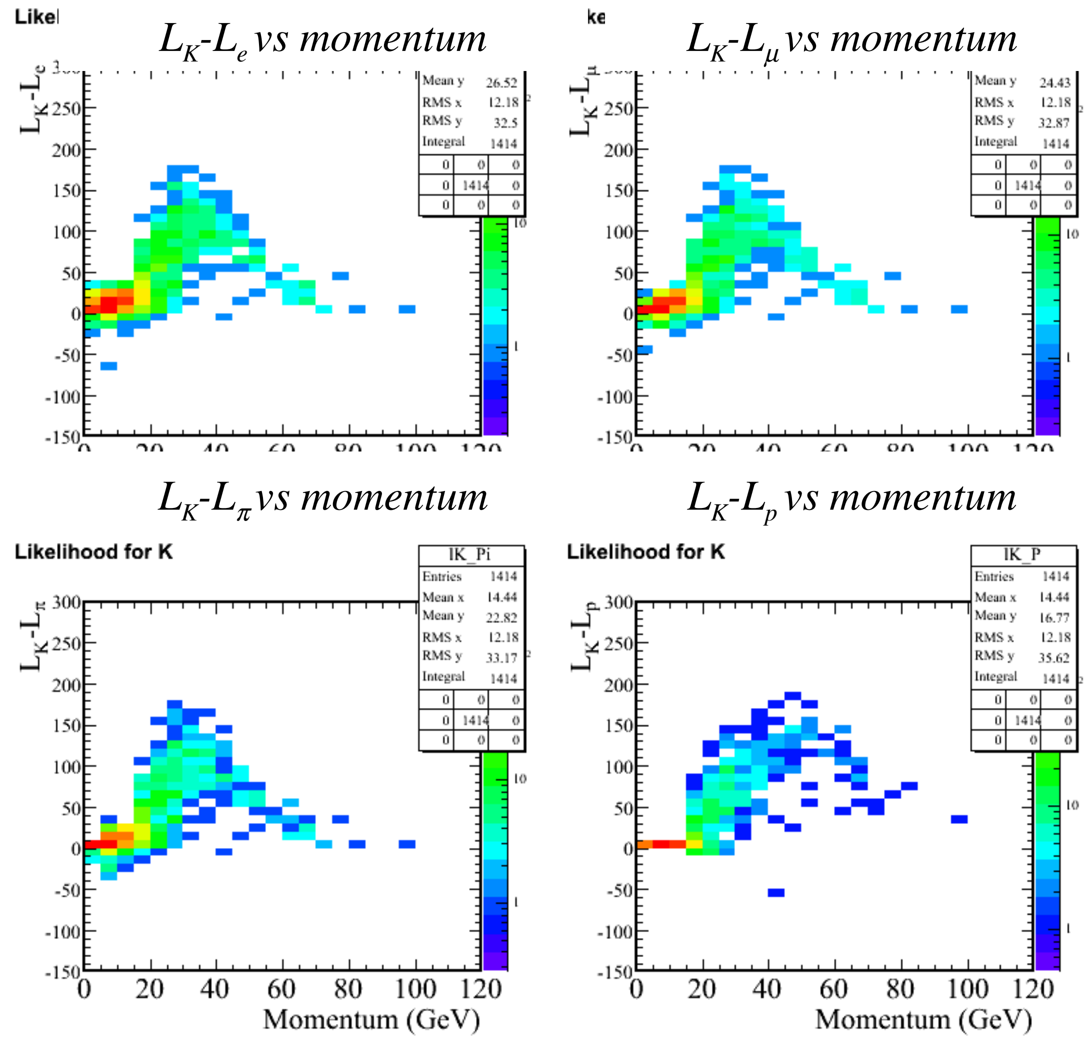
- Most misidentified particles have 0 non-shared digits
- Big improvement after cut on number of non-shared digits

True K Identification

RICH Reco PID
Electron
Muon
Pion
Kaon
Proton

Efficiency of K identification vs momentum





Monte Carlo Status

- GEANT side is nearly complete. Need to
  - Configure target location, etc based on real run
  - Compress calorimeter hits
  - Add Scinteraction counter to active volumes
  - Perhaps a few more little technical details
- Particle production is simulated outside of GEANT
  - Fluka, dpmjet are ready to go
- Within 2-3 weeks could start generating GEANT hits

MIPP

Monte Carlo Status (cont)

- Digitization of idealized hits is done in C++ JobCModule's
  - Wire chambers are nearly complete
  - TPC is almost there
  - RICH is being tuned
  - Ckov/TOF/Calo's need work
- Making MC resemble our data is not trivial; needs a dedicated effort

MIPP

Thoughts on upgrade

- Focus of offline needs to be analysis of MIPP-I data.
- Expect to start data processing in ~2 weeks. DST's available ~1 month following
- Once that processing has started, focus will shift toward MC tuning

- Places where people can get plugged in:
  - Kalman Reconstruction
  - MC tuning
  - Data cross-checks (invariant mass studies

Thoughts on upgrade

Thinking ahead to the upgrade we will need

[1] Revisit and revise the online monitoring

[2] Identify 1-2 people to form an alignment group

- Need better thought out plan for alignment. For example, magnet off running has been extremely valuable and not plentiful enough in MIPP I analysis
- Need better coordination with survey crews

[3] Identify 1-2 people to form TPC analysis group

- Begin working with Mike Hefner to develop magboltz simulation of distortion effects
- Need better thought out plan for calibration of distortion effects.

Thoughts on upgrade

Thinking ahead to the upgrade we will need

[4] Need calibration plan

TOF/DCKOV in particular are difficult to calibrate with data collected in MIPP-I

[5] Batch processing. We cannot count on LLNL farms and the people managing the batch processing are leaving. Need to identify a batch manager and possibly more computing resources at or away from FNAL

[6] Need to continue reconstruction effort, especially in directions required by new detector systems

[7] Need data handling plan: 1 spill ~= 12,000 events ~= 2+ of what we now call a sub-run. Long subruns have been problematic for us due to data size.