WA105 Data / Online Reconstruction

Vyacheslav Galymov
IPN Lyon

DUNE Software and Computing Meeting
15.12.2015

WA105 Data model overview

Run header

Contains some general information about specific MC/Data file

Run type, beam type (μ, π, …), beam energy, information about parameters used in MC production

Beam info (filled only from test beam data)

Information from beamline monitors

E.g., secondary target type, Cherenkov counter/TOF settings

Integrate beam info into the WA105 data stream

Event info

Event number/UID

Some basic items about the event: e.g., total energy deposited, (EM and hadronic parts)

Particle info from beamline monitors

Raw data

Raw 12 bit ADC data (compressed / uncompressed)

Light readout data

Raw 12bit ADC values from digitized 36 PMTs

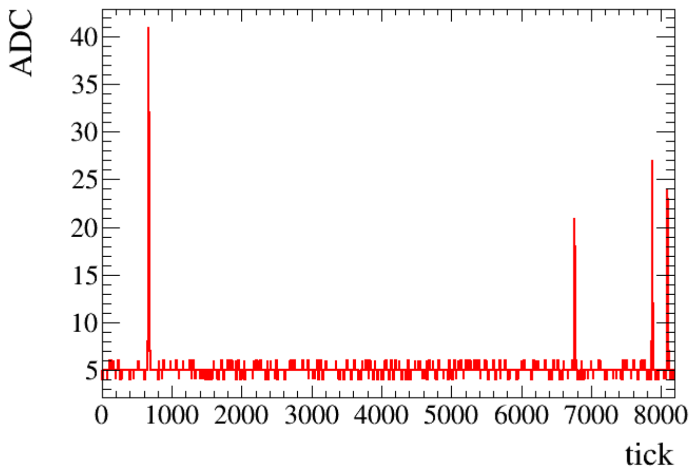
Step (voxel) info (MC only)

Information from GEANT steps

Edep, pid, light, charge at anode, diffusion spread

Raw Data Compression

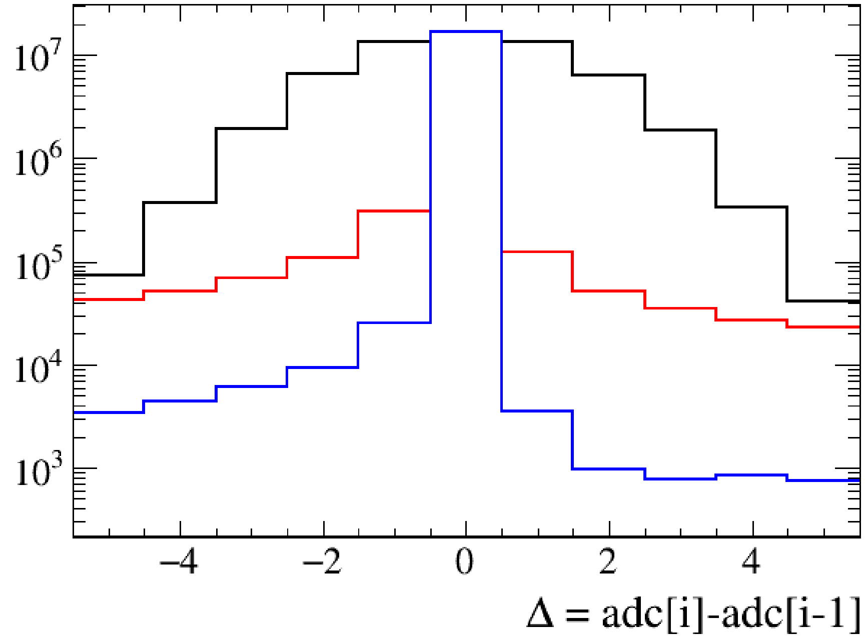
Example of simulated waveforms from one of the channels with cosmic background events (gain per view = 10)




```
Huffman coding scheme:
  00000001    3   8 bits code length
   0000001   -3   7 bits code length
    000001    2   6 bits code length
     00001   -2   5 bits code length
      0001    1   4 bits code length
       001   -1   3 bits code length
        01    0   2 bits code length
 >>-- number of repeated values to encode --<<
         1    4   1 bits code length
```

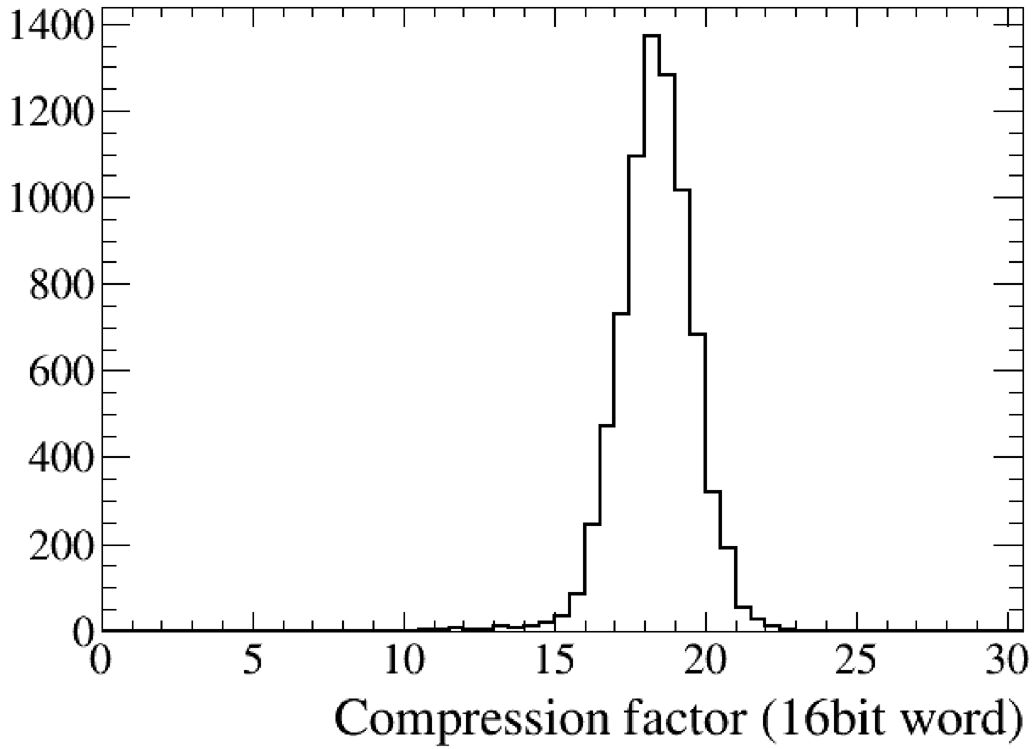
First step of compression → compute Δ

Cosmics w/ noise
Cosmics w/o noise
Single mu w/o noise



Huffman coding scheme to encode $\Delta(\mathrm{ADC_i} - \mathrm{ADC_{i-1}})$ same as in uBooNE

Raw Data Compression



Compression of simulated muon background w/ expected noise

Average compression factor on the 16bit words ≈ 18

Actual compression factor (for 12 bit data) 18 x 12/16 ≈ 14

For noiseless data can get compression factors ~50

As explained by Elisabetta, this is sufficient to store full data “online” for ~week

Online analysis

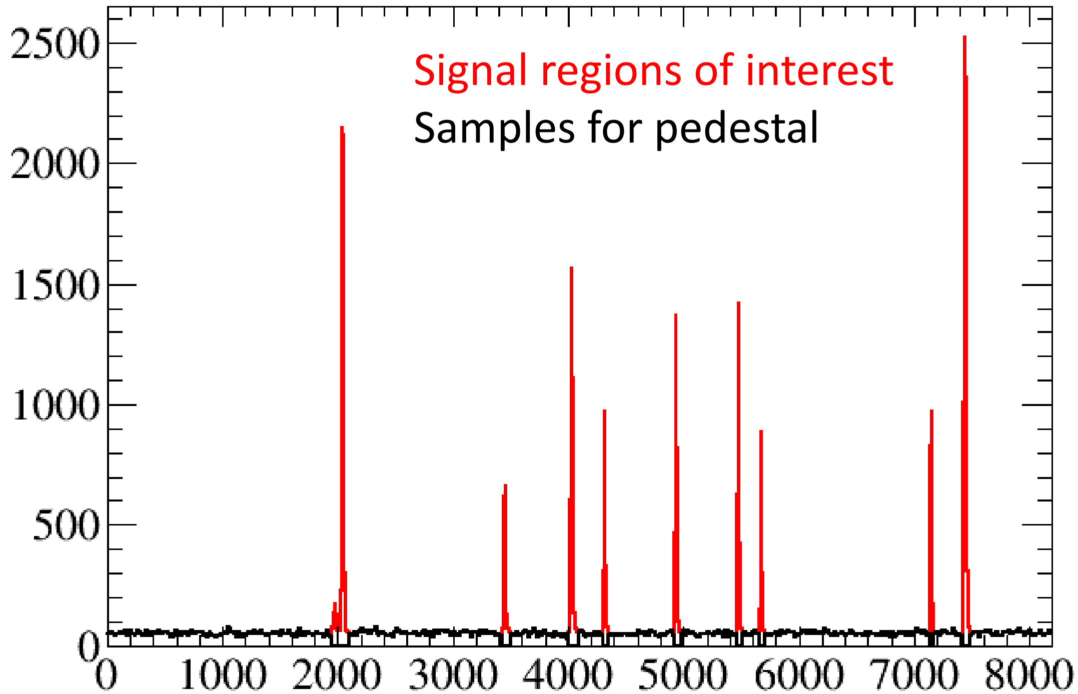
Online monitoring tasks

- Total deposited charge
- Hit rate per channel / view / CRM
- Pedestals
- From cosmic tracks
  - Purity monitoring
  - Gain stability

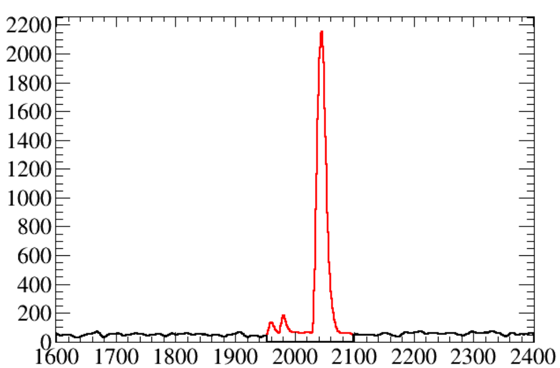
Hit reconstruction (total deposited charge, hit rate per channel / view / CRM, pedestals)

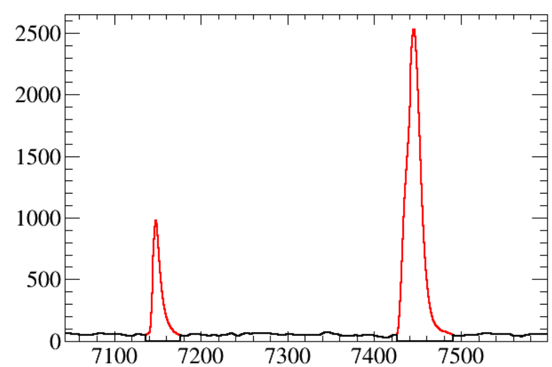
Track reconstruction (from cosmic tracks: purity monitoring, gain stability)

Hit reconstruction



1st step to find region of interest and pedestal for a given channel
In addition some FFT noise filtering can also be performed

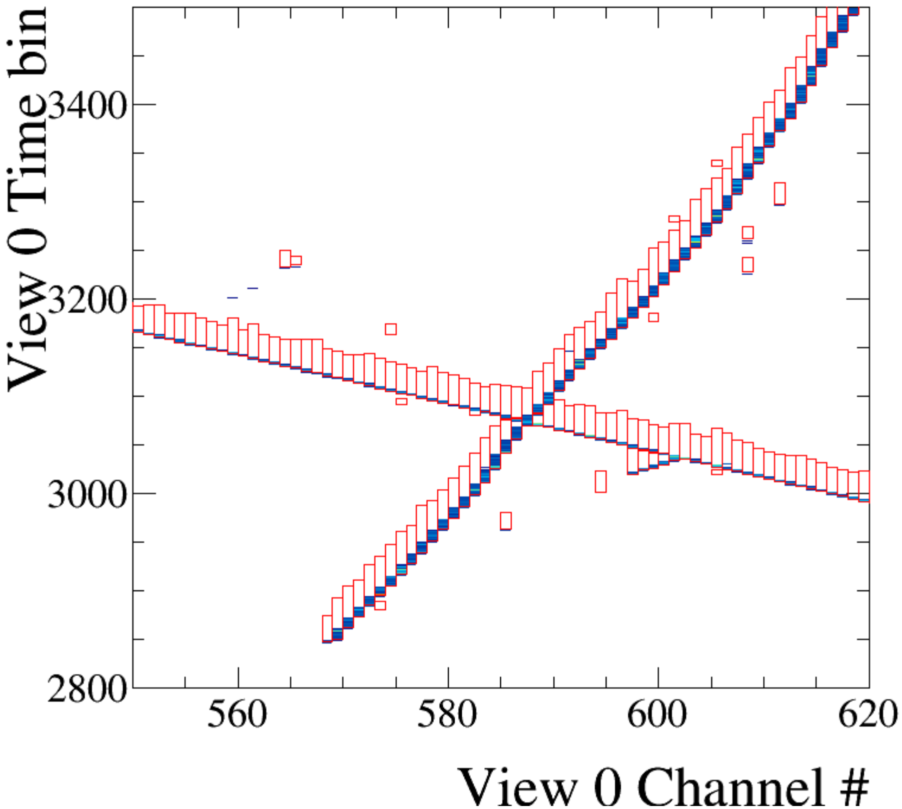




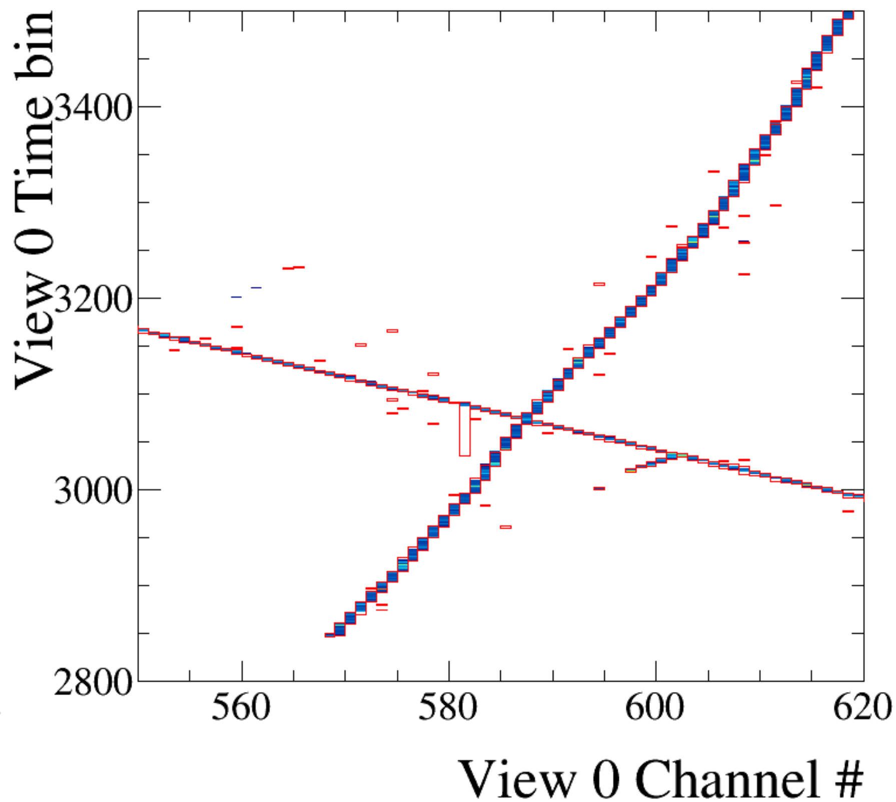
Hit reconstruction

Red boxes are regions of reconstructed hits
Distributions underneath is true charge

AlgoMultiHit



AlgoTDeconHit



Split ROI into several regions if several peaks are detected. Integrate waveform in each region to find charge. ← simple & fast : suitable for online processing

Deconvolve true charge distribution from response using least-squares fit ← slow (good for specialized offline studies e.g., diffusion)

Track reco & cosmic selections

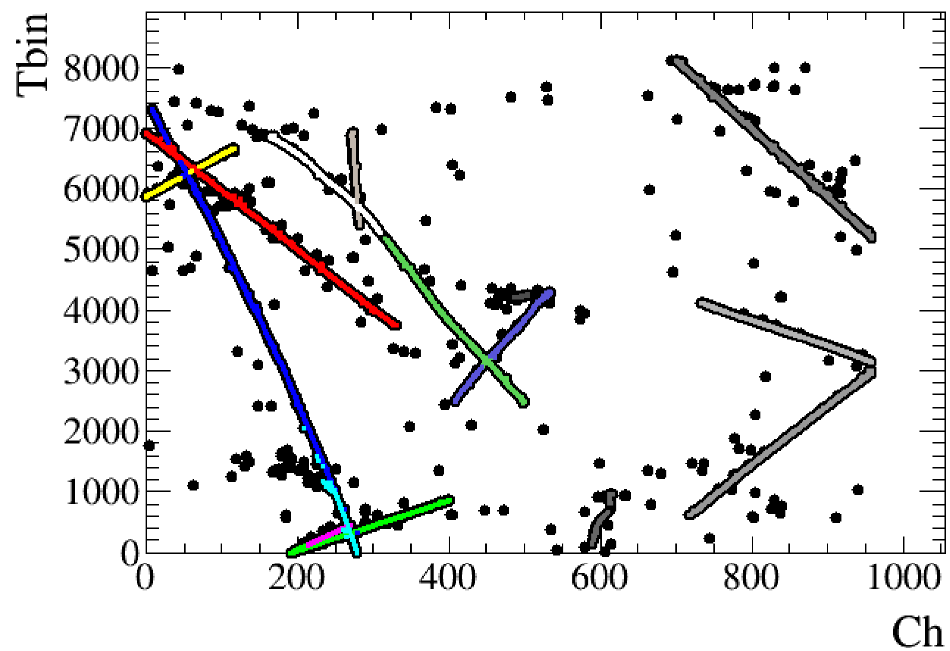
Many track segments from chopped on-surface cosmic ray background
In addition muon halo from pion decays in the beamline

Hough transform (HT) for global search for track candidates

Post HT processing:

- Identify & merge track segments which are part of a single track
- Resolve hits assigned to overlapping tracks / regions
- Tag delta rays

Different colours mark hits associated with track candidates in a given cluster



Simple selection (start & end points) of a few long cosmic track segments for online analysis

Summary

- Plan to take full data without zero-suppression
  - With Huffman coding scheme possible to get ~14 compression
- Online computing farm to monitor detector performance
  - Fast reconstruction needed
    - Hits
    - Tracks and cosmics selection
  - Analysis of purity & gain stability