SBS software, simulation/analysis interface, tracking

Eric Fuchey
University of Connecticut

SBS winter '18 collaboration meeting
Jefferson Lab, February 23-24, 2018

UCONN Jefferson Lab

Overview

SBS Software:

- Introduction: project overview and milestones.
- Simulation digitization;
- Interface with analysis;

Progress on subsystems:

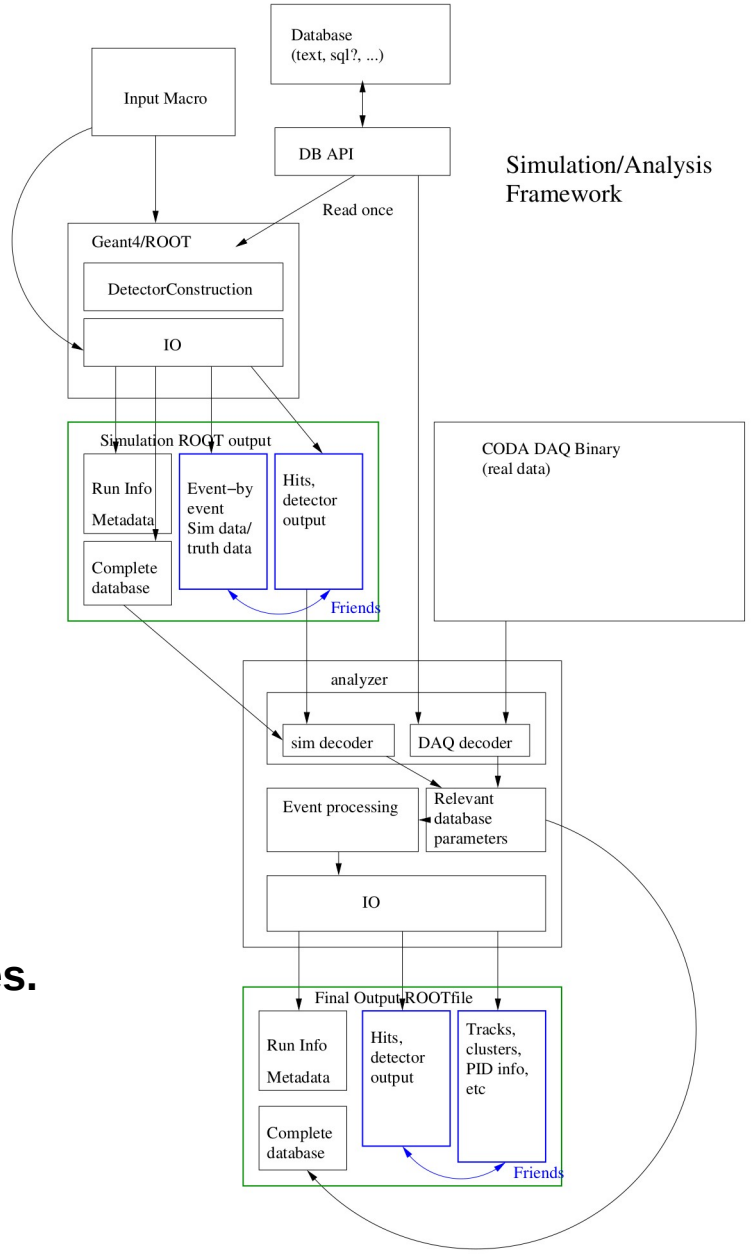
- GEMs / Tracking;
- GRINCH;
- HCal;
- Summary and next steps.

Summary

UCONN Jefferson Lab

SBS Software:
Project overview (see Seamus' slides)

*** Major goal: "End-to-end" simulation:**
production of pseudodata + simulation of data sizes;
***** Both simulation and analysis framework need to be:
→ *modular* (ease configuration changes);
→ *accessible* (ease handling for new people);
→ *flexible* (ease inclusion of new configurations);
***** Also need:
- Well defined IO formats and standards
- Flexible database to accomodate both MC and data (SQL ?);
***** Requires significant coordination between working subgroups
→ 1 dedicated software meeting every 2 weeks
(in addition to SBS weekly meeting).
+ About to migrate to e.g. Redmine for project management
**** Well defined responsibilities and milestones*** (next slide)



Strong requirement:
Online and offline analysis both need to be ready and tested, and pseudo-data sets have to be analyzed before data taking
(likely spring/fall **2020**).
=> critical given high luminosities / high detectors and DAQ rates.

G4SBS simulation is up and running, and has already produced useful and compelling results (e.g. G_M^n ERR last summer *).
We have since then set our focus on its interface w/ analysis.

(***** see my slides from Summer '17 Hall A/C collaboration Meeting)

SBS Software:
Project milestones (see Seamus' slides)

Rebaselined milestones, taken into account accelerator planning change, actual progress on software and manpower:

- Nov 2016: Software review
- Jan 2017: Start digitized simulation output (first focus on GEMs)
- Apr 2017: Decoder for all DAQ modules written
- Jul 2017: Each detector system in analyzer, experiment configurations, basic reconstruction algorithms
  => can analyze channel-level raw data at this point
- **Jun-Jul 2018**: **Simulation interfaced to analysis**; Have detector event displays, calibration scripts
- **Jul 2018**: Start simulated analysis for detector reconstruction
- **Jan 2019**: Begin simulated experimental analysis for core Form Factors (FF) experiments
- **Jun 2019**: Ready for beam for FF experiments, start simulated expetimental analysis for SIDIS and TDIS

- Spring 2020: likely earliest start of neutron experiments
- Spring 2021: likely earliest start for Gep

UCONN Jefferson Lab

SBS Software Project:
Simulation digitization

Purpose: Convert simulation files produced by G4SBS to files containing digital information in a similar format as ADCs and TDCs

G4SBS file:
* Global channel number (and position, for GEMs);
* Energy deposit / Number of photoelectrons (for PMT like detectors);
* Time;
* Other info which may not be important directly for our purpose;

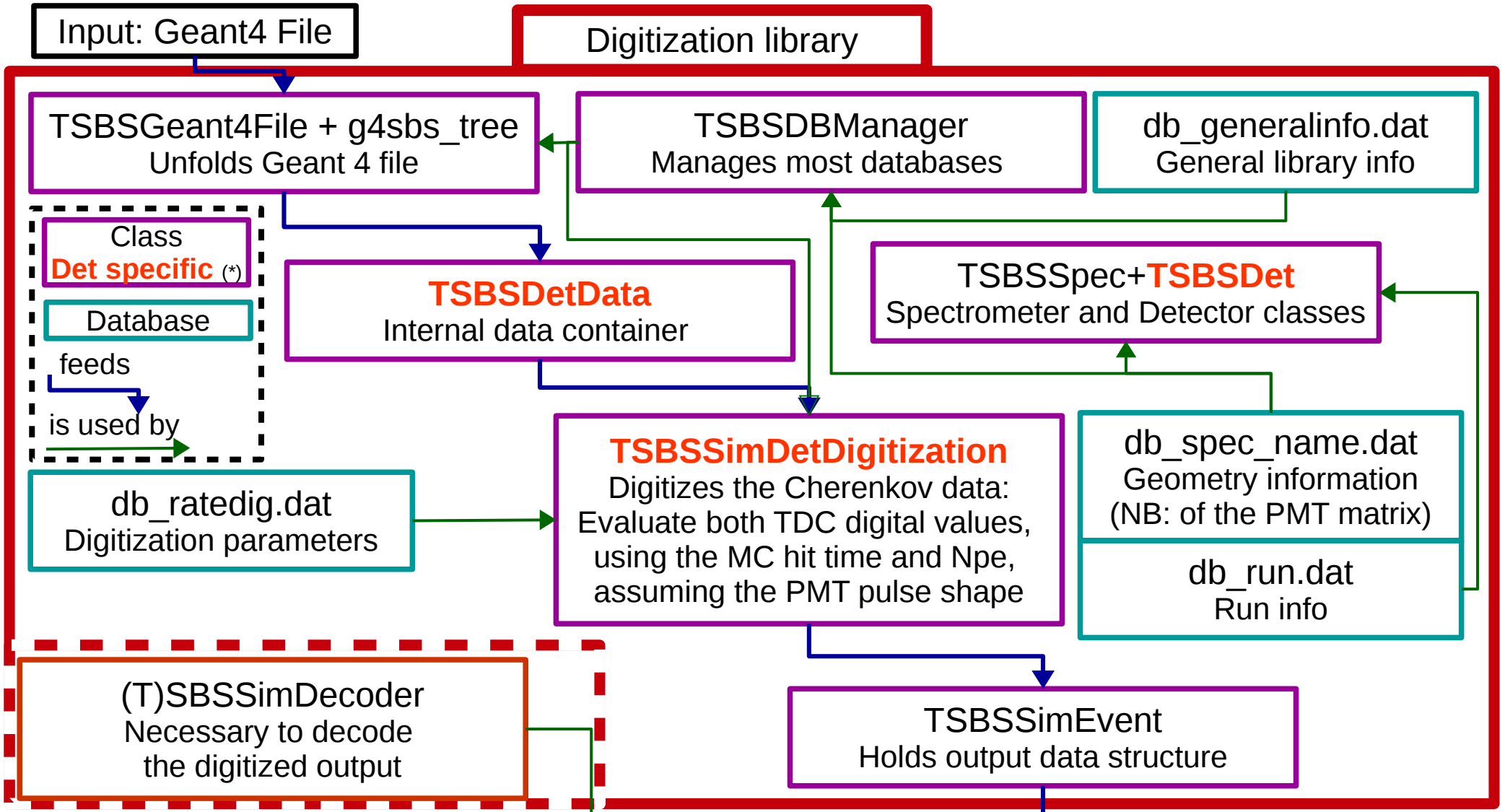
Digitization: introduction of electronics imperfections:
intrinsic resolutions, pedestals, etc;
Addition of pile-up (background);

Digitized file:
* Module number, local channel number;
* ADC;
* TDC.
* Other info which may be important for MC/data comparison;

Analyzed file:
* Global channel number;
* Energy deposit / Number of photoelectrons (if PMT like detectors);
* Timing.
* "Advanced" reconstruction objects and variables:
Clusters, Tracks, momentum, PID, etc.

UCONN Jefferson Lab

SBS Software Project:
Digitization: library Class structure



(*) : may inherit, but a template must be written

SBS Software:
Digitization Interface with analysis

Purpose: Convert simulation files produced by G4SBS to files containing digital information in a similar format as ADCs and TDCs

G4SBS file:
* Global channel number (and position, for GEMs);
* Energy deposit / Number of photoelectrons (for PMT like detectors);
* Time;
* Other info which may not be important directly for our purpose;

Digitization

Digitized file:
* Module number, local channel number;
* ADC;
* TDC.
* Other info which may be important for MC/data comparison;

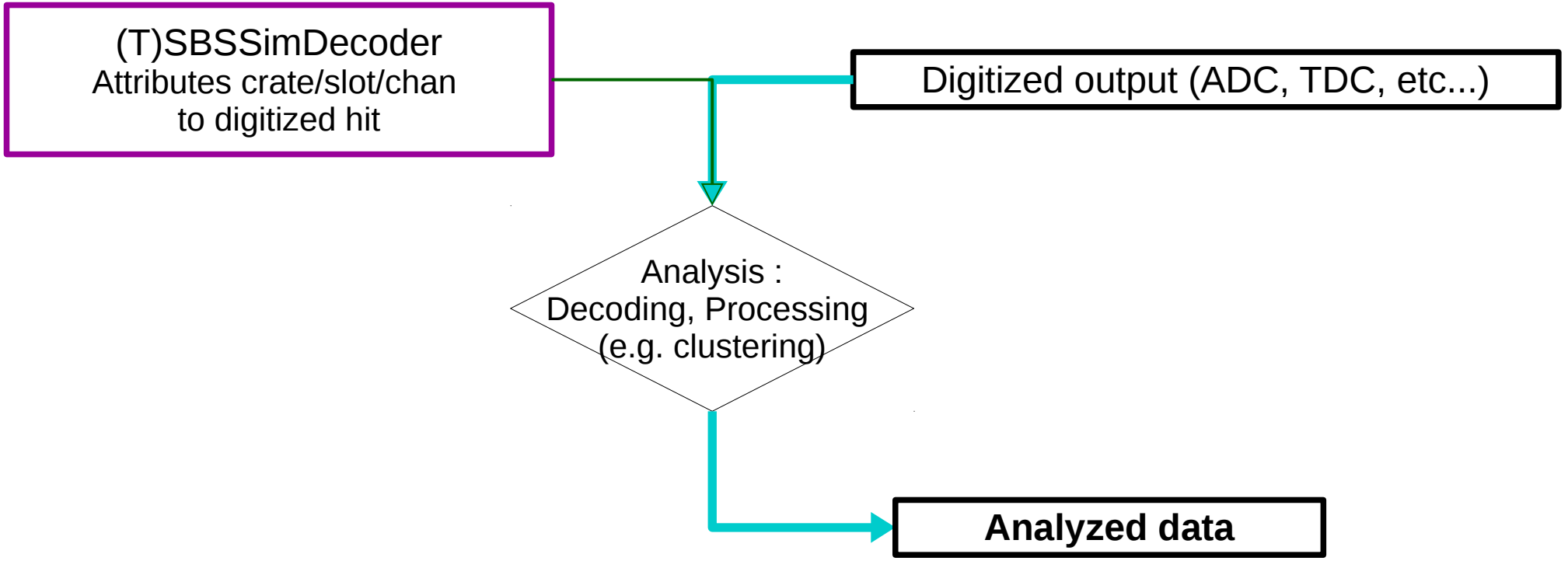
Interface with analysis: assignment of crate/slot #, decoding, processing (reconstruction)

Analyzed file:
* Global channel number;
* Energy deposit / Number of photoelectrons (if PMT like detectors);
* Timing.
* "Advanced" reconstruction objects and variables:
Clusters, Tracks, momentum, PID, etc.

SBS Software:
Digitization Interface with analysis

Required for the interface:

* (T)SBSSimDecoder (inside or outside the Digitization library);
* Databases :
  - global parameter databases (theta, pcentral);
  - virtual cratemap (db_sbssim_cratemap.dat);
  - detector geometry database (in a format supported by analyzer 1.6, etc...)
* An analysis script which puts everything altogether ;



Overview

SBS Software Project:

- Introduction: Overview, Milestones.
- Simulation digitization;
- Interface with analysis;

Progress on subsystems:

- GEMs / Tracking;
- GRINCH;
- HCal;
- Summary and next steps.

Summary

Progress on subsystems:
GEMs / Tracking (Danning + myself)

Reminder about GEM trackers :

Gas ionizing detectors

GEM readout by 400 µm strips

=> 40.96 x 51.20 cm^2 GEMs = 1024 x 1280 strips;

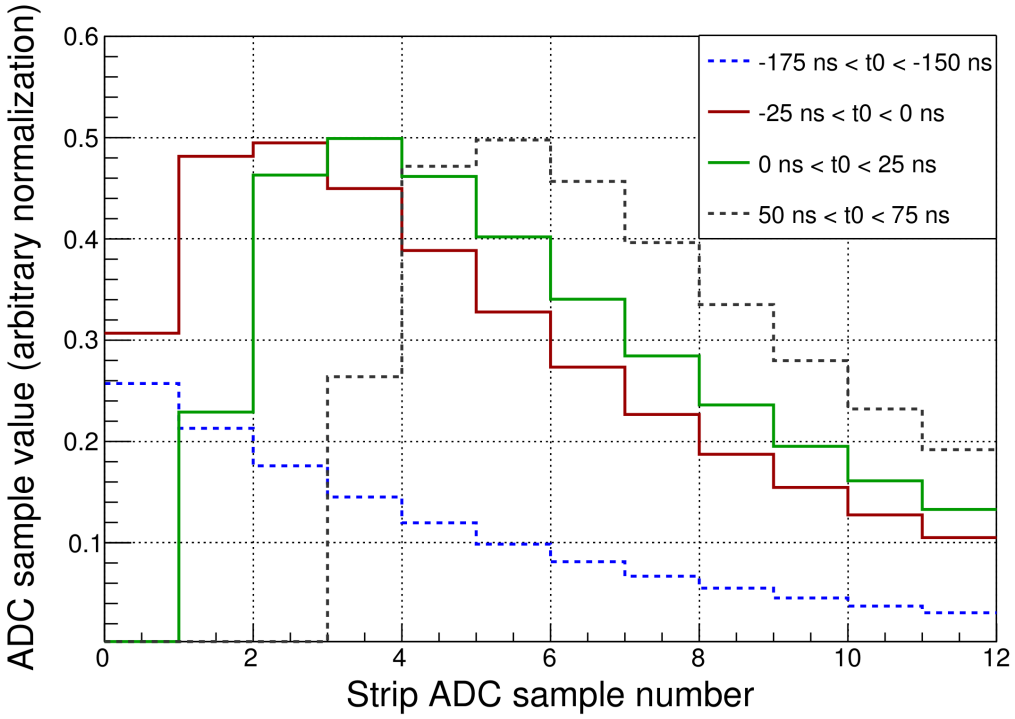
The strips will be recorded by APV25 (128 strips/chip)

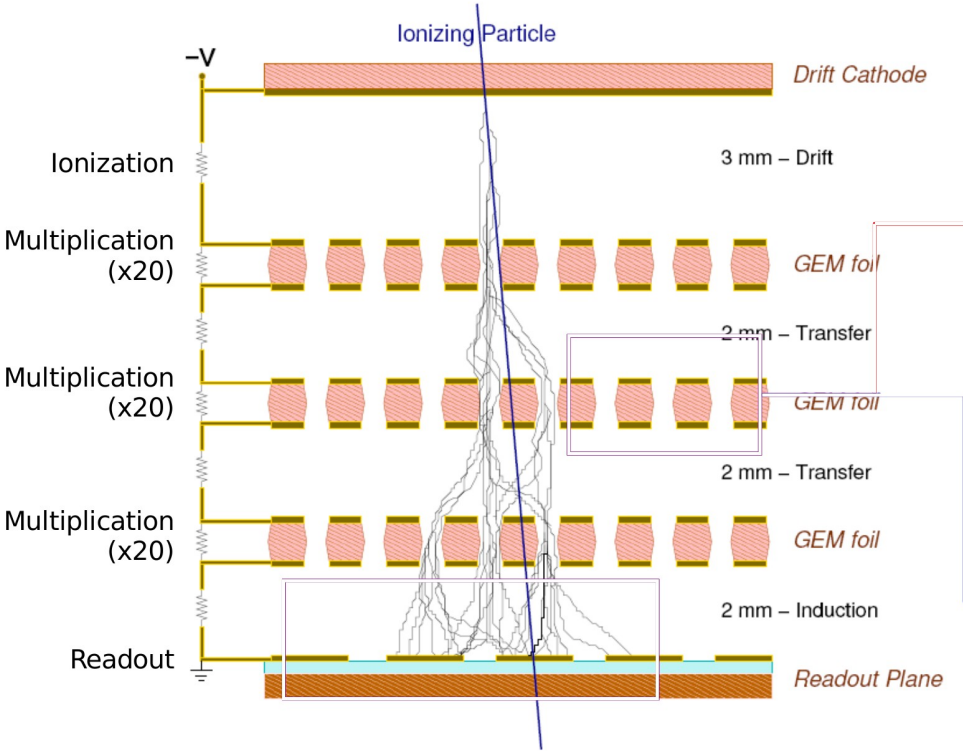
Each APV 25 has a pedestal shift (common mode).

6 ADC samples of 25ns each are recorded for each strip

→ pulse shape;

Of course, 1 hit spreads on many strips => clustering





UCONN Jefferson Lab

Progress on subsystems:
GEMs / Tracking

Digitization: basic idea

Use Energy deposit and Position of the hits

* model avalanche with energy deposit;

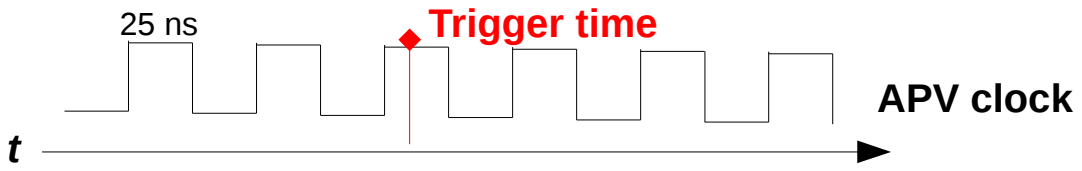
* for all X, Y strips on which avalanche spreads, calculate the 6 ADC samples as function of hit time and fraction of avalanche energy;

Refinements for electronics :

* Pedestal noise + Common mode;

* Cross talk;

* APV clock / trigger time not synchronized together: needs to be taken into account.

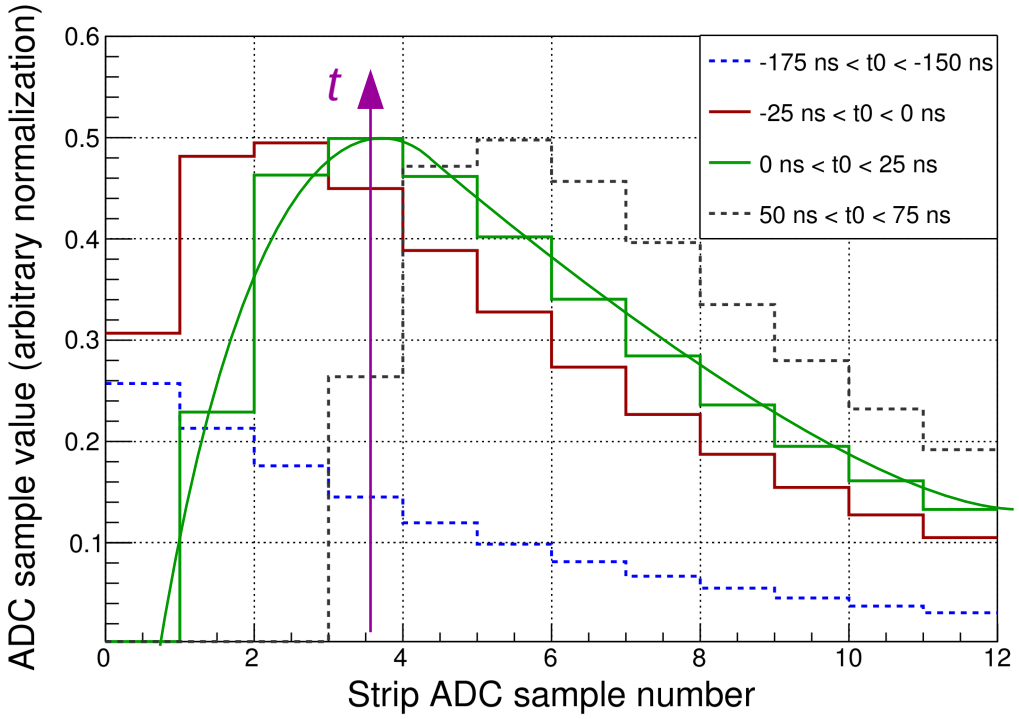


UCONN Jefferson Lab

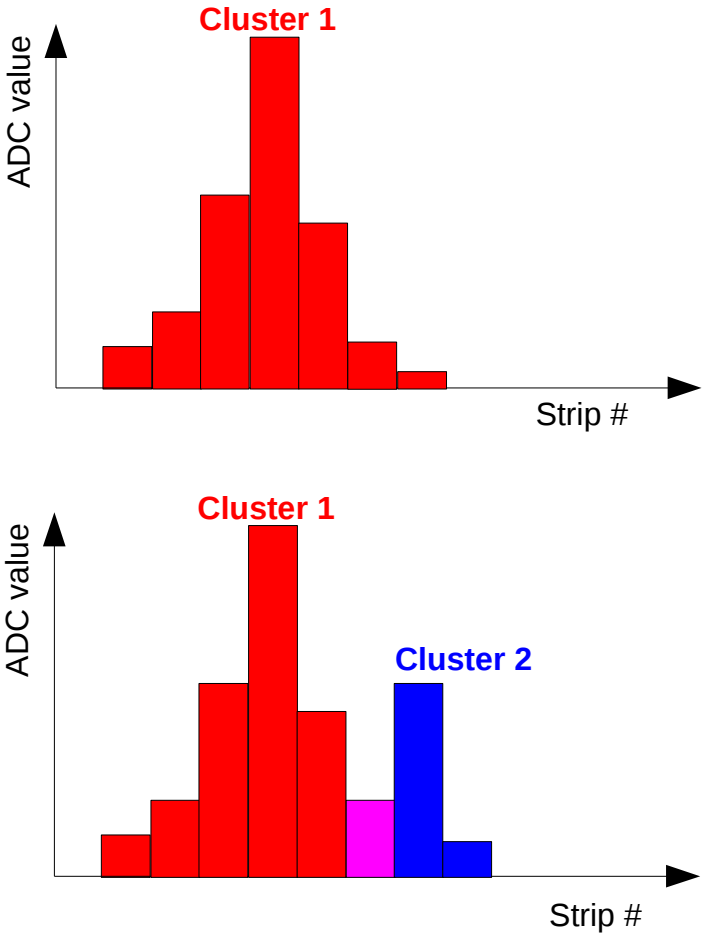
Progress on subsystems:
GEMs / Tracking

"Reminder" about GEM analysis:

Strips samples readout:
- Strip ADC estimated with sum of ADC samples **or** ADC deconvolution;
- pulse shape fit for time information;



Clustering:
Cluster : group of adjacent firing strips.
If peak-valley-peak configuration : cluster split.
In practice : cluster timing:
fit of the sum of the ADC samples of all strips

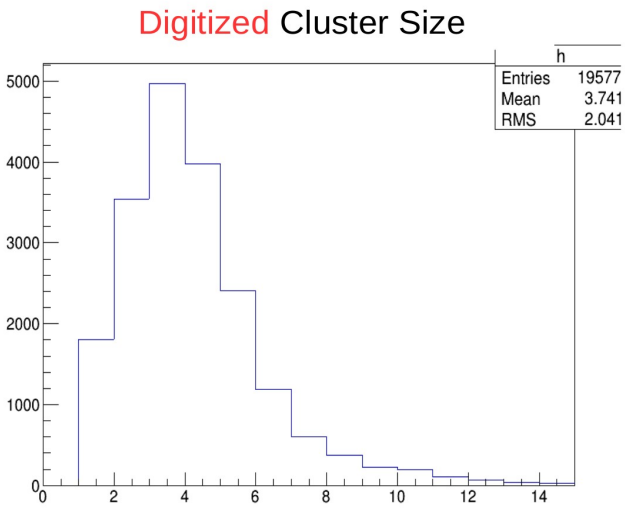


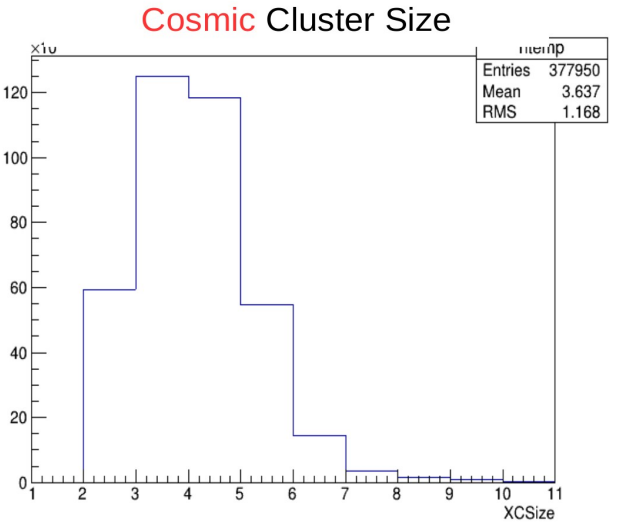
Progress on subsystems:
GEMs / Tracking

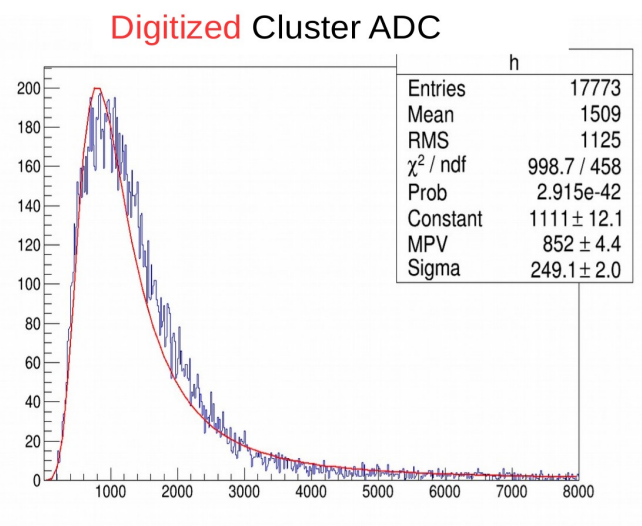
Tuned digitization algorithm and database to match ADC spectrum and hit size with cosmic data;

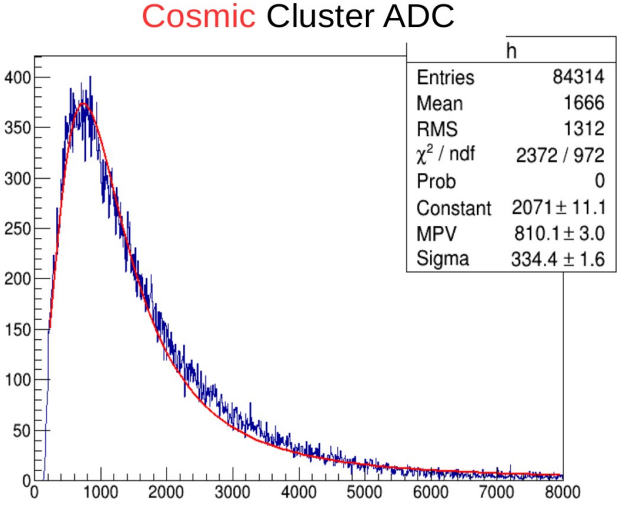
(right now, we use simple sum of ADC samples for strip ADCs, but deconvolution is implemented)

See Danning's slides



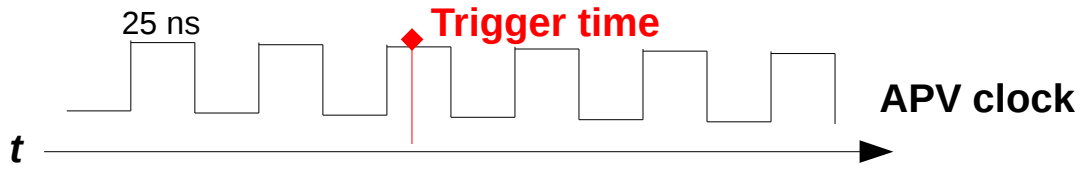






Progress on subsystems:
GEMs / Tracking

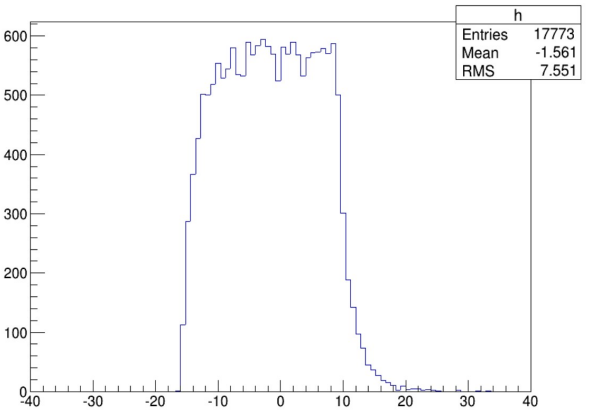
Added a 25 ns APV-trigger "jitter" to the strips timing, to translate that APV chip and trigger time are not in sync;



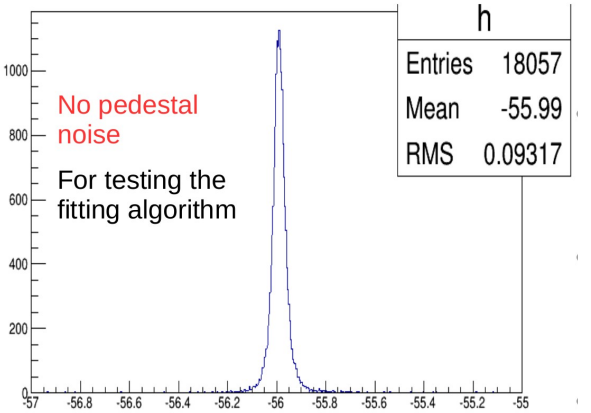
Impletemented pulse shape fitting to determine the timing for each strip;

See Danning's slides

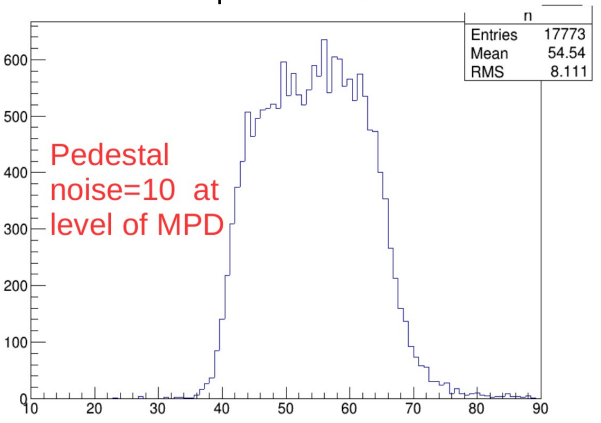
True(MC) Cluster peak time



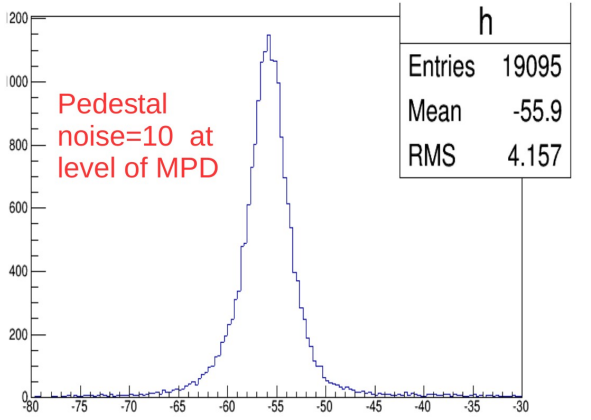
Peak time difference between fit and true time



Cluster peak time from fit

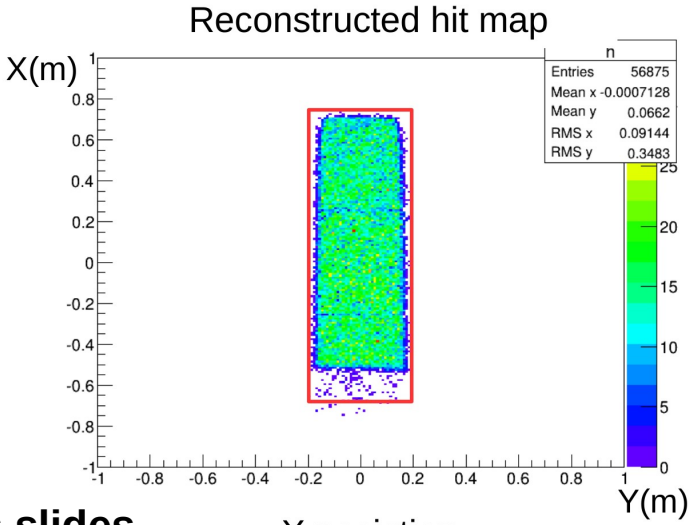


Peak time difference between fit and true time



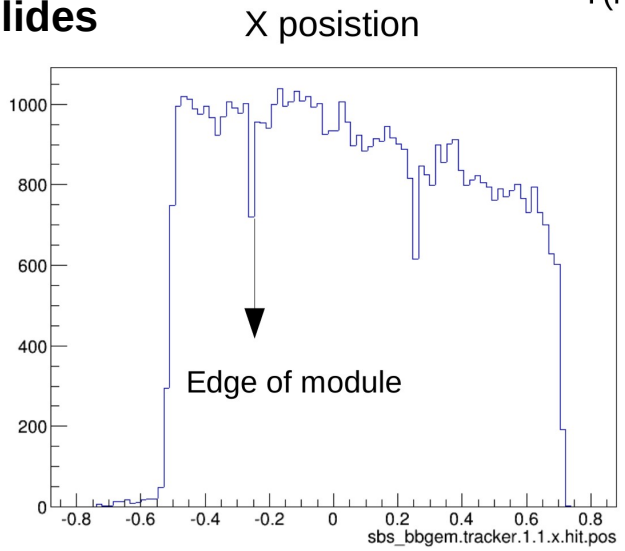
Progress on subsystems:
GEMs / Tracking

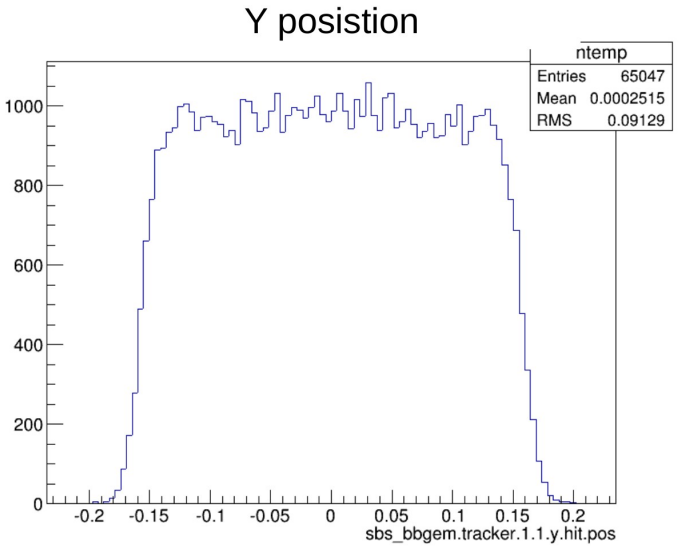
Cluster reconstruction position



- INFN Plane size: (3x50cm)x40cm
- Edge of modules can be seen due to lost strips at the edge of module

See Danning's slides



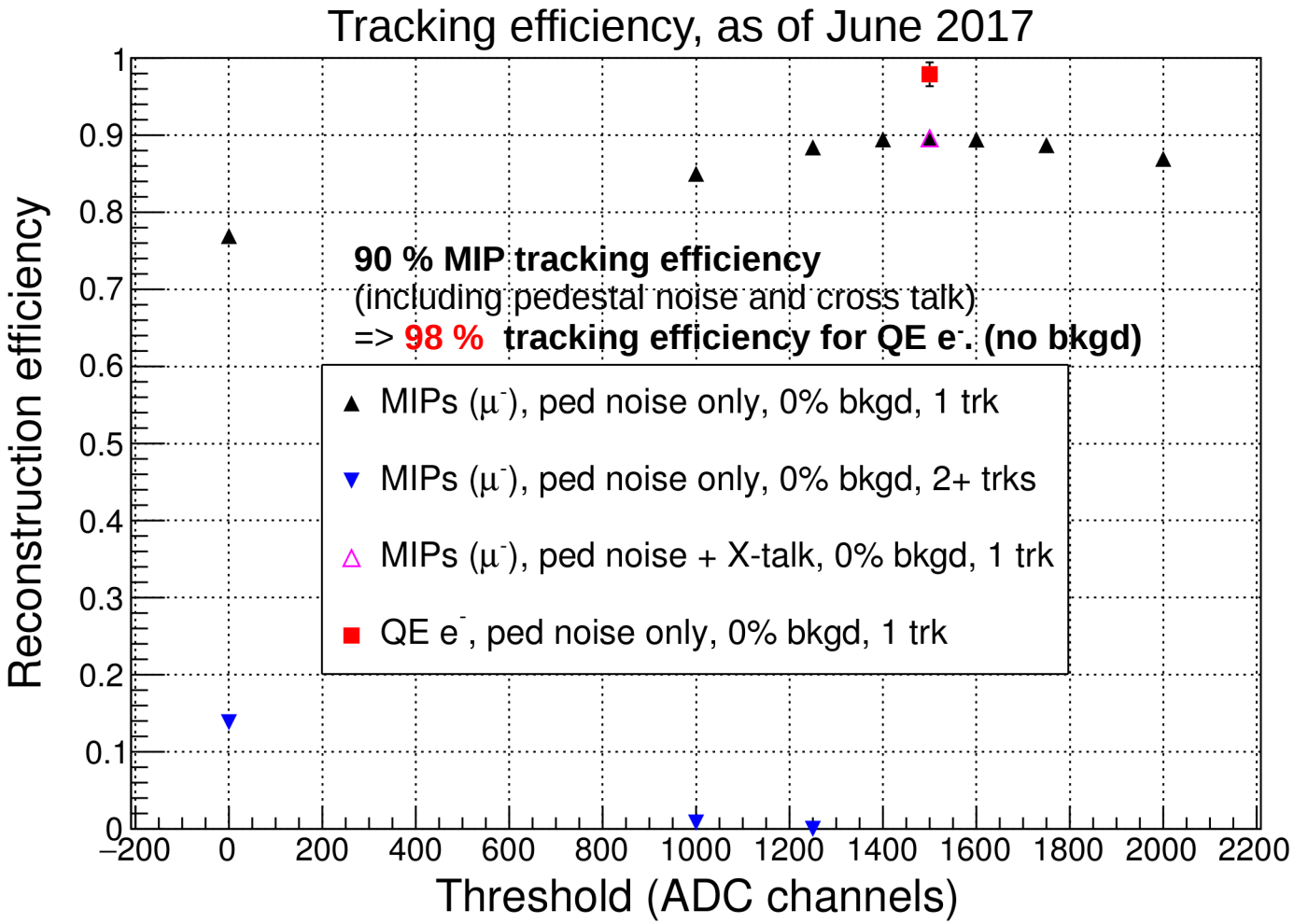


(Plot credit: **Danning**, SBS weekly, 2017/11/15)

Progress on subsystems:
GEMs / Tracking

Tracking status:

Uses the TreeSearch algorithm developed by Ole for the Hall A tracking (HRSs, SoLID, etc).



UCONN Jefferson Lab

Progress on subsystems:
GEMs / Tracking: Status (see Danning's slides summary)

Summary for GEMs :

- The full simulation → digitization → analyis → analyzed data (clusters, tracks) has been functional since roughly 9 months

- Tuned digitization database and digitization code to have simulated digitized hits similar to cosmic data ;

- Added subdivision of “GEMPlane” (called "module"), to *allow for transverse segmentation* of one plane within one tracker, in both digitization and Treesearch library (did not exist before)

- Implemented pulse fitting on sampled ADC to extract timing;

- Implemented pedestal noise, cross talk.

- Implement common mode in digitization and common mode subtraction in Treesearch library;

TO-DO:

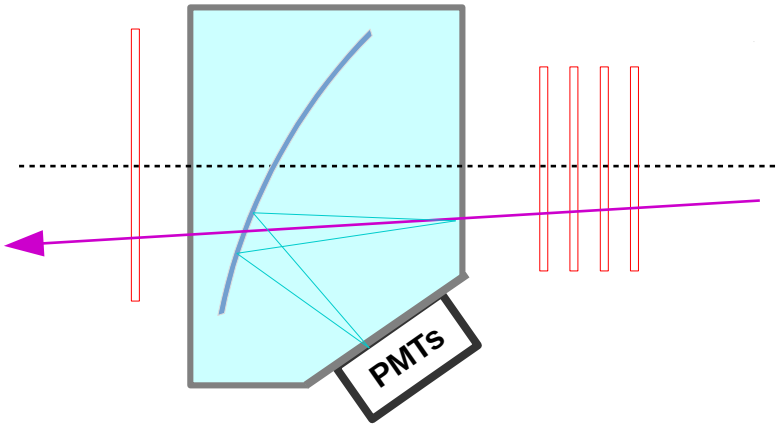
- Improve cluster separation method: Add timing information to currently existing peak-valley-peak method ;
  => This is very important to analyze data with huge backgrounds.

- Cross talk removal in Treesearch library;

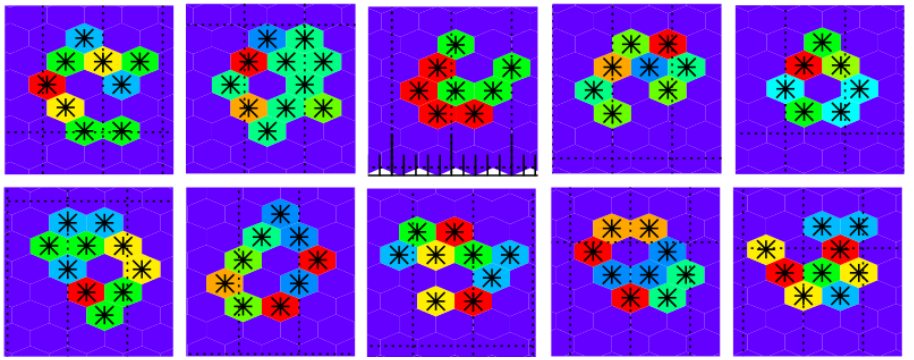
- Check Tree Search algoirthm under multiple track conditions.

Progress on subsystems:
GRINCH

Cherenkov detector readout by 510 PMTs (see Todd's talk)



6



Light generated by particles will be reflected to the PMT matrix, and leave signal such as the ones seen above (right).

GRINCH PMTs are readout by NINO cards:
- record TDC values for PMT rise time and fall time (pulse amp. goes resp. over/under NINO thr ; estimation of pulse amp. with Time-over threshold).

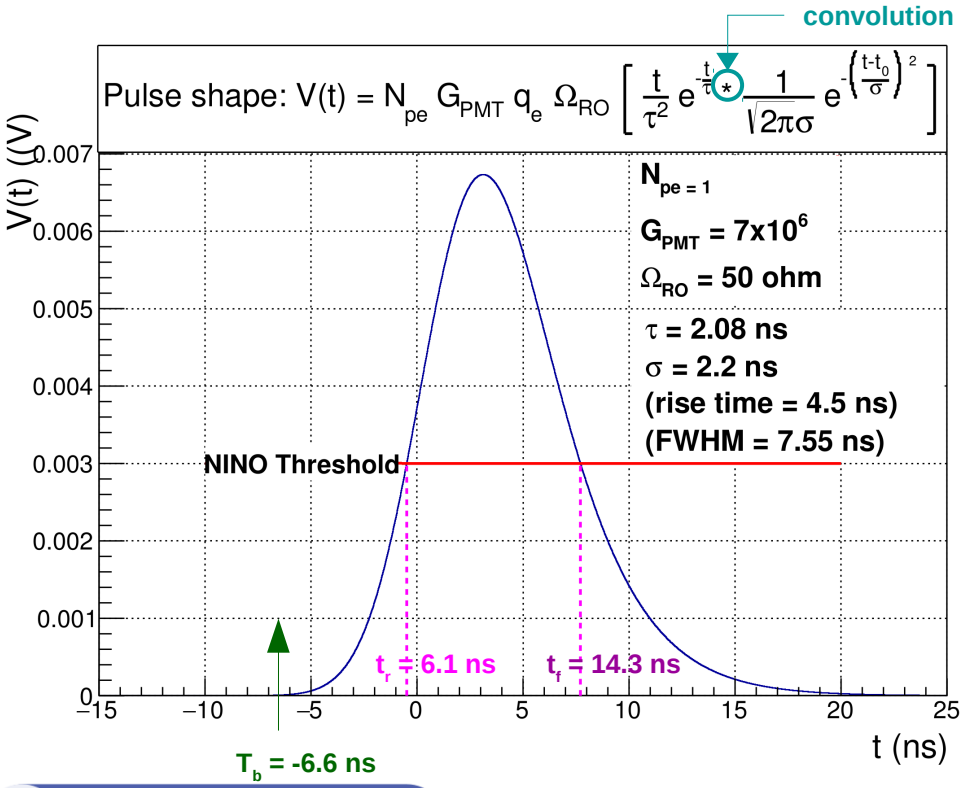
Refinements for electronics :
* TDC channel size (1ns uncertainty for NINO);
* Pedestal noise;
* Cross-talk.

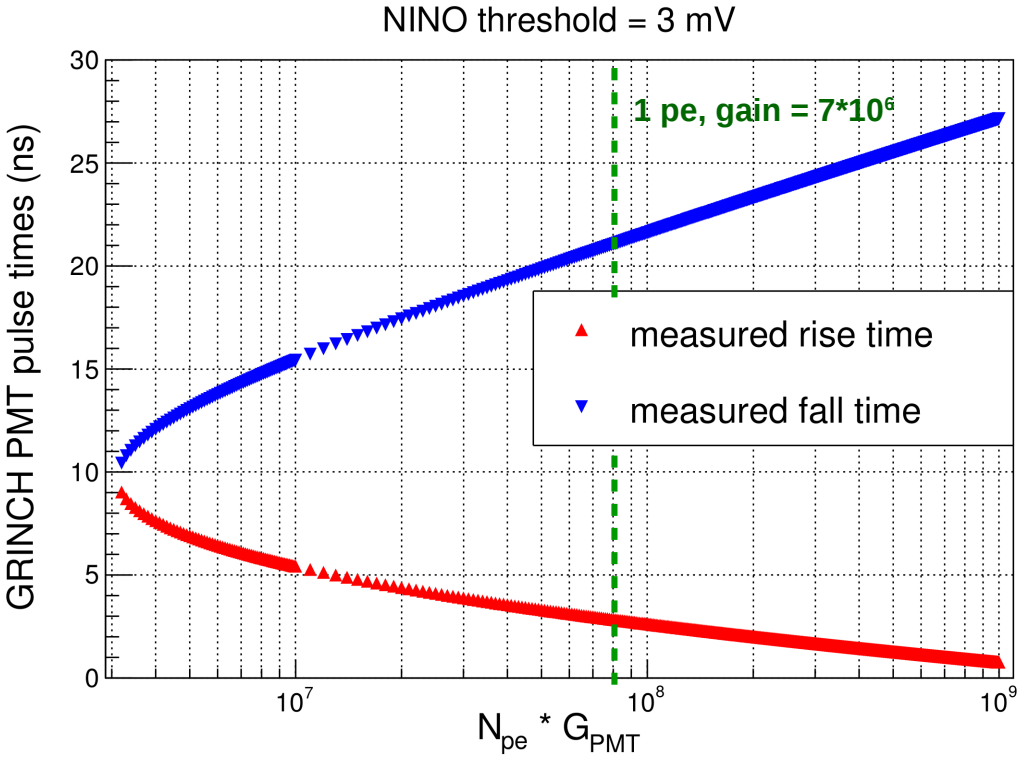
UCONN Jefferson Lab

Progress on subsystems:
GRINCH

GRINCH (and RICH) digitization

To evaluate the times t_r and t_f at which the pulse will go over/under the threshold, we have modeled a normalized pulse shape (left) which reproduces the PMT specs (9125B) ;
We calculate tables for a given threshold, over a wide span of amplitudes (right)





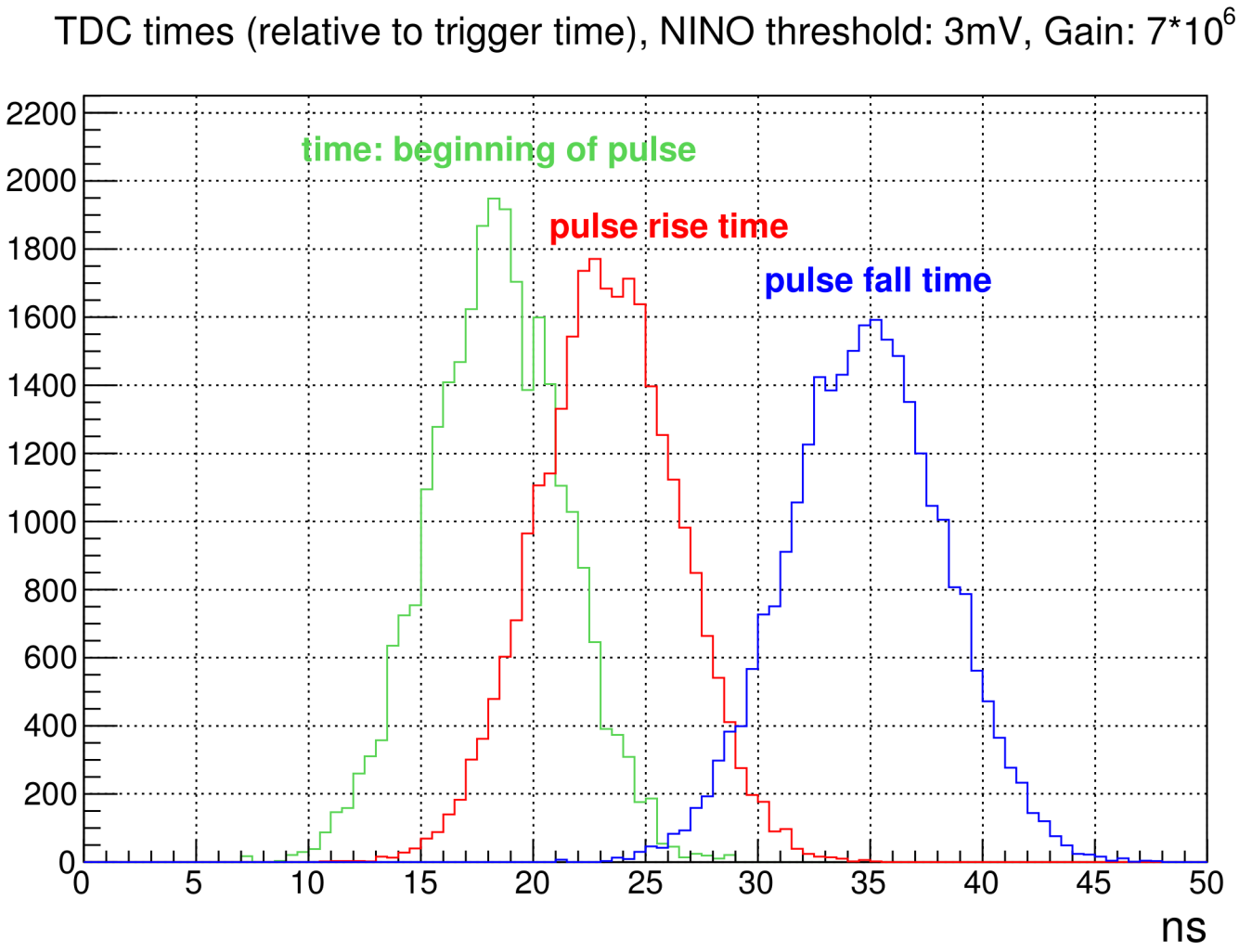
6 characteristics (ET entreprises 9125B)

	unit	min	typ	max
timing:				
single electron rise time	ns		4.5	
single electron (fwhm)	ns		7.5	
single electron jitter (fwhm)	ns		4	
transit time	ns		33	

* Values plugged in the digitization library.
* Exponential extrapolation between 2 points to calculate t_r, t_f for any given amplitude;
* Easy to redo these tables provided a new pulse shape function or threshold;

Progress on subsystems:
GRINCH

Digitized pulse timings
(GRINCH hits from QE electrons generated with G4SBS, processed with digitization library)



No background

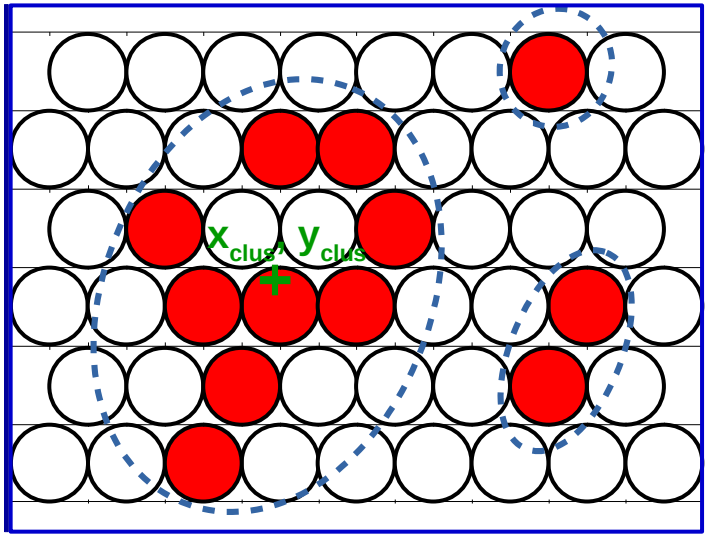
UCONN Jefferson Lab

Progress on subsystems:
GRINCH

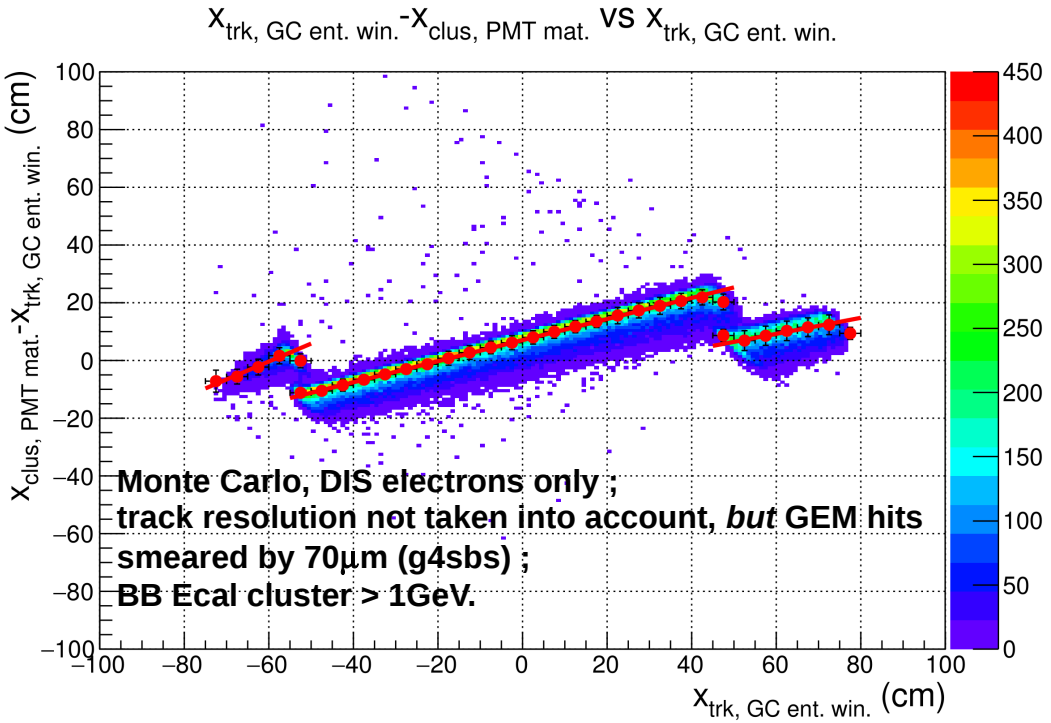
GRINCH analysis:

- Clustering (*rewritten*): all neighboring hits PMTs are considered to be part of the same cluster (*);

- Association with tracks: correlation between cluster position in GRINCH and charged track position;



(*) this will have to be different for the RICH...

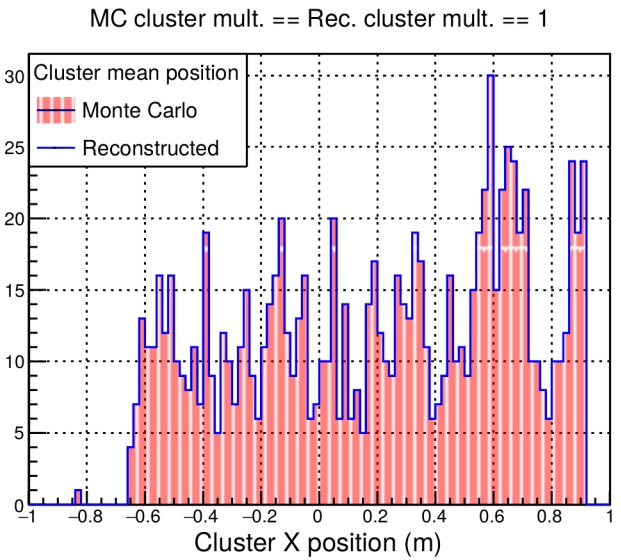


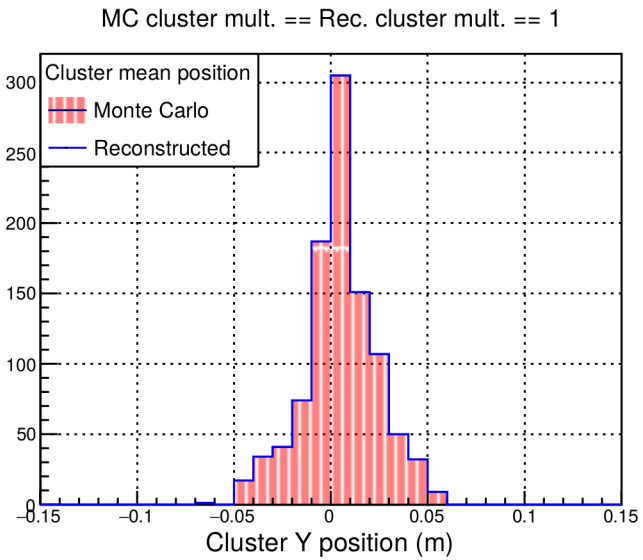
UCONN Jefferson Lab

Progress on subsystems:
GRINCH

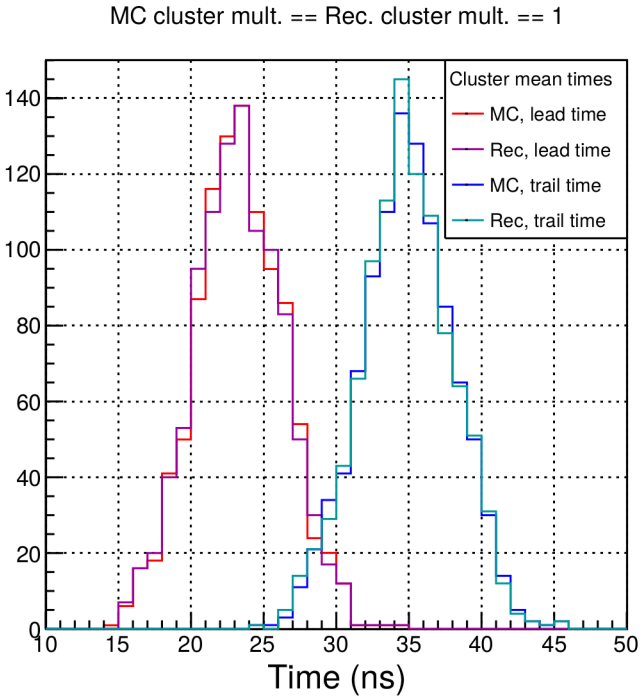
Clusters : comparison with MC clusters (PMTs which hit come from the same particle)

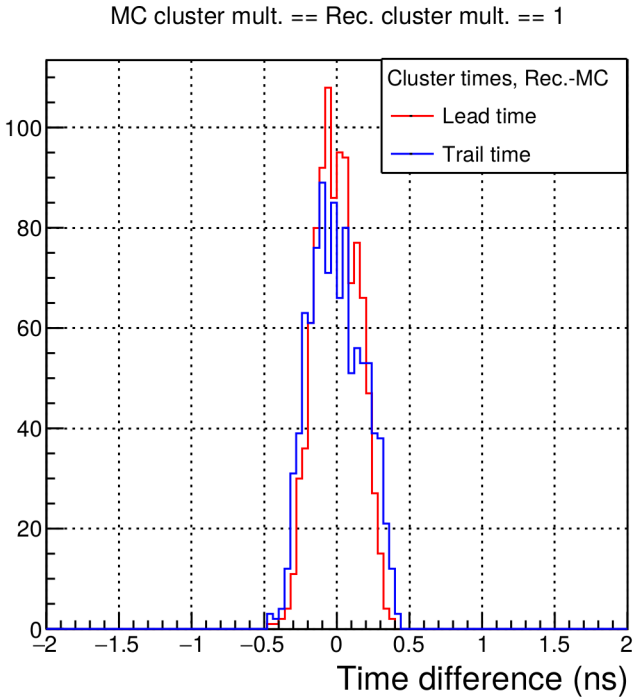
Mean cluster position





Mean cluster times





UCONN Jefferson Lab

Progress on subsystems:
GRINCH: Status

Summary for GRINCH :

- The full simulation → digitization → analyis → analyzed data (clusters) has been setup very recently

- Data flow goes correctly: PMT numbers are retrieved correctly, positions of "unmodified clusters" is constructed correctly, .

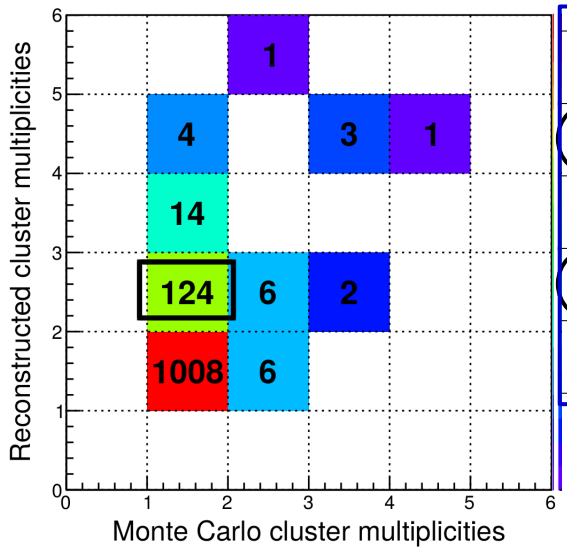
- A large part of this work can straightforwardly be used for RICH.

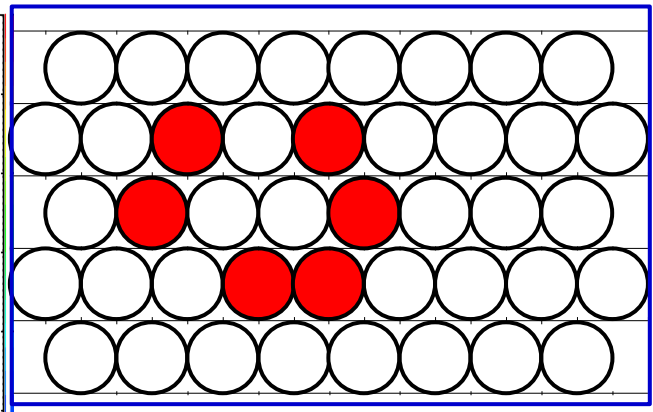
TO-DO:

- Implement pedestal noise, cross-talk

- Improve clustering : we lose ~10% clusters, most of them because they do not spread on all adjacent PMTs (see below).

- Ensure a good handling of the pile-up.





Progress on subsystems:
HCal (Juan Carlos)

HCal digitization is a special case: readout by FADC 250;
G4SBS HCal hits integrated to 1 hit/chan/evt;might be too "rough" for FADC250 i.e. we might lose some information. => extra attention to be paid digitizing HCal.

Digitizing HCAL PMT signal in G4SBS

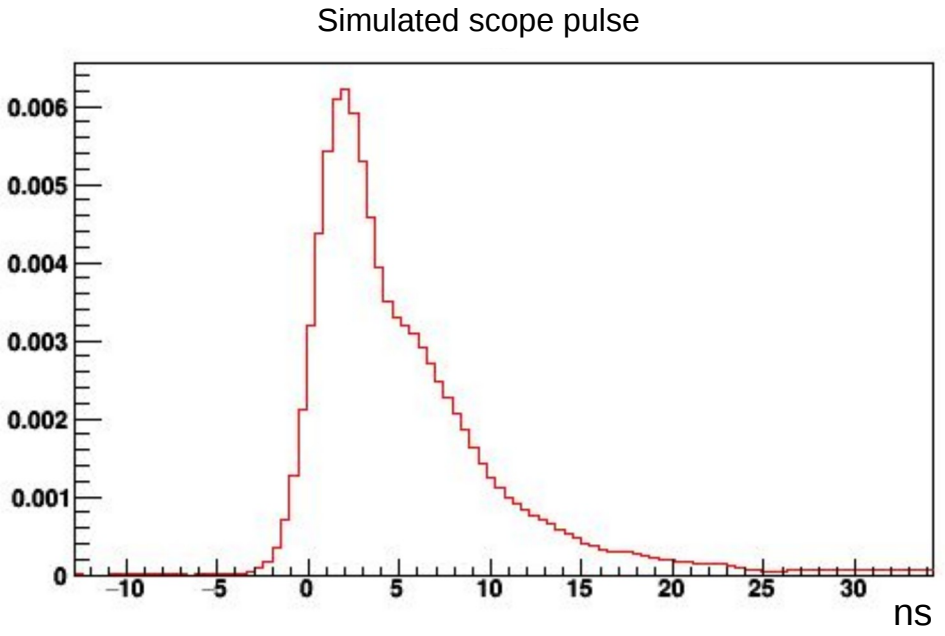
Build up simulated PMT signal for N photo-electrons using model of single-photo-electron.

- Use a model of an **ideal** single photo electron pulse.
- Not as realistic, but no timing corrections need to be performed.
- Easier to work with a TSpline.

I found this model that Vahe when he was working on the Light Guide design.

I don't know the details of exactly how he made this one.

It can either be fit to many single-pe or just used Polya statistics to make it.



(Slide from **Juan Carlos Cornejo**, SBS soft/simu, 2017/12/13)

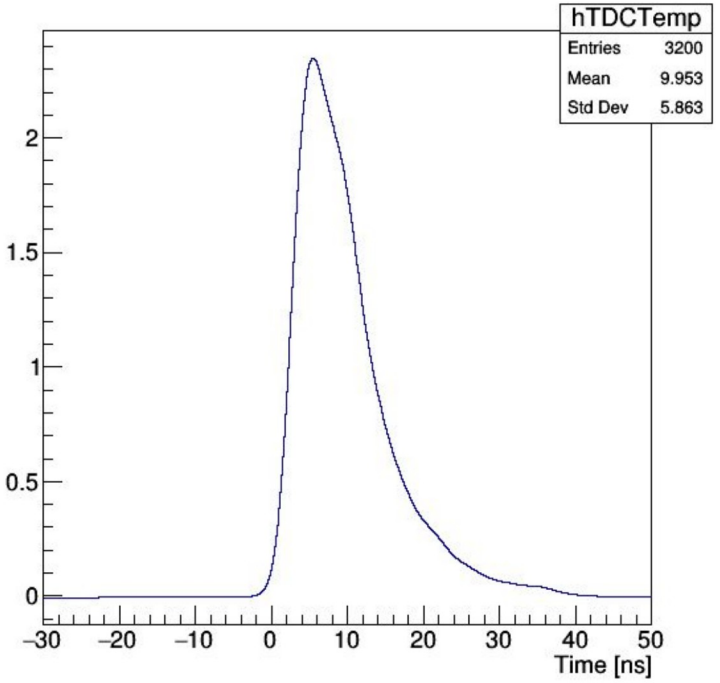
UCONN Jefferson Lab

Progress on subsystems:
HCal

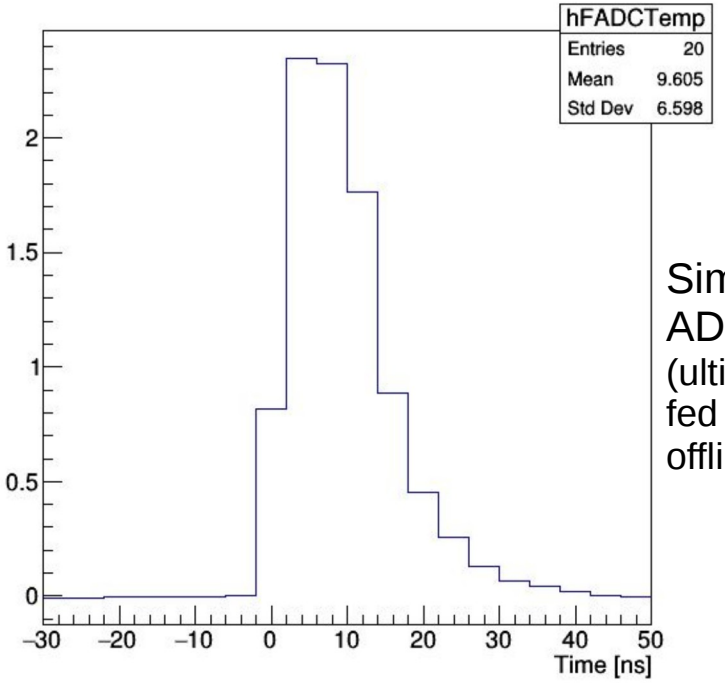
Early comparisons of HCAL digitization with real cosmics

This is what a simulated signal looks like when built out of 309 photo-electrons

Simulated scope pulse, i=3 from NPE=309



FADC i=3 from NPE=309



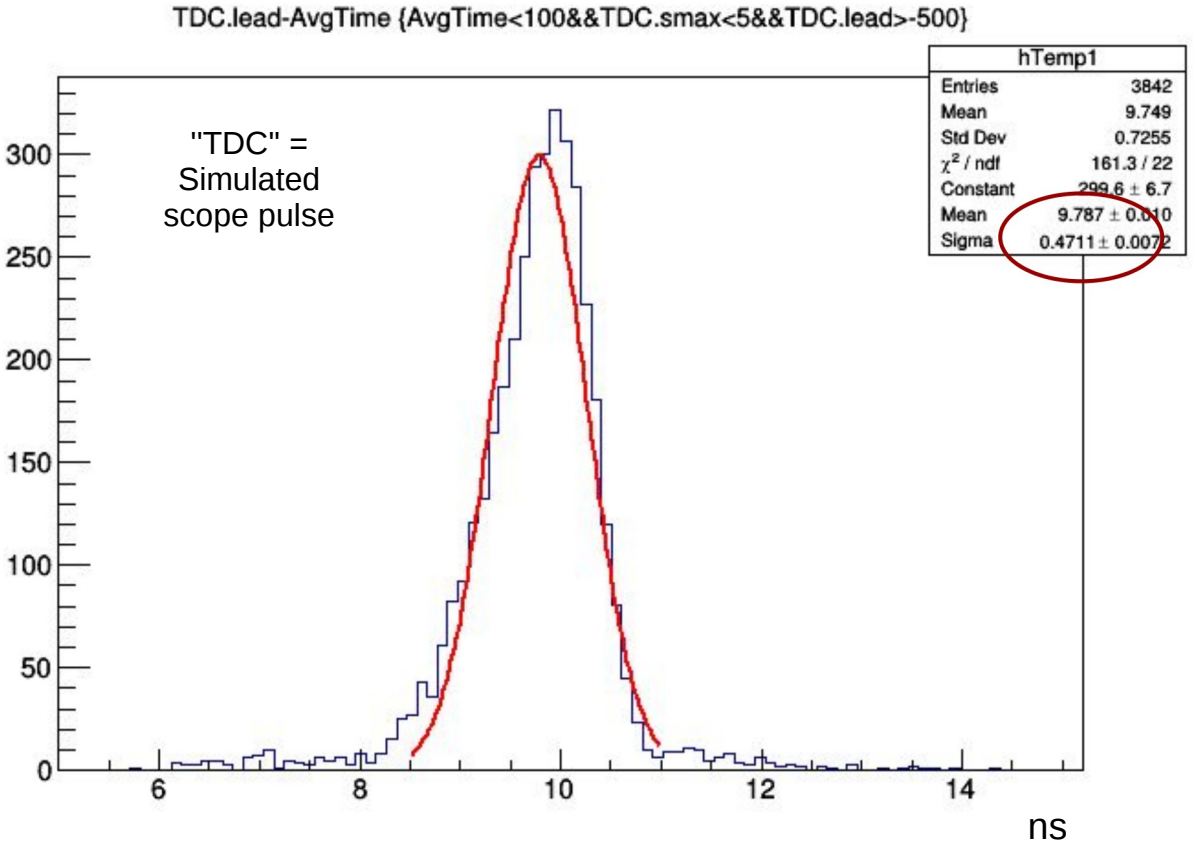
Simulated ADC sample (ultimately to be fed to SBS-offline)

(Slide from **Juan Carlos Cornejo**, SBS soft/simu, 2017/12/13)

UCONN Jefferson Lab

Progress on subsystems:
HCal

First attempt at timing with G4SBS (cosmic muon)



- Very basic timing by getting time of first bin above threshold.
- Need to do time-walk correction, and hope to get better timing with that.

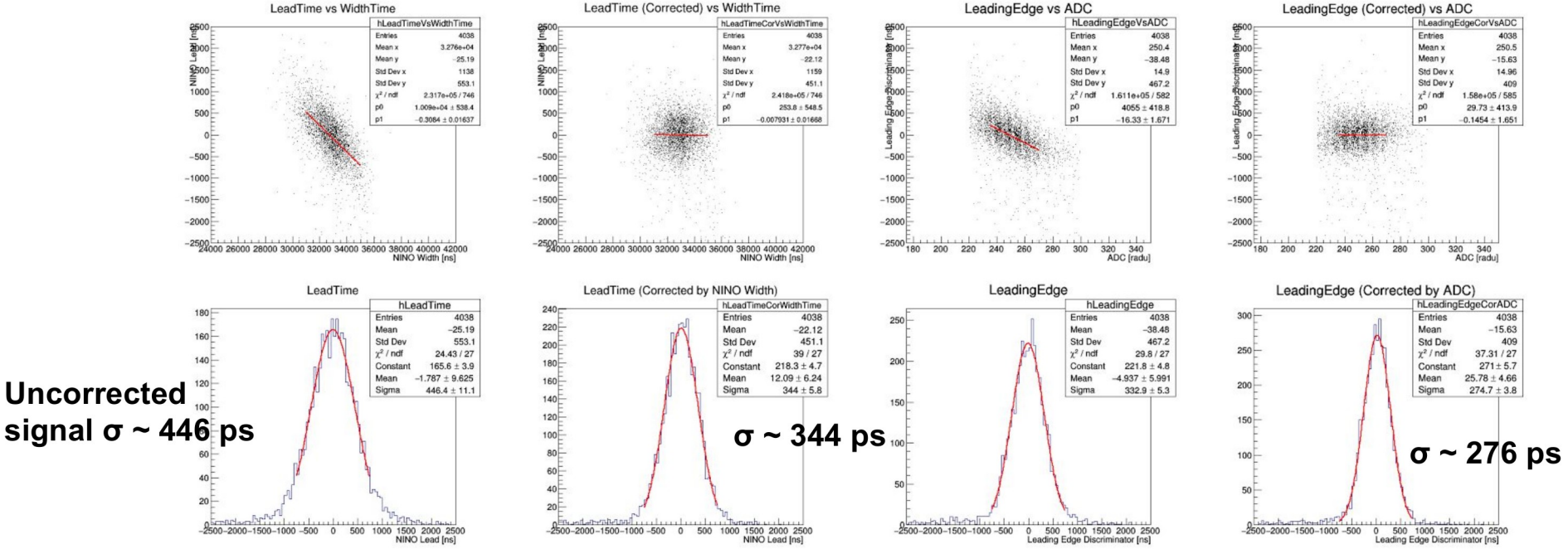
(Slide from **Juan Carlos Cornejo**, SBS soft/simu, 2017/12/13)

UCONN Jefferson Lab

Progress on subsystems:
HCal

Cosmic data timing (using NINO and Leading Edge)

Here is an example from Cosmic data (Left to graphs are using NINO, Right two graphs are using a leading edge discriminator



(Slide from Juan Carlos, SBS soft/simu, 2017/12/13; see also Brian's talk from yesterday).

UCONN Jefferson Lab

Progress on subsystems:
HCal: Status

Summary for HCal :

- Digitization is work in progress, what was shown here is preliminary;

TO-DO:

- The full simulation → digitization → analyis → analyzed data (clusters) to be worked out

- Implement pedestal noise, cross-talk;

- Pile-up handling.

Overview

SBS Software Project:

- Introduction: Overview, Milestones.
- Simulation digitization;
- Interface with analysis;

Progress on subsystems:

- GEMs / Tracking;
- GRINCH;
- HCal;
- Summary and next steps.

Summary

Summary

*** The main focus of the software is the digitization of all subsystems and interface with analysis;**

* The GEM digitization and TreeSearch algorithm (by far the most advanced) keeps being improved (see Danning's slides);

* GRINCH simulation is now been interfaced with analysis
(+ most of this work can be reused "as is" for the RICH analysis);

*** Main focus is now on the digitization and interface of other subsystems (Hodoscopes, ECals, HCal);**
NB: for e.g. Hodoscopes, most of the work made for one subsystems will apply straightforwardly to the others of the same kind. Same story for Ecals.

* At one point, one wants to: merge GEM digitization library with the main digitization library, and the TreeSearch Library to the SBS-offline.

Ultimate goal: Under experimental (primary focus on G_M^n) conditions, digitize simulation results and be able to efficiently reject backgrounds, do reconstruction and determine efficiencies (tracking, PID, etc) and resolutions.

Useful links
(if one wants to participate;))

Digitization library repository:
https://github.com/efuchey/libsbsdig

SBS-offline analysis library:
https://github.com/JeffersonLab/SBS-Offline

NB:
https://github.com/JeffersonLab/SBS-Offline/tree/allsystems_rewrite

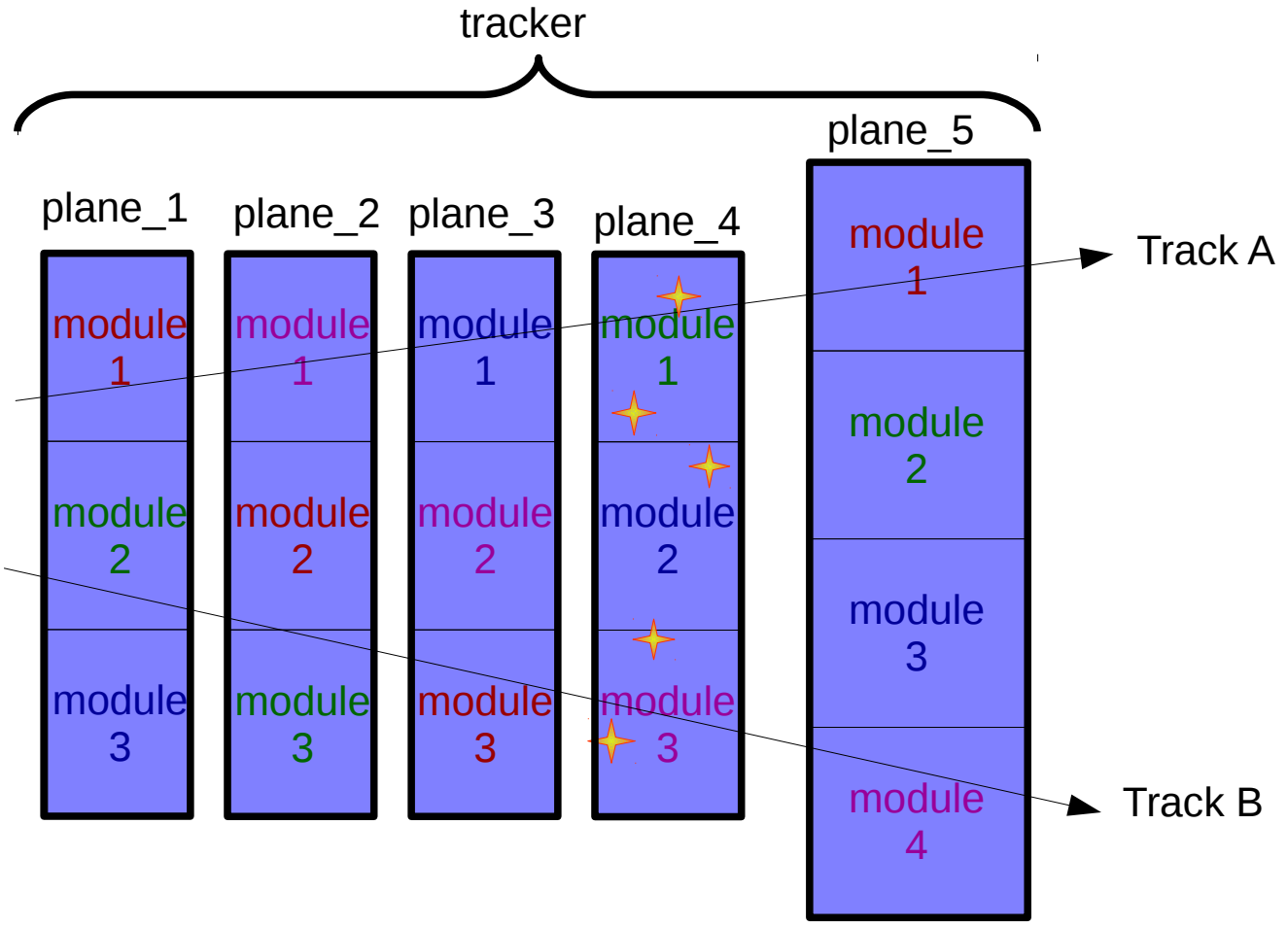
GEM digitization library for SBS:
https://github.com/JeffersonLab/libsolgem/tree/libsbsgem

TreeSearch library for SBS:
https://github.com/JeffersonLab/TreeSearch/tree/sbs

Thank you for your attention !

Tracking progress
Improvements and adjustments

Modified the GEM geometry hierarchy in both digitization library and TreeSearch algorithm:
Added a subdivision of the GEM plane called "module";



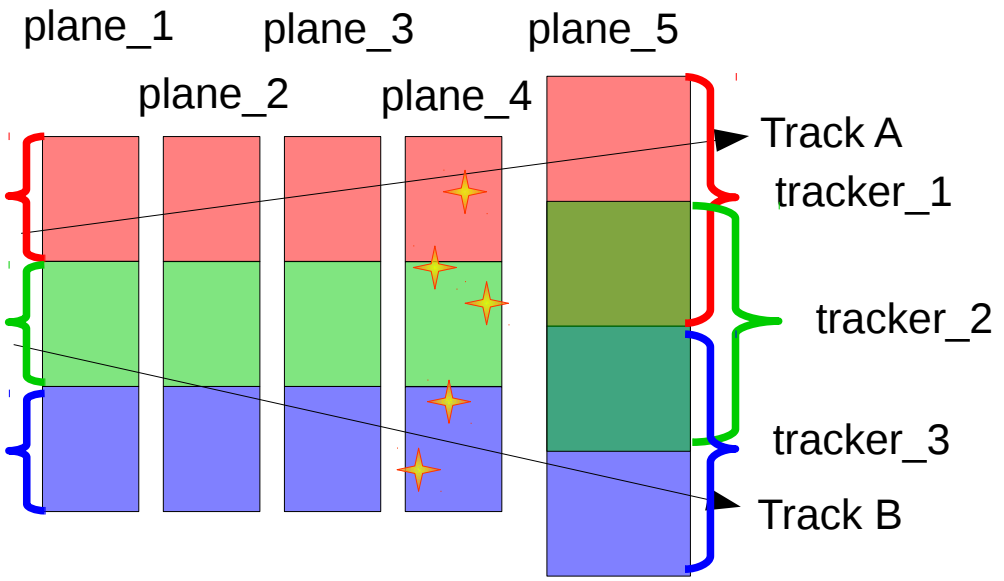
* Save “moduleID” with each hit for later 2D track matching
* New database format

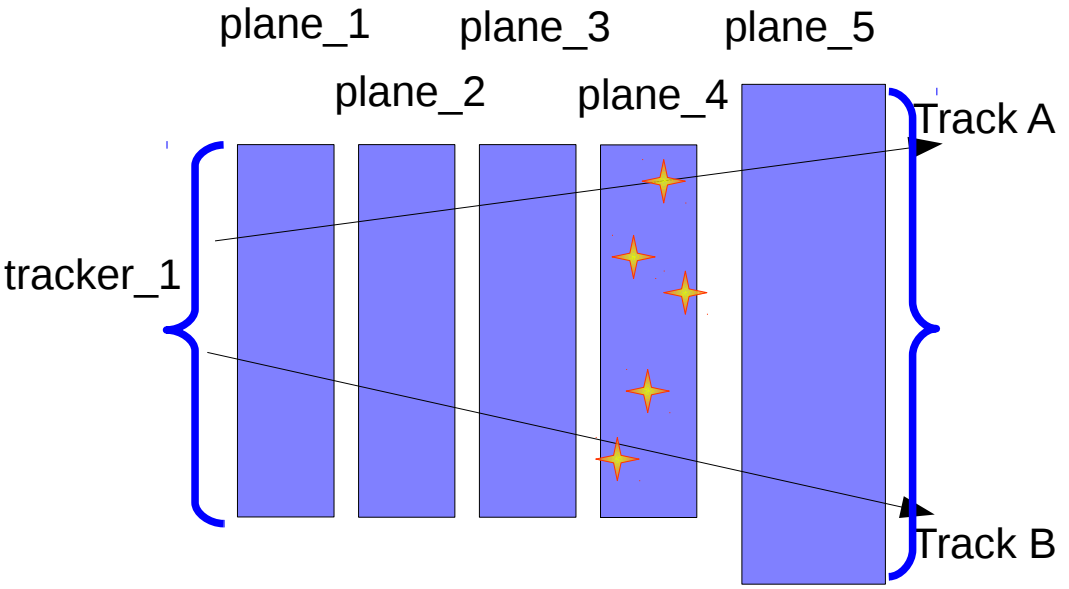
UCONN Jefferson Lab

Tracking progress
Digitization adjustements

Modified the GEM geometry hierarchy in both digitization library and TreeSearch algorithm:
Added a subdivision of the GEM plane called "module";

There was no such possibility before, leading to two options for SBS GEMs (none being satisfactory):
* dividing the detector in 3 trackers, leading to efficiency losses (tracking is made within a tracker) ;
* using one tracker with large planes, leading to exaggerated occupancies when adding background





=> Track A reconstructed;
=> Track B lost;
=> No anomaly occurs when adding background.

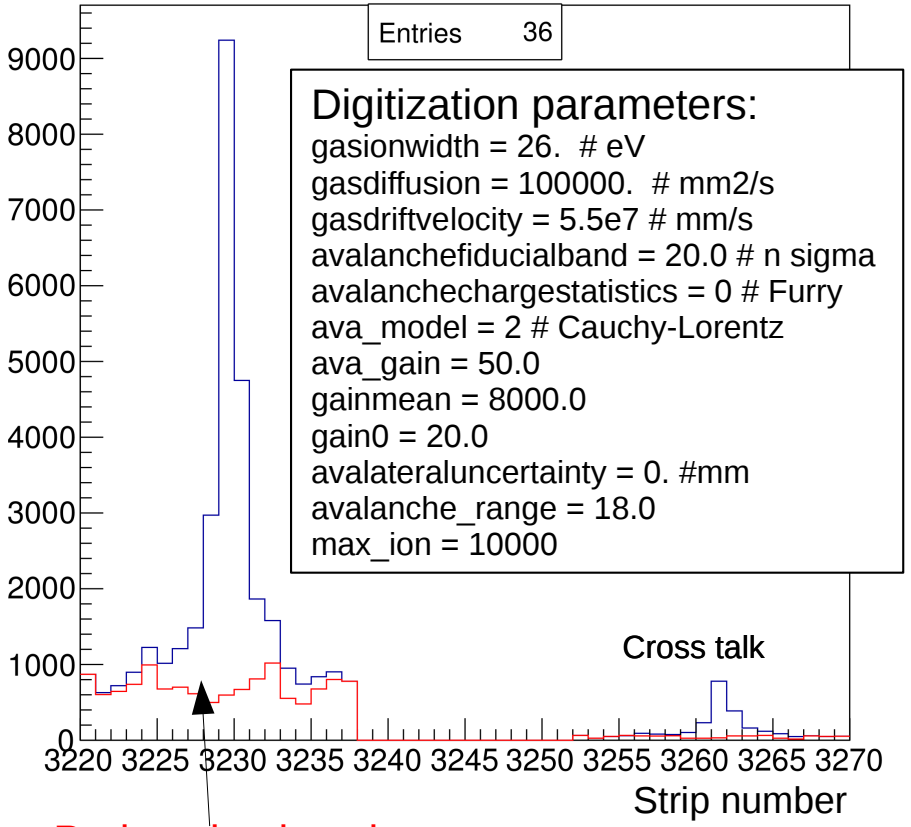
=> Track A and B reconstructed;
=> when adding background, the horizontal coordinate integrates background on the full plane
=> exaggerated occupancy
+ tracking efficiency damage

UCONN Jefferson Lab

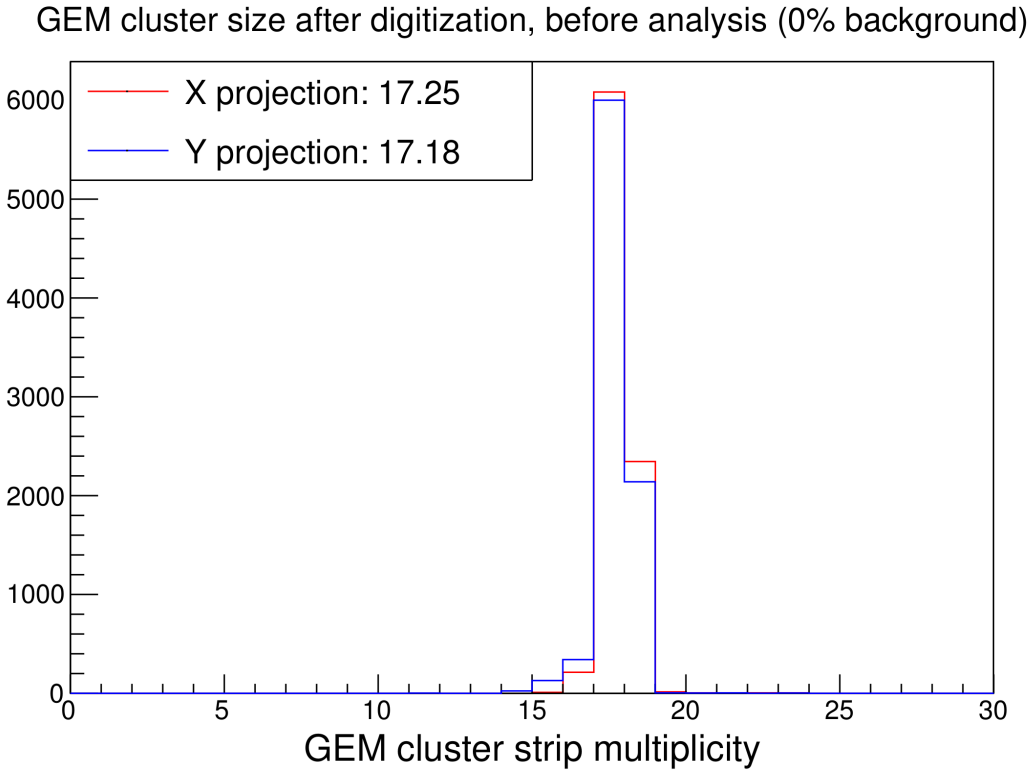
Tracking progress
Former cluster size

No background

One **elastic electron** event
sbs_bbgem.tracker.1.x1.adcmap



Pedestal noise alone
(generated sample by sample, then deconvoluted)



Mean cluster size in X, Y: ~17
(event though there are mostly elastic electrons...)

Digitization / analysis interface
GRINCH

In detail: for each PMT hit, TSBSSimCherDigitization produces 2 VETROC words (one per TDC value) with the format posted on the right (provided by Evan, courtesy of Ben Raydo).

TSBSSimDecoder associates both VETROC words with a crate/slot/channel number*

VETROC words are then deciphered by the SBS-offline GRINCH analysis class (SBSGRINCH), and PMT number and times are retrieved.

PMT hits are then reconstructed as one "leading" and one "trailing" TDC time associated for the same channel.

Data Type: TDC Hit

Type: 0x8
Size: 1 word
Description: This data type provides the time, channel, and edge of TDC hits

Word 1:

31	30	29	28	27	26	25	24
1	1	0	0	0	EDGE	0	0
23	22	21	20	19	18	17	16
CHANNEL							
15	14	13	12	11	10	9	8
TIME							
7	6	5	4	3	2	1	0
TIME							

EDGE:

Edge:

0	TDC measure of rising edge
1	TDC measure of falling edge

CHANNEL:

TDC channel number

TIME:

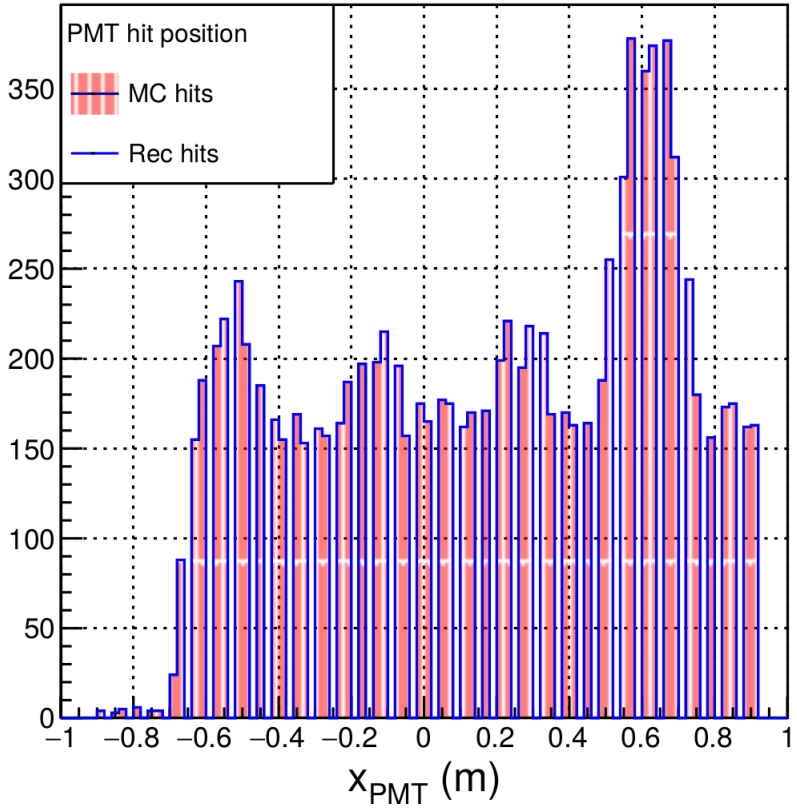
TDC hit time: 1ns resolution measured from beginning of trigger window

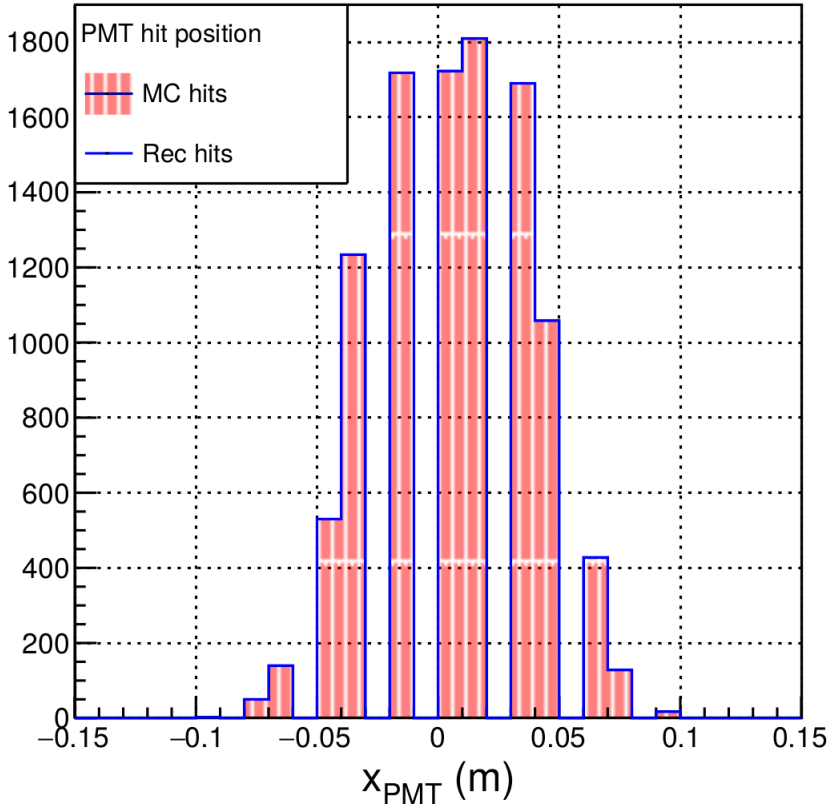
* slot number corresponds to VETROC number => retrieve PMT number

UCONN Jefferson Lab

Digitization / analysis interface
GRINCH PMT hit reconstruction, NO bkgd

PMT hit positions: depend solely on the PMT number and the geometry database : should be the same for MC and reconstructed

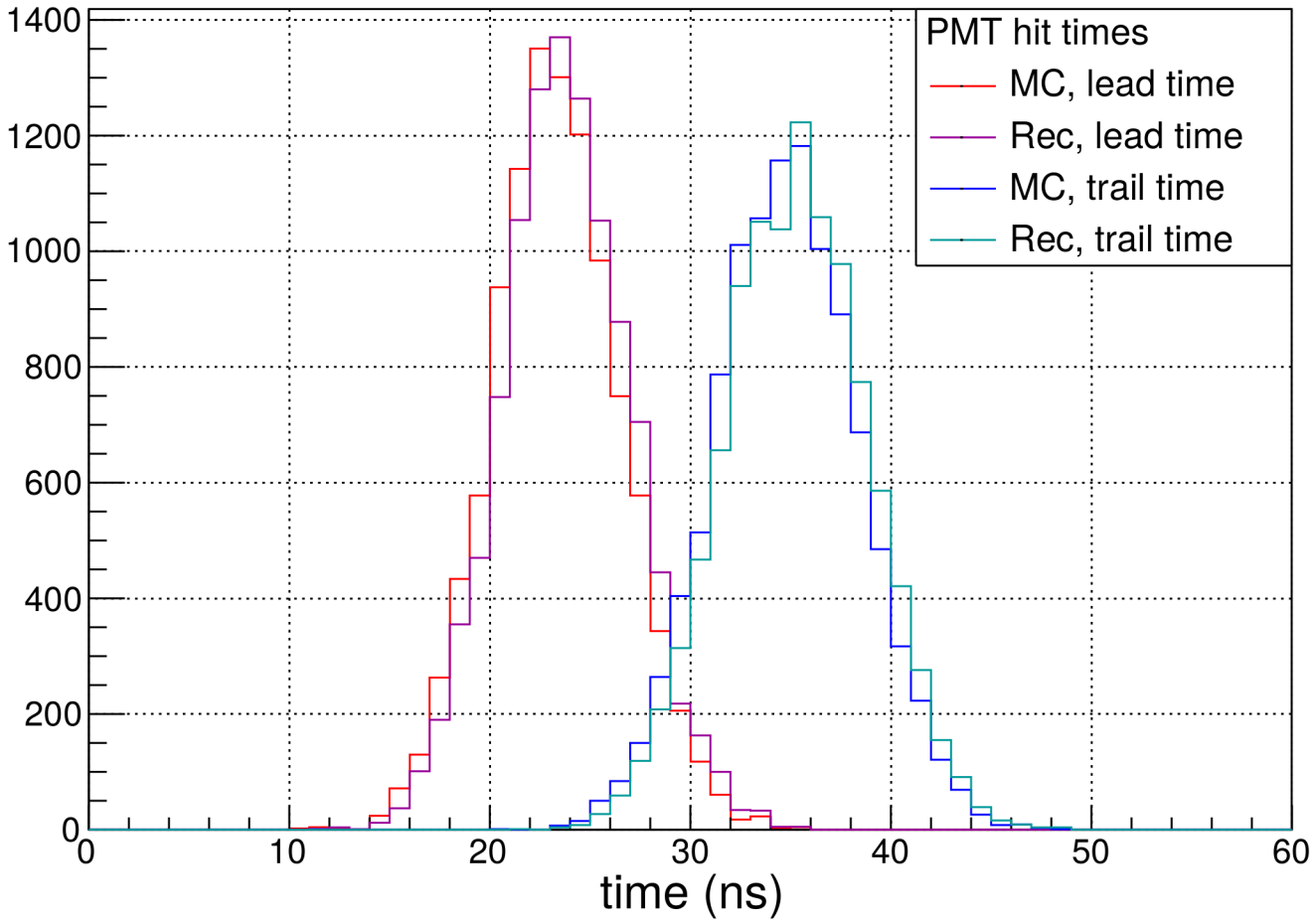




UCONN Jefferson Lab

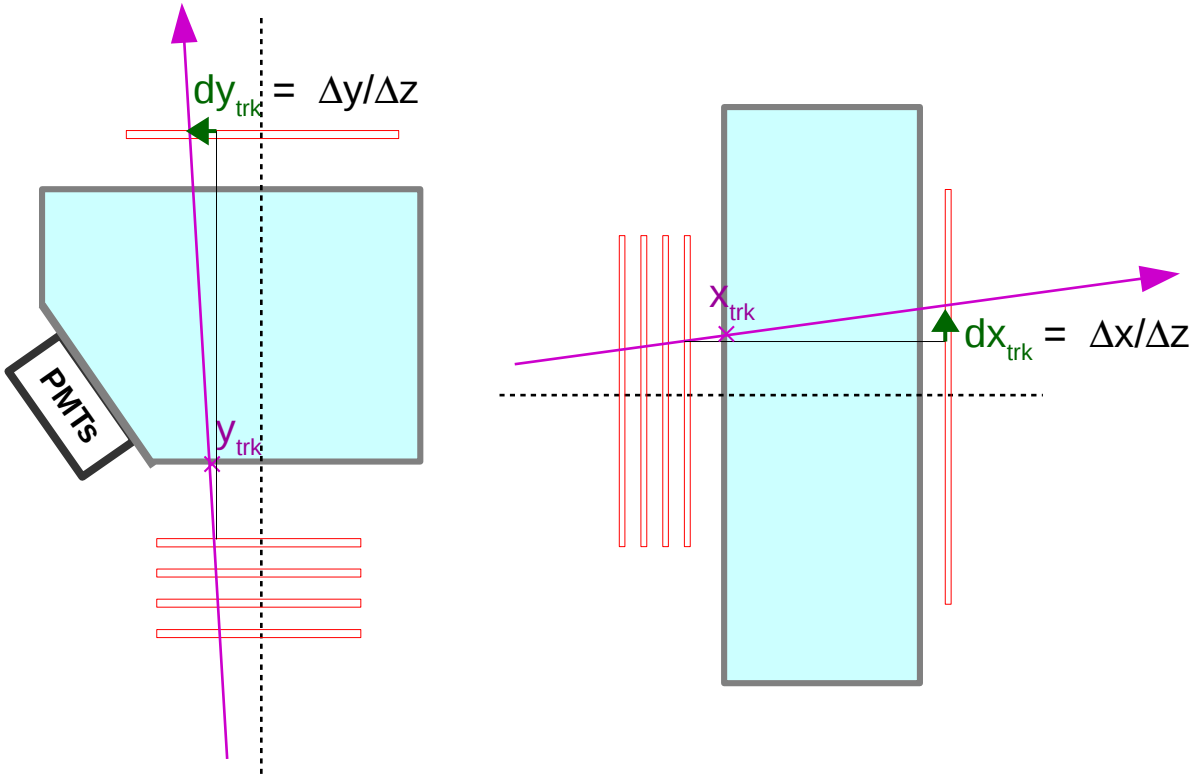
Digitization / analysis interface
GRINCH PMT hit reconstruction, NO bkgd

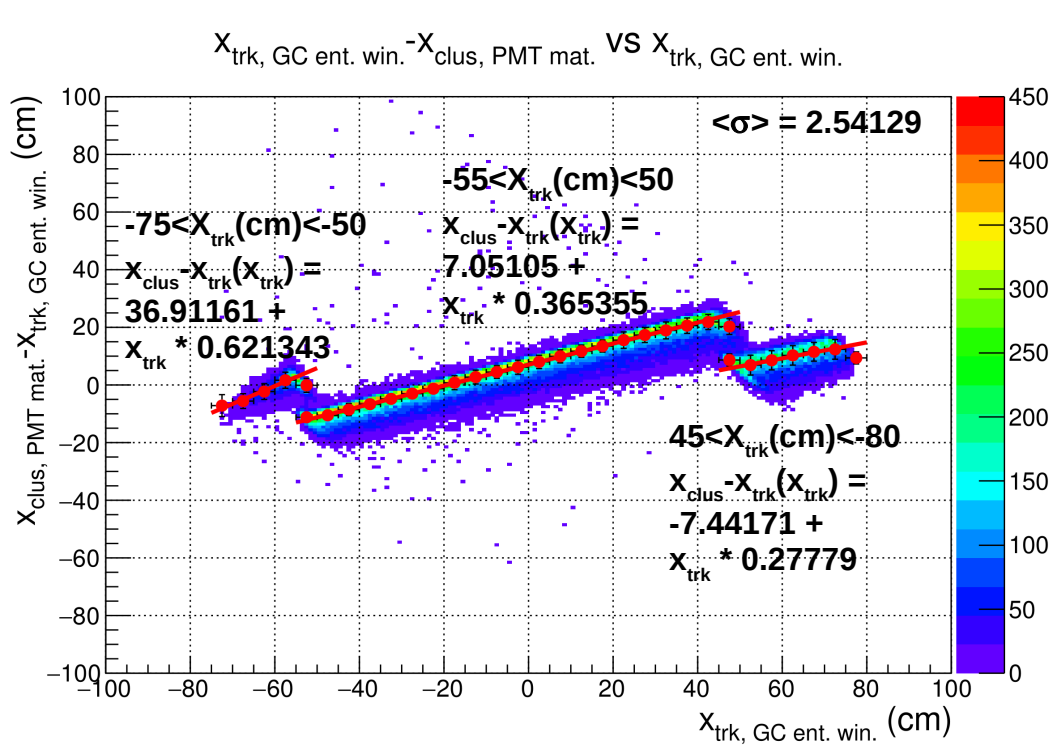
PMT times: "smeared" by the TDC "resolution" (1ns) (in practice, the TDC value is rounded to the nearest integer).

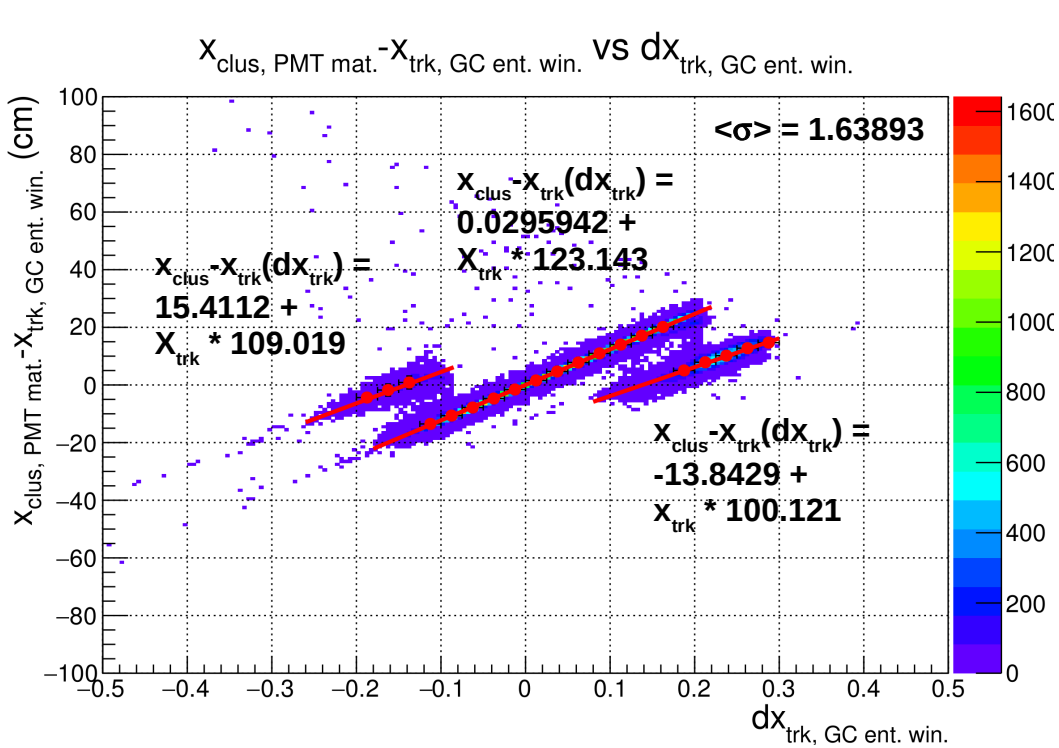


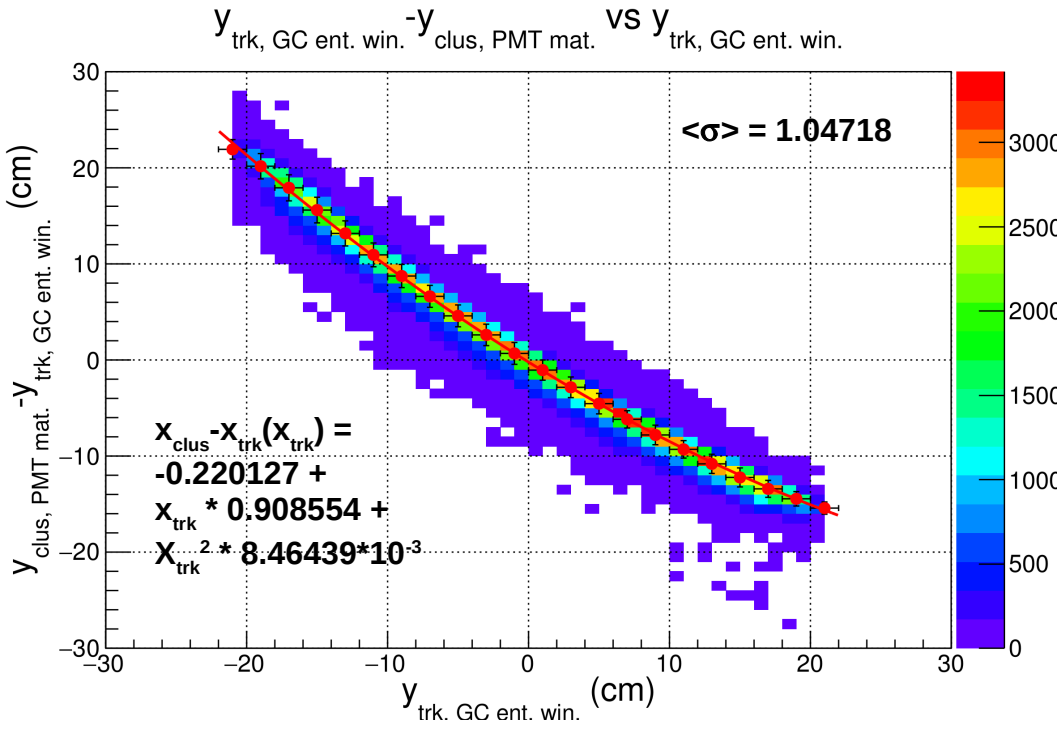
UCONN Jefferson Lab

Definition of x_{trk}, y_{trk}, dx_{trk}, dy_{trk}









Monte Carlo, DIS electrons only ;
track resolution not taken into account, *but* GEM hits smeared by 70µm (g4sbs) ;
BB Ecal cluster > 1GeV.

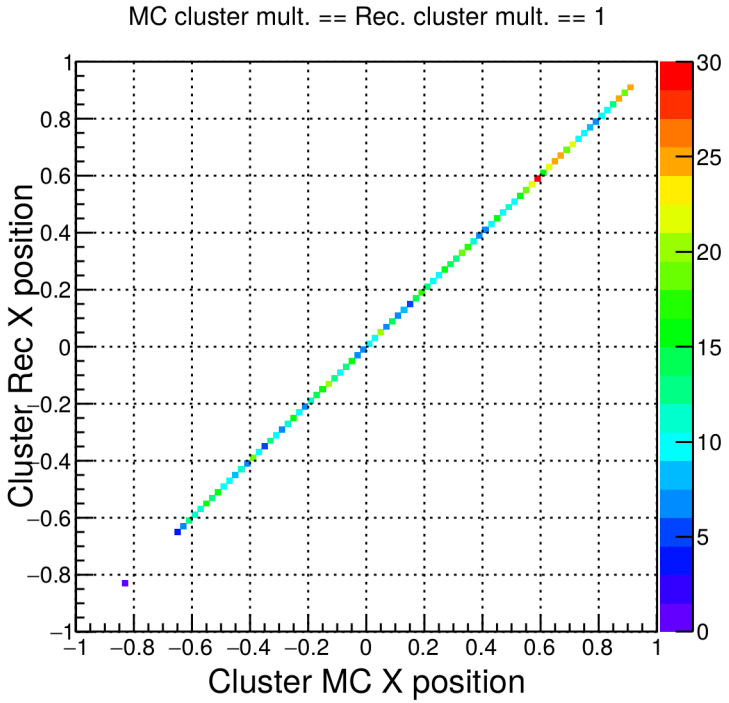
A correlation function is fitted between two sets of variables, and the events are selected within 5 σ of this correlation function

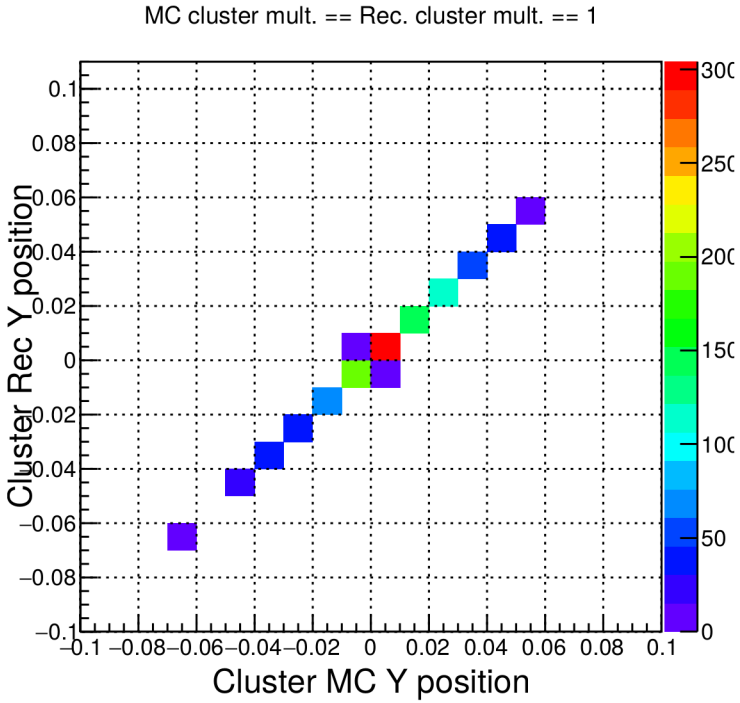
=> 98 % signal selection efficiency.
/!\ estimated with a non digitized simulation file, and does not include effects such as background on cluster position

Todo : put params in DB ??? (makes a lot of them...)

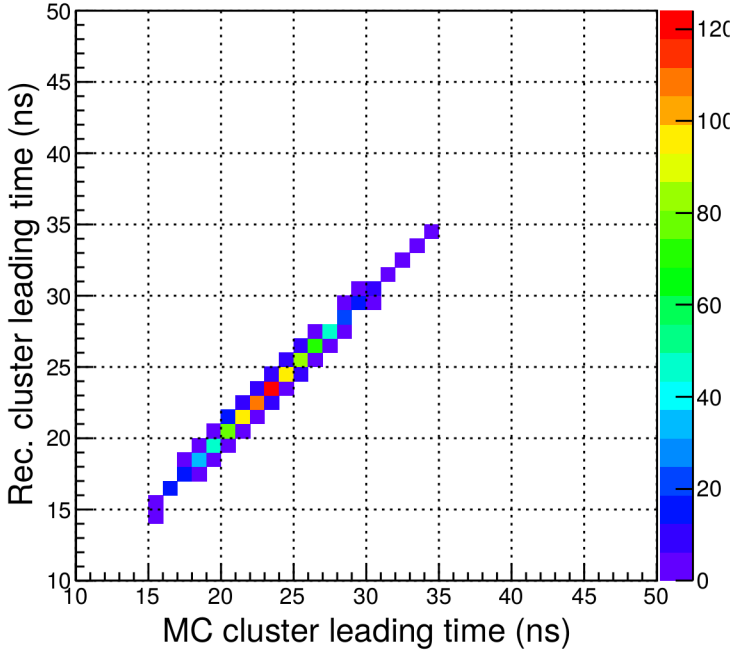
Digitization / analysis interface
GRINCH cluster hit reconstruction, NO bkgd

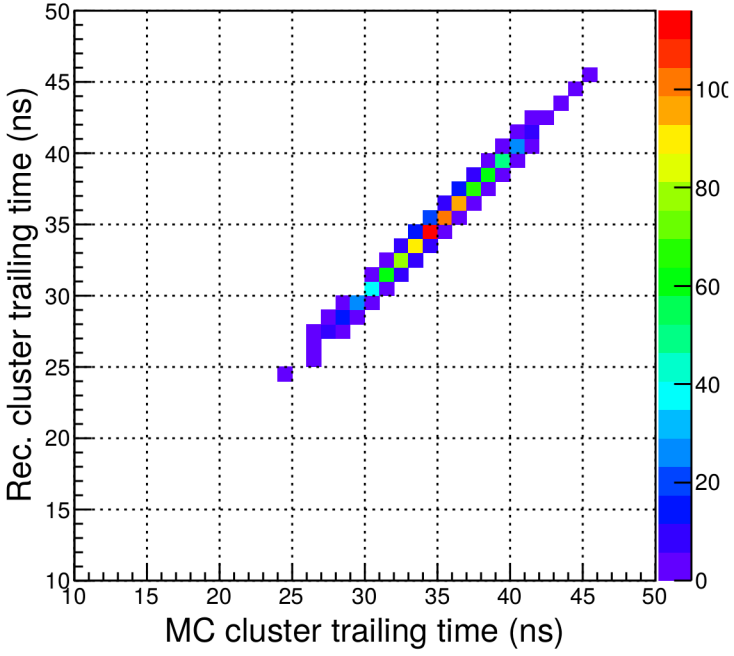
1 MC cluster, 1 rec. cluster





Cluster mean position

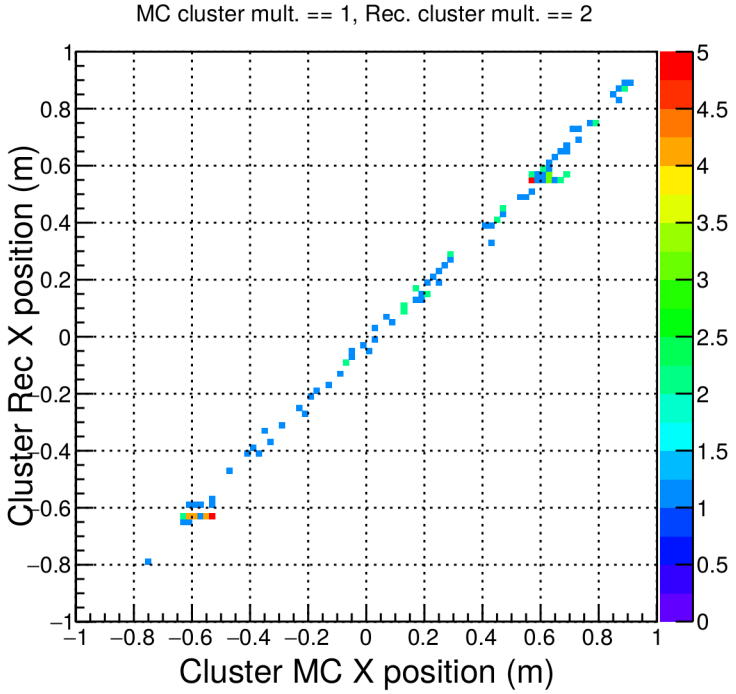


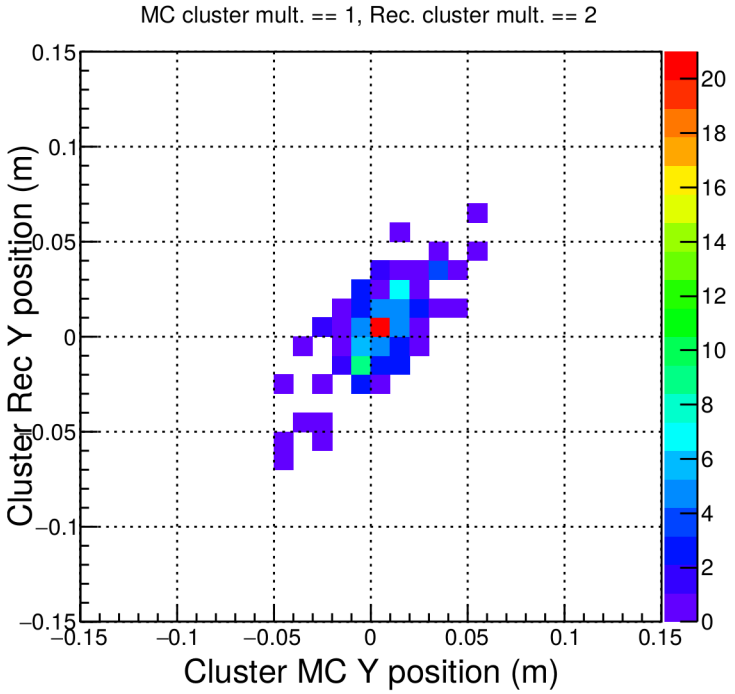


Cluster mean leading and trailing time

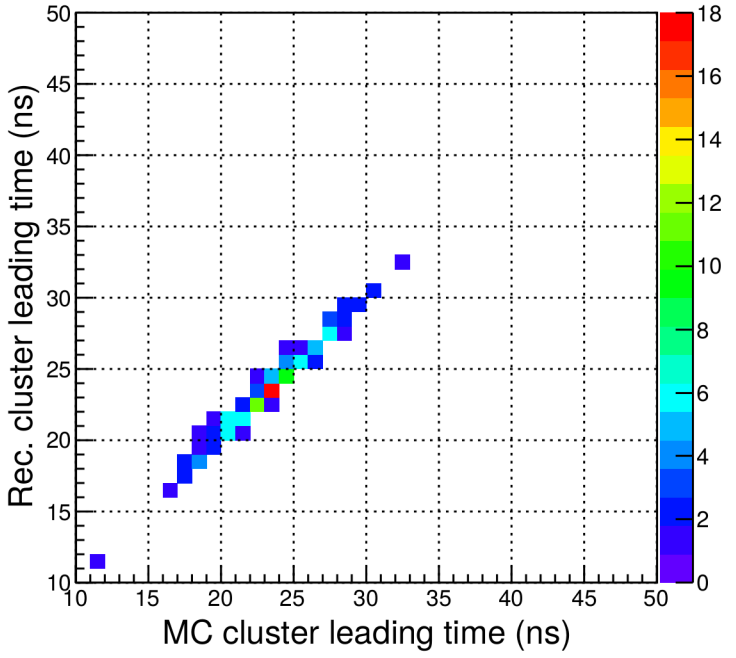
Digitization / analysis interface
GRINCH cluster hit reconstruction, NO bkgd

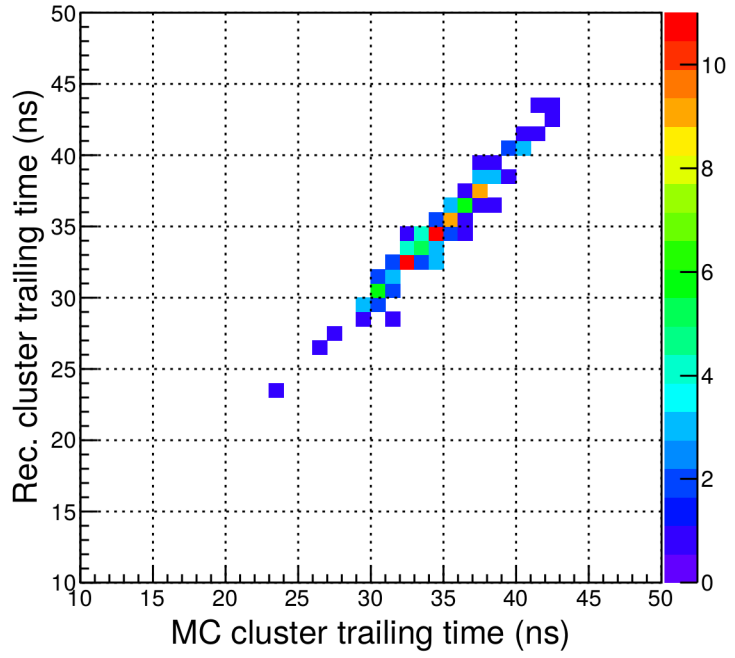
1 MC cluster, 2 rec. cluster





Cluster mean position



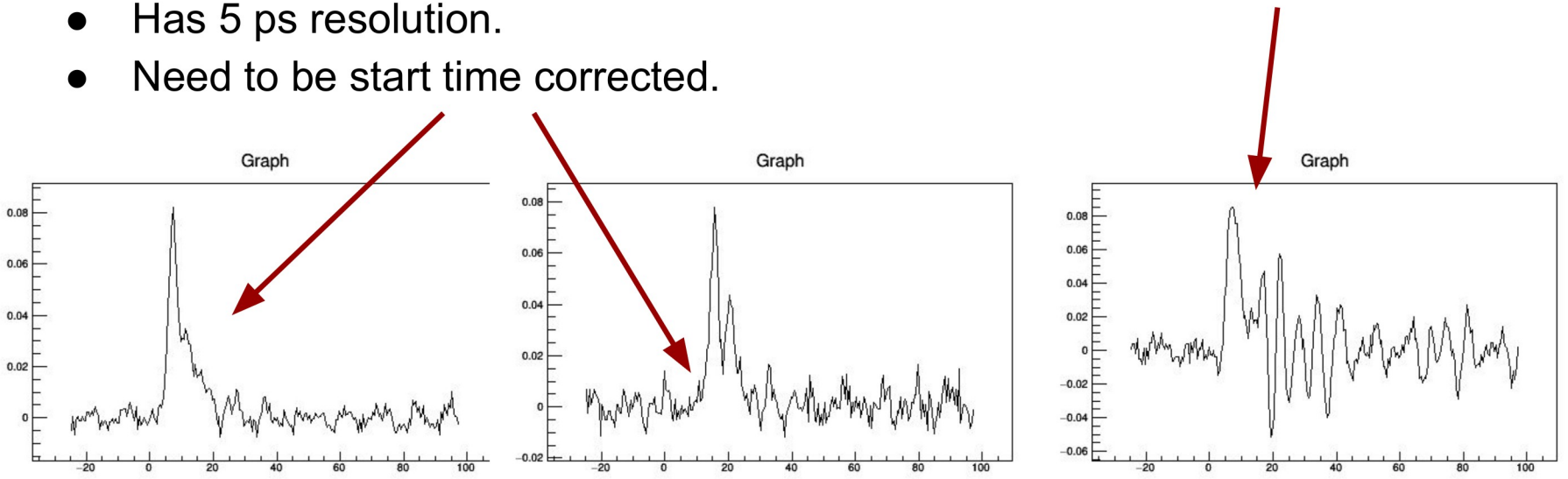


Cluster mean leading and trailing time

Progress on subsystems:
HCal

Option 1 for digitizing HCAL PMT Signal in G4SBS

- Build up simulated PMT signal for N photo-electrons using model of single-photo-electron.
- Use scope images for single photo-electrons (Vahe took this using scintillator and one PMT).
- Can be more realistic since at single-pe pulses look very jagged
- Has 5 ps resolution.
- Need to be start time corrected.



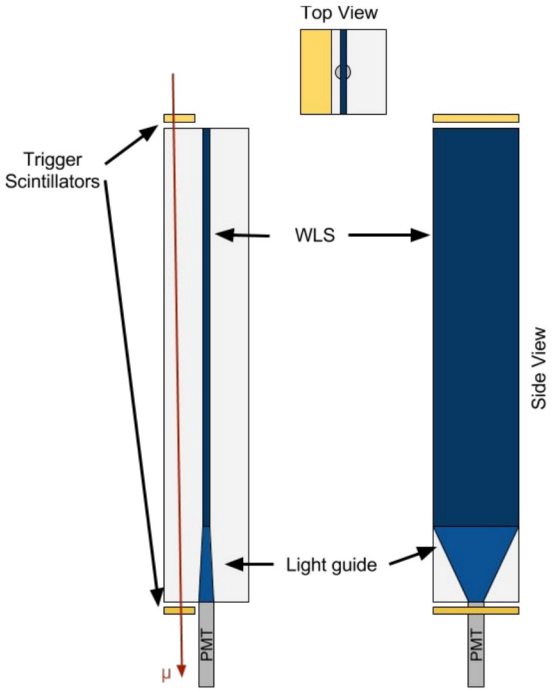
(Slide from Juan Carlos, SBS soft/simu, 2017/12/13)

UCONN Jefferson Lab

Progress on subsystems:
HCal

Early comparisons of Option 2 and real cosmics

- I am benchmarking using Option 2 (since it proved to be easier to work with) to simulate real cosmic data on a single HCAL module.



- Vertical module with two trigger scintillators above and below.
  - Positioned so that muons do not go through Wavelength Shifter
- Used similar setup in G4SBS and used gun to generate 4GeV muons.
- Took only muons that passed through similar region defined by trigger scintillators.

(Slide from Juan Carlos, SBS soft/simu, 2017/12/13)

UCONN Jefferson Lab