SBS Software and Analysis

Andrew Puckett
University of Connecticut
Hall A Winter Meeting
February 11, 2022

UCONN

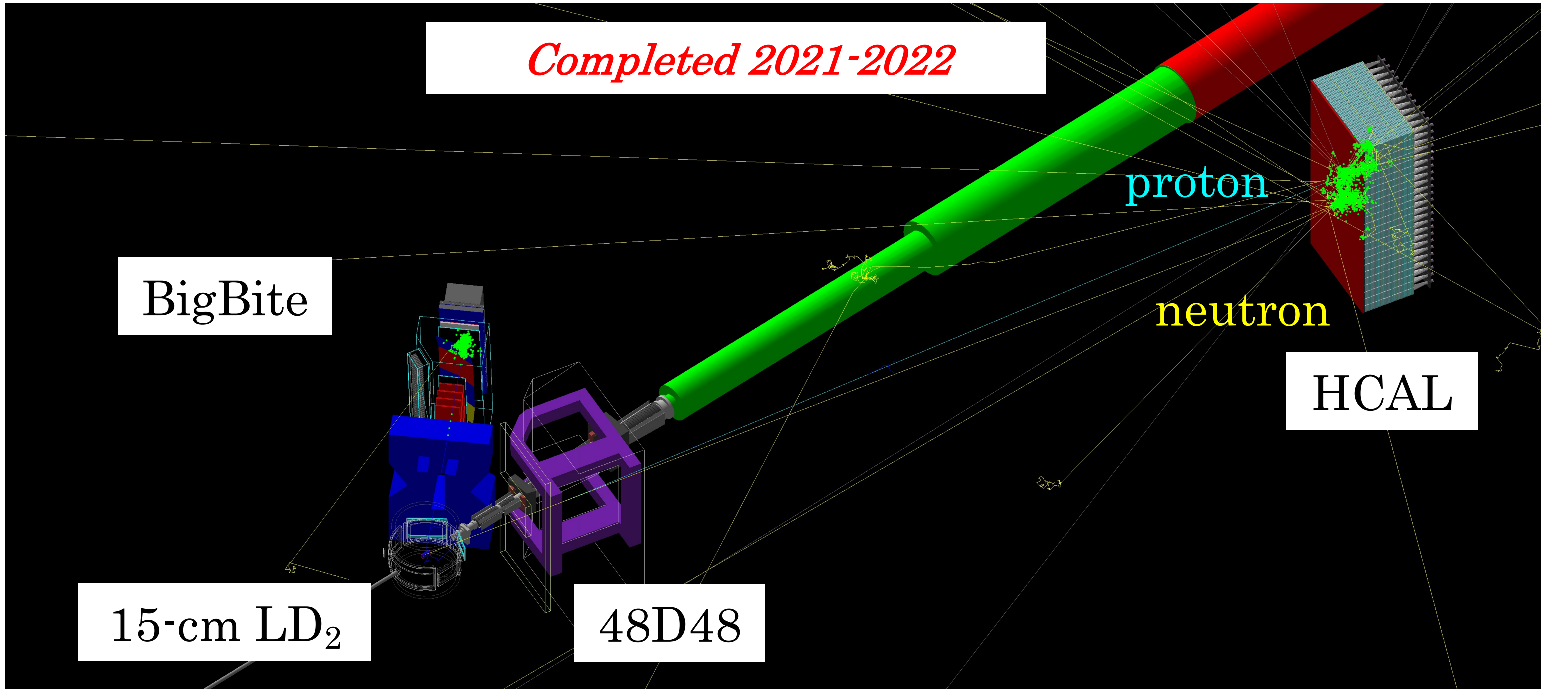
Existing SBS Software

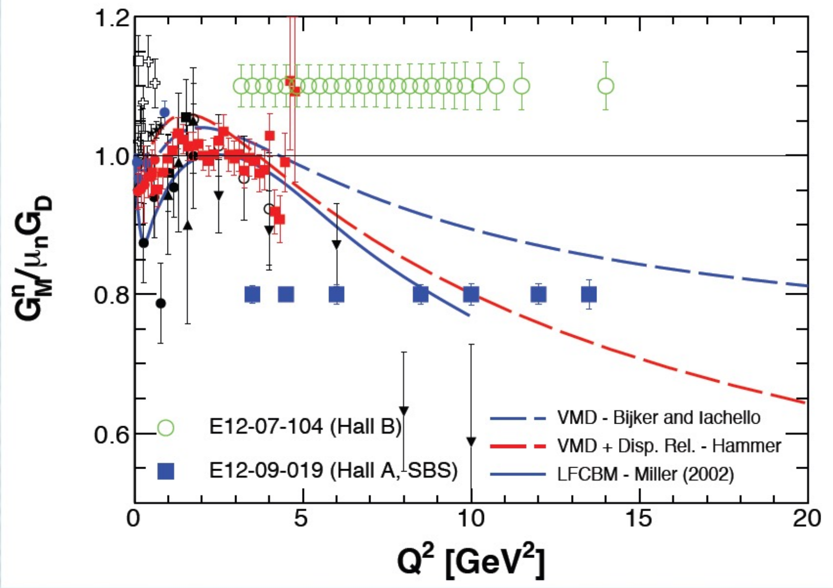
- SBS online/offline analysis software is based on Podd, the standard C++/ROOT-based Hall A analysis framework, and uses the ROOT-based "panguin/OnlineGUI" for online monitoring plots for shift workers.
- Existing repositories:
  - **SBS-offline**: (primary authors: S. Riordan, A. Puckett, E. Fuchey, O. Hansen, J. C. Cornejo, M. Jones, R. Montgomery, D. Hamilton, *et al.*) https://github.com/JeffersonLab/SBS-Offline Main software repository of SBS-specific libraries and source code. Includes raw data decoders that aren't yet standardized under Podd for new readout modules such as MPD w/VTP and VETROC
  - SBS-replay: (principal authors: A. Puckett, E. Fuchey, O. Hansen, S. Seeds, P. Datta, D. Hamilton, others) https://github.com/JeffersonLab/SBS-replay Repository for analyzer database files, replay scripts, analysis and calibration macros, online GUI configuration files, etc. No build system. Just a collection of files. This repo is needed to analyze GMN/nTPE data.
  - **Libsbsdig**: (principal author Eric Fuchey) https://github.com/JeffersonLab/libsbsdig Main library for digitization of simulation output; translates *g4sbs* output (hit time, position, energy deposit) into simulated raw detector signals ("pseudo-data"), populates "hit" data structures used by reconstruction (ADC, TDC, crate, slot, channel, etc); purpose is to test and develop reconstruction algorithms on simulated events using identical algorithms to those used for real data: **Crucial for high-rate tracking studies done with simulation so far**
  - **G4sbs**: (principal authors Andrew Puckett, Seamus Riordan, Eric Fuchey, many, many contributors) https://github.com/JeffersonLab/g4sbs GEANT4-based simulation of all of the major SBS experiments. Documentation at https://hallaweb.jlab.org/wiki/index.php/Documentation_of_g4sbs
  - **SBSGEM_standalone**: (principal author A. Puckett) https://github.com/ajpuckett/SBSGEM_standalone standalone GEM reconstruction code, takes decoded raw data (after common-mode/pedestal subtraction and zero suppression), does clustering, tracking, and alignment. Still useful here and there, but mostly superseded by analyzer/SBS-offline. No longer under active development.

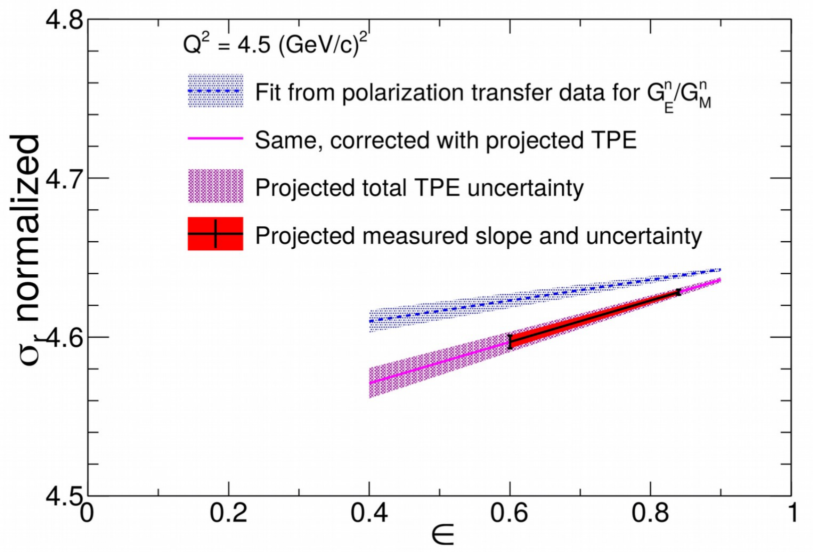
SBS software working group

- Mailing list: https://mailman.jlab.org/mailman/listinfo/Sbs_software
- Standing weekly meeting; currently Fridays at 1:00 PM
- SBS Software Coordinator: Andrew Puckett
- Core software working group members/participants:
  - JLab: J.-O. Hansen, A. Camsonne, M. Jones, S. Barcus, D. Flay, H. S.-Vance, R. Michaels, B. Wojtsekhowski, etc.
  - Glasgow U: R. Montgomery, D. Hamilton, R. Marinaro
  - Syracuse: W. Xiong
  - UConn: A. Puckett, E. Fuchey, P. Datta, S. Seeds
  - Hampton U: M. Kohl, T. Gautam, M. Suresh, others…
  - Northern Michigan U: W. Tireman
  - W&M: M. Satnik, E. Wertz, *et al.*
  - UVA: S. Jeffas, A. Rathnayake, J. Boyd *et al.*

E12-09-019: Neutron magnetic form factor G_M^n to $Q^2 = 13.5$ GeV^2

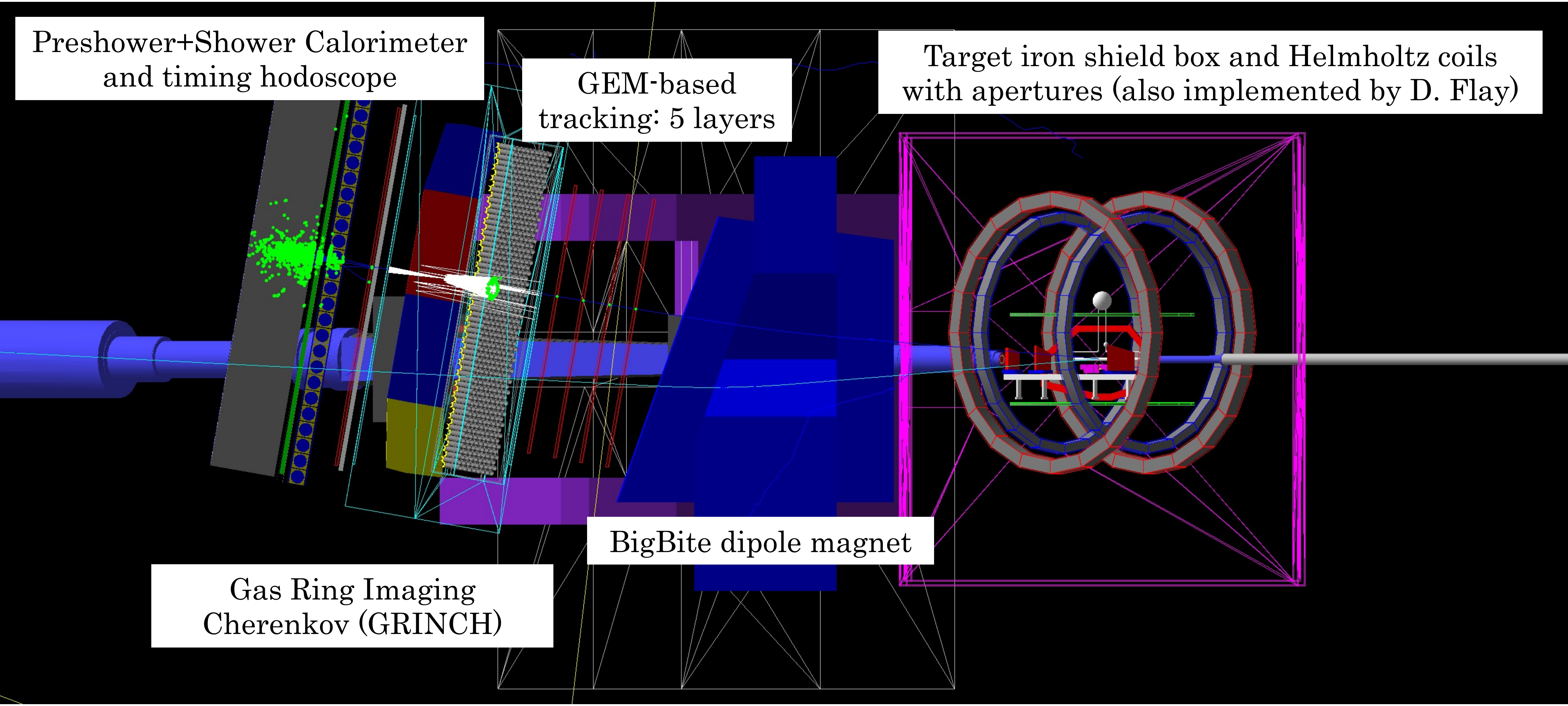




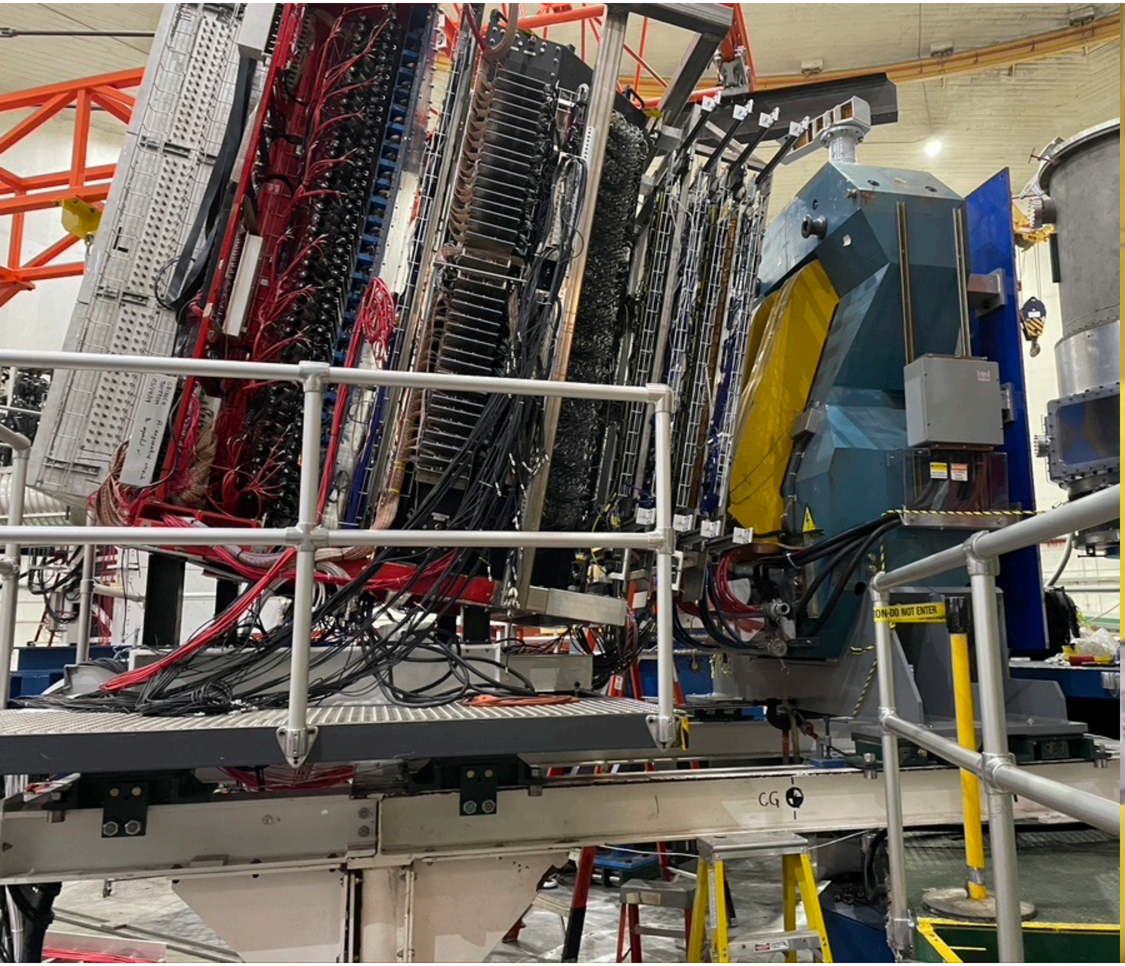


- E12-09-019 measured the neutron magnetic form factor G_M^n to 13.5 GeV^2 using the "ratio" method on deuterium.
- E12-20-010, a recently approved "add-on" measurement, determined the Rosenbluth slope in elastic *en* scattering for the first time at $Q^2 = 4.5\ GeV^2$
- Uses hadron calorimeter for efficient nucleon detection; magnetic deflection for nucleon charge ID
- BigBite detects electron, defines $\vec{q}$ vector, vertex for selection of quasi-elastic scattering

BigBite Spectrometer in Monte Carlo (w/GEN-II target)



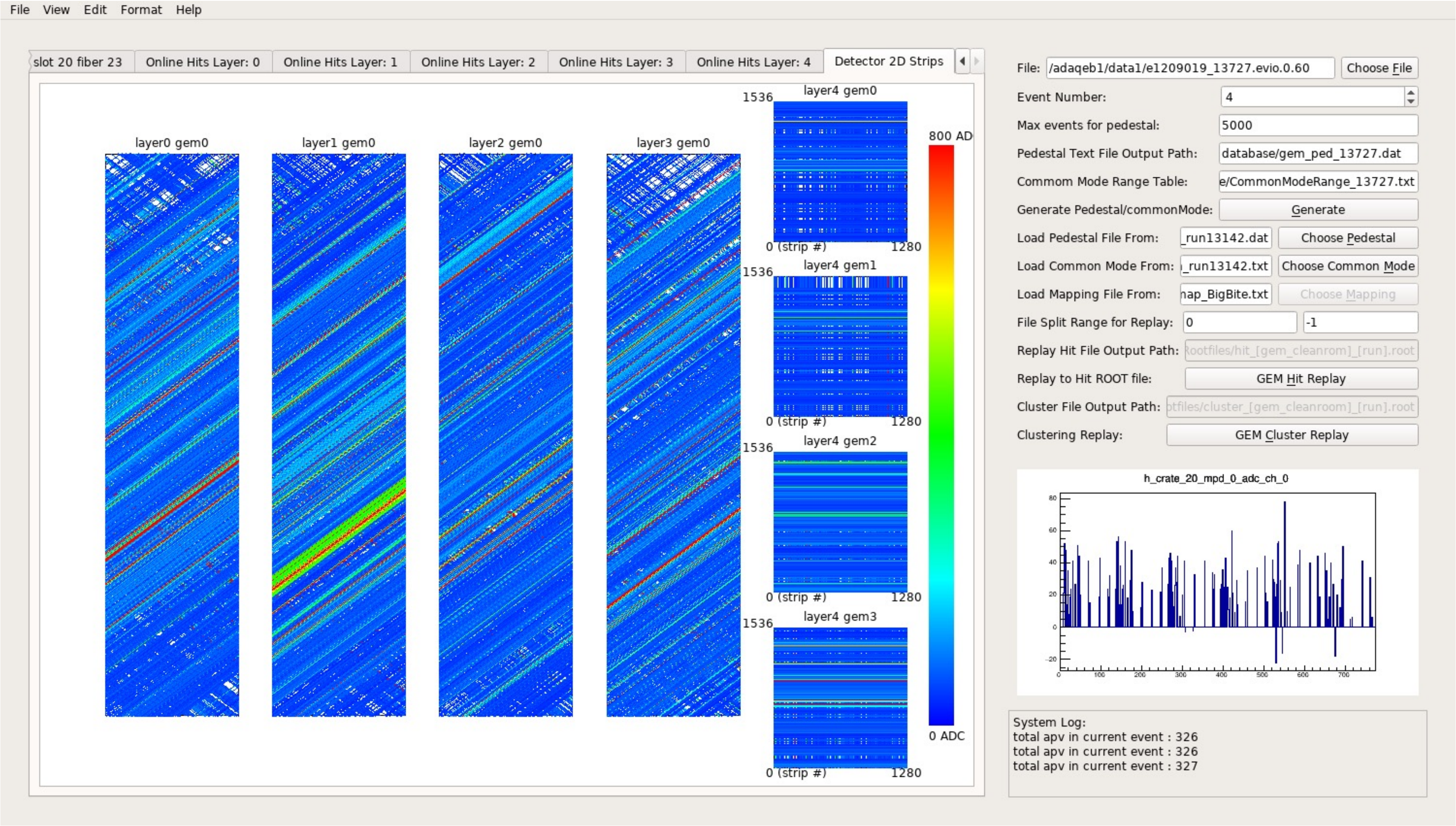
BigBite Spectrometer in Hall A



GMN/GEN/nTPE analysis software requirements

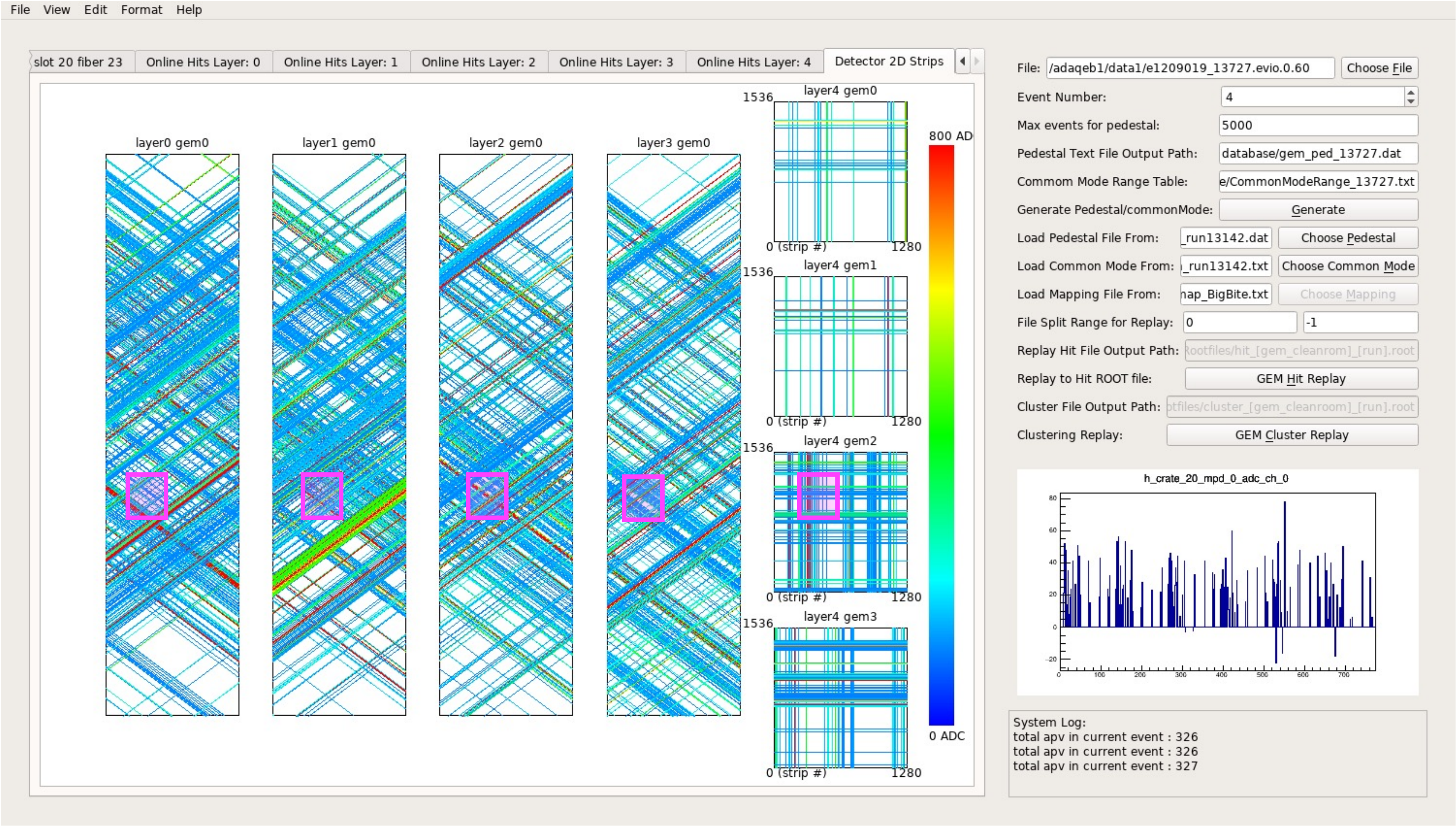
- BigBite spectrometer—reconstruct full kinematics of scattered electron:
  - Momentum reconstruction: $\frac{\sigma_p}{p} \approx 1 - 1.5\%$
  - Angular resolution: $\sigma \approx 1 - 2\, mrad$ (in-plane and out-of-plane)
  - Vertex resolution: $\sigma_z \leq 1\, cm$
  - Predict $\vec{q}$ direction and neutron position at HCAL from Q.E. kinematics
  - Suppress charged pions using preshower calorimeter (plus GRINCH?)
  - Straight-line tracks in field free regions → simple and reliable data analysis!
- Hadron Calorimeter HCAL—reconstruct neutron kinematics; nucleon charge ID:
  - Time resolution $\sigma_t \approx 0.5 - 1\, ns$
  - Angular resolution ~2 mrad
  - Reconstruct missing parallel and perp. momenta, reject protons and inelastics

What we're up against, I (run 13727, 12 uA LD2)



- **Event display credit: Xinzhan Bai (UVA).** This shows all strips passing (online) zero suppression for one BigBite trigger, color coded by ADC value
- **This is where most of the ~1 GB/s of SBS raw data volume comes from!**

What we're up against, II (run 13727, 12 uA LD2, $Q^2 = 4.5\ GeV^2, E = 4\ GeV$)



- This is the same event as previous slide, but requiring max ADC sample on a strip greater than 100, a typical offline threshold for cluster maxima that is higher than online threshold

□ = approximate size of calorimeter-constrained track search region at each layer

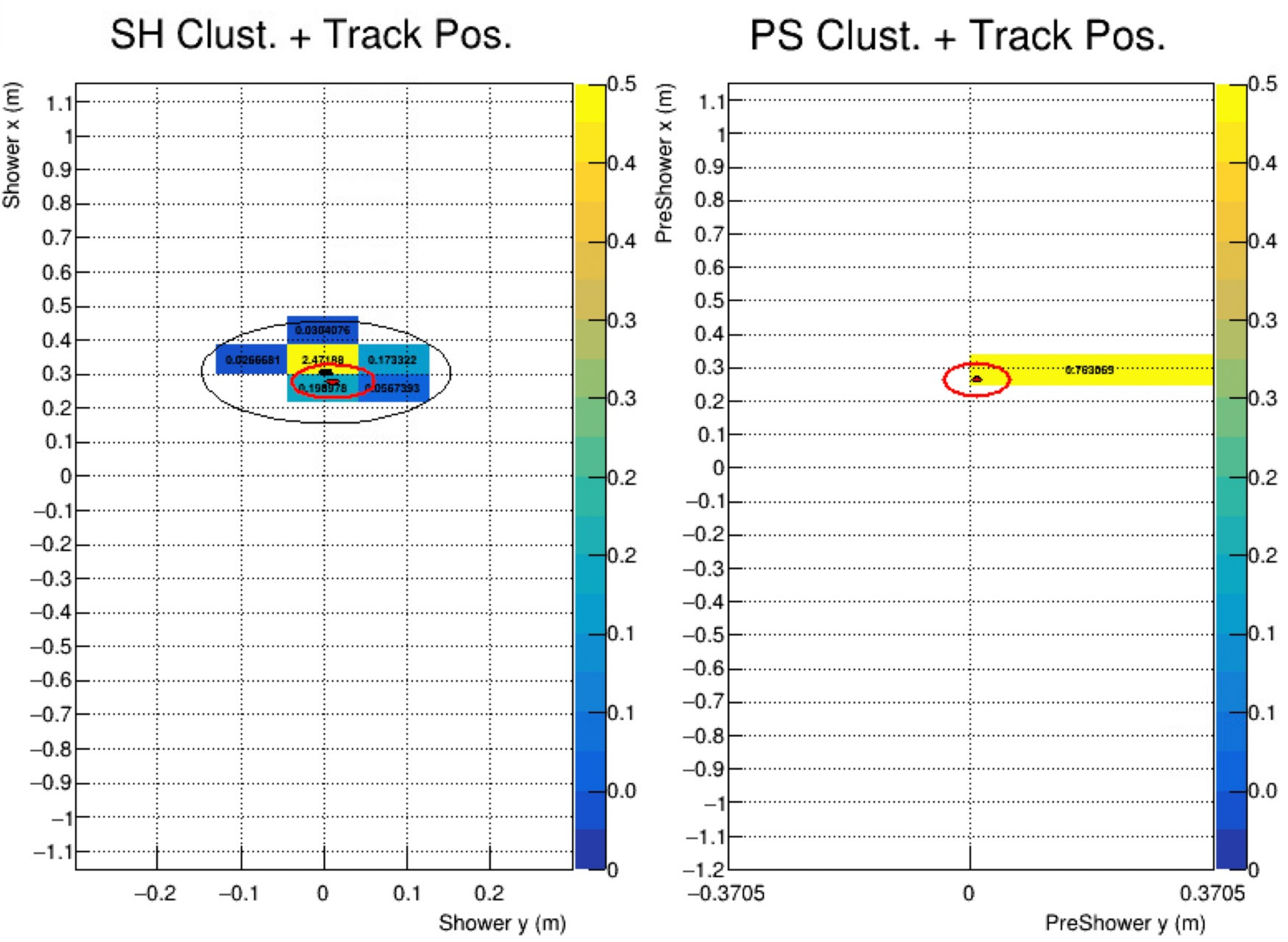
Benchmarking SBS-offline/analyzer performance

puckett — a-onl@aonl2:~/sbs/logs — ssh -X hallgw.jlab.org — 97×36

```
RawDecode_master          1                         61891      61891 (100%)

BLOCK: Decode
Decode_master             1                         61882      61882 (100%)

BLOCK: CoarseReconstruct
CoarseReconstruct_master  1                         61882      61882 (100%)

BLOCK: Physics
GoodPreShower             bb.ps.ngoodADChits>0      61882      61824 (99.9%)
GoodShower                bb.sh.ngoodADChits>0      61882      61863 (100%)
GoodBBCAL                 GoodPreShower||GoodShower 61882      61882 (100%)
CutPreShower              bb.ps.e>0.15              61882      16215 (26.2%)
GoodTrack                 BB.gold.index>-1          61882      7977  (12.9%)
GoodElectron              GoodTrack&&CutPreShower   61882      2658  (4.3%)
Physics_master            GoodTrack                 61882      7977  (12.9%)

Timing summary:
Init             : Real Time =    3.66 seconds Cpu Time =    3.64 seconds
Begin            : Real Time =    0.02 seconds Cpu Time =    0.02 seconds
RawDecode        : Real Time = 1552.63 seconds Cpu Time =  315.09 seconds
Decode           : Real Time =  857.77 seconds Cpu Time =  857.82 seconds
CoarseTracking   : Real Time =    0.20 seconds Cpu Time =    0.16 seconds
CoarseReconstruct: Real Time =   13.24 seconds Cpu Time =   12.96 seconds
Tracking         : Real Time = 1566.26 seconds Cpu Time = 1564.06 seconds
Reconstruct      : Real Time =    2.46 seconds Cpu Time =    2.33 seconds
Physics          : Real Time =    0.18 seconds Cpu Time =    0.17 seconds
End              : Real Time =    0.18 seconds Cpu Time =    0.18 seconds
Output           : Real Time =  222.04 seconds Cpu Time =  221.27 seconds
Cuts             : Real Time =    2.50 seconds Cpu Time =    2.82 seconds
Total            : Real Time = 4223.47 seconds Cpu Time = 2982.46 seconds
[a-onl@aonl2 logs]$
[a-onl@aonl2 logs]$
[a-onl@aonl2 logs]$
[a-onl@aonl2 logs]$
[a-onl@aonl2 logs]$
```

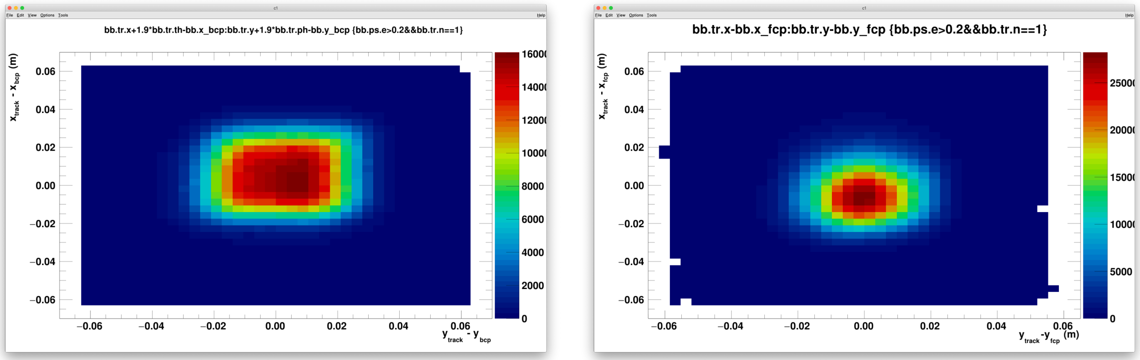
- This is a replay of one 20 GB EVIO file from 13798 (2nd-to-last production run of GMN), at 12 uA on LD2 (about 20 s of data taking)
- Analysis time in this example is ~50% "Raw Decode" plus "Decode" and ~50% Tracking (Note: "Raw Decode" is dominated by reading data from disk at least for this particular job, that ran from the counting house, but large non-CPU time in "RawDecode" is not typical on e.g., the batch farm)
- **Processing rate (per single-thread analysis job) under these conditions is ~15 Hz per unit real time or ~21 Hz per unit CPU time.**
- **In a 2019 review for GEN-RP, we promised 8 Hz for GMN, so our replay is running nearly 3X faster per single-thread job than we promised**
- **There is plenty of scope for further improvement in the tracking speed (I have less confidence in the prospects for speeding up raw data decoding)**

Shower constraint on GEM track search region



- BBCAL event display **by Provakar Datta (UConn)** with shower/preshower cluster position and energy, and track projections

- Calorimeter cluster position and energy used to define track search region
- Search region optimized for electrons
- Position and energy from calorimeter combined with first-order optics of BigBite to constrain both the position and slope of good electron tracks.
- This allows us to define a small search region for tracking; presently 12x12 cm^2 at the front and back of the GEM stack or about **2-3% of the GEM active area**



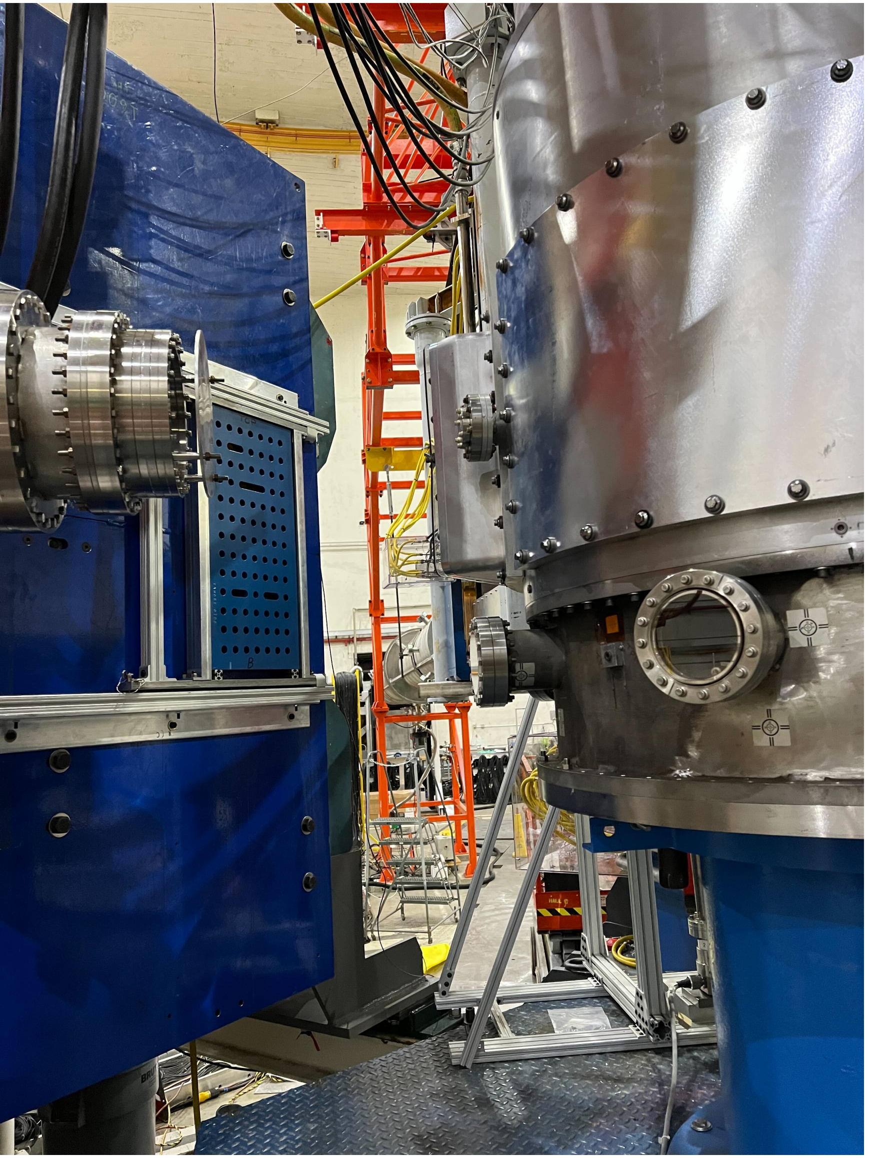
Back Front

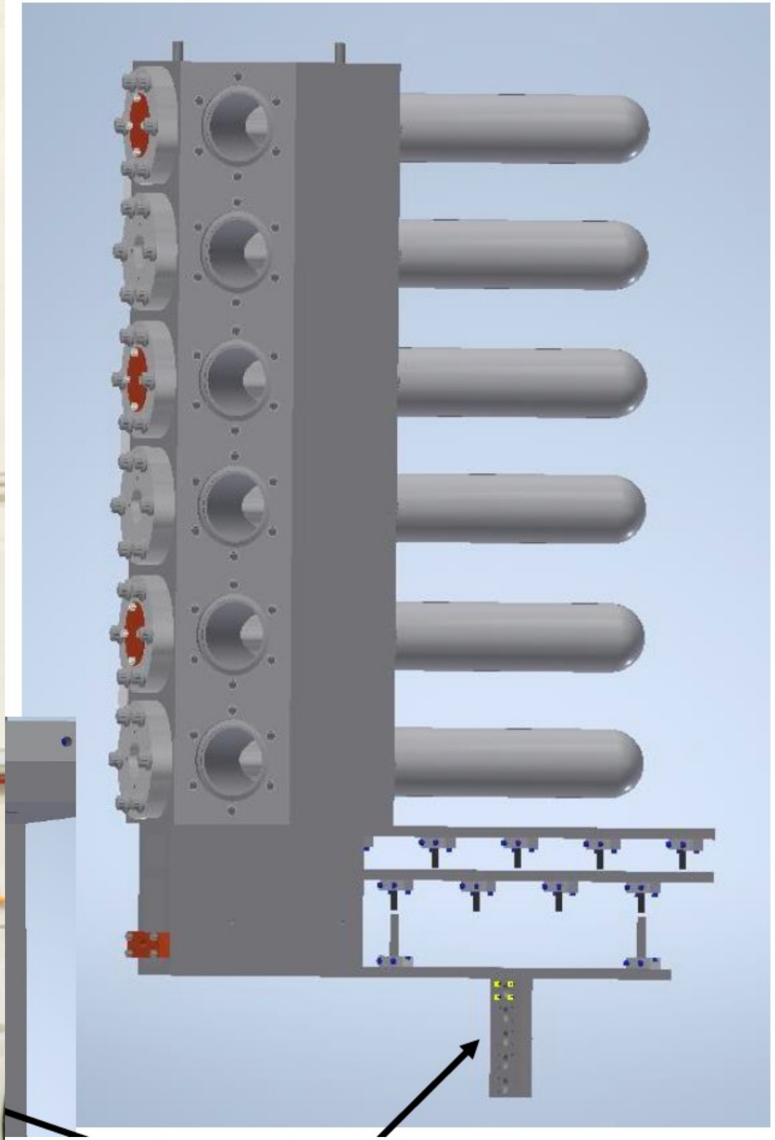
- Distributions of reconstructed track position within calorimeter-defined search region at back (left) and front (right) of GEM stack, with cut on pre-shower energy > 0.2 GeV (reject pions)

How BigBite Tracking Works (Simplistically):

- Perform 1D clustering of strips along each dimension in each GEM chamber
- Form all possible 2D combinations from 1D clusters within calorimeter-defined search region.
  - **Number of fake 2D hit candidates is proportional to the *square* of the number of real hits!**
- Filter 2D hits according to criteria such as cluster ADC sum, correlation coefficient, X/Y (or U/V) ADC asymmetry, timing, etc.
- Divide each tracking layer into a uniform 2D rectangular grid, accumulate a list of (2D) hit candidates in each grid bin.
  - When GEMs are aligned internally, we currently use grid bin width of 1 x 1 cm^2 (bin size ~100X spatial resolution along each dimension)
- Loop on all possible combinations of one hit from the two outermost layers (within search region)
- Form straight-line projection from hits in outermost layers to inner layers.
- Loop on all possible combinations of one hit from each inner layer, in grid bins consistent with straight line projection from outer layers, find the hit combination with best $\frac{\chi^2}{ndf}$ (and possibly other criteria)
- We also impose some basic track quality/track slope/optics-based constraints on "track candidates" to reject obviously bad hit combinations within the search region
- Initially, we require hits on all 5 layers. If we don't find a track at the maximum hit requirement, we decrease the hit requirement by one and repeat for all possible combinations of 4 out of 5 layers, then all possible combinations of 3 out of 5. We do not consider two-hit "tracks" as they have no degrees of freedom (we can always draw a straight line between any two points). We repeat until we run out of "unused" hits in the search region.
  - With a 3-hit minimum, the "fake track" probability is still rather significant at high rates
  - More hits = more confidence that the "track" is real
  - At a given minimum hit requirement, we treat all possible layer combinations on an equal footing, but this could be questionable given the tracking geometry.

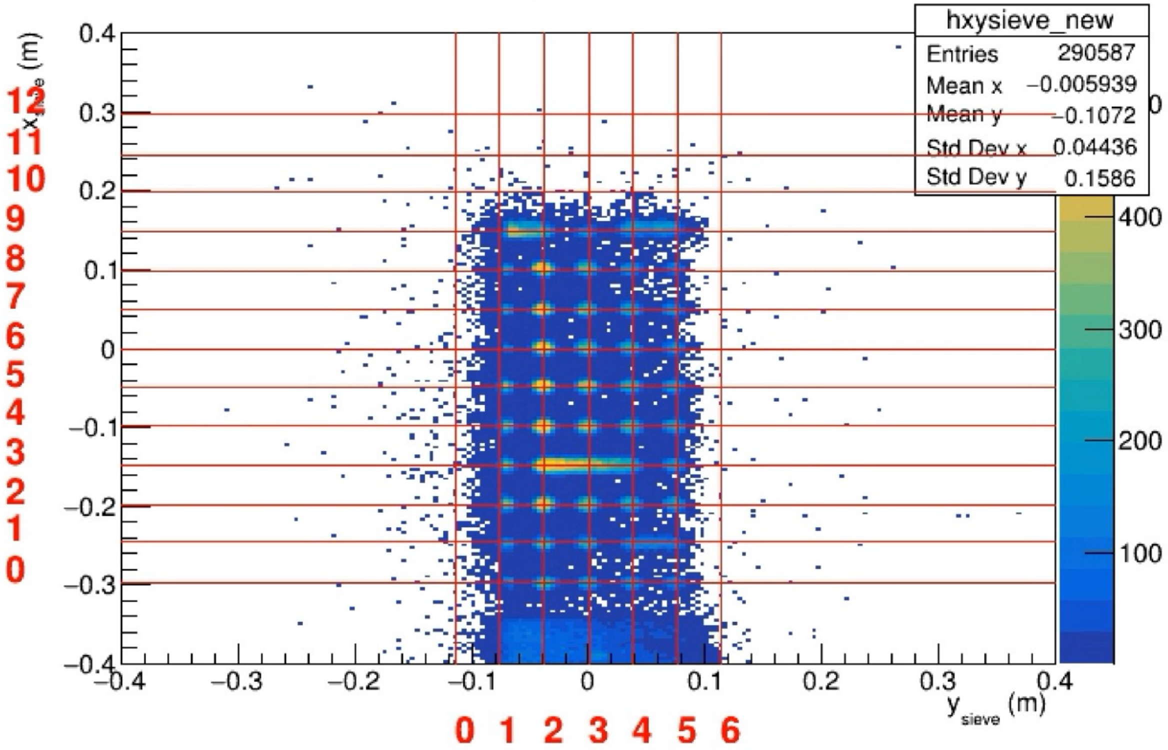
BigBite Optics, I: Hardware





BigBite, Optics II: Zero-field alignment

Post fit results, and application to the track data:



GEMX0 = -0.213869
GEMY0 = 0.0394128
GEMZ0 = 3.12109
GEMtheta = 8.82378 deg
GEMphi = 179.40567 deg
ZSIEVE = 1.59027

- See, e.g.: https://sbs.jlab.org/DocDB/0001/000157/001/str_trk_sbs8.pdf
- With all magnets off, use single-foil carbon target as point source of straight-line tracks from the origin of Hall A through the sieve slit and GEMs.
- Use known dimensions of sieve slit and GEMs to determine absolute position and orientation of GEMs and position of sieve in Hall
- Critical to relating internal GEM/detector coordinate system to ideal optics coordinate system.

BigBite Optics, III: Formalism

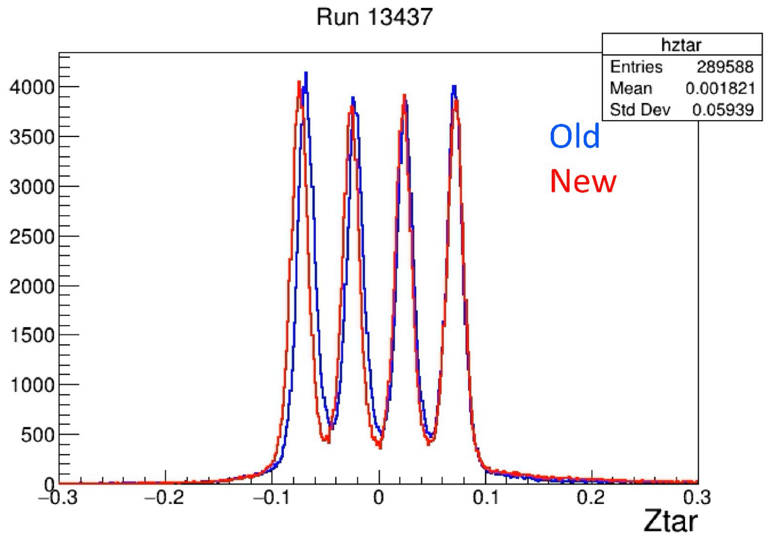
- Generate starting optics model by tracking electrons through *g4sbs* (GEANT4) Monte Carlo with calculated BigBite field map, fit 2nd-order polynomial expansion of target variables in terms of detector ("focal plane") variables
- Absolute trajectory bend angle determination for electrons of known momentum is achieved by comparing LH2 elastic electrons going through the central sieve hole in BigBite with straight-through tracks going through the central sieve hole. This measurement is independent of any optics model or absolute knowledge of the BigBite magnetic field.
- The calibration above determines the appropriate scale factor for the simulation field map to generate the starting matrix elements

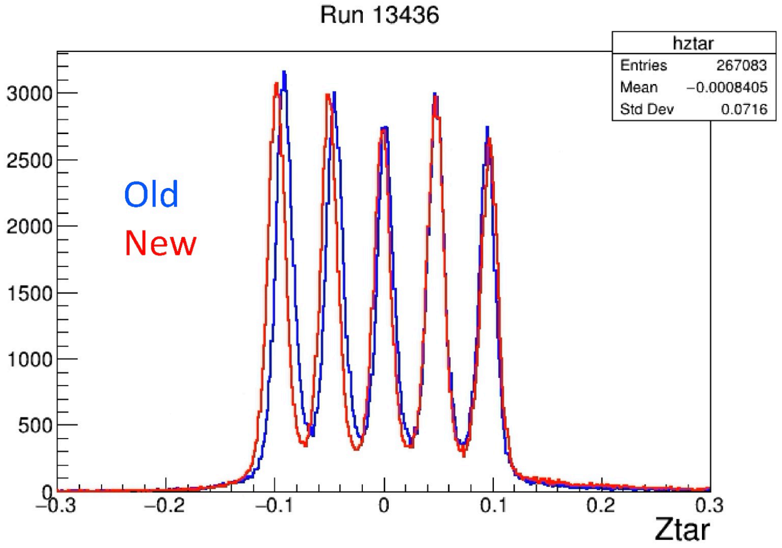
```
#bb.gemphi = 180.8818

# straight-track results from Holly:
#bb.gemorigin_xyz = -0.213869 0.0394128 3.12109
#bb.gemtheta = 8.82378 # degrees
#bb.gemphi = 179.406 # degrees
# new straight-track results from Andrew (maybe slight improvement):
bb.gemorigin_xyz = -0.21323 0.03872 3.11940
bb.gemtheta = 8.81847
bb.gemphi = 179.534

# NOTE: ZSIEVE = 1.59074

# starting model for SBS-8 from g4sbs: updated with scale factor 0.9644 for BB field map
bb.optics_order = 2
bb.optics_parameters =
   -0.013951072 -6.7189159e-05  0.00022486239     0.26957803  0 0 0 0 0
     0.43275107   0.0010961293  -0.0022068665     0.12271782  1 0 0 0 0
   -0.029264943   0.0089086215   -0.018497904   -0.018303949  2 0 0 0 0
    0.015286973  -0.0090579402      1.0091097    0.006524678  0 1 0 0 0
     -0.3771755    0.097156428    -0.23342715    0.018574322  1 1 0 0 0
   -0.014208137    0.023096216   -0.032298685    0.078452456  0 2 0 0 0
    -0.34635355 -0.00090695503   0.0018848963   -0.068108122  0 0 1 0 0
    0.037279672   -0.040196986    0.085799908    0.059415542  1 0 1 0 0
     0.28933729     -0.2041523     0.62906153   -0.048050144  0 1 1 0 0
  0.00081558551    0.042209088   -0.099658514   0.0017245797  0 0 2 0 0
   -0.041171438     0.93418238      -2.868918   -0.014706424  0 0 0 1 0
      1.0957664    -0.32511236     0.72772292   -0.045808129  1 0 0 1 0
     0.11481975    -0.44512229      1.0695019    0.036626935  0 1 0 1 0
    -0.87794966     0.84190651     -2.3706859     0.11360024  0 0 1 1 0
      -0.103686     0.94946168     -2.4442919    -0.28814815  0 0 0 2 0
 -3.6600546e-15   1.209214e-14 -3.8572524e-14  -1.527841e-14  0 0 0 0 1
  3.4661576e-15 -1.1778173e-14  2.9222354e-14 -4.0242914e-15  1 0 0 0 1
  6.4569628e-16  7.5649958e-16 -2.2941854e-15 -3.0794741e-15  0 1 0 0 1
  6.1103084e-16 -2.2602392e-15  5.9912025e-15 -2.2093524e-16  0 0 1 0 1
  3.4916311e-16 -5.7801037e-16 -8.8795124e-16 -1.1214327e-14  0 0 0 1 1
 -5.5736366e-16  2.2786736e-16  4.0535244e-17  5.8646239e-15  0 0 0 0 2

```

$$(x'_{tar}, y'_{tar}, y_{tar}, (p\theta_{bend})) = \sum_{i,j,k,\ell,m=0}^{i+j+k+\ell+m\leq 2} (C_{x'}, C_{y'}, C_y, C_{p\theta})(x_{fp})^i (y_{fp})^j (x'_{fp})^k (y'_{fp})^\ell (x_{tar})^m$$

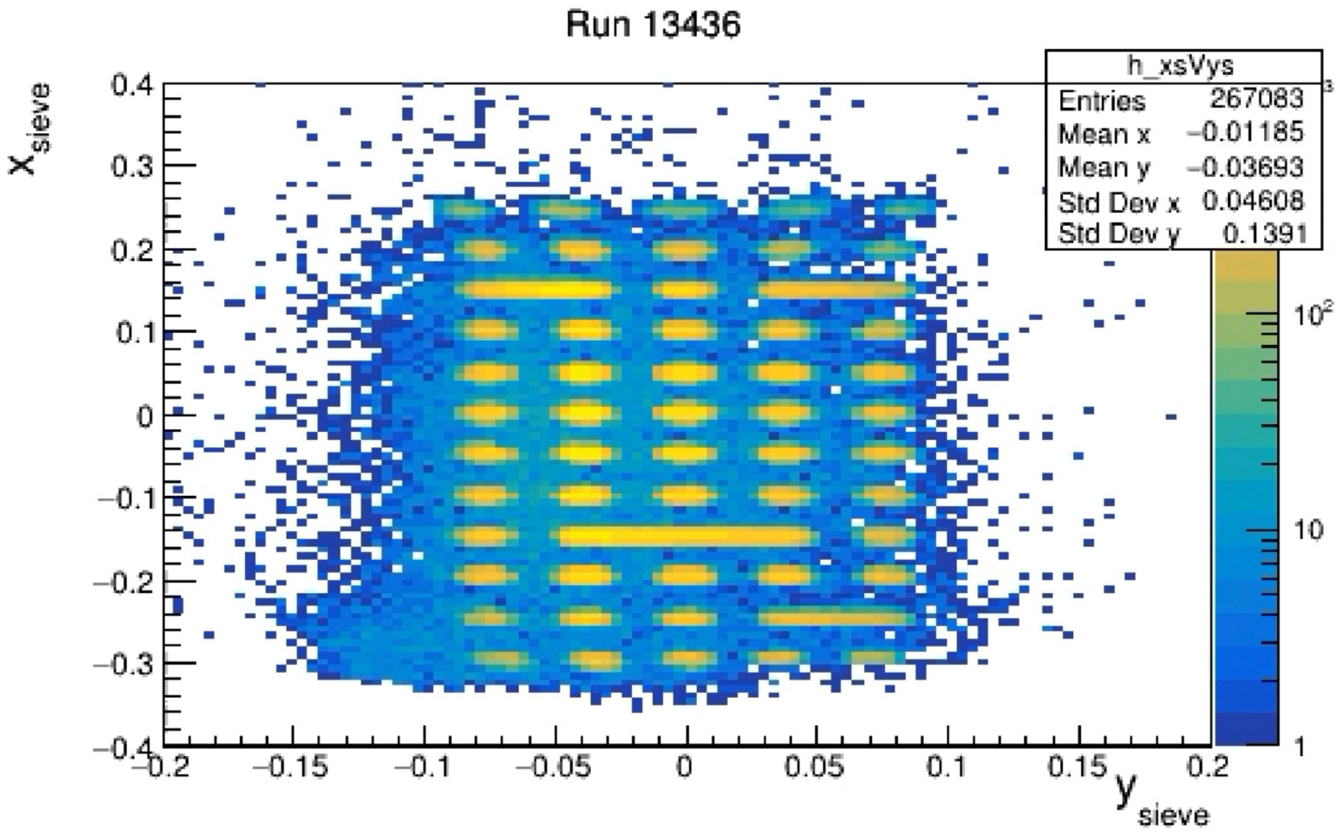
$$x_{tar} \approx 0$$

BigBite Optics, III: Angle and vertex reconstruction



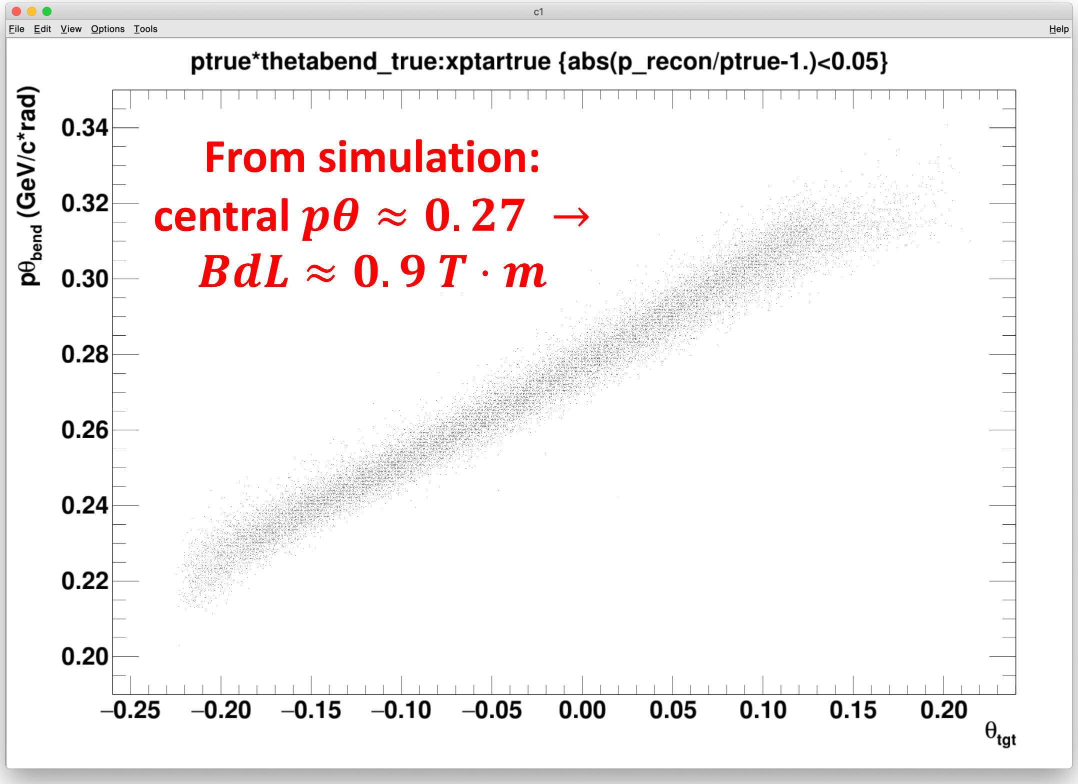


New sieve reconstructed (all foils)



- **Optics analysis credit: Holly Szumila-Vance (JLab)**
- 4-foil and 5-foil optics targets with sieve slit: “Old” = starting optics from simulation, “New” = fitted optics.
- https://sbs.jlab.org/DocDB/0001/000159/001/sbs8_optics_optimization.pdf

BigBite Optics, IV: Momentum



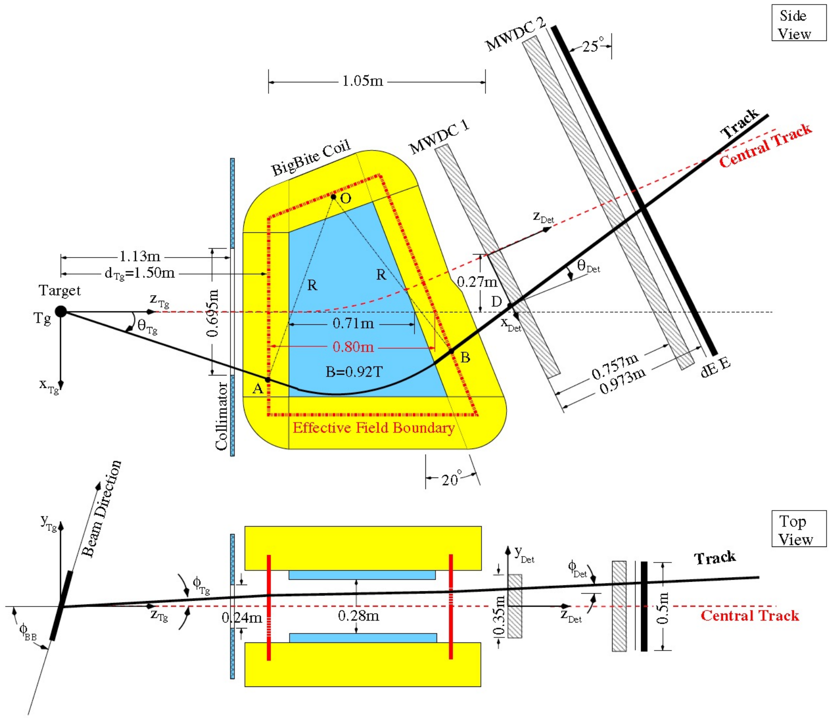


Fig. 4. The schematic of the dispersive (top) and non-dispersive (bottom) planes of the BigBite spectrometer. Small angular deflections in the non-dispersive plane occur if the particle trajectory is not perpendicular to the effective field boundary [2,19,20]. At the entrance to the magnet, they are at most 18 mrad (close to the acceptance boundaries in the dispersive direction). At the exit field boundary, the effect acts in the opposite sense and partially cancels the deflection at the entrance.

Schematic of BigBite in an earlier experiment from NIM A, 686, 20 (2012)

- As a simple dipole, the product of the momentum and the trajectory bend angle is proportional to the field integral seen by the particle along its trajectory:

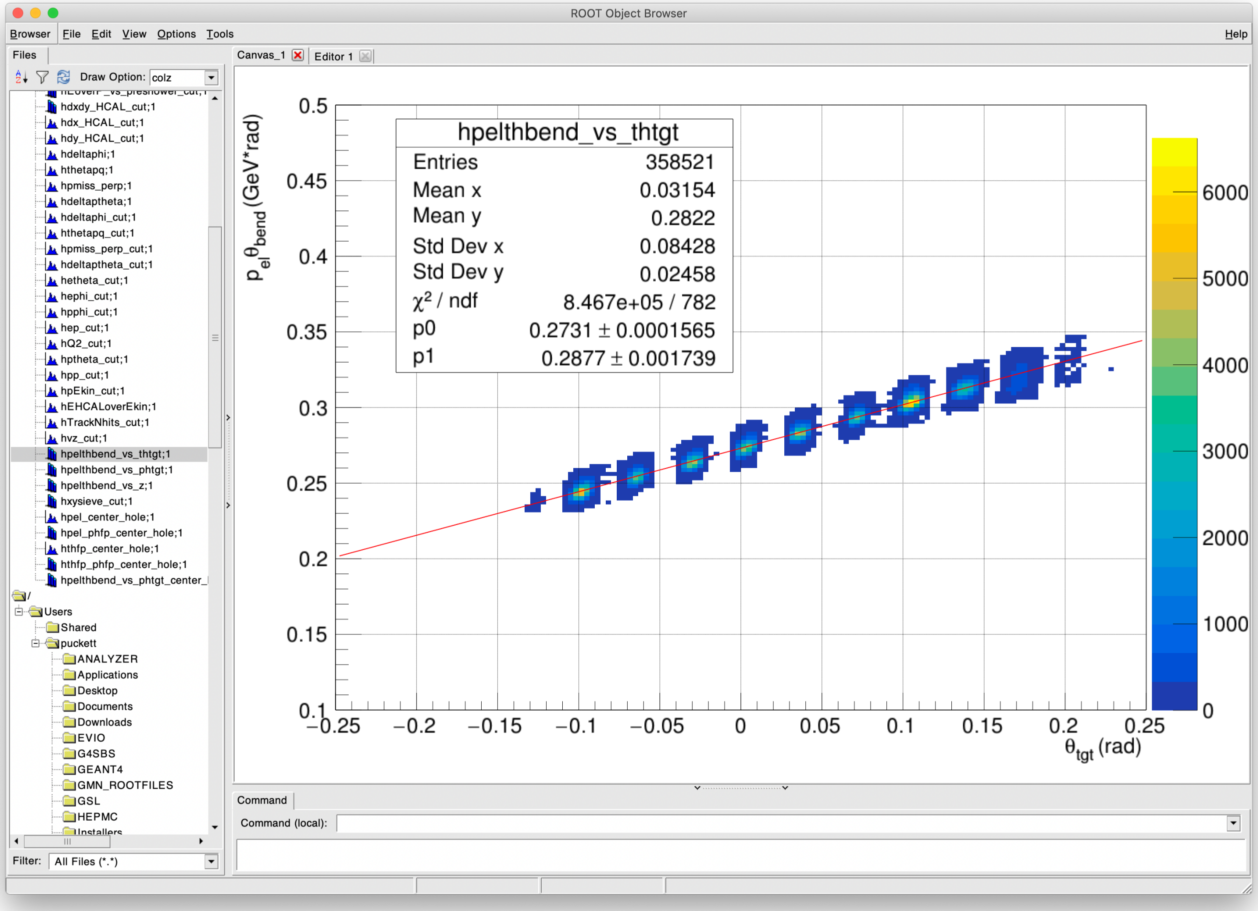
$$p\theta_{bend} = 0.3 \int B_{\perp} dL$$

For any given magnet distance from the target, the BdL (and therefore $p\theta_{bend}$) depends mainly on the dispersive plane angle θ_{tgt}, and the dependence is roughly linear. **Deviations from this first-order model are small**

Absolute bend angle and momentum calibration, SBS-1

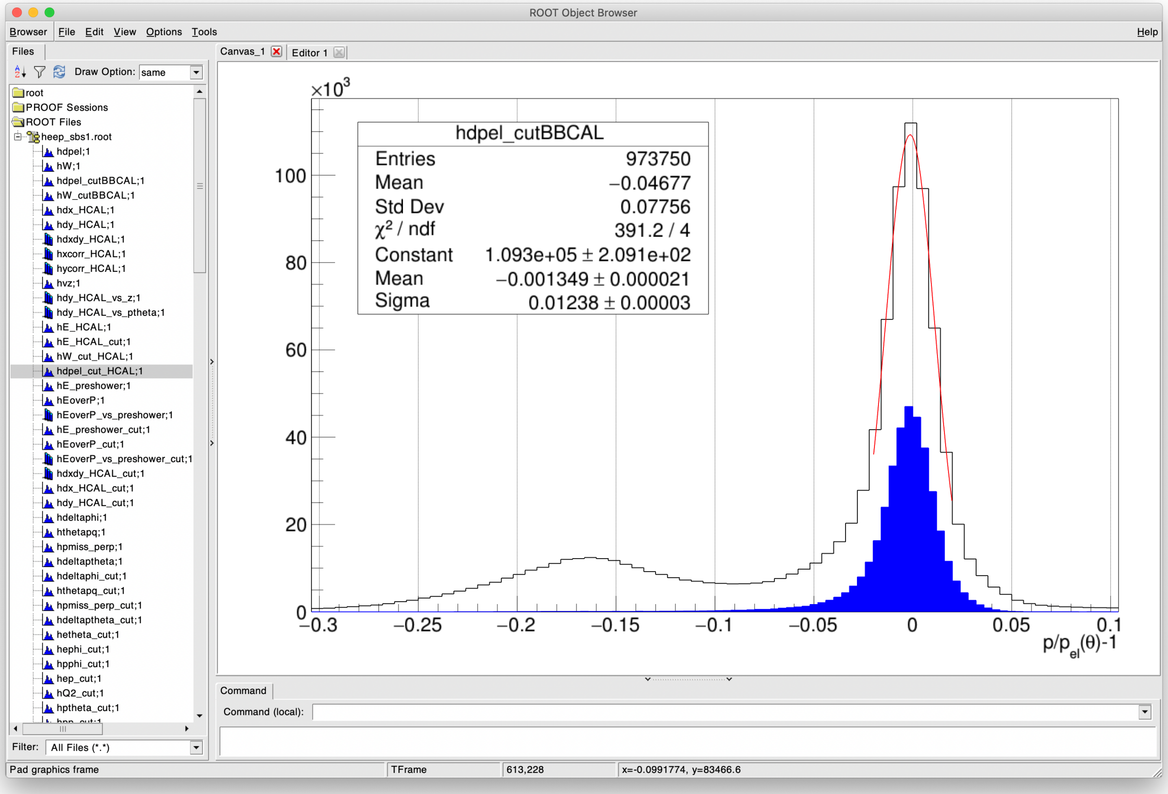
$$E = 1.916 \text{ GeV}$$
$$\theta_{BB} = 51 \text{ deg}$$

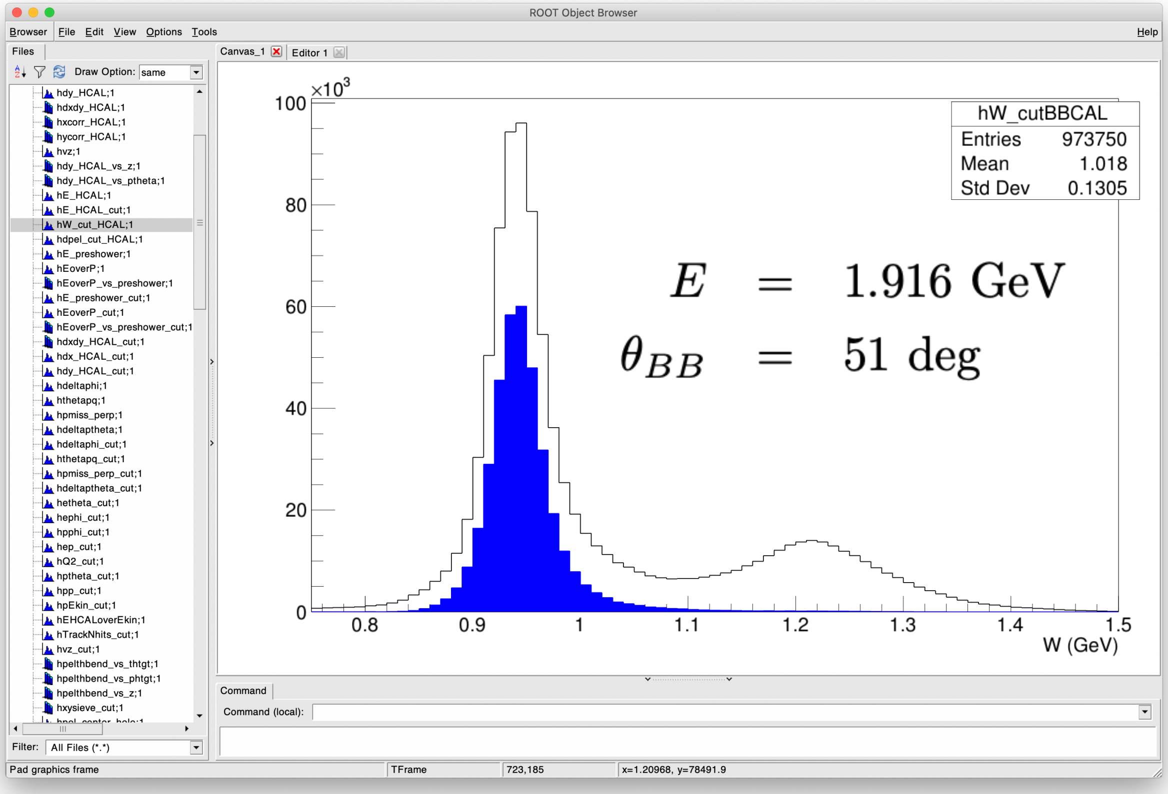
- Central BdL x 0.3 ~= 0.273
- Momentum for an elastic electron going through the center sieve hole is 1.09 GeV
- Bend angle for a 1.09 GeV electron on a central trajectory is 14.35 degrees.
- This implies the best scale factor for the simulation field map to match the actual field strength for production magnet current of 750 A is ~0.96-0.97 (some uncertainty due to beam energy, scattering angle, sieve and magnet survey, etc)



$$p\theta_{bend}(\text{first-order}) = 0.28615 \cdot s\left[1 + (0.1976 + 0.4764 \cdot d_{BB}(\text{m})) \cdot \theta_{tgt}\right] \text{ GeV/c} \cdot \text{rad}$$
$$s \equiv \text{Magnetic field scale factor relative to map\_696A.dat}$$
$$p\theta_{bend}(\text{first-order}) = A\left[1 + (B + C d_{BB})\theta_{tgt}\right]$$

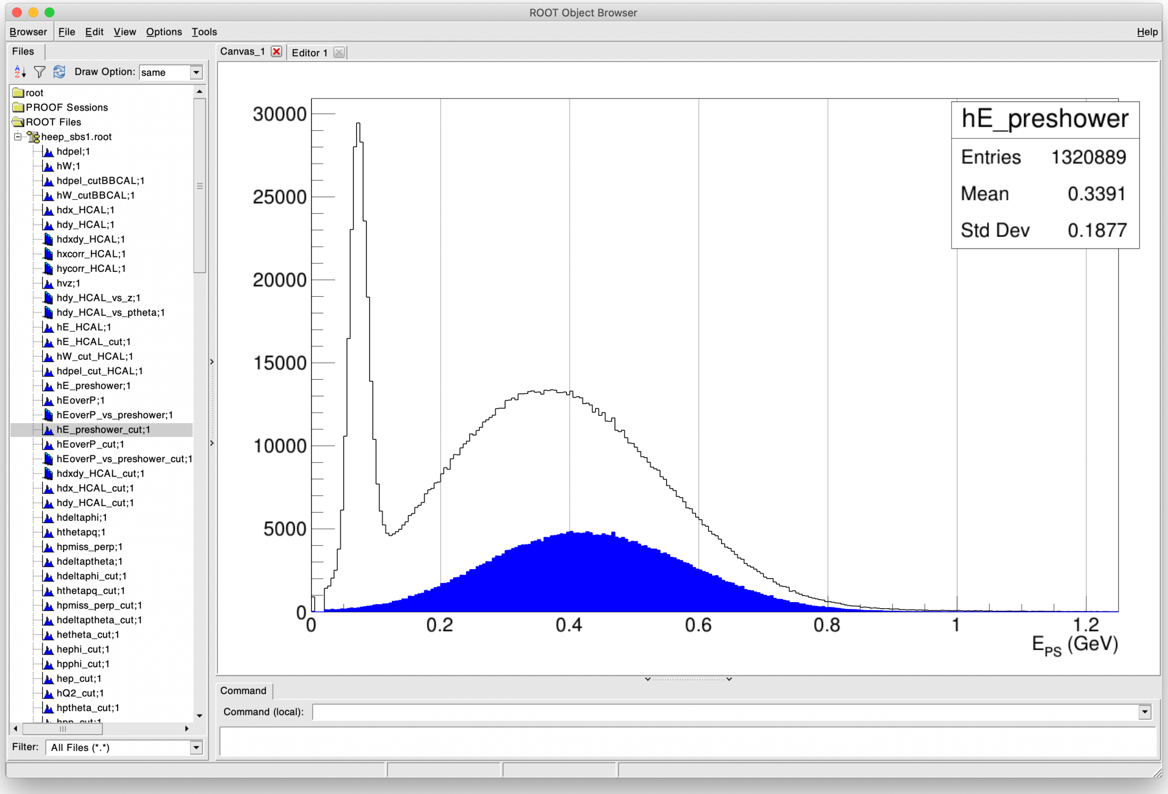
Momentum and Invariant Mass resolution, SBS-1

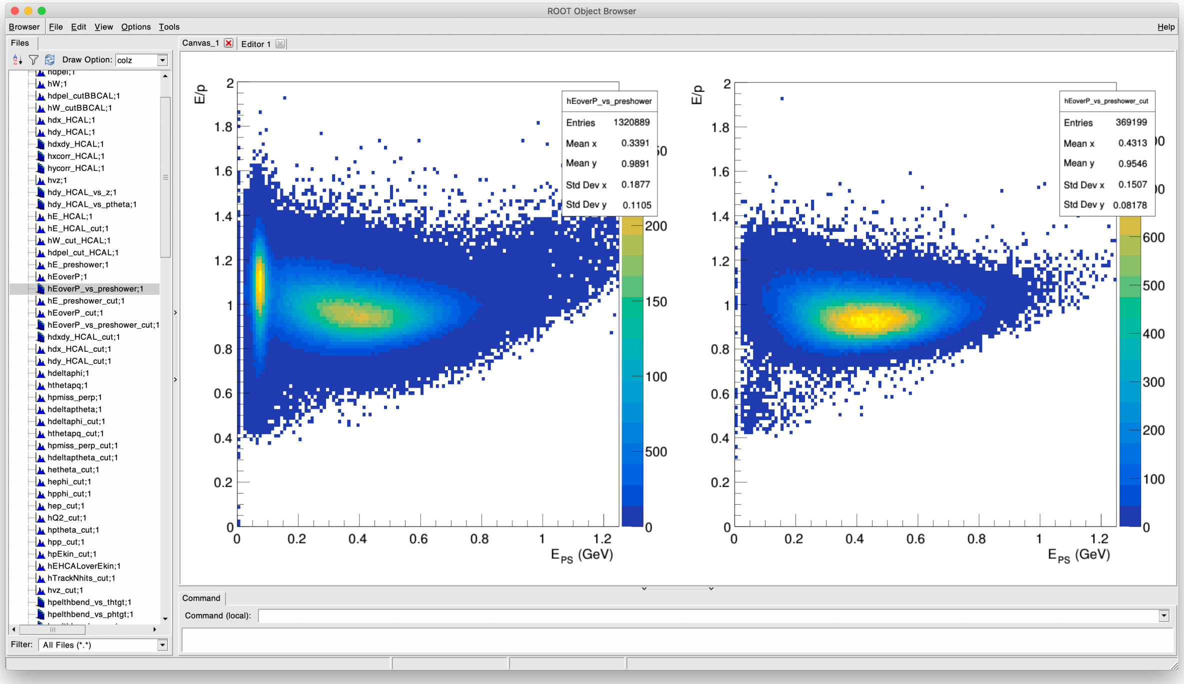




- Momentum resolution about 1.2% (consistent with simulation prediction)
- **Events passing angular correlation cut with HCAL**

BBCAL performance, SBS-1





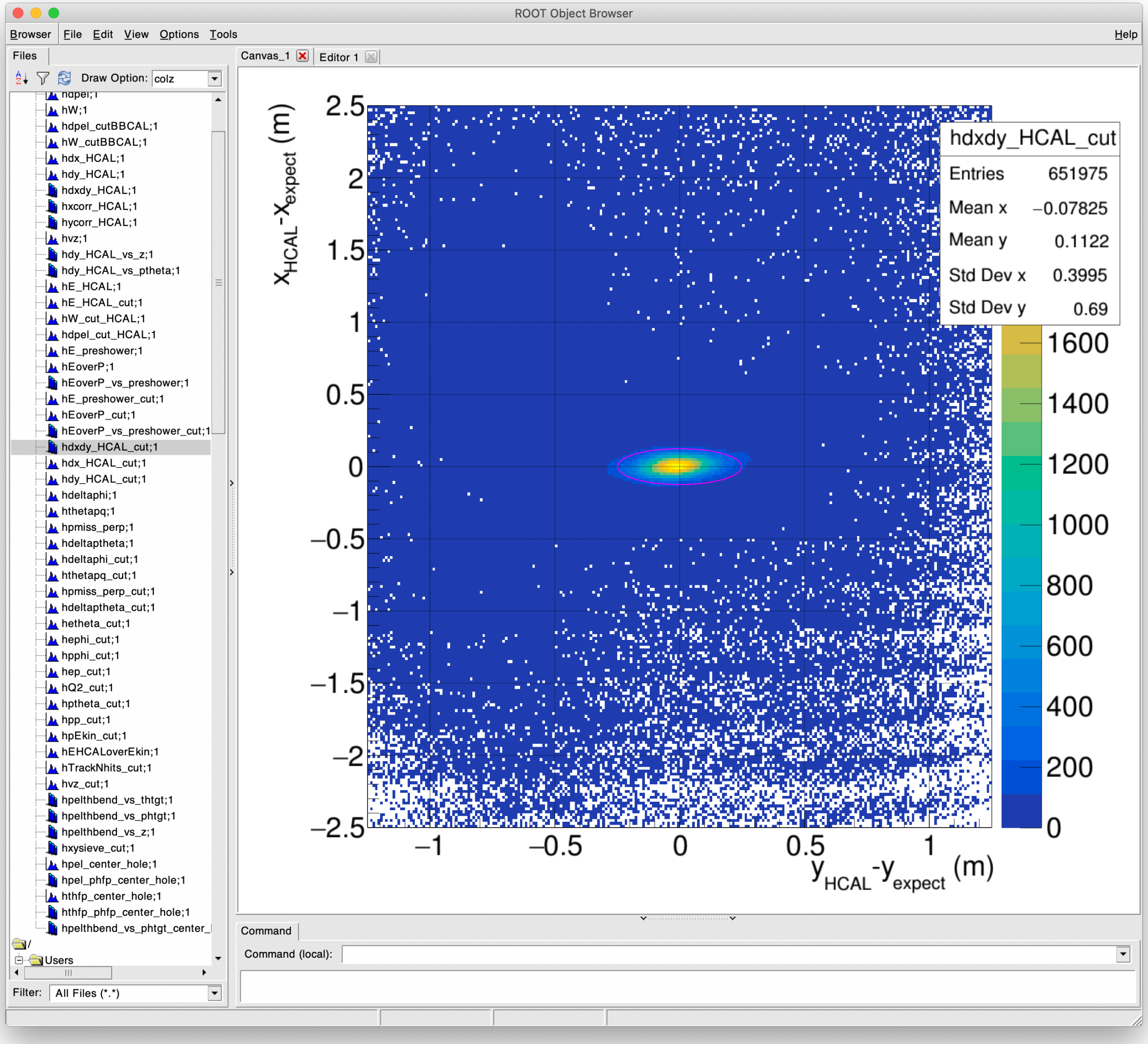
Note: plots above (E/p vs preshower) used initial calibration of BBCAL from cosmic

Preshower energy spectrum

- **Black = All events**
- **Blue = Elastic events including HCAL cuts**

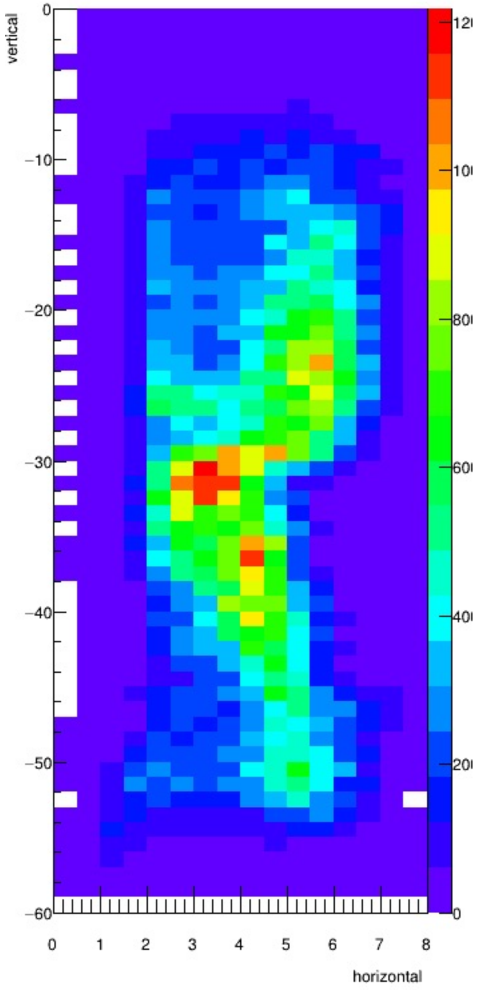
- **Left: all events**
- **Right: elastic events**

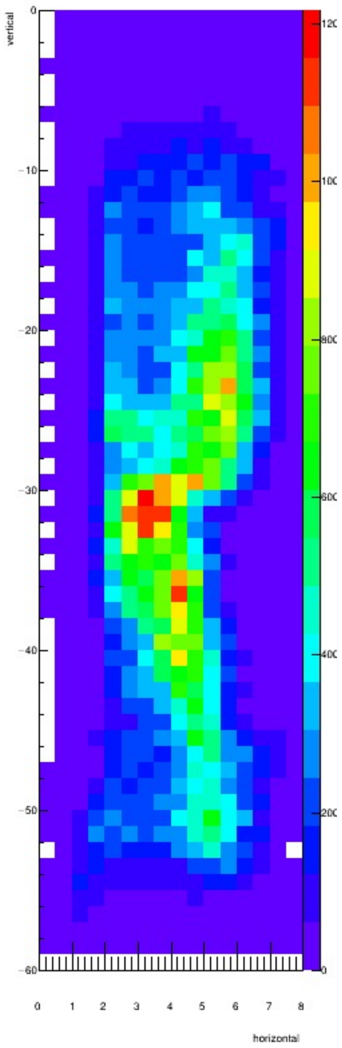
Elastic proton spot in HCAL (SBS field OFF), SBS-1 data

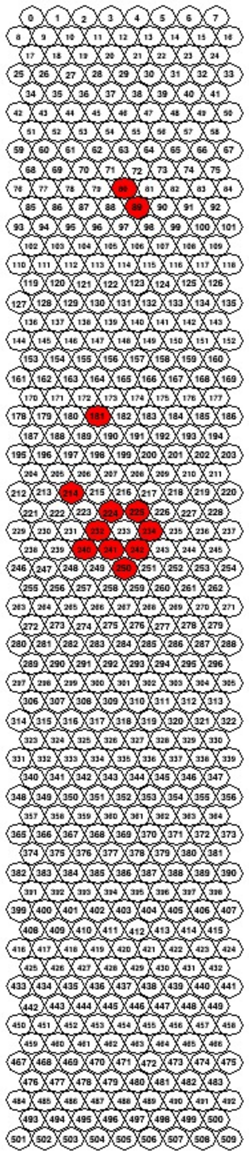


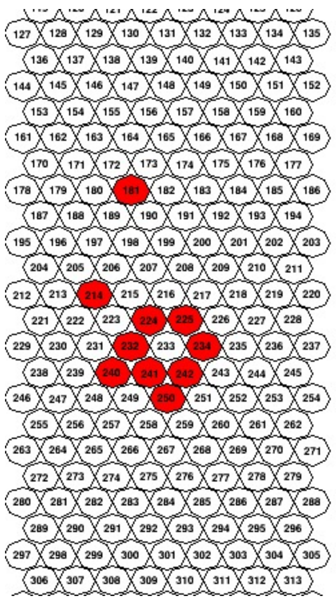
- Pink ellipse is the cut applied to select elastic (e,e'p) coincidences in previous plots

GRINCH single-event displays and cluster heatmaps









- Credit: Maria Satnik (W&M)

Near-future plans

- Implement ideas for improving GEM analysis:
    - Use full-readout events (1/100) to measure rolling average of common-mode using more robust sorting method offline, detect events with large negative bias in the online calculated common-mode due to negative pulses and correct the ADC values for these events—expect SOME improvement in efficiency and resolution but potentially large improvement in tracking speed by removing “signals” that are really just baseline/noise from clustering/tracking analysis.
    - Implement BBCAL trigger time correction to enable tighter strip timing cuts
    - Implement time sample deconvolution to reliably reconstruct good signals riding on the tail of earlier background hits
    - Implement cross-talk rejection/filtering
    - Implement software ”gain match” to improve ADC U/V (or X/Y) correlation, tighter asymmetry cuts
    - Use timing hodoscope to achieve tighter track search region constraint
    - Get everything calibrated (Calorimeters, hodoscope, optics) across all kinematics, and optimize all the many, many dials at our disposal to improve tracking performance
- Improve the speed of the tracking algorithm by rewriting parts of the code that use (inefficient) std::map containers (a choice that was made for programming convenience without thinking carefully enough about the speed overhead)
- More fully integrate the GRINCH data into the analysis of 2022 GMN data (unfortunately the GRINCH data taken in 2021 is mostly not usable due to use of CO2 instead of heavy gas, DAQ and operating HV issues)
    - Aid electron ID and better understand the GRINCH for future experiments like SIDIS for which it is essential.
- Start first pass of mass production on the batch farm and get preliminary physics results