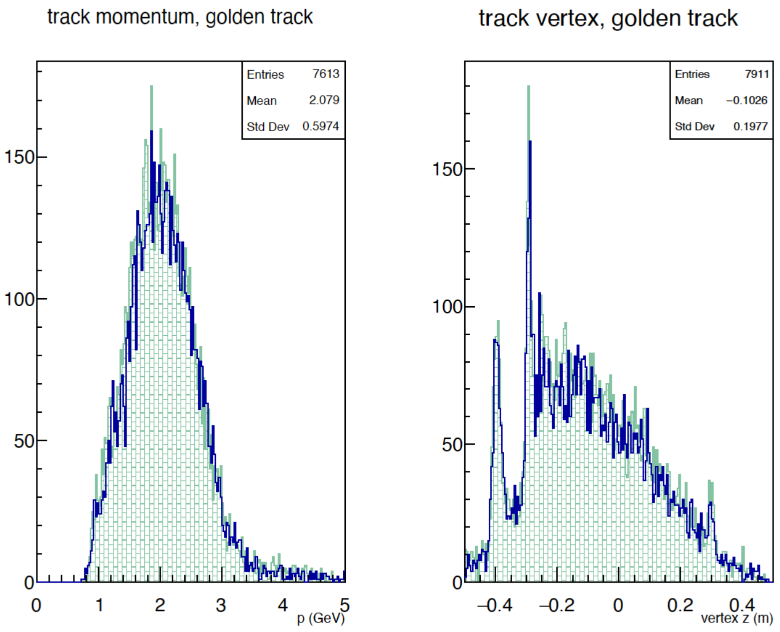
SBS Software and Analysis System

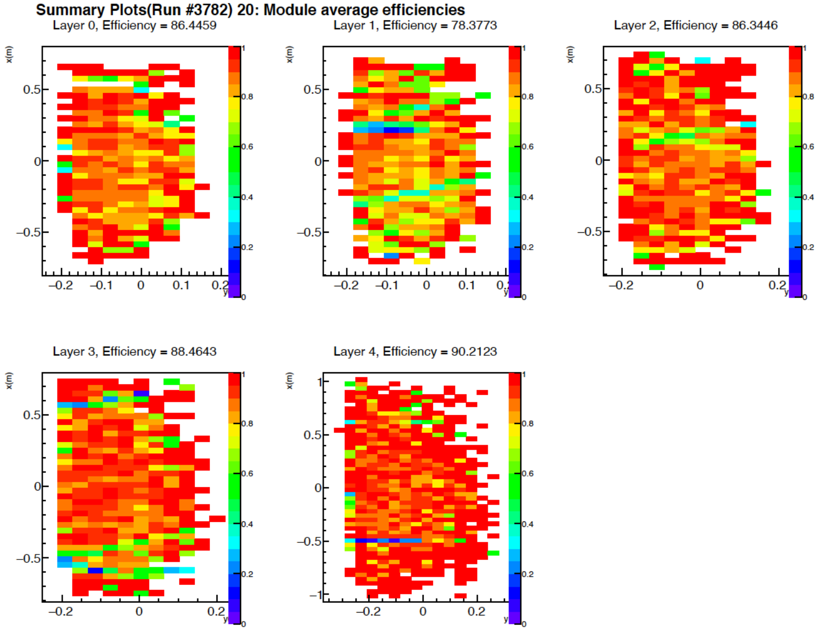
Andrew Puckett
University of Connecticut
SBS Collaboration Meeting
July 18, 2023

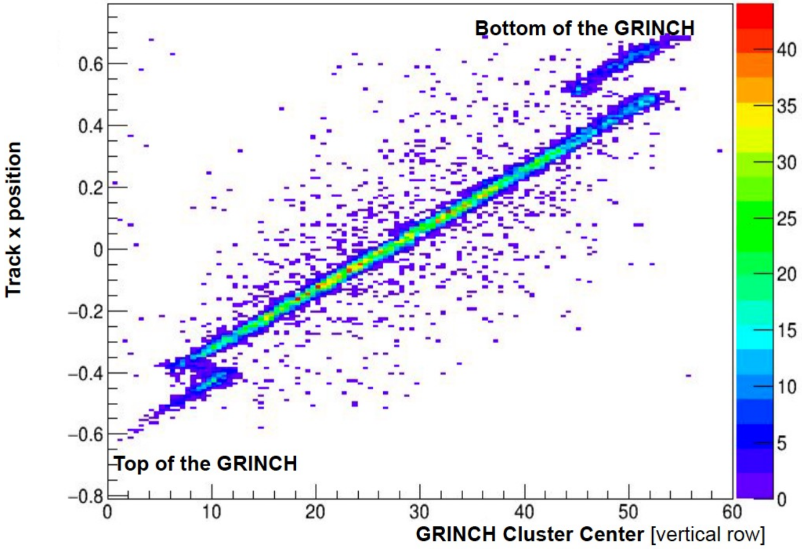
UCONN

Summary Plots(Run #3788) 5: Golden track momentum, vertex



Summary Plots(Run #3782) 20: Module average efficiencies





Negative x is “up” in this transport coordinate system.

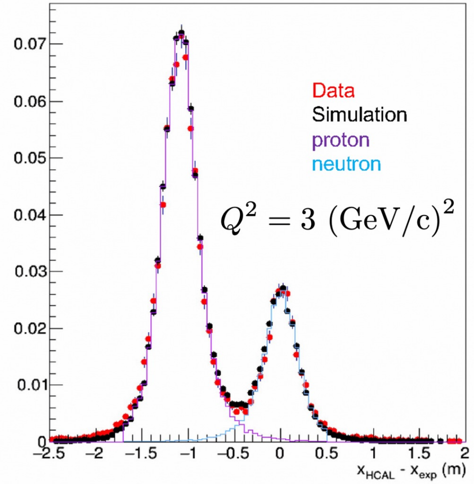


Figure 2: Difference between the vertical coordinate of the particle detected by the SBS hadron calorimeter and its predicted position from the measured electron kinematics in BigBite, assuming (quasi-) elastic scattering.

Meme-smithing by Jack Jackson

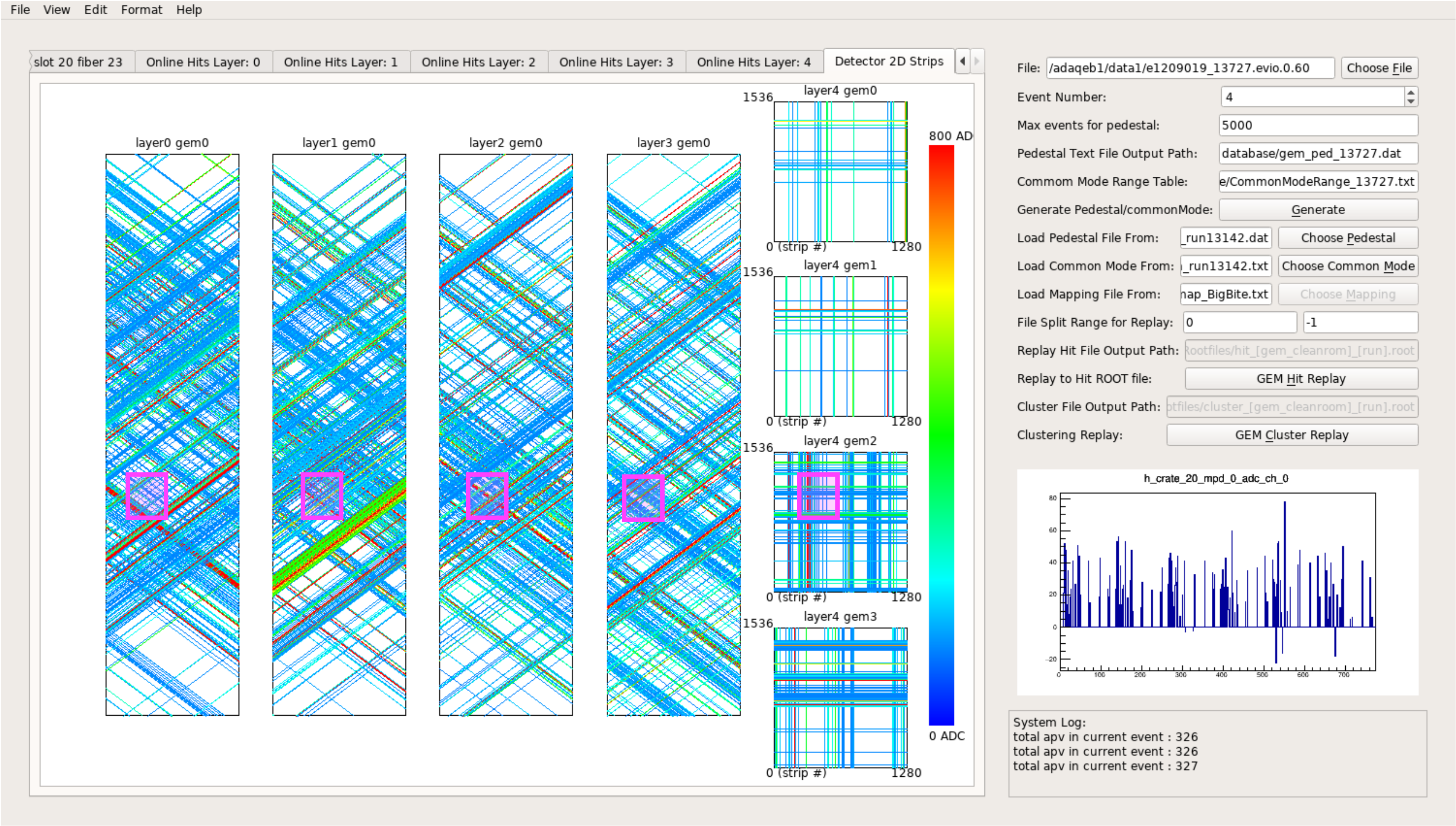
Existing SBS Software

- 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, S. Jeffas, *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.

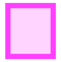
SBS software working group

- Mailing list: https://mailman.jlab.org/mailman/listinfo/Sbs_software
- Standing weekly meeting; currently Fridays at 1:00 PM
- SBS Software and Analysis Coordinator: Andrew Puckett
- SBS software/analysis wiki page: https://sbs.jlab.org/wiki/index.php/SBS_Software/GMN_analysis_meeting_agendas_and_minutes
- Dedicated GMN/nTPE analysis meetings; currently Fridays at 10:00 AM
- GEN analysis effort ramping up now
- Upcoming GEN-RP and GEP experiments

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



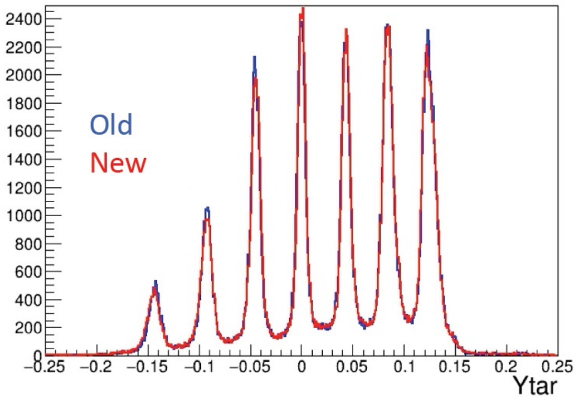
- Single event display for BigBite GEMs; all fired strips color-coded by ADC values

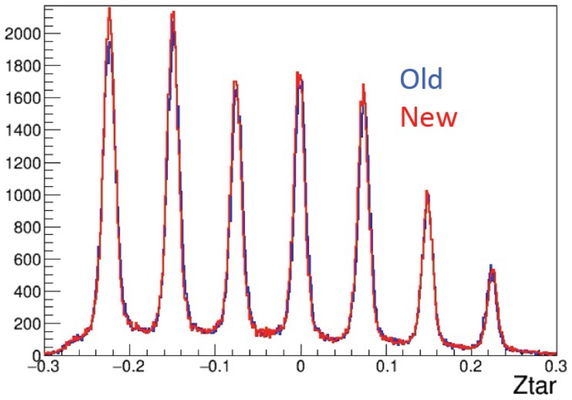
 = approximate size of calorimeter-constrained track search region at each layer

GEN Optics calibration (by Holly)

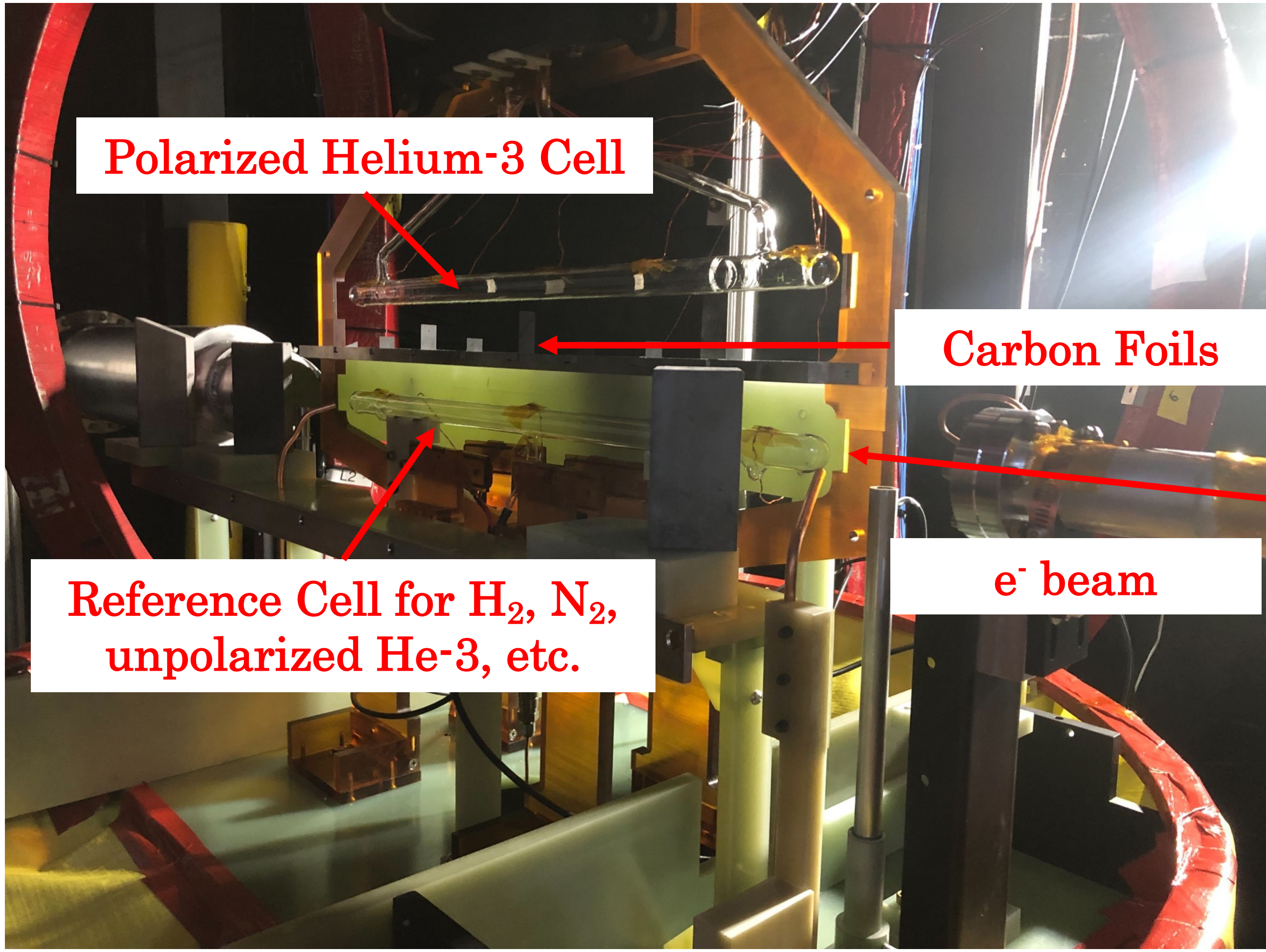
https://sbs.jlab.org/cgi-bin/DocDB/private/ShowDocument?docid=344

Comparing the old and new reconstruction.





- Lessons learned from GMN experience allow "pretty good" starting optics model for BigBite from simulation



GRINCH software

Plots and analysis credit: Maria Satnik

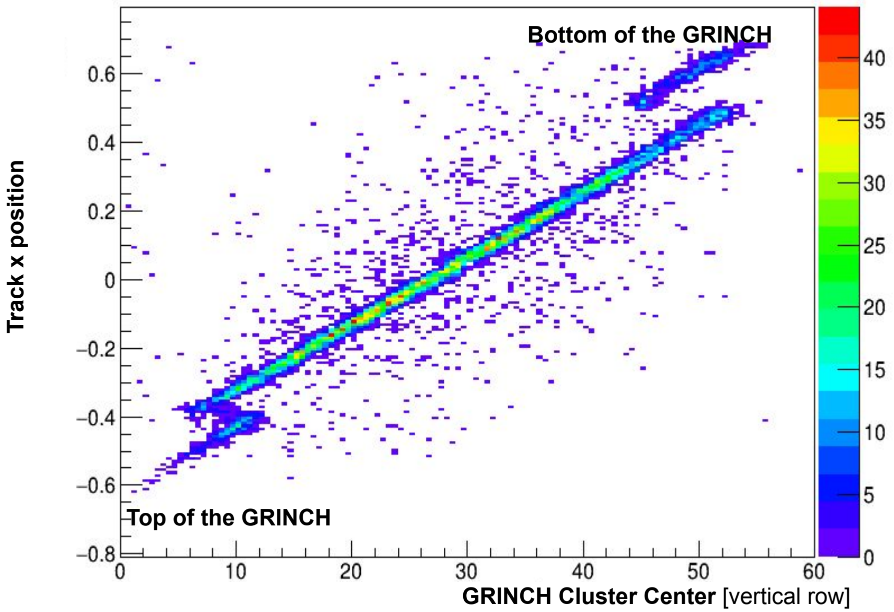
Maira is working on implementing multiple clusters into SBS offline.

90% - 97% electron detection efficiency during GEn and 80% - 90% during GMn.

Clear correlation between track position and GRINCH cluster position.

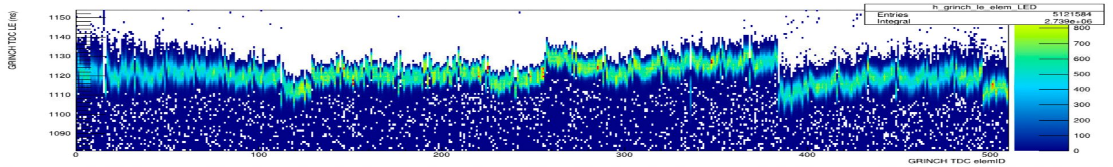
Timing correction script created and implemented.

Satnik, 9 Jan 2023

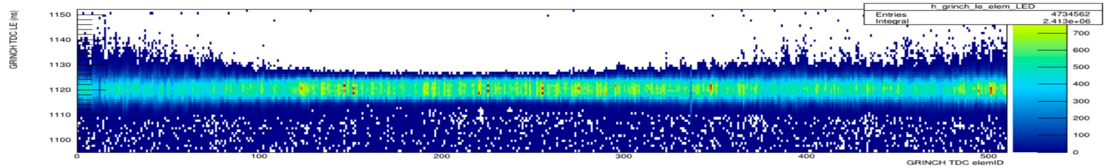


Negative x is “up” in this transport coordinate system.

Before



After

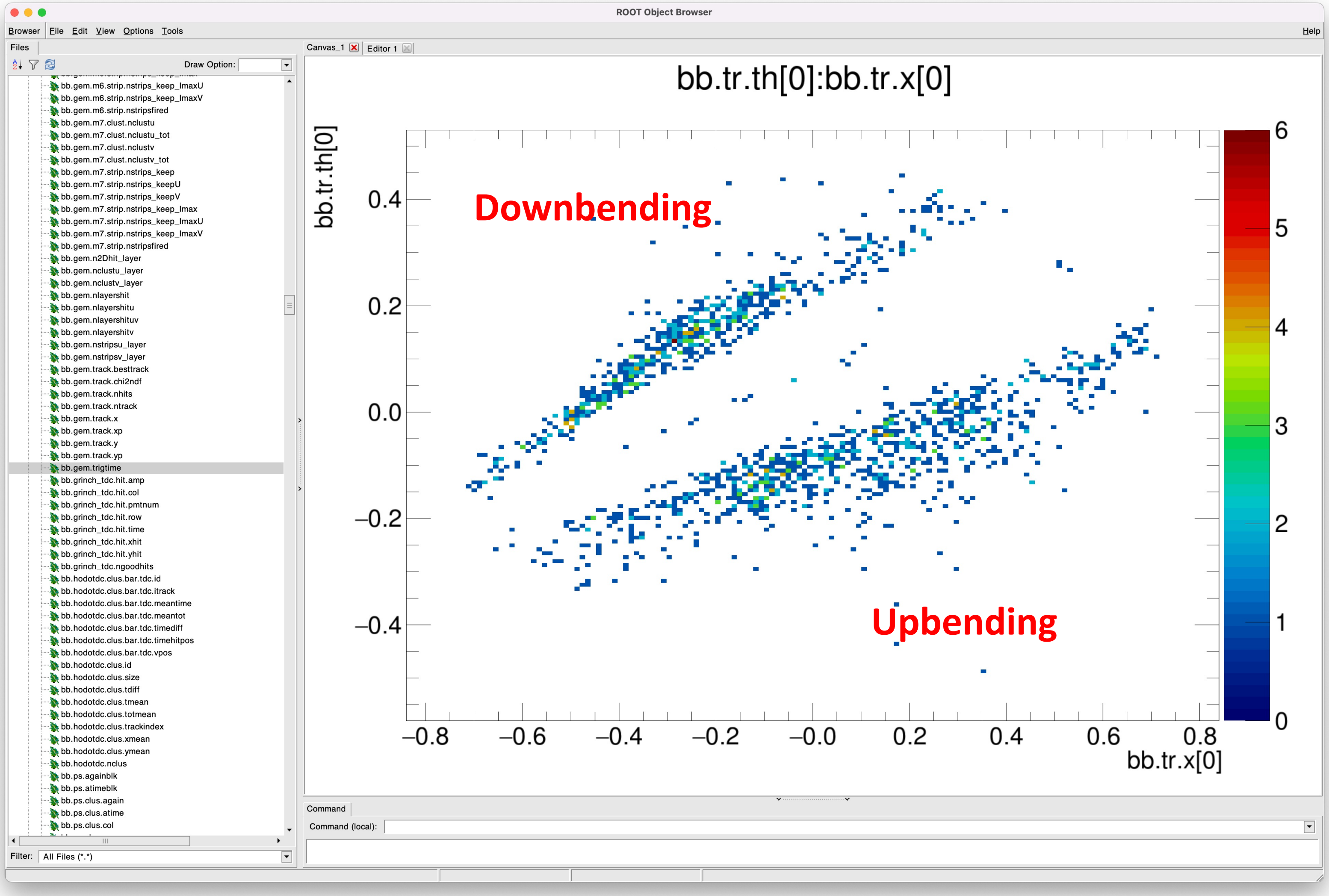


Timing correction using GRINCH LED

Analyzing downbending tracks in BigBite

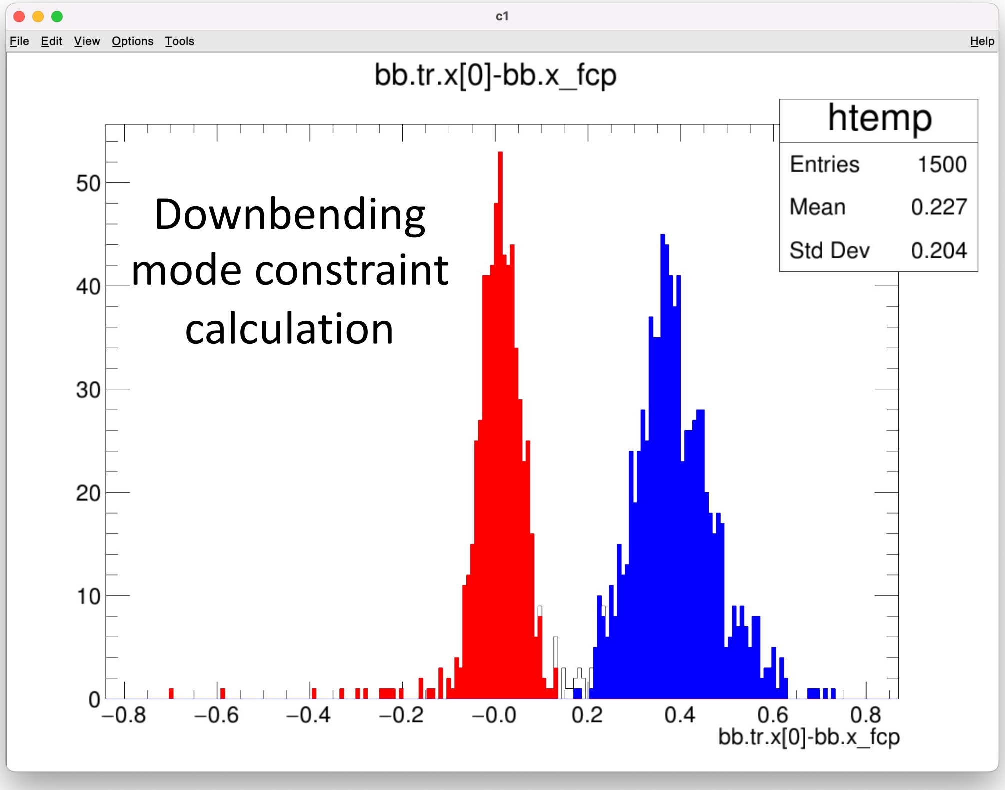
- Motivation: obtain tagged neutron sample in $\gamma p \rightarrow n\pi^+$ with pion in BigBite and neutron in HCAL using endpoint method
- Anuruddha Rathnayake (UVA) is working on downbending tracking analysis for GMN SBS-9 kinematics.
- SBS-4 also looks promising.
- We might also look to take dedicated opposite-polarity data with BigBite during GEN-RP with LH2 target and radiator (but polarimeter elements (analyzers) in the path of the neutrons would likely need to be removed)

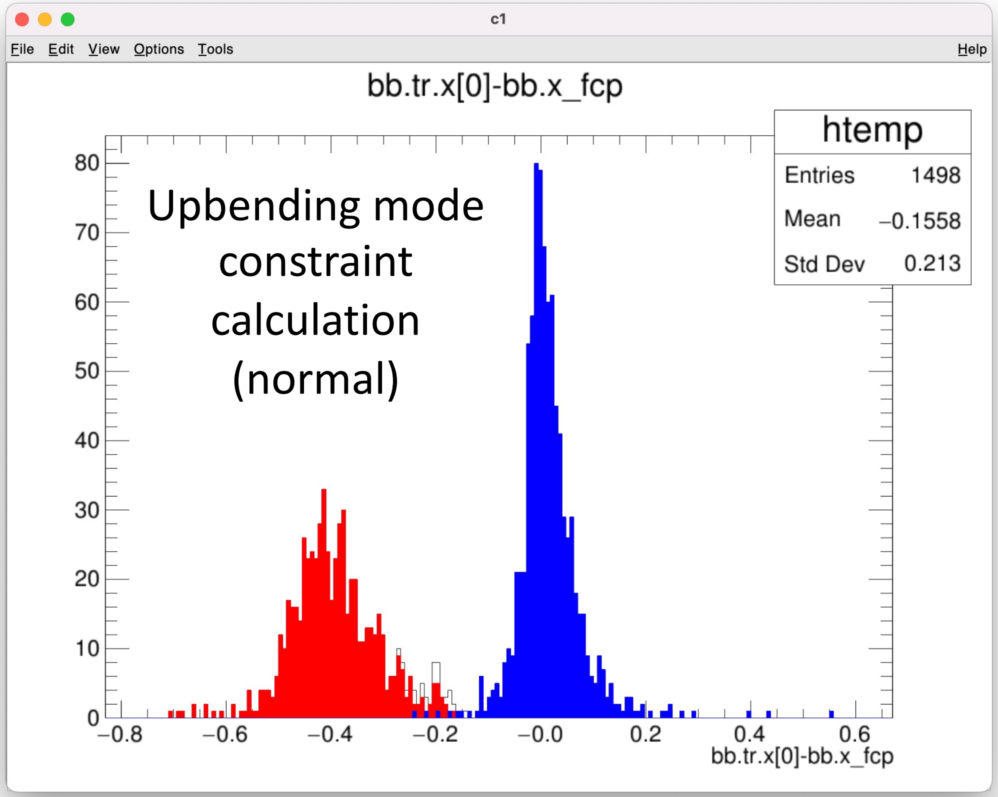
Selecting downbending tracks



- Replay data with no/loose tracking constraints at the FRONT of the GEMs
- Plot correlation between track vertical angle “bb.tr.th” and track vertical position “bb.tr.x”
- Upbending and downbending tracks separate into two distinct “stripes”

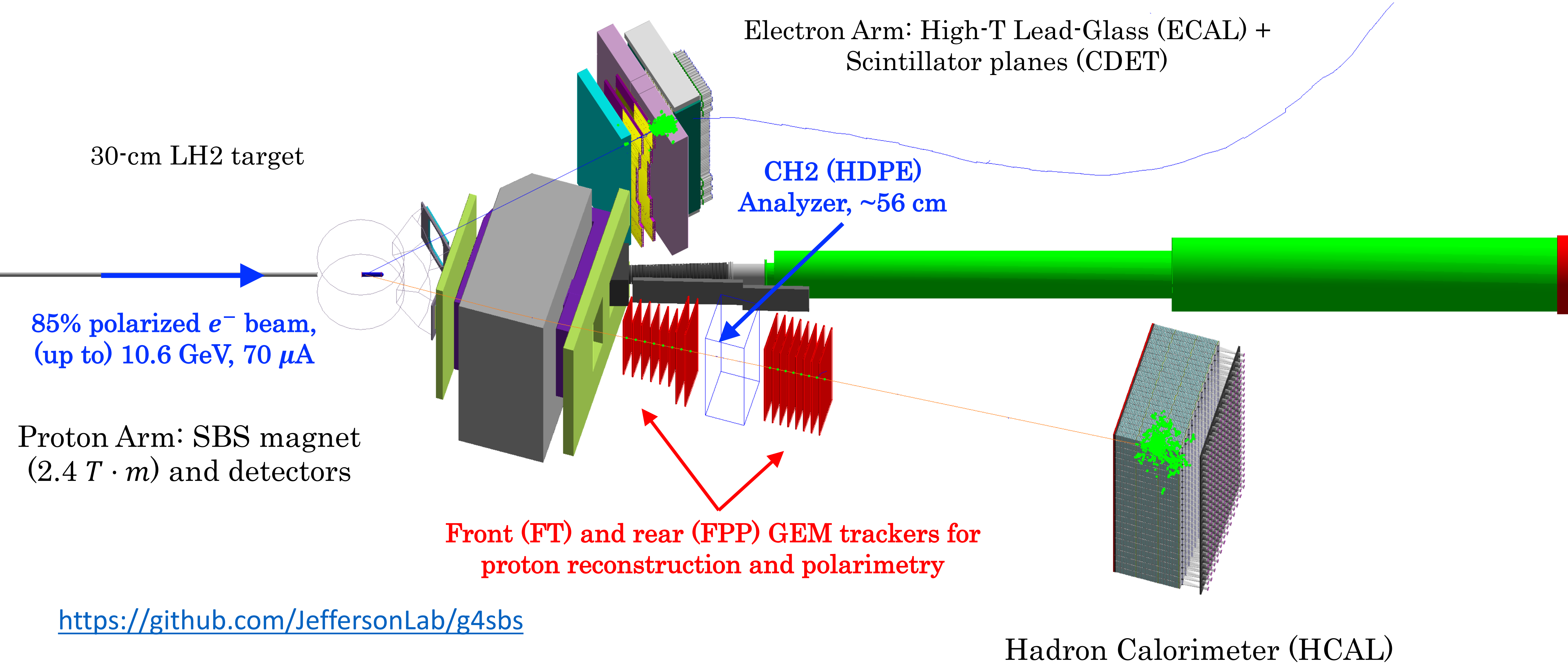
Checking constraint calculation for downbending tracks





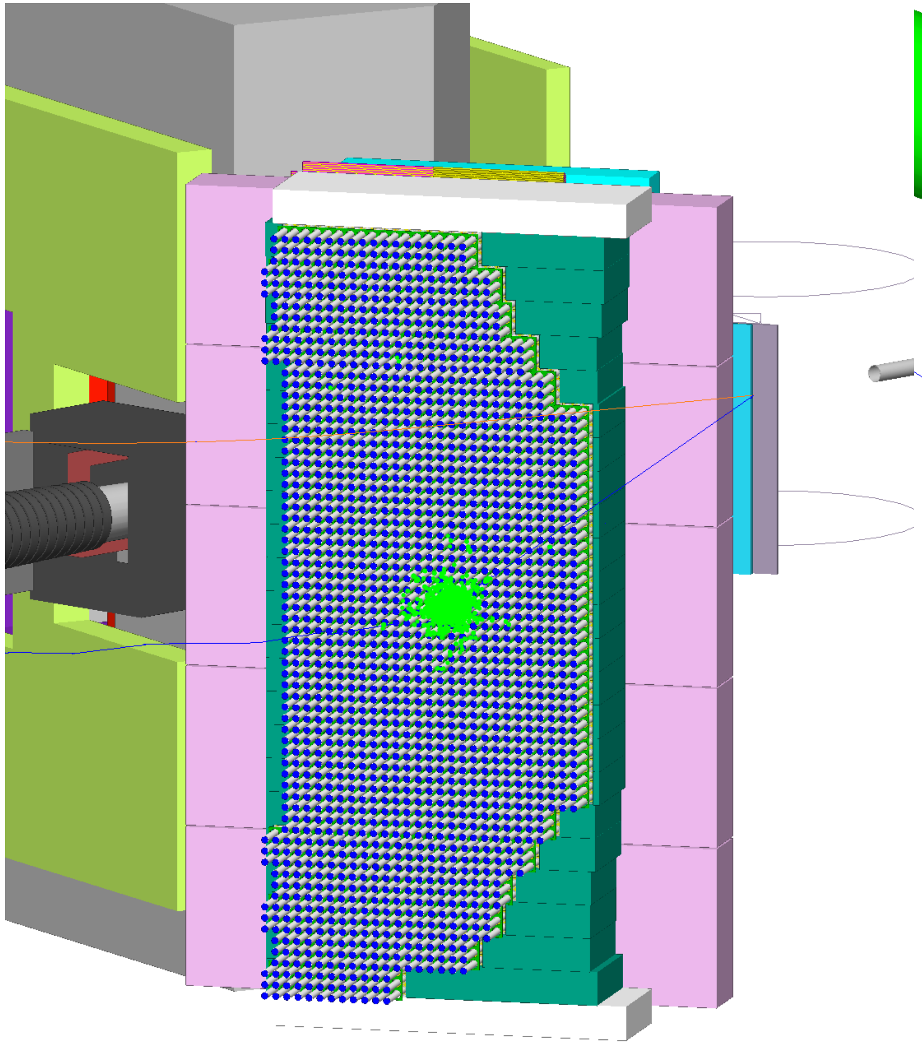
- **Red** (**blue**) **is downbending** (**upbending**) tracks when replaying data in downbending (left) and upbending (right) mode

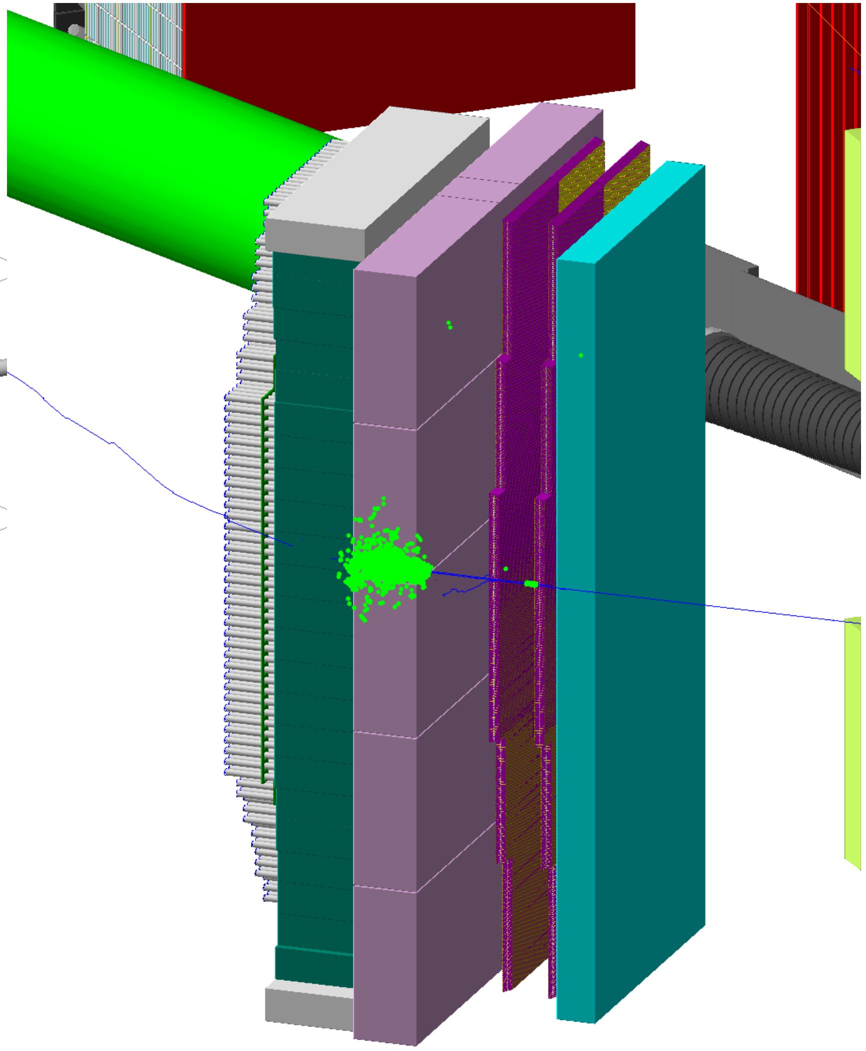
GEP geometry updated in *g4sbs*



https://github.com/JeffersonLab/g4sbs

Updated ECAL/CDET geometry (by Kip Hunt (UConn))





- Kip extracted ECAL geometry details from JT file provided by Hall A engineers
- Updated supermodule layout to (near) final
- Added inactive lead-glass, glass-foam insulation, and CH2 absorber
- CDET geometry has correct z position, but still a placeholder (waiting for CNU to release their CDET model into “official” g4sbs)

SBS Software Early Successes—A Brief History

JEFFERSON LAB ELECTRONIC LOGBOOK

Add content | Logbooks | Tags | Useful Links | Preferences | Help/About

View | Link Downtime | Post Follow-up Entry

<<Prev | Index | Next>>

Follow-up Re: Run 11107 optics with 50k

Lognumber 3924706. Submitted by hszumila on Thu, 10/14/2021 - 18:04.

▾ There is 1 comment...

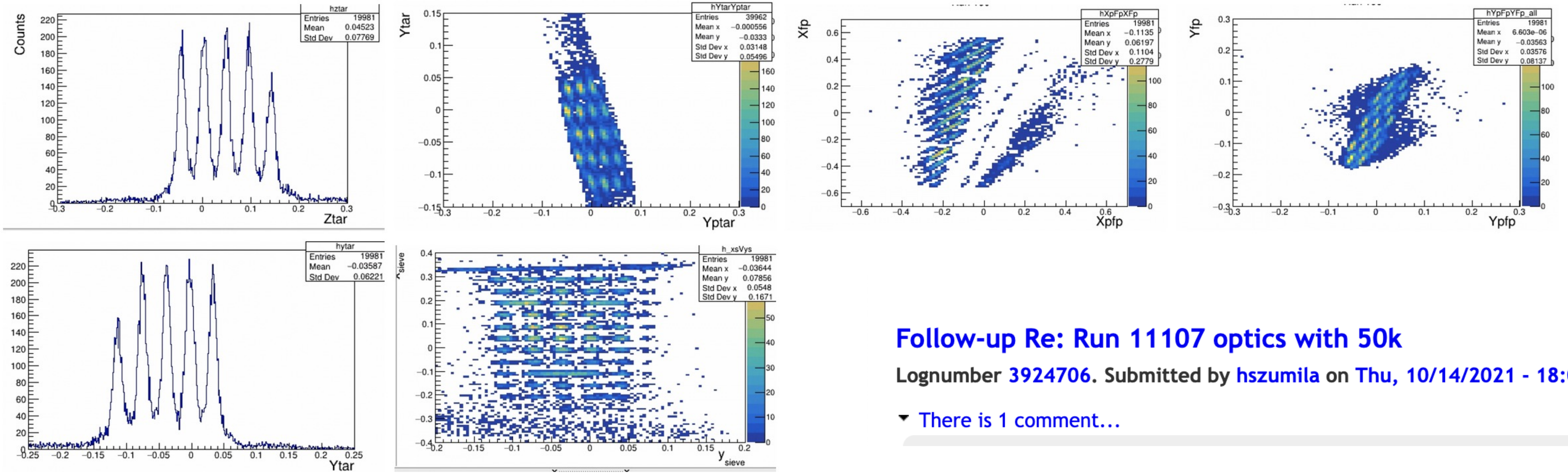
I THINK these results are

by puckett on Thu, 10/14/2021 - 18:57

I THINK these results are consistent with what we see in the survey results, as the survey thinks that y = 0 of the GEM coordinate system is approximately 3 CM to the left of y = 0 of the optics coordinate system. That means that if we don't correct for this offset between the GEM internal coordinates and the optics coordinates, i.e., if we assume that yGEM = yOPTICS, then the center sieve hole should show up about 3 CM to the ***right*** of the center of the GEM active area, or at about -3 cm, as we see in the data. Similarly, we see a roughly +4-5 cm offset in z vertex, that is also reflected in y target.
The survey also predicts a roughly 6 mrad offset in y' at the focal plane, in the sense that y' = 0 in the GEM coordinate system corresponds to roughly y' = -0.006 in the magnet coordinate system. Since the detector origin is roughly 3 meters from the target, we would expect naively that rays coming from the central optics foil through the center sieve hole (y' ~= 0 in optics coordinates) would have roughly y' ~= +0.006 in the GEM coordinate system (before alignment corrections).

Logbooks: HALOG
Entry Makers: puckett, pking
References: 3924344 - Run 11107 optics with 50k

I attach here the plots I showed at the RC meeting of the optics from run 11107. This seems to be consistent with the offsets we observed in the survey data (not included in the optics model here). This is not the full run but only a larger portion than the 50k. We probably have at least double this once everything completes the replay.



▸ Comment Form

Follow-up Re: Run 11107 optics with 50k

Lognumber 3924706. Submitted by hszumila on Thu, 10/14/2021 - 18:04.

▾ There is 1 comment...

https://logbooks.jlab.org/entry/3924706

- BigBite tracking and kinematic reconstruction ready and working on day 1 of GMN

JEFFERSON LAB ELECTRONIC LOGBOOK

Add content | Logbooks | Tags | Useful Links | Preferences | Help/About

View | Link Downtime | Post Follow-up Entry

BB GEM HV Scan Results w/ beam

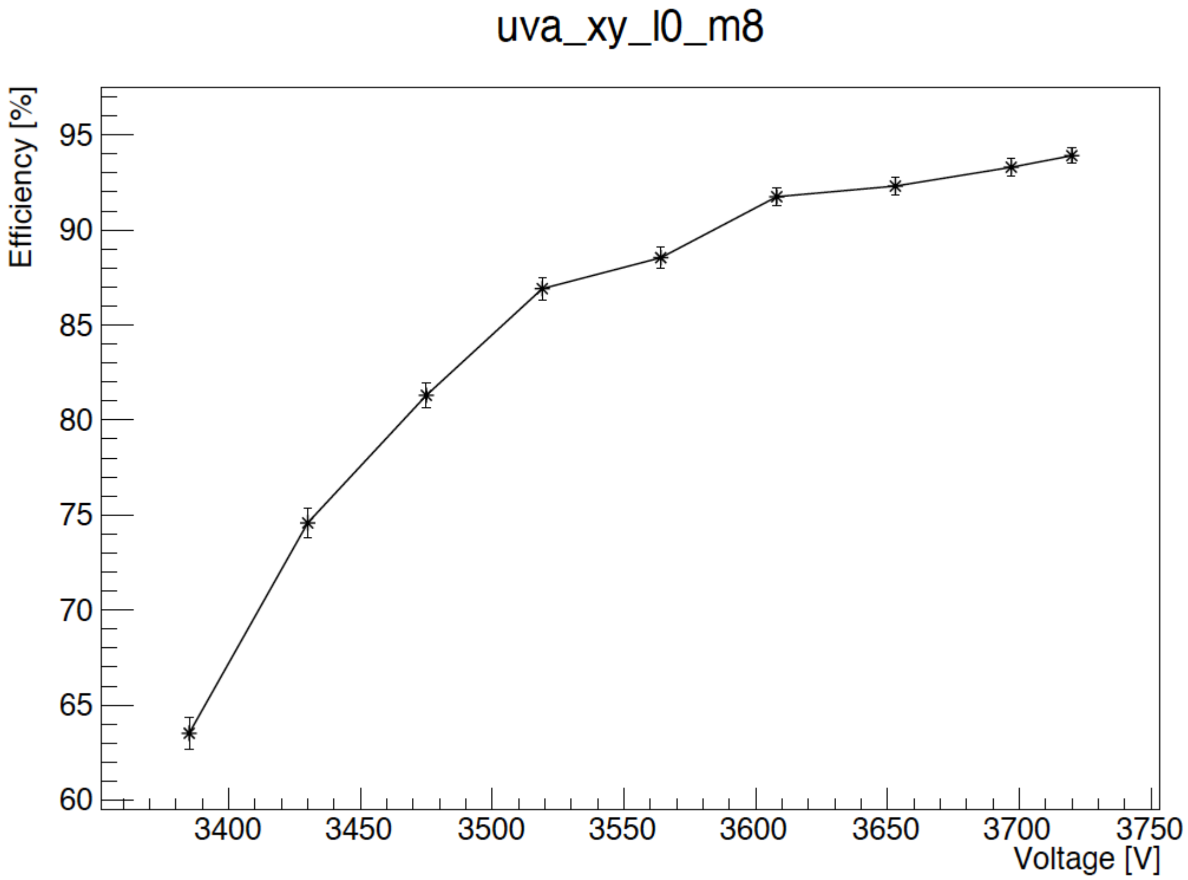
Lognumber 3924320. Submitted by adr on Thu, 10/14/2021 - 05:58.

Logbooks: HALOG SUPERBIGBITE
Tags: GEM
Entry Makers: adr, puckett, hszumila, ewertz
Attached Files: BB GEM HV Scans with beam.pdf

Attached are the HV scan results for BB GEM layers we obtained with beam today morning

▸ Comment Form

https://logbooks.jlab.org/entry/3924320
https://logbooks.jlab.org/entry/3925085



- GEM track-based efficiency plateaus taken on first shift with beam

LH2 elastic electrons in BigBite

Lognumber 3925220. Submitted by adaq on Fri, 10/15/2021 - 07:13.

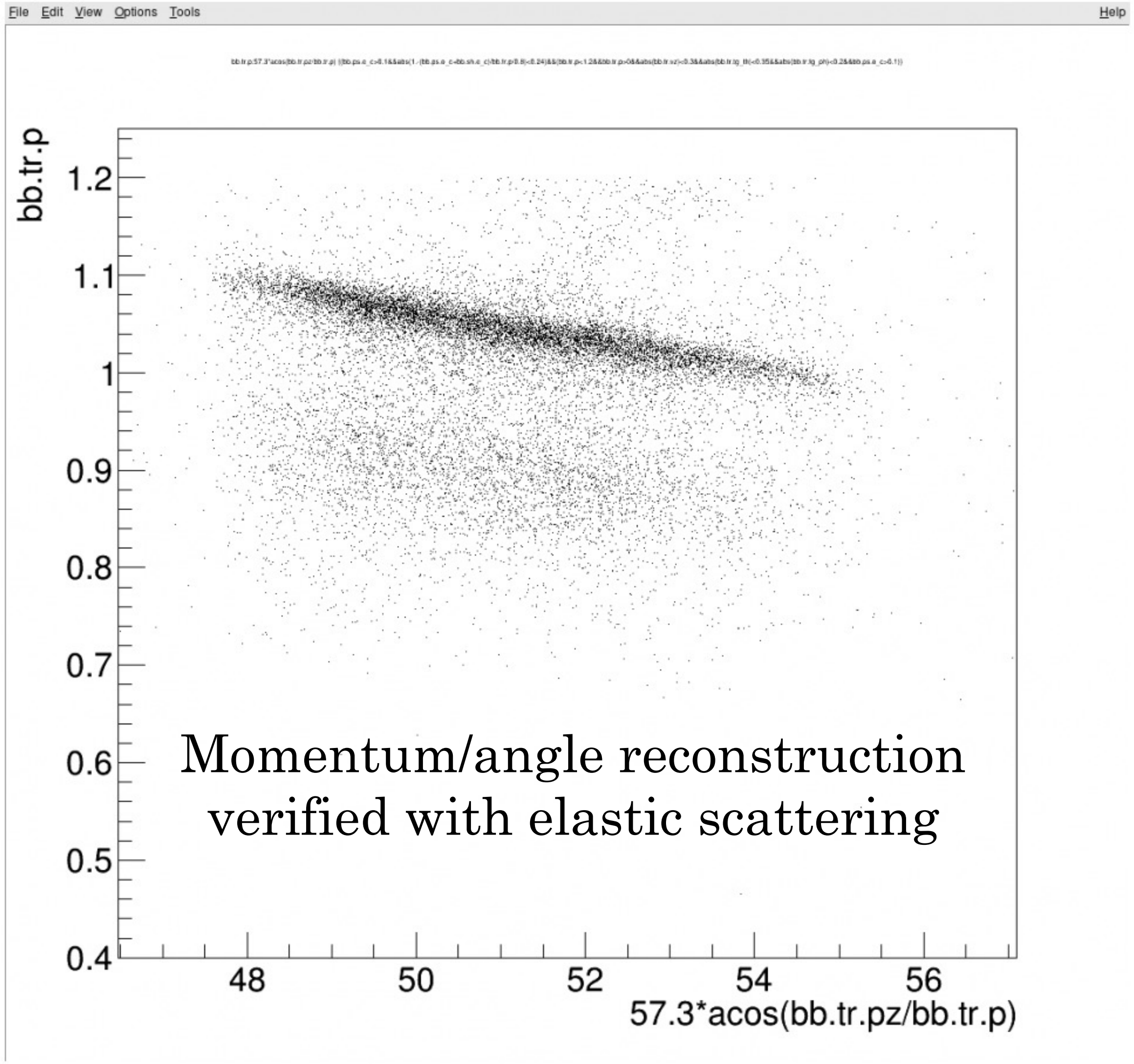
https://logbooks.jlab.org/entry/3925220

Logbooks:	HALOG
Entry Makers:	puckett
Backlinks:	Follow-up Re: LH2 elastic electrons in BigBite

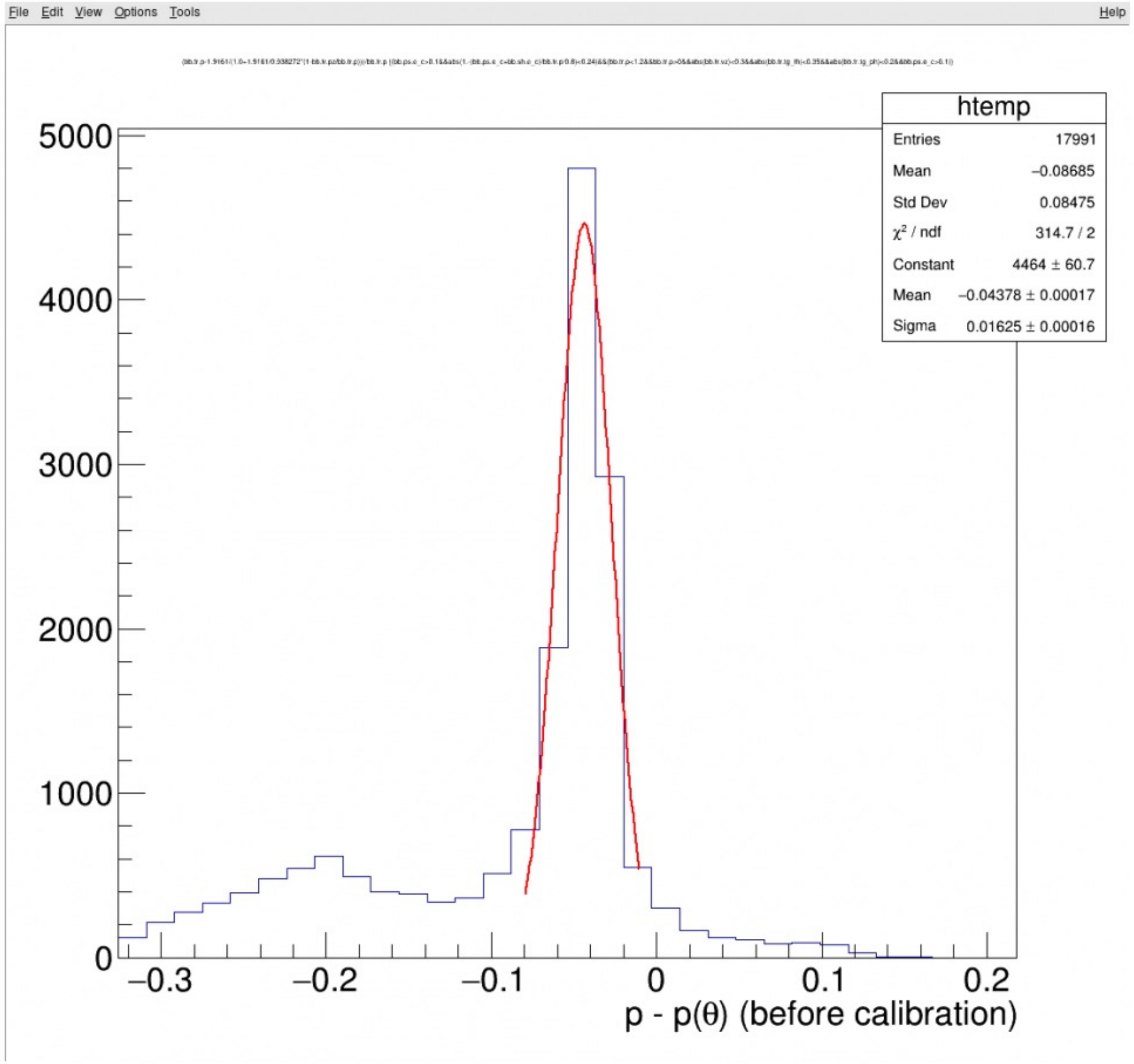
With uncalibrated optics from the starting model, we can clearly see elastically scattered electrons in BigBite, in the angle-momentum correlation and in the p-pelastic(theta) distribution.

We have kept the current on the LH2 target conservatively at 5 uA for the time being, as we work to roughly calibrate things well enough to ready the tracking software for the higher occupancy we expect after removing the sieve slit.

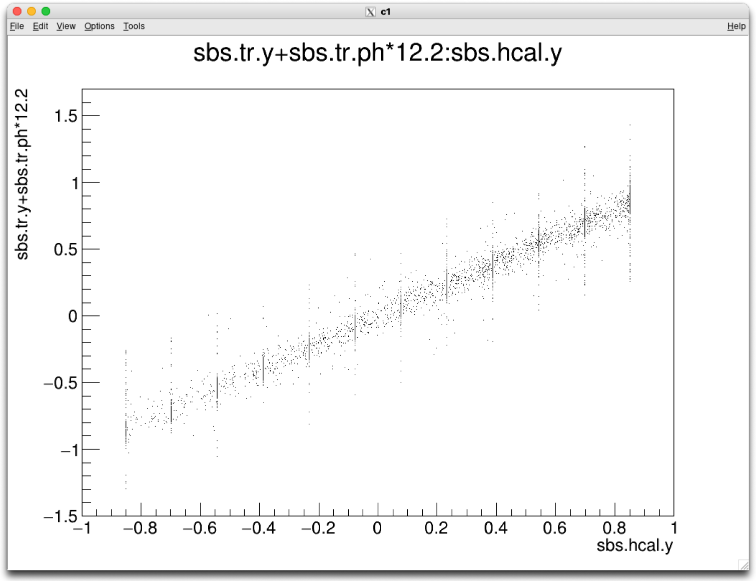
snapshot_1

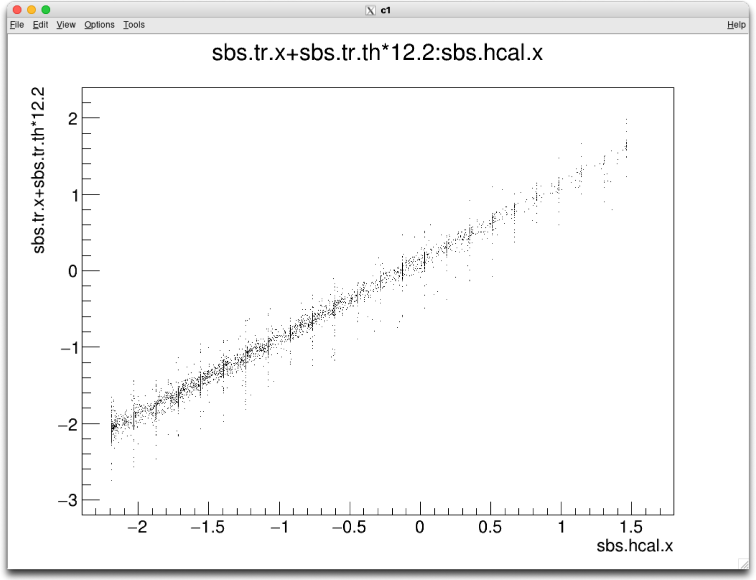


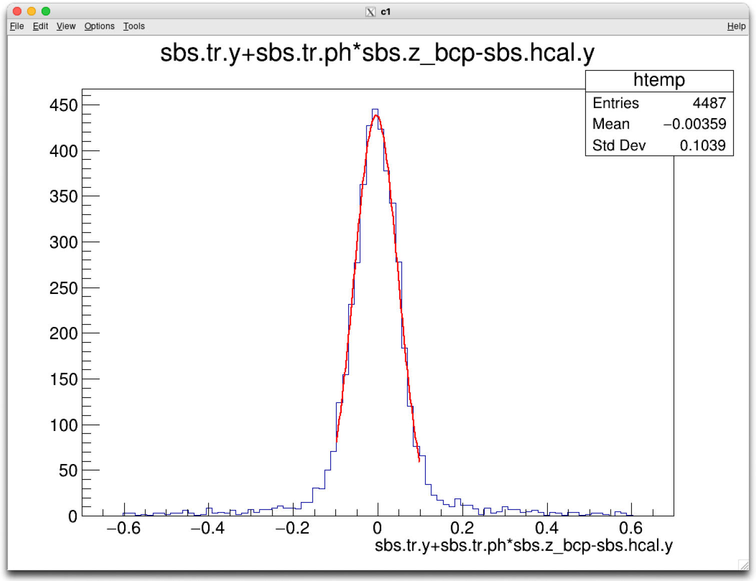
snapshot_2

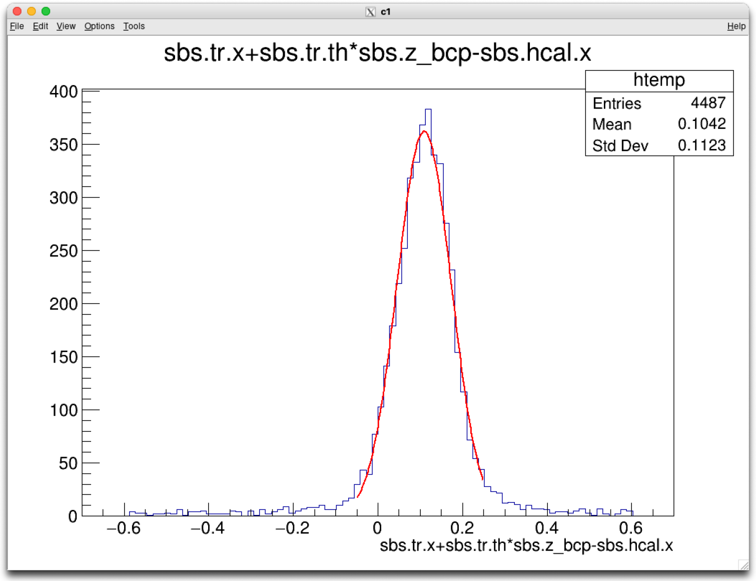


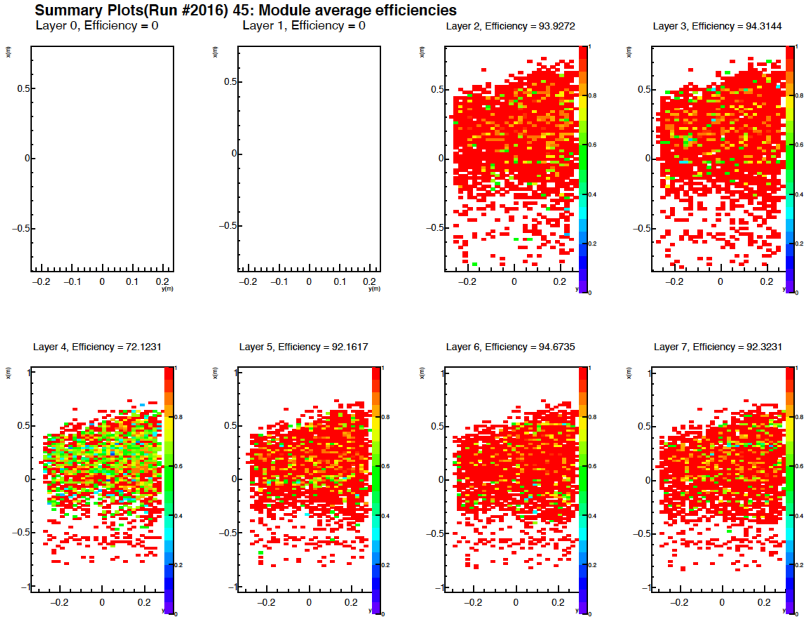
SBS GEM commissioning in GEN: https://logbooks.jlab.org/entry/4061825

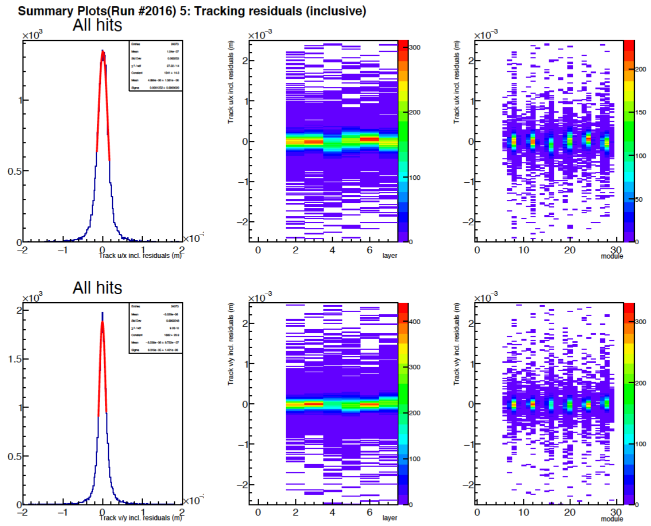






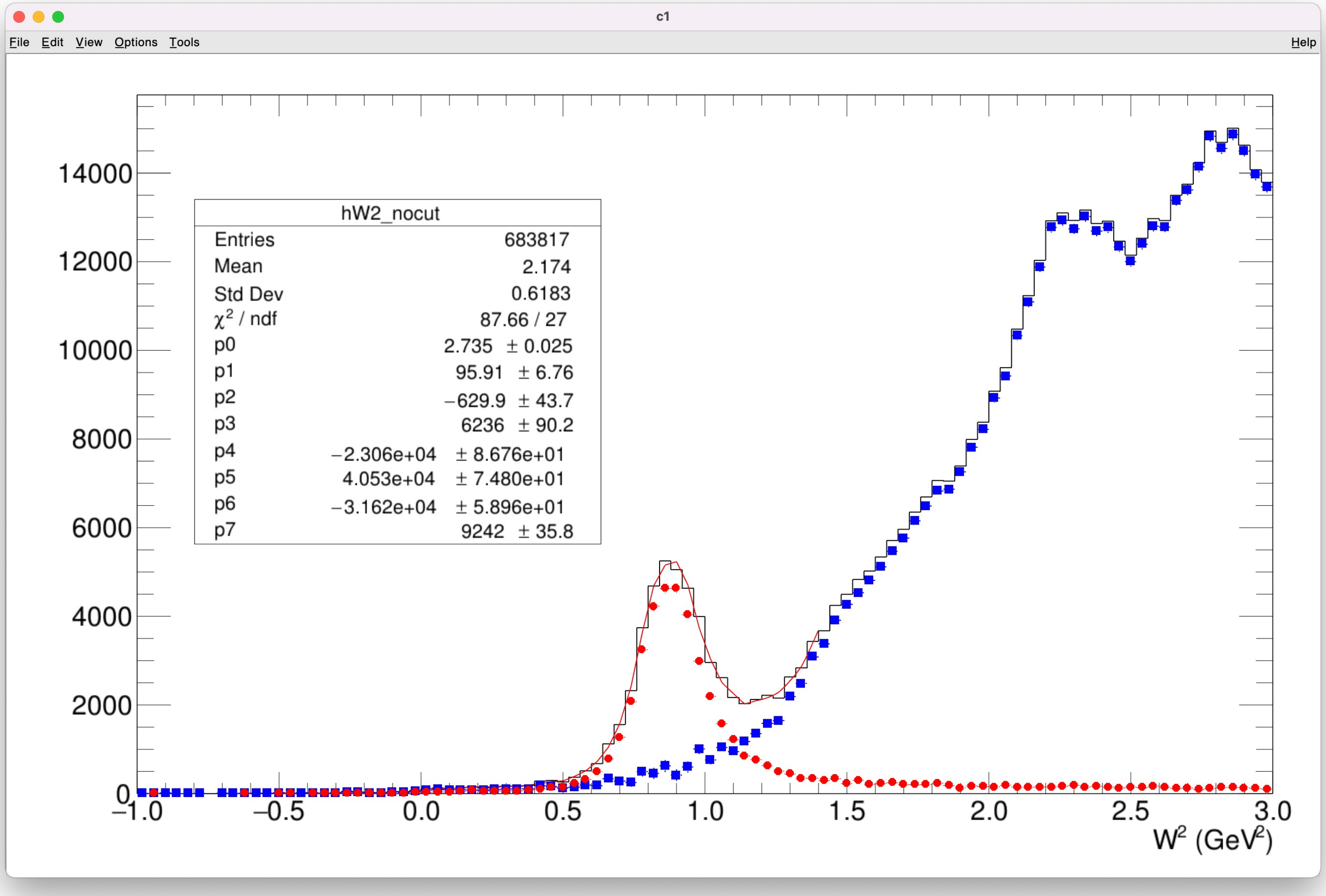






GEP ERR—SBS GEM Tracking Efficiency Studies

Elastic Yield Estimation from GMN data



"Yield" is charge-normalized, live-time corrected elastic peak integral → most reliable proxy for overall tracking efficiency

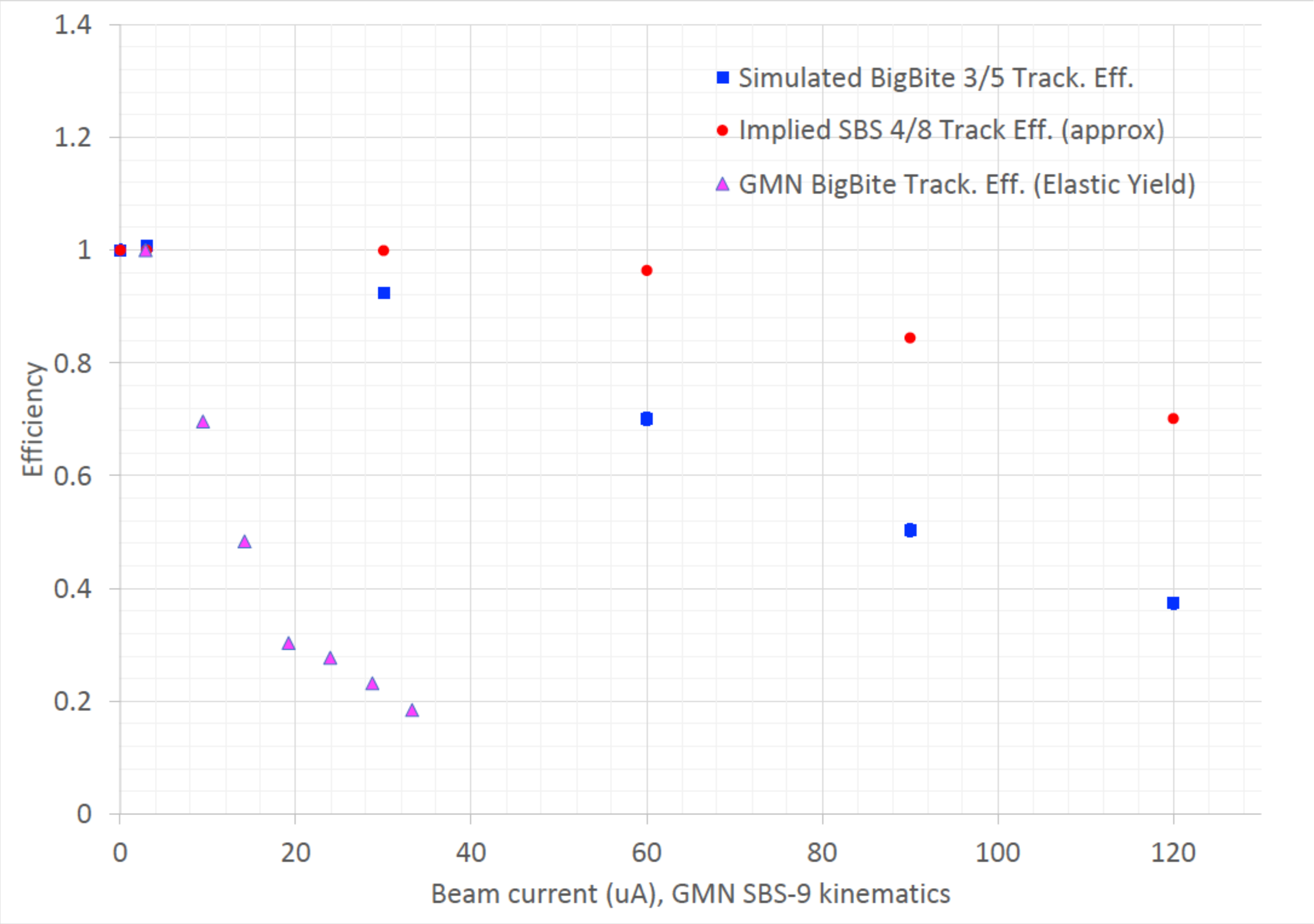
Elastic yield estimation from data

- **Black = all events**
- **Red = scaled elastic from coincidence with HCAL**
- **Blue = estimated background**

GMN SBS9 Kinematics:

$$E = 4\text{ GeV}$$
$$\theta = 49\text{ deg}$$
$$Q^2 = 4.5\text{ GeV}^2$$
$$\text{Target} = 15\text{-cm } LH_2$$

GEM Reconstruction Efficiency at High Rate



- *g4sbs* reproduces observed BigBite GEM rate/occupancy at "low" beam current (3 uA)
- "High current" study done at end of GMN went up to 34.5 uA on LH2, LD2 in previous slide's kinematics.
- Elastic yield drops rapidly with current (effect of GEM gain/efficiency drop → Holly's talk)
- **Simulated BigBite tracking efficiency** shows much slower drop-off (without any fine-tuning or optimization)
- **GEP-equivalent** beam current for this configuration is ~120 (50) μA for Front Tracker (Back Tracker)
- Implied 4/8 efficiency for SBS FT in GEP is ~70%, consistent with assumption in PAC47 uncertainty projections (rough, preliminary)

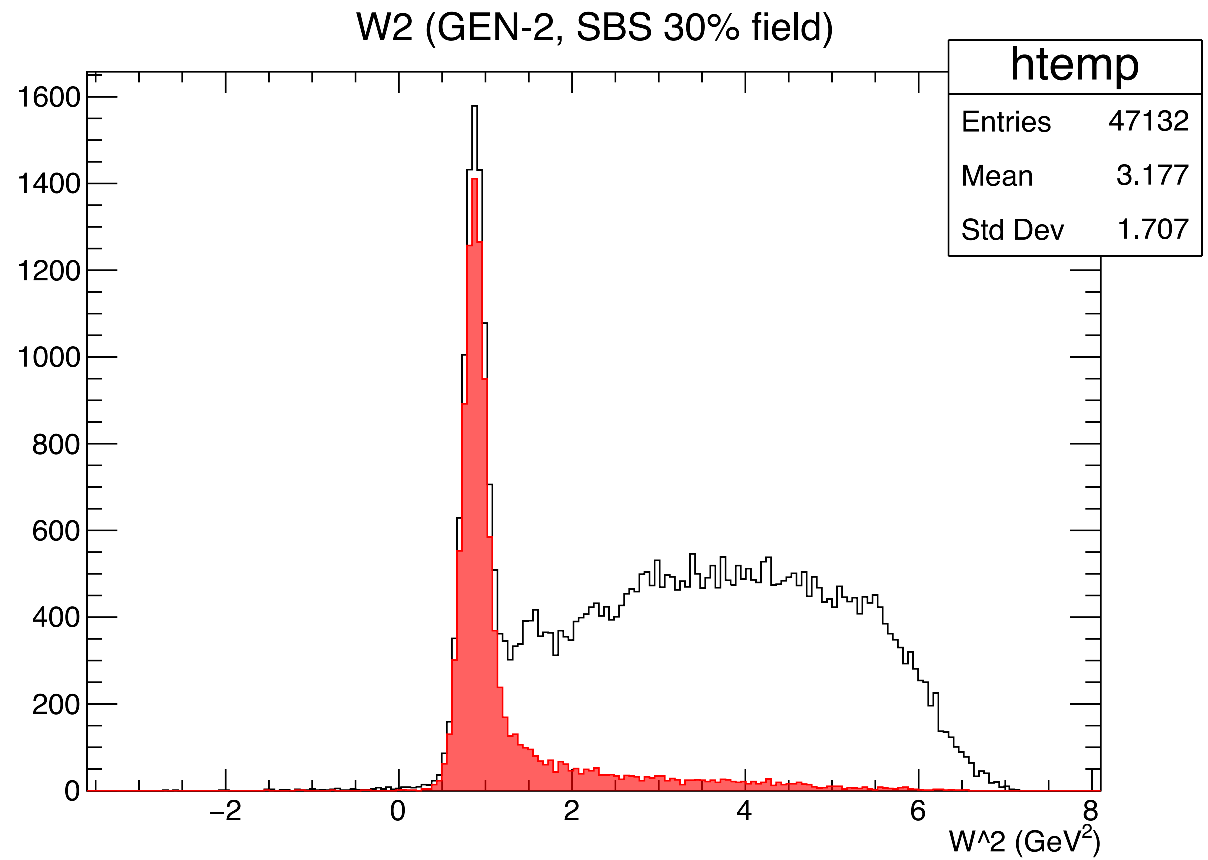
PRELIMINARY conclusions on GEM gain/efficiency during GMN

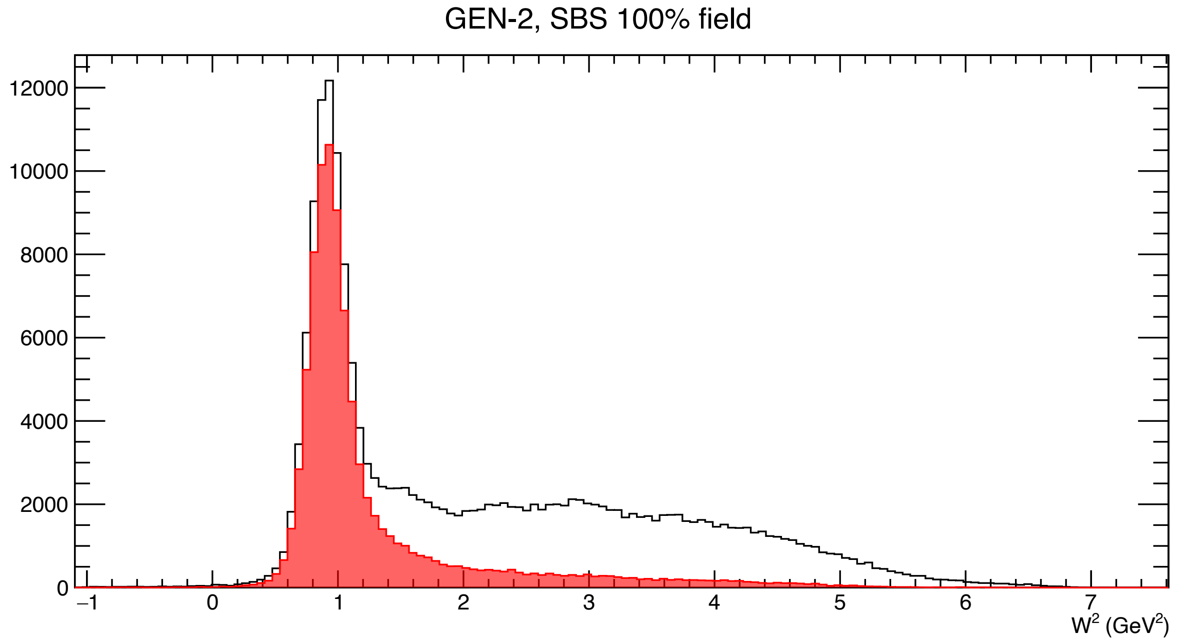
- The observed deviations from linearity of the excess divider current and the hit rate/occupancy are qualitatively consistent with the observed reductions in tracking efficiency*
- BigBite tracking was/is particularly vulnerable to this issue due to using only 5 GEM layers in tracker → gain/efficiency drop required us to run GMN significantly below proposal luminosity (but we still got the physics!)
- Simulated BigBite tracking at various background levels (under stable gain assumption) shows that the efficiency reduction seen during GMN comes mostly (but not entirely) from the hardware.
- "Parallel divider" concept will mitigate this issue for GEP front tracker
- GEP polarimeter redesign with more redundant 8-layer tracking assemblies in front and back trackers will make the overall tracking far more robust against individual layer inefficiencies; small dead areas, intermittent localized hardware/electronics issues, etc.
- Hardware fix for gain drop will be verified during upcoming GEN-RP run
- **as measured by elastic yield, requiring 3/5 layers to form a track*

NEW--SBS optics calibrations from GEN data w/SBS GEMs

- At GEN Kin-2 we have H2 reference cell data at 30% and 100% SBS field with SBS GEMs in the data stream
- Q^2 is relatively low, elastic event selection with W^2 and HCAL is extremely clean
- Can calibrate angle and vertex reconstruction extremely well
- Momentum calibration still a work in progress due to "overfitting" issues (too few independent constraints)—formalism needs some development

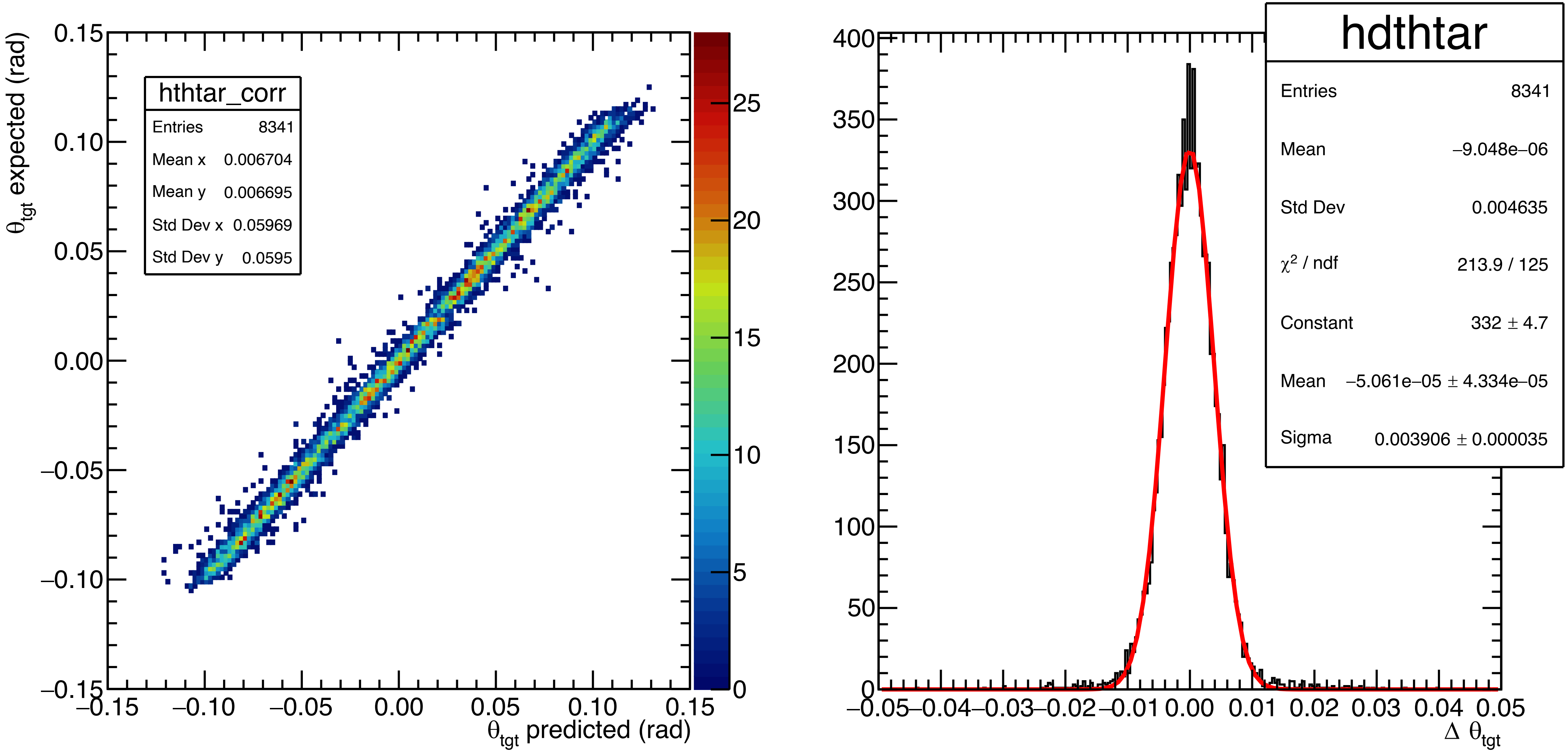
GEN-2 Invariant Mass (H2 Ref. Cell, SBS 30% and 100% field)



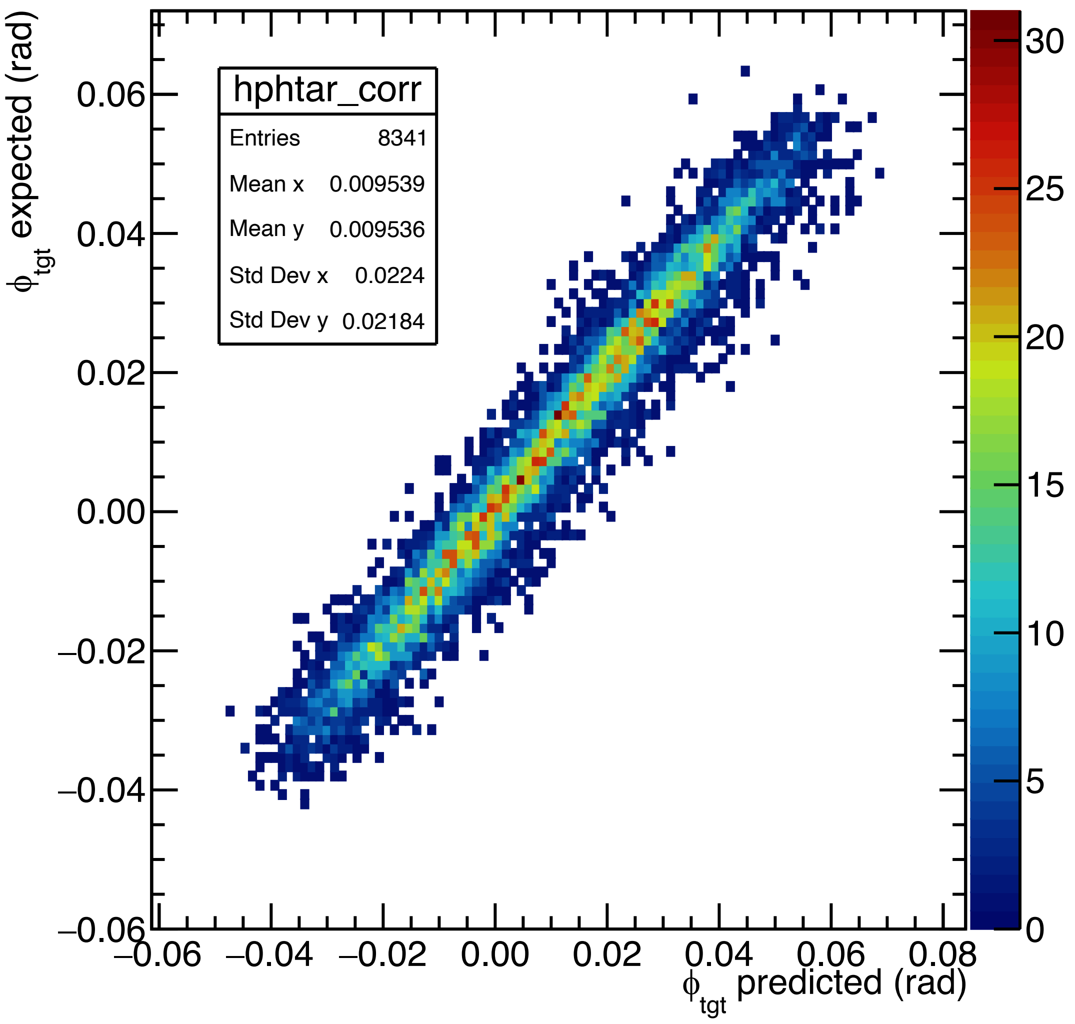


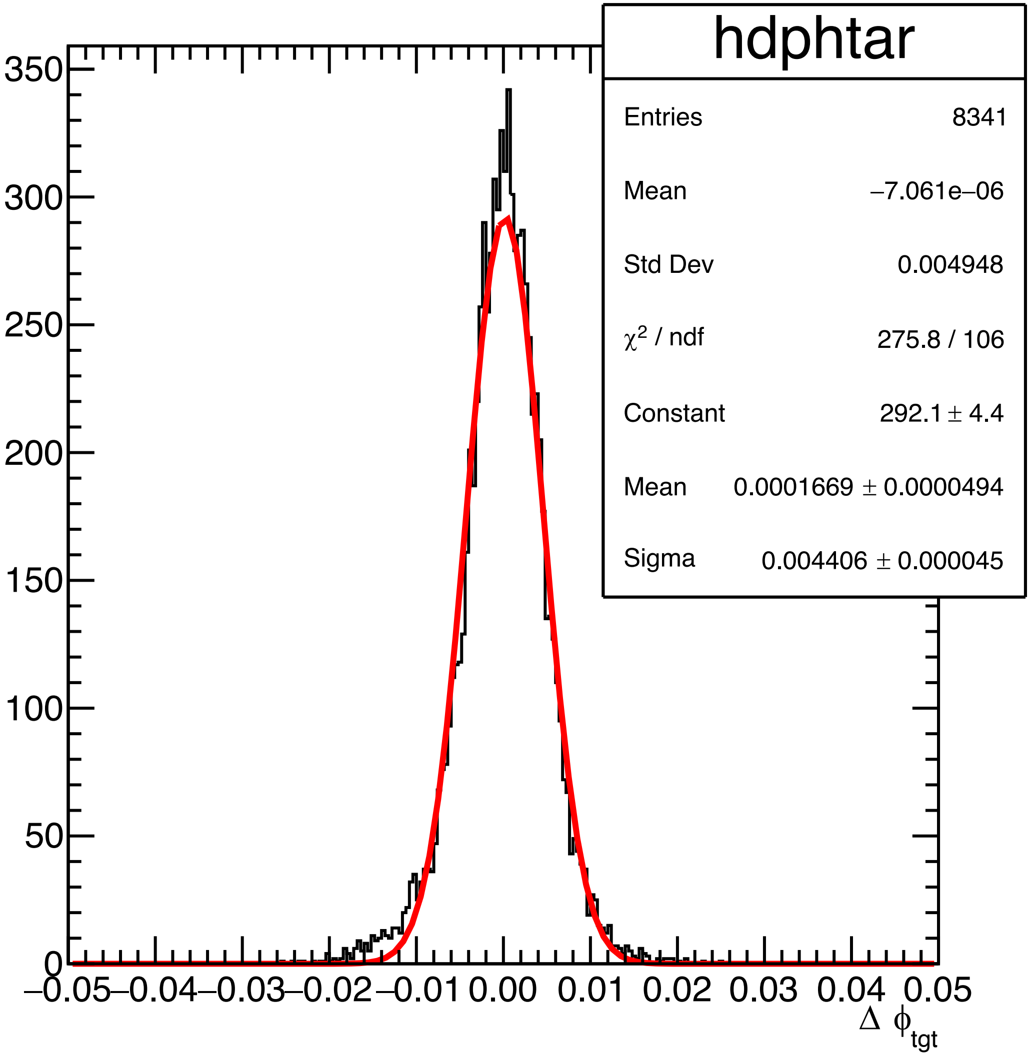
Background lower with 100% field due to more sweeping of inelastic particles

Target θ reconstruction (SBS 30% field)

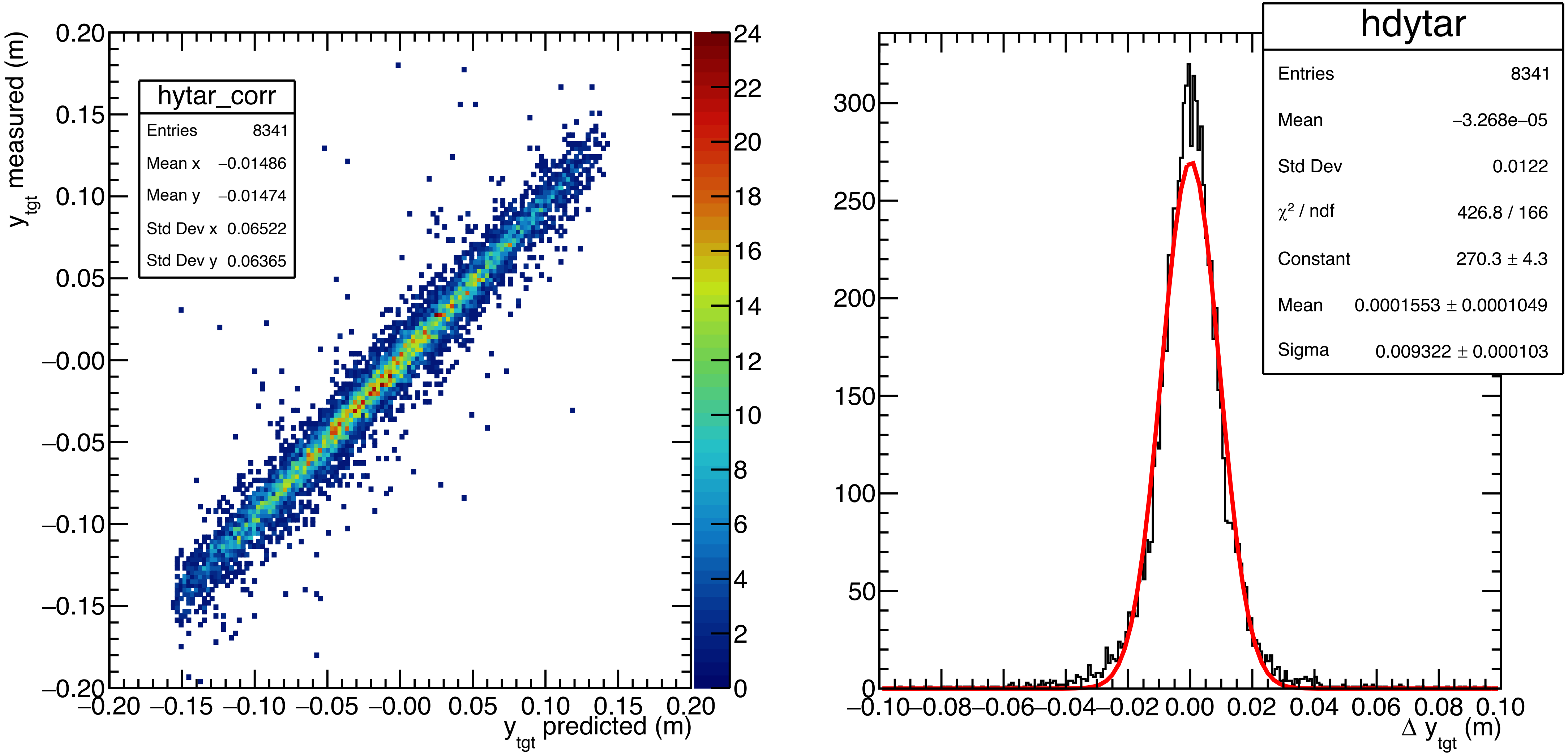


Target ϕ reconstruction (SBS 30% field)

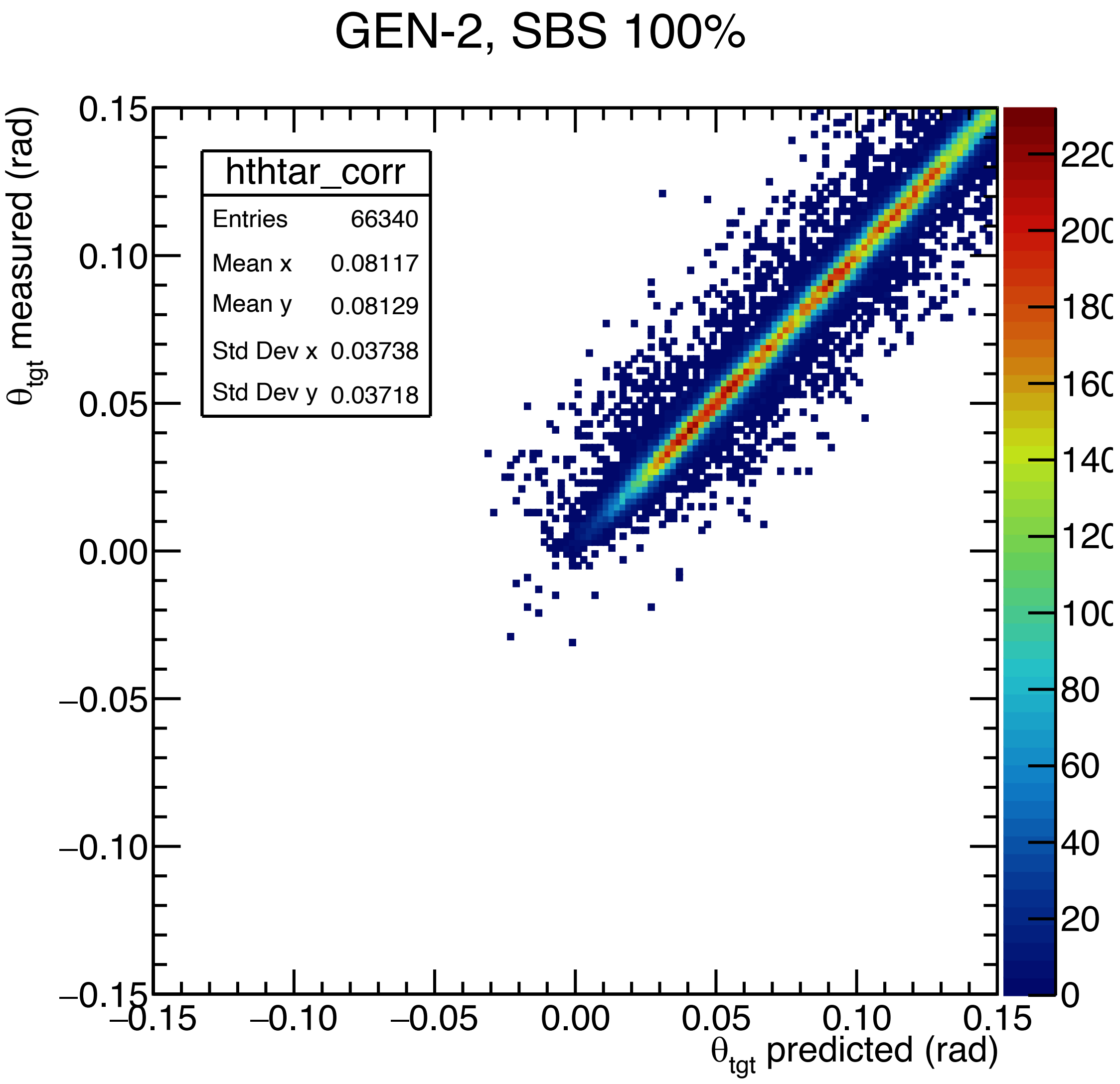


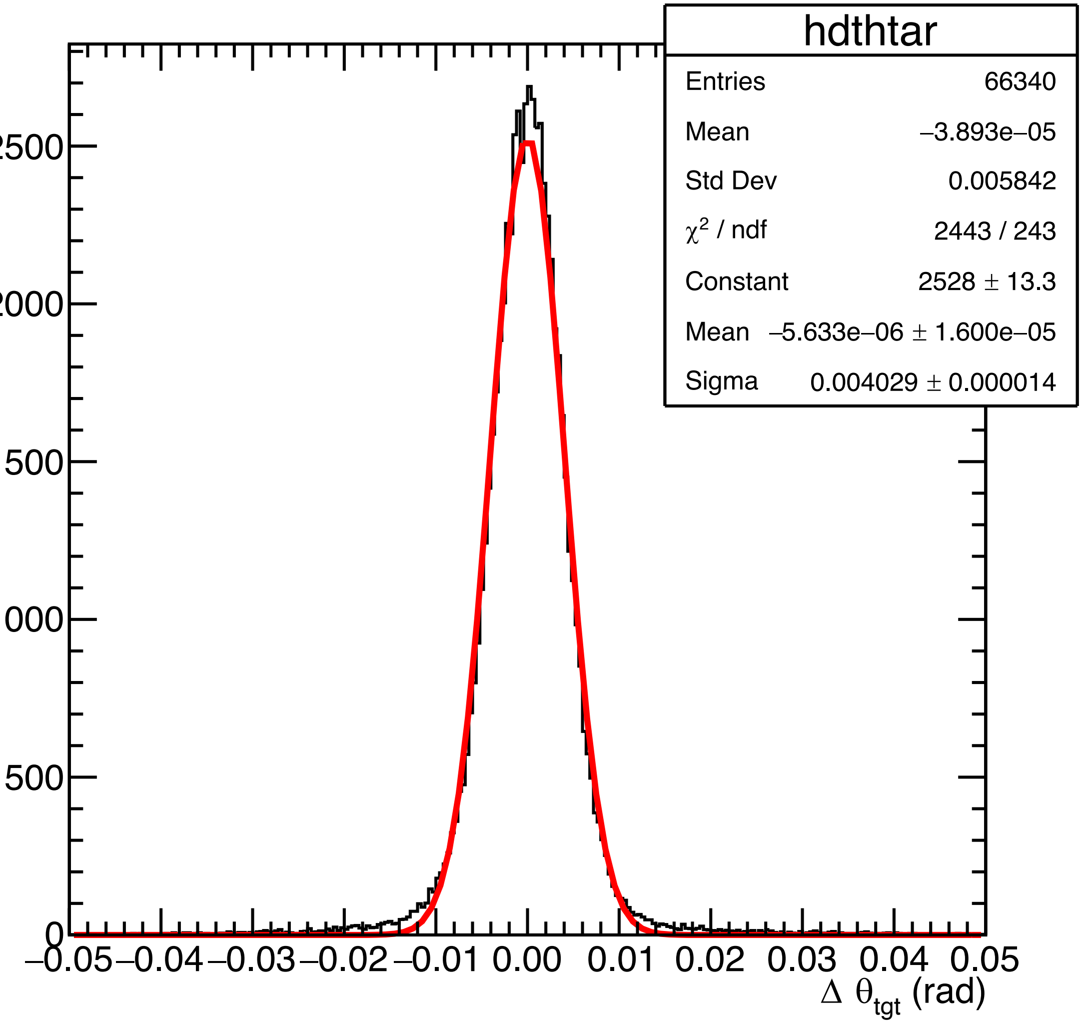


y target reconstruction (SBS 30% field)

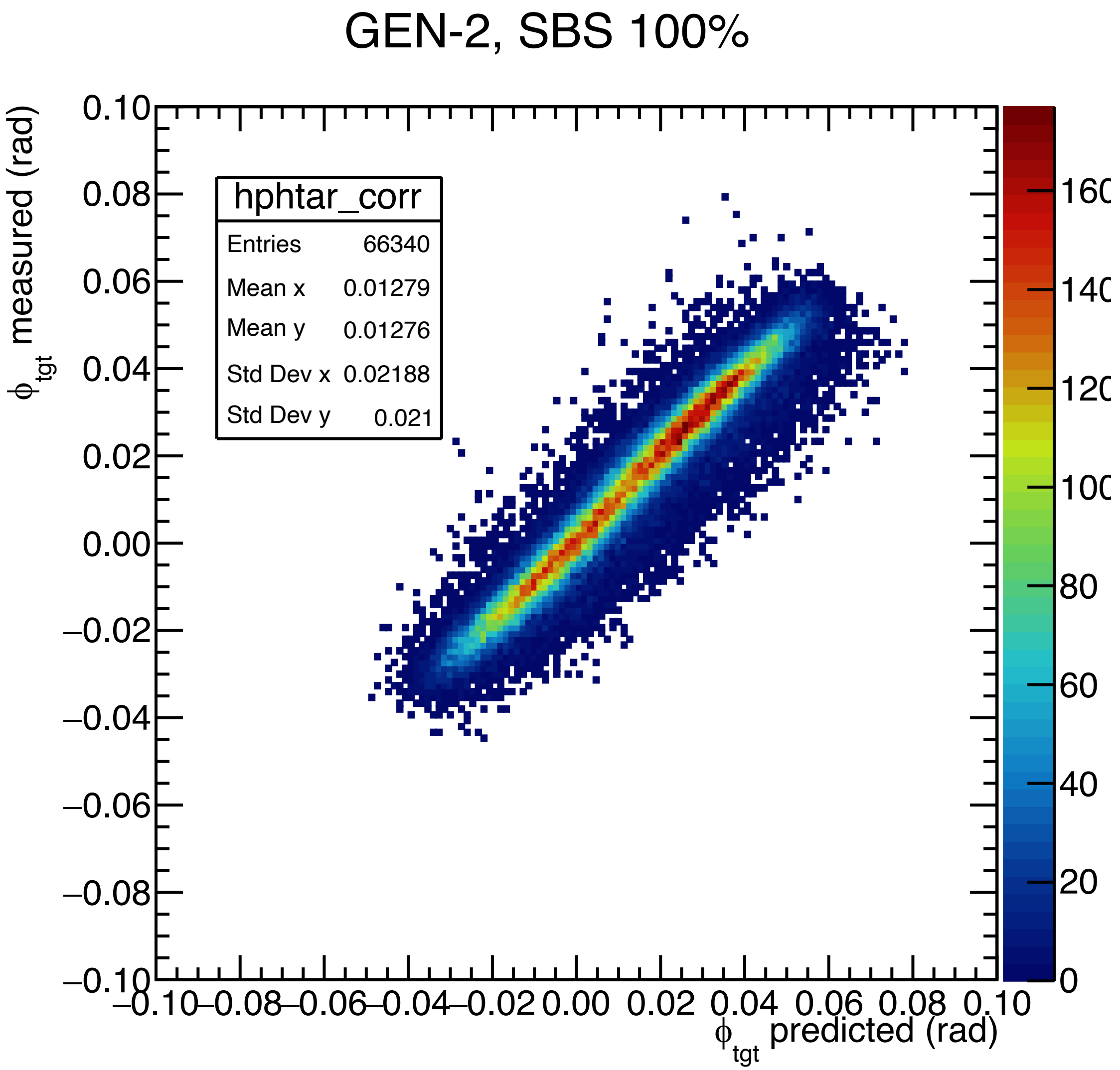


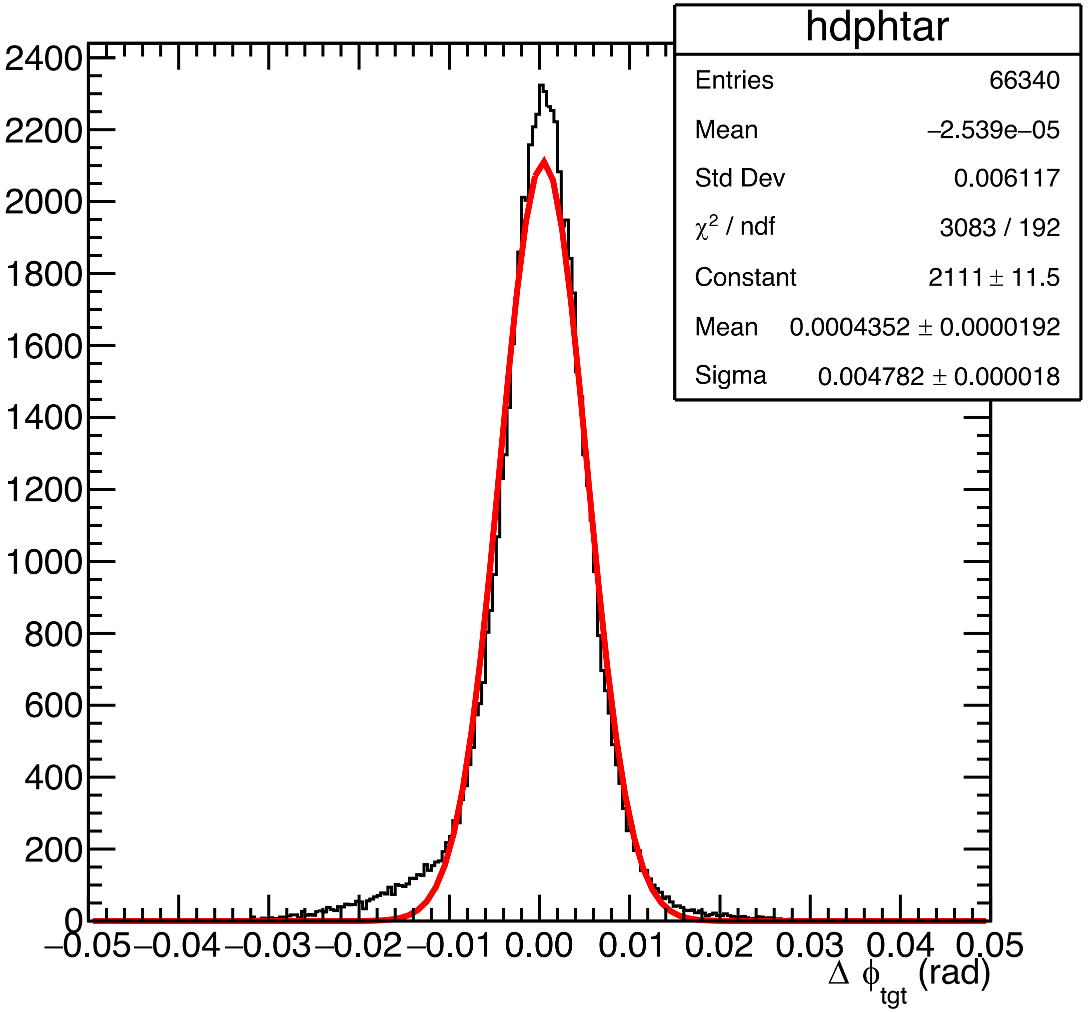
Target θ reconstruction (SBS 100% field)



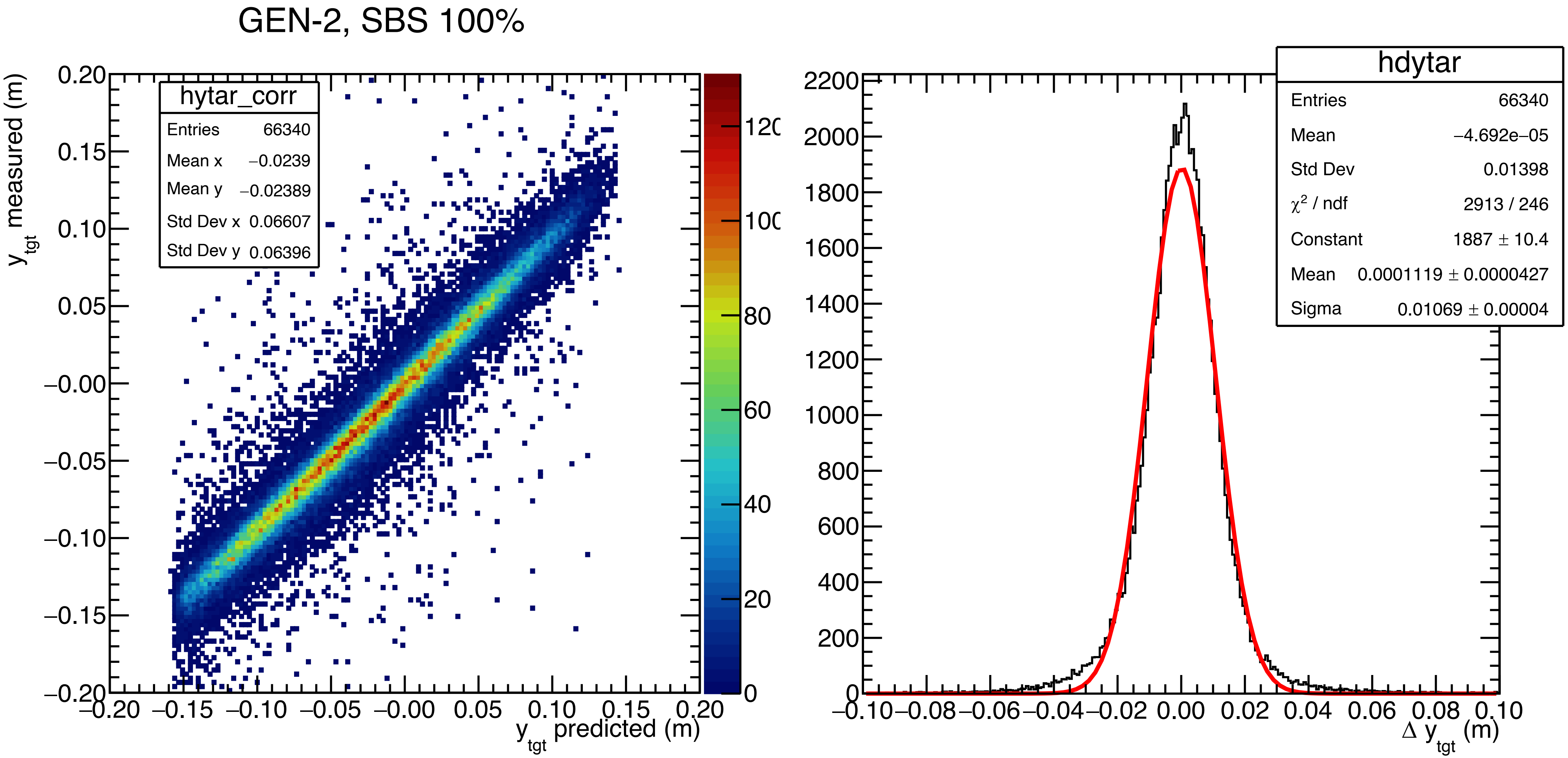


Target ϕ reconstruction (SBS 100% field)





y target reconstruction (SBS 100% field)



Summary/conclusions

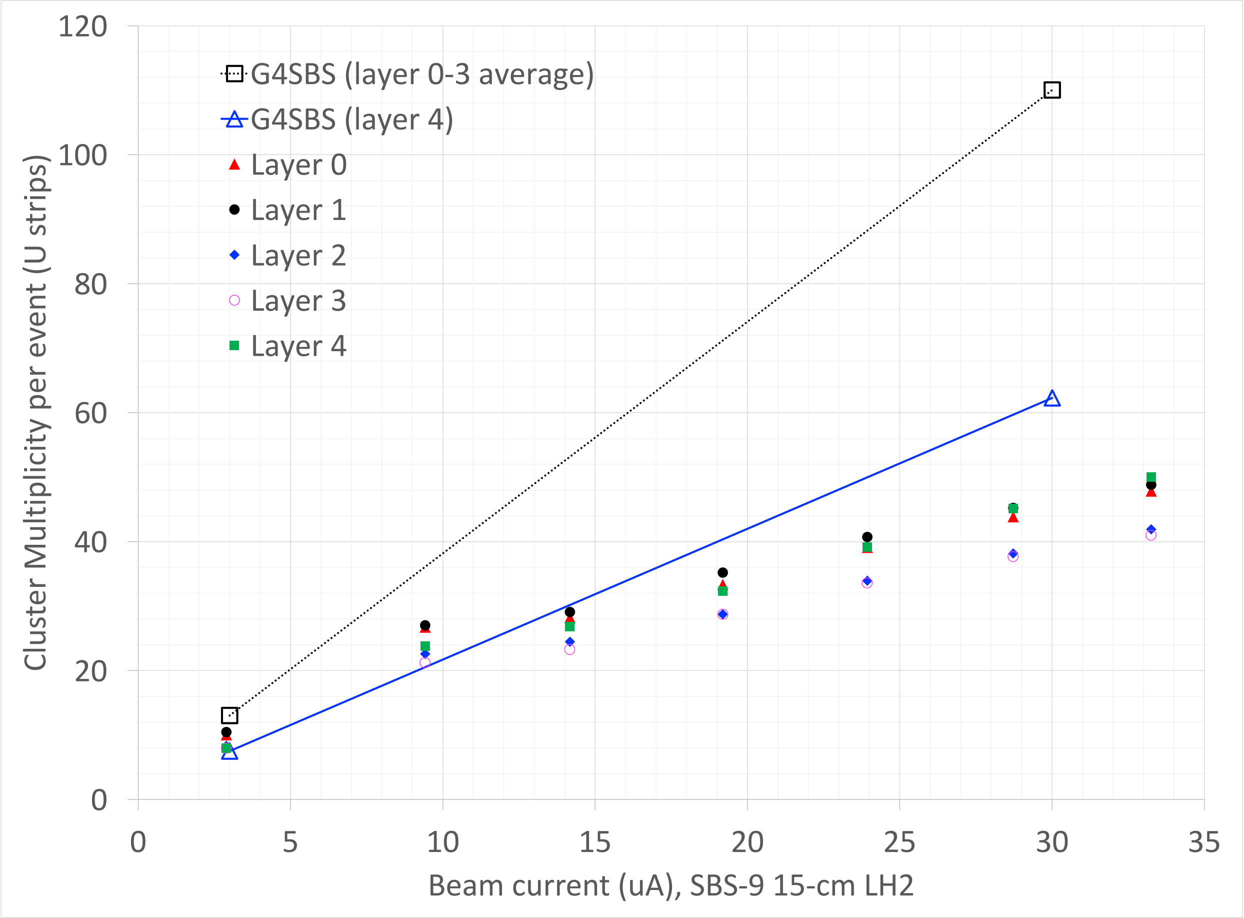
- GMN/GEN software effort was an unqualified, smashing success:
  - Ready on day 1 of beam for detector commissioning—rapid understanding of GEM performance
  - Rapid analysis progress for entirely new detector systems
  - First large-scale deployment of GEMs in Hall A high-luminosity environment
  - Rapid, efficient (in memory/CPU) processing of "big" (2 PB) dataset (down to <5 TB in first reconstruction pass)
- Well-developed, mature simulation and reconstruction infrastructure, with working machinery to robustly benchmark algorithm performance under realistic background rate conditions (and test changes/improvements)
- New code(s) needed for GEP/GEN-RP are mostly incremental additions to codes developed for GMN/GEN, which are flexible, modular and reusable
- Much of the remaining work is in the development of reconstruction/calibration/monitoring tools for ECAL/CDET and SBS optics/spin tracking
- We have a clear solution path forward to achieve acceptable high-rate performance of GEM hardware and software
- Thank you for your attention: questions?

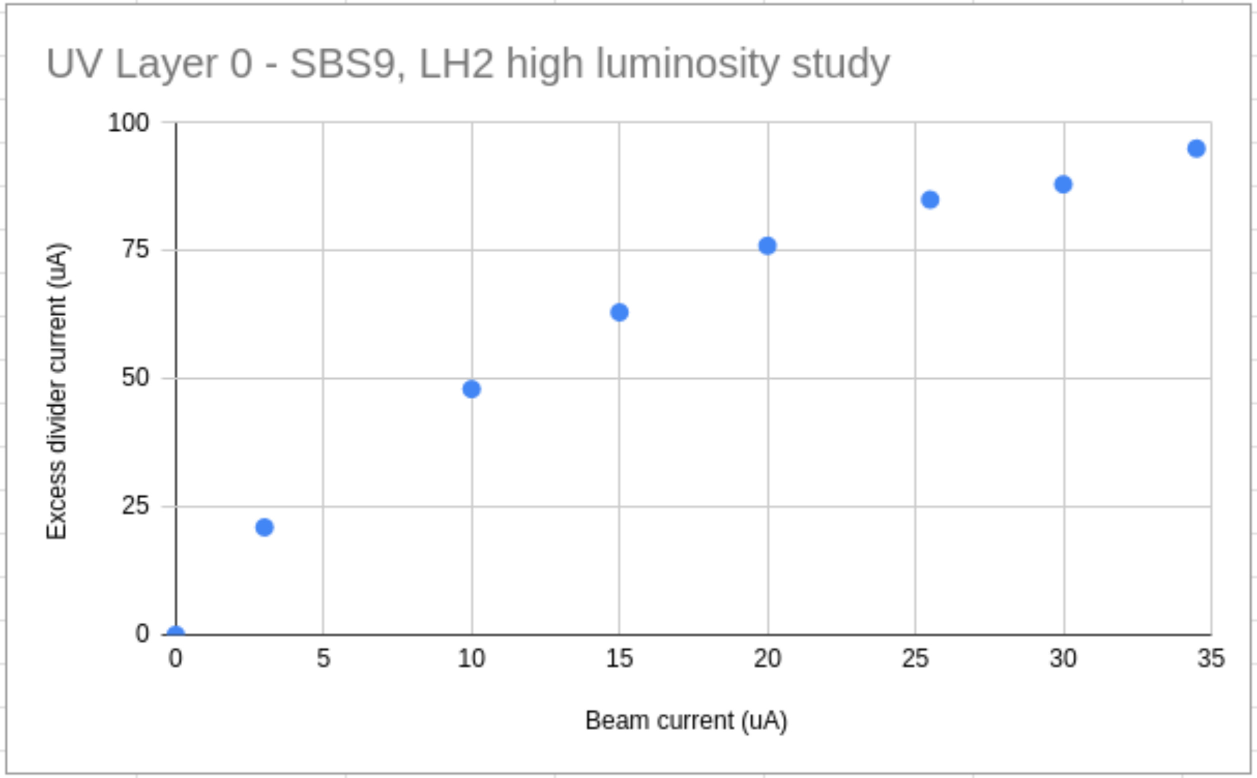
Backups

How Generic SBS GEM Tracking Works (Simplistically):

- All SBS GEMs are planar with two non-parallel readout strip orientations; each GEM layer measures one point along the track in 3D space
- First step is baseline estimation and subtraction and zero suppression (online and/or offline)
- 1D clustering of strips along each dimension in each GEM chamber
- Form all possible 2D combinations from 1D clusters within calorimeter-defined search region, subject to some basic quality cuts
  - **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, X/Y (or U/V) ADC asymmetry and correlation coefficient, 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 region of interests defined by other detectors)
- 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). Definition of "good track" is based on a user-defined maximum $\frac{\chi^2}{ndf}$
- We also impose some basic track quality/track slope/optics-based constraints on "track candidates" internally to the track-finding algorithm 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 these iterations until we run out of "unused" hits in the search region.
  - With a 3-hit minimum, the "fake track" probability is 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.

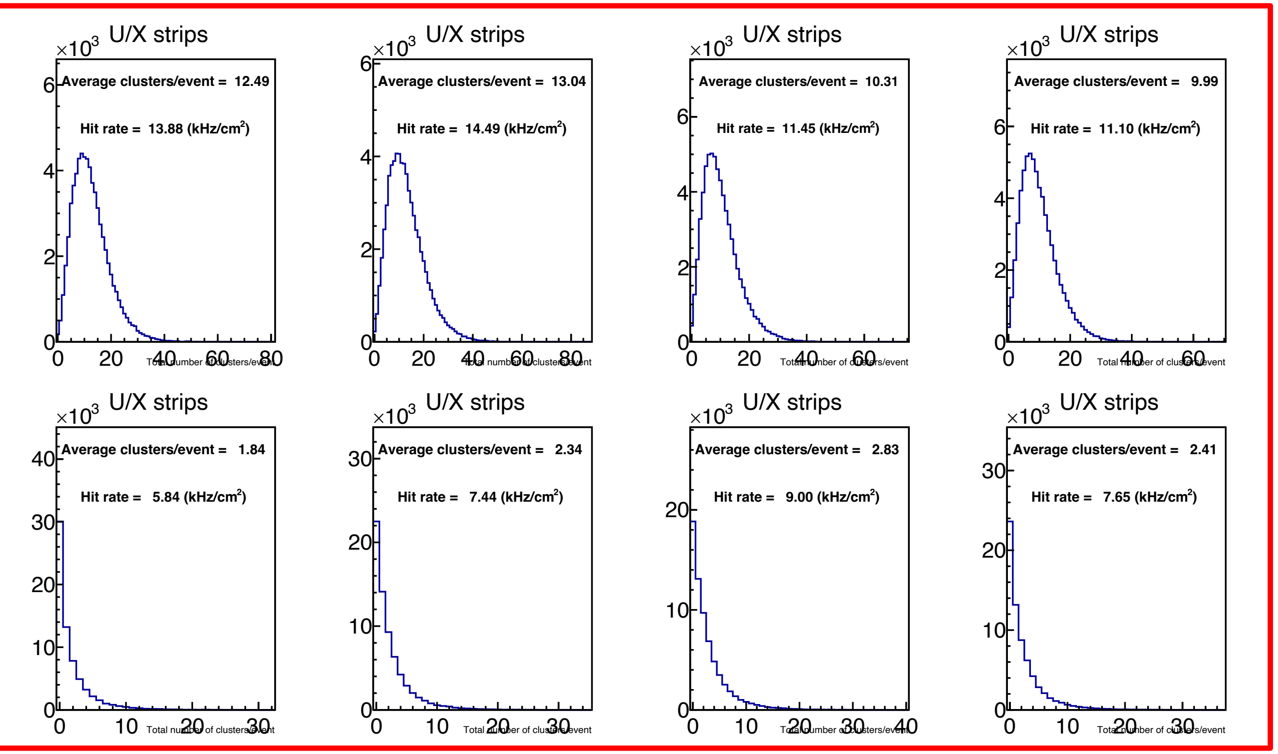
GEM hit rates and divider currents during GMN



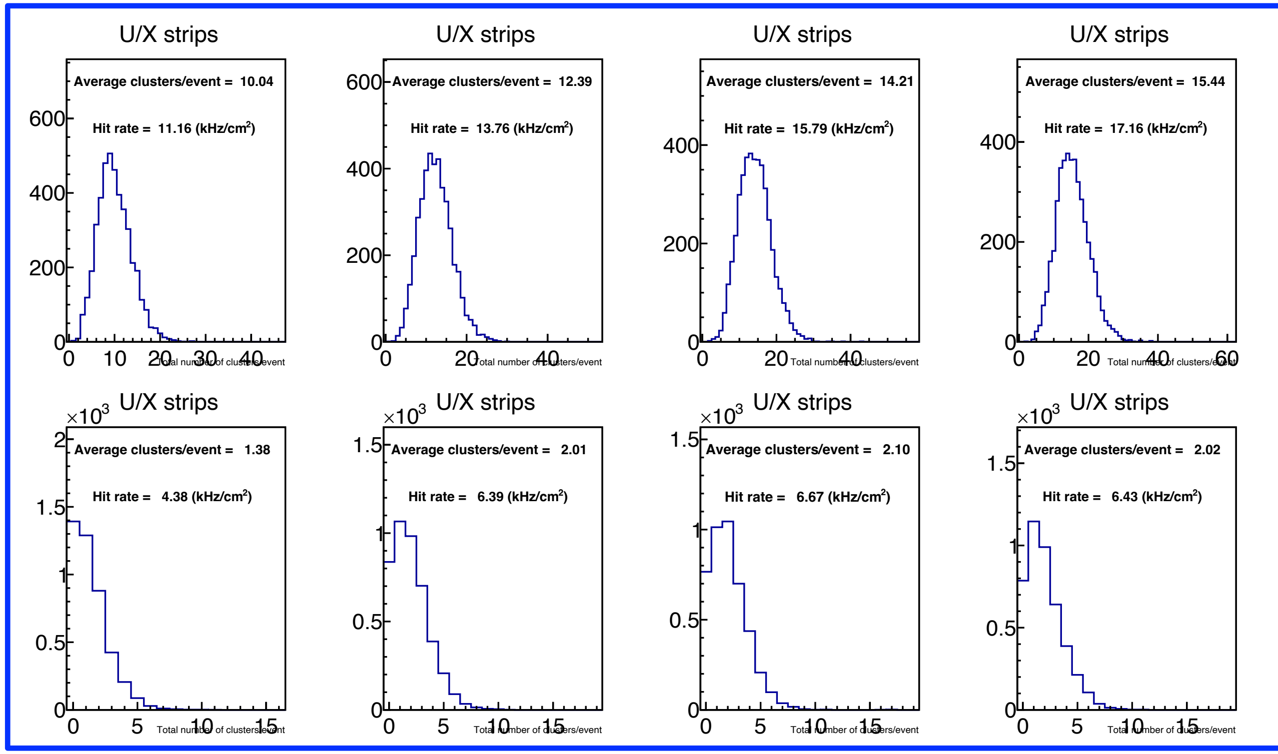


Benchmarking *g4sbs* GEM background rates

Run 13656, 3 uA, LH2 target (no timing cuts)



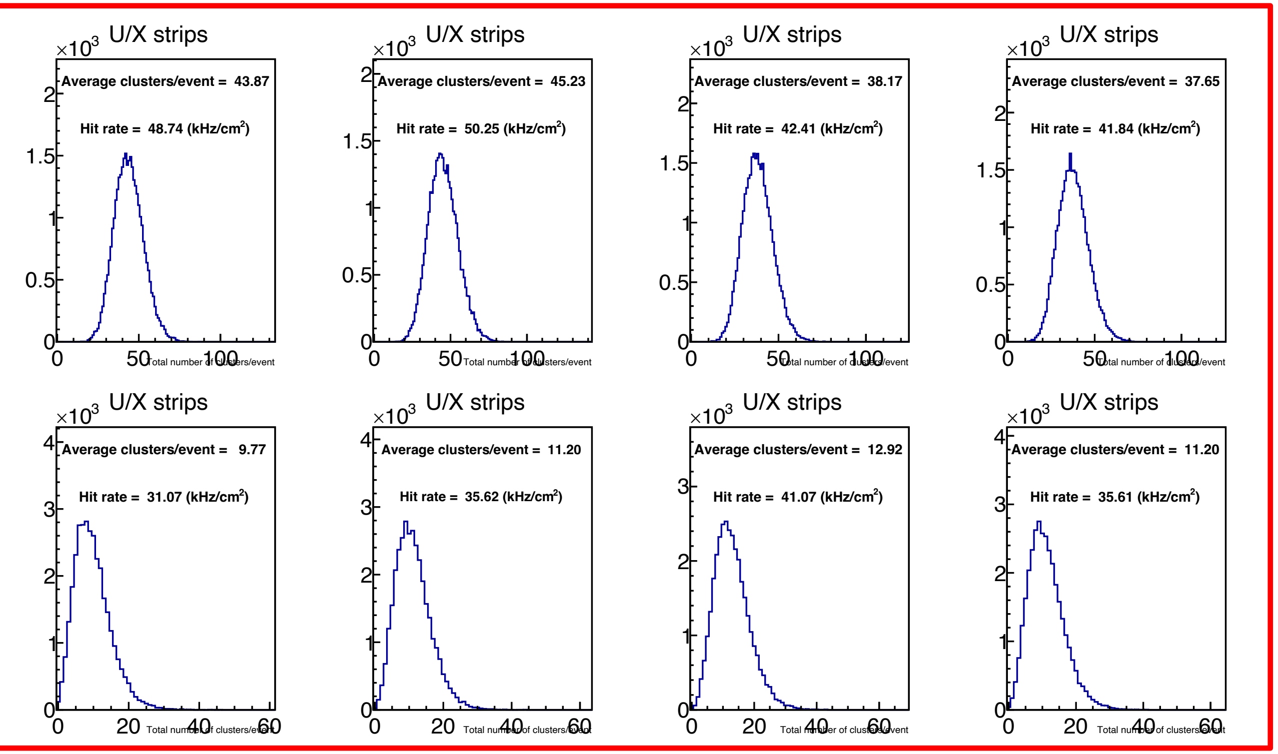
G4SBS simulated background @3 uA, LH2 target



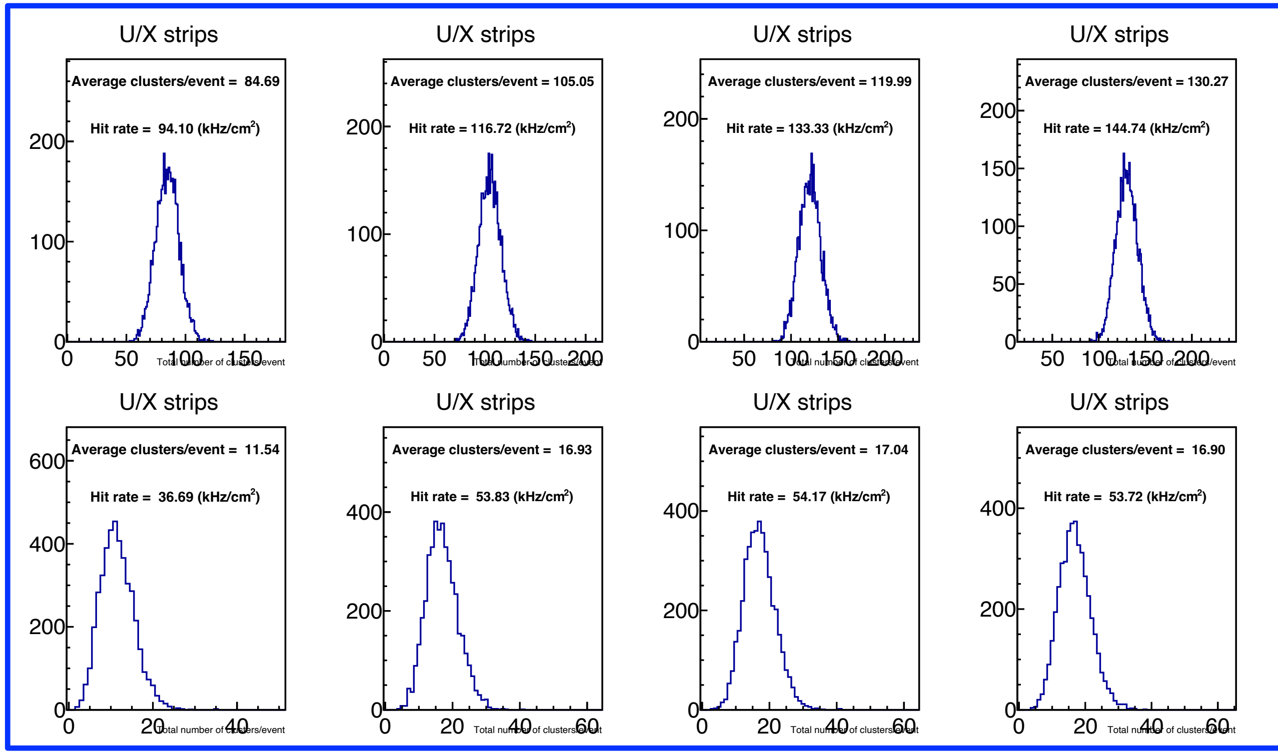
- At low beam current (3 uA on 15-cm LH2), reconstructed cluster multiplicities per event agree between real data and Monte Carlo to within a factor ~1.5 without any fine-tuning of simulated pedestal/common-mode noise, GEM gain, zero suppression thresholds, electronics effects like crosstalk, etc.
- Suggests *g4sbs* rate estimate is more reliable than initially feared, after full digitization and reconstruction with simulated beam-induced background, *under conditions where GEM gain/efficiency drop were not significant during GMN!*

Comparing *g4sbs* and real data at high beam current

Run 13789, 30 uA, LH2 target (no timing cuts)



G4SBS simulated background @30 uA, LH2 target



- At "high" beam current (30 uA LH2), the cluster multiplicities per event still *roughly* agree with *g4sbs* for the back GEM layer, but are much lower (2-3X) for the front 4 layers, indicating a significant gain/efficiency drop for the first four layers, roughly consistent with the observed reduction in elastic yield.
- (Note: GEM gain drop versus hit rate is *not* included in the simulation)