Offline Software: ongoing projects status

June 25, 2024

Recent developments and ongoing work

A non-comprehensive list of ongoing tasks:

- Event reconstruction
  - New FD tracking in Coatjava 10.1.0, more to come
  - Test of the new charged particle swimmer
  - AI/ML-based track finding for CVT
  - Support to ALERT
- Geometry
  - Fixes to DC geometry
- Simulations
  - Optimization of detector digitization to improve MC-data agreement
  - Run-dependent simulations
  - New code for the analysis of GEMC-generated background
- Analysis
  - IGUANA
  - New fiducial cuts based on trajectory information

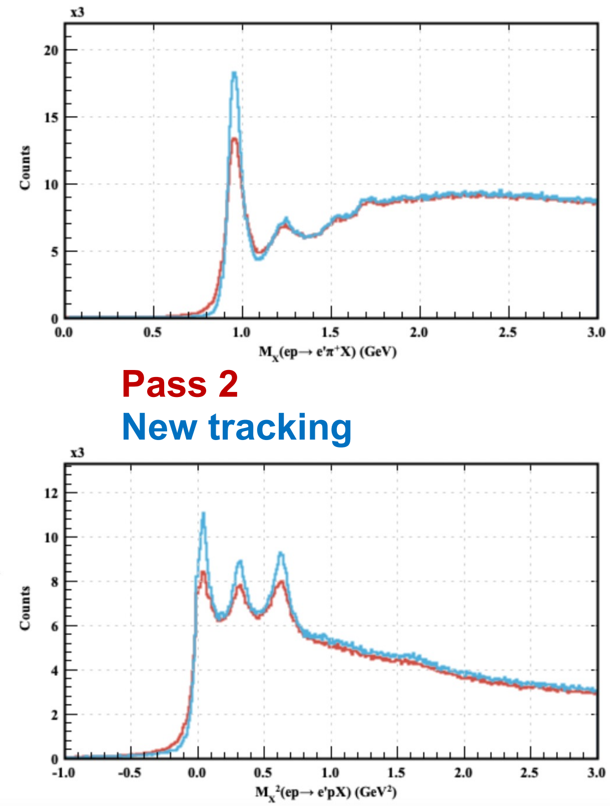
Updated FD tracking

- New track fitting based on Deterministic Annealing Filter (DAF):
  - Very significant improvement in resolution
  - Sizable increase of efficiency
  - See Tongtong's presentation in March
  - Deployed in coatjava 10.1.0, to be used for upcoming data processing
- In progress:
  - Support for new DC time-walk correction
  - New DC clustering
  - Tracking for the high-lumi configuration
  - ...

RG-A Fall2018, outbending (run 5445)

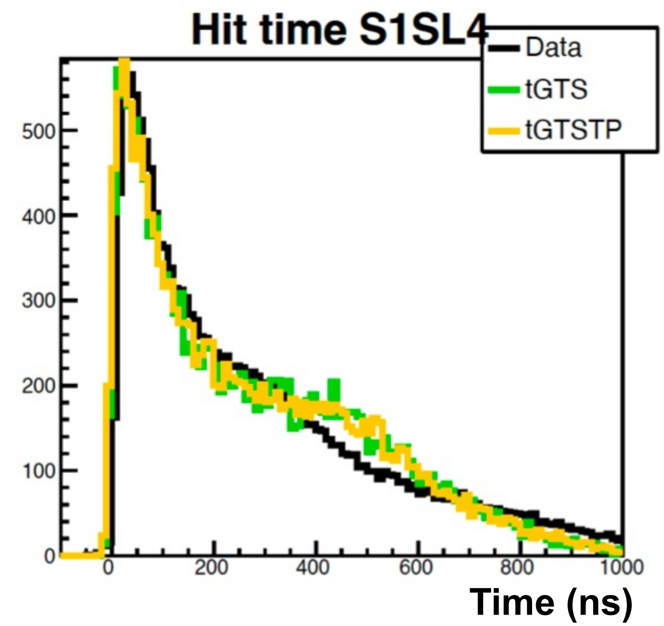
$$ep \rightarrow e'\pi^+ X$$

$$ep \rightarrow e'pX$$



Optimization of detector digitization

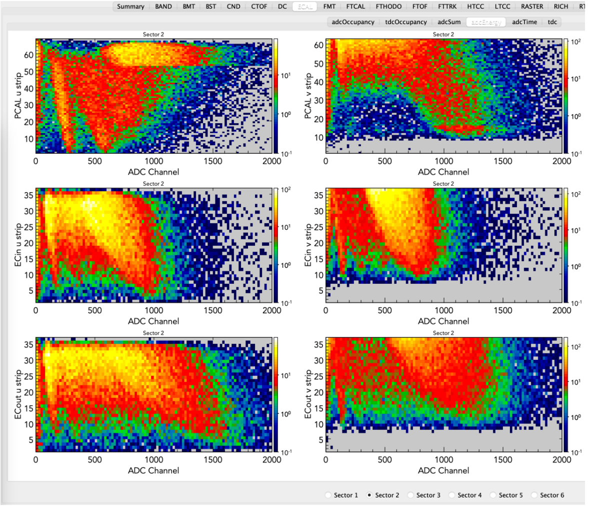
- Current simulations do not reproduce the resolution (and efficiency) observed in data
- Large discrepancy in DC tracks, other/all detectors involved too
- Systematic check/update of detector digitization algorithm:
  - DC:
    - Missing terms in simulated time added
    - Optimization of smearing function
    - Intrinsic efficiency next
    - Important findings:
      - Bias in current calibrated time offsets
      - Incorrect DC cell geometry
      - More wires than in reality…



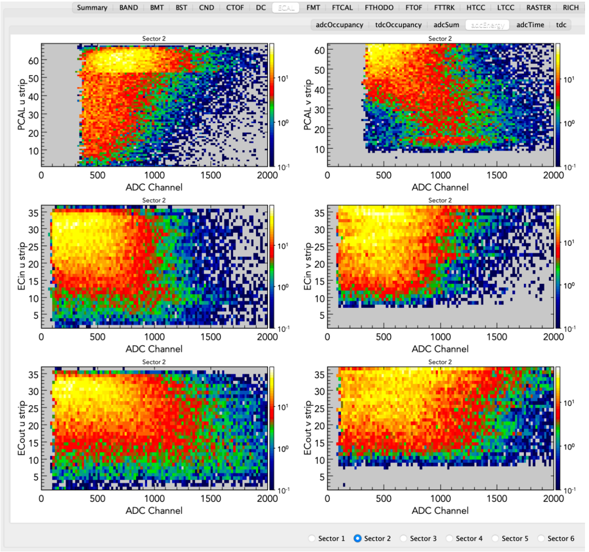
Comparison data and MC hit time including particle flight time (IGTS) and propagation time along the wire (IGTSTP), courtesy of M. Tenorio Pita

Optimization of detector digitization

- Current simulations do not reproduce the resolution (and efficiency) observed in data
- Large discrepancy in DC tracks, other/all detectors involved too
- Systematic check/update of detector digitization algorithm:
  – ECAL:
    - Using measured attenuation length
    - New, more accurate hit time-to-TDC algorithm
  – FTCAL:
    - Using measured thresholds and calibration constants
    - New smearing function in ADC calculation to account for calibration accuracy
  – FTOF:
    - Adding threshold



Old

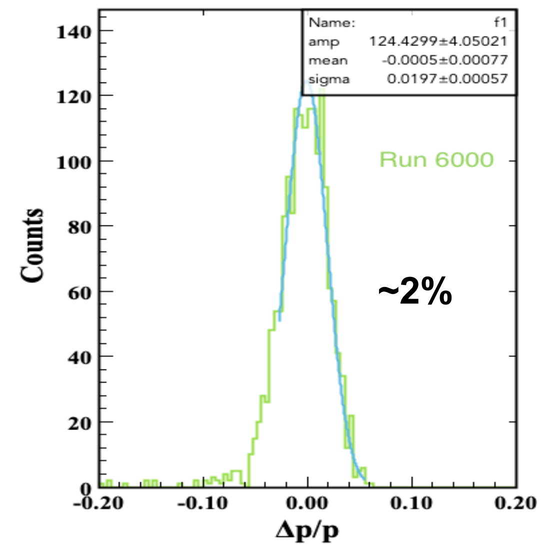


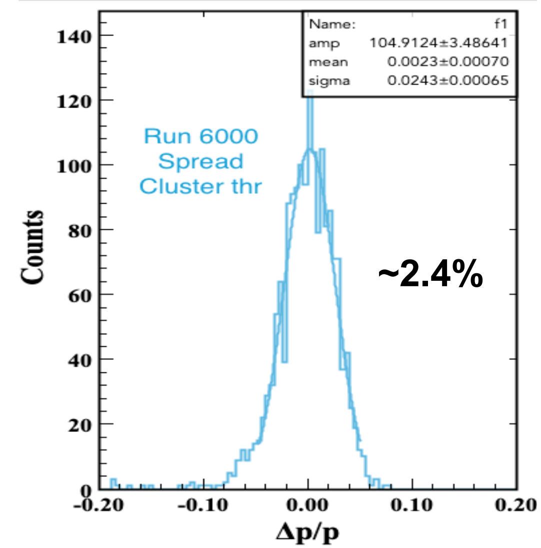
New

Digitized ECAL signals, courtesy of C. Smith

Optimization of detector digitization

- Current simulations do not reproduce the resolution (and efficiency) observed in data
- Large discrepancy in DC tracks, other/all detectors involved too
- Systematic check/update of detector digitization algorithm:
  - ECAL:
    - Using measured attenuation length
    - New, more accurate hit time-to-TDC algorithm
  - FTCAL:
    - Using measured thresholds and calibration constants
    - New smearing function in ADC calculation to account for calibration accuracy
  - FTOF:
    - Adding threshold





FTCAL energy resolution, courtesy of M. Spreafico

Run-dependent simulations

- Current simulations use a fixed run number and get run-group dependent conditions (e.g. list of malfunctioning elements) reading CCDB tables from different variations
  - Sufficient to account for main or average conditions in a data set
  - High maintenance because of multiplication of CCDB variations, gcards, and yamls
  - Impractical for implementing run-by-run changes such as temporarily dead elements

- **Ongoing efforts toward running simulations with "real run numbers", to pick up changes in detector configuration and performance (resolutions and efficiency) from CCDB tables**

Requirements:

- CCDB geometry, status, resolutions, and efficiency tables need to be properly populated for the relevant run ranges
- GEMC should load run-dependent geometry (NEW) and detector response constants (already in place based on CCDB calibration tables
- Reconstruction of MC events should work when using "real run number" calibration constants, i.e. the GEMC digitization algorithm has to be fully consistent with the reconstruction algorithm used for real data

Mechanism (on OSG):

- The user will provide a list of run numbers (or choose a data set, i.e. a predefined list of runs) and the total number of events
- Run numbers from the list will be assigned to the simulation jobs
- The number of jobs per run will be automatically determined based on accumulated charge information
- The resulting batch will reproduce the features of the data set

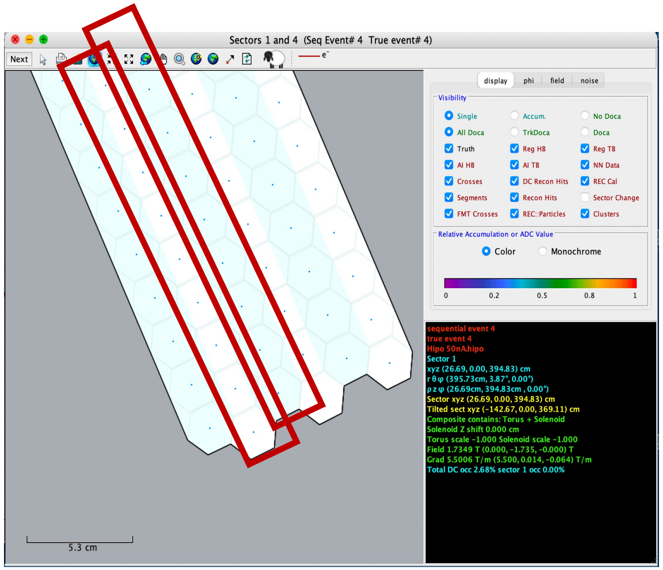
Status:

- GEMC source code, geometry and material interface developed for run dependent indexing
- Run dependent database is sqlite file (was ASCII files)
- DC geometry moved to run dependent database, other detectors to come
- Consistency of reconstruction-simulation in progress, done in conjunction with digitization update

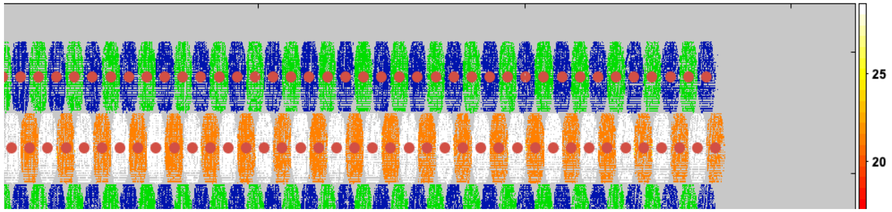
DC geometry

Work on simulation, alignment, calibration and analysis hs led to a number of findings

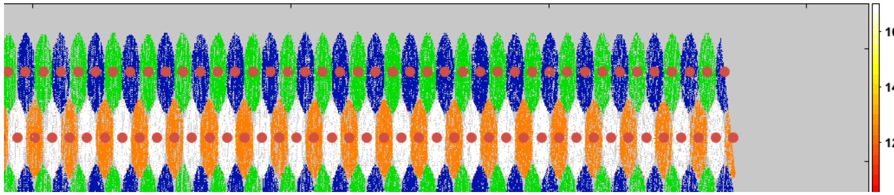
- DC cells in GEMC are rectangles instead of hexagons:
  - Affects resolution and efficiency
  - Fix in place, to be deployed
- First wire in the first two layers of superlayers 1, 2, 3, and 5 does not exist
- Phi coverage of R2 larger in simulation than in data
  - still being investigated
- Superlayer 5 and 6 wire positions do not account for ministagger of reference wire
  - 300 um shift
- (Not really a geometry issue) In R3, drift time is different in the left and right side of the cell because of the wire ministagger
  - Appears as an "effective", larger ministagger size
- Wires position does not account for the size of the wire feed-throughs
  - Shifts in the radial direction and change of stereo angle (dependent on the wire length)
  - Fix in progress



Before:



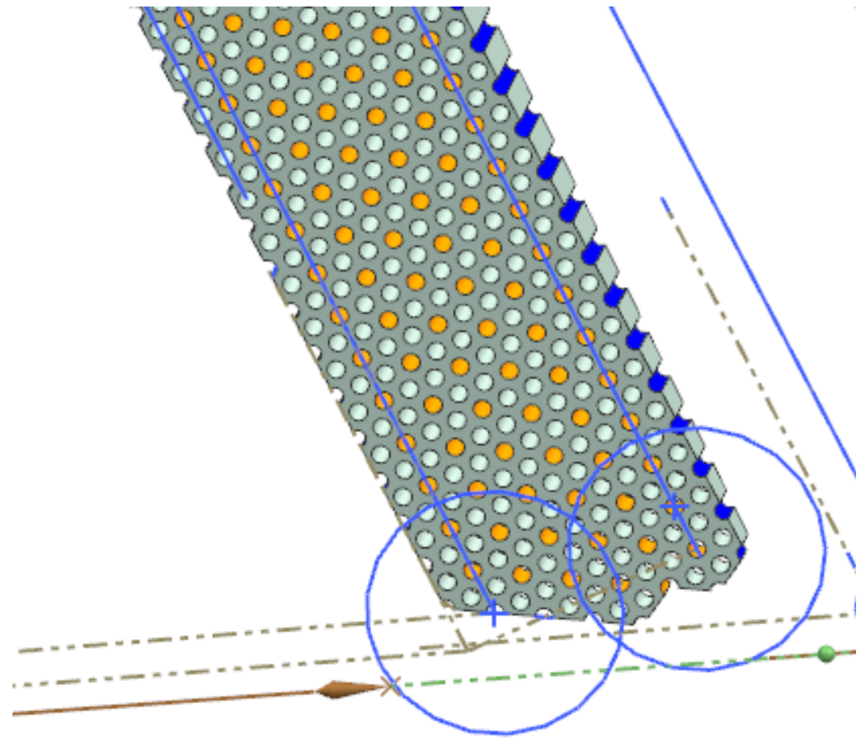
After:



DC geometry

Work on simulation, alignment, calibration and analysis hs led to a number of findings

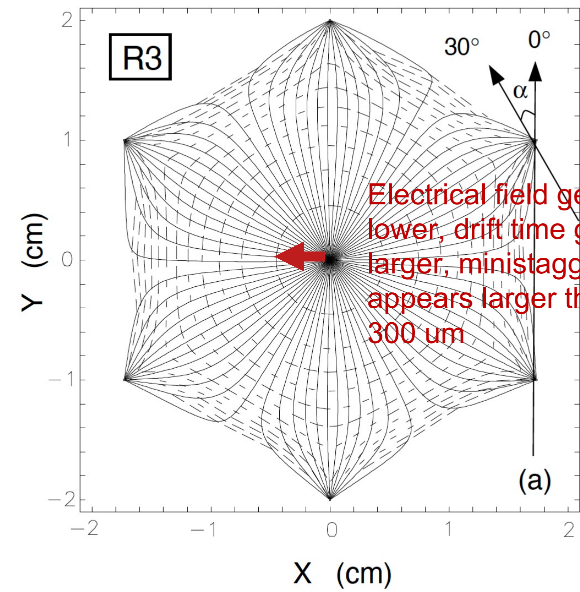
- DC cells in GEMC are rectangles instead of hexagons:
  - Affects resolution and efficiency
  - Fix in place, to be deployed
- First wire in the first two layers of superlayers 1, 2, 3, and 5 does not exist
- Phi coverage of R2 larger in simulation than in data
  - still being investigated
- Superlayer 5 and 6 wire positions do not account for ministagger of reference wire
  - 300 um shift
- (Not really a geometry issue) In R3, drift time is different in the left and right side of the cell because of the wire ministagger
  - Appears as an “effective”, larger ministagger size
- Wires position does not account for the size of the wire feed-throughs
  - Shifts in the radial direction and change of stereo angle (dependent on the wire length)
  - Fix in progress

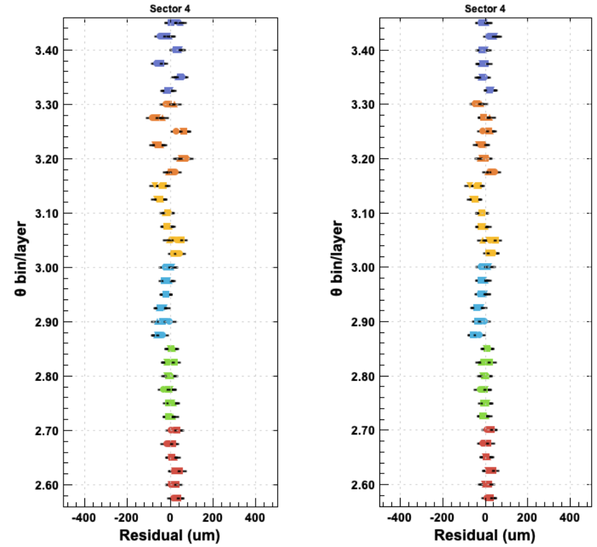


DC geometry

Work on simulation, alignment, calibration and analysis hs led to a number of findings

- DC cells in GEMC are rectangles instead of hexagons:
  - Affects resolution and efficiency
  - Fix in place, to be deployed
- First wire in the first two layers of superlayers 1, 2, 3, and 5 does not exist
- Phi coverage of R2 larger in simulation than in data
  - still being investigated
- Superlayer 5 and 6 wire positions do not account for ministagger of reference wire
  - 300 um shift
- (Not really a geometry issue) In R3, drift time is different in the left and right side of the cell because of the wire ministagger
  - Appears as an "effective", larger ministagger size
- Wires position does not account for the size of the wire feed-throughs
  - Shifts in the radial direction and change of stereo angle (dependent on the wire length)
  - Fix in progress



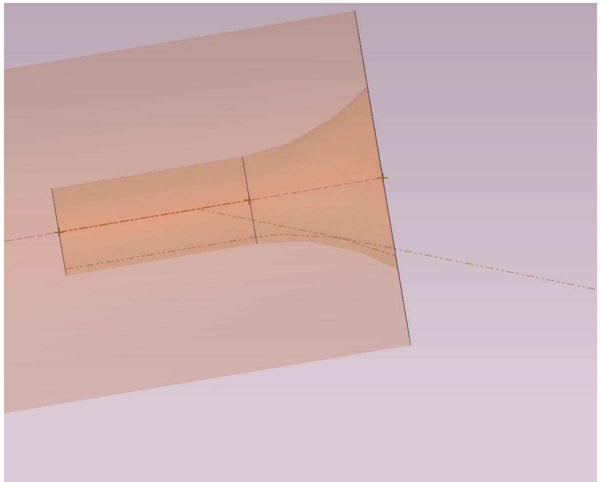


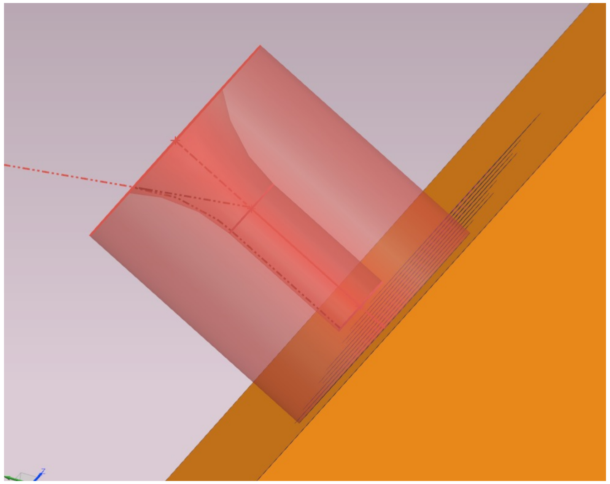
Alignment residuals with 300 um and 350 um ministagger

DC geometry

Work on simulation, alignment, calibration and analysis hs led to a number of findings

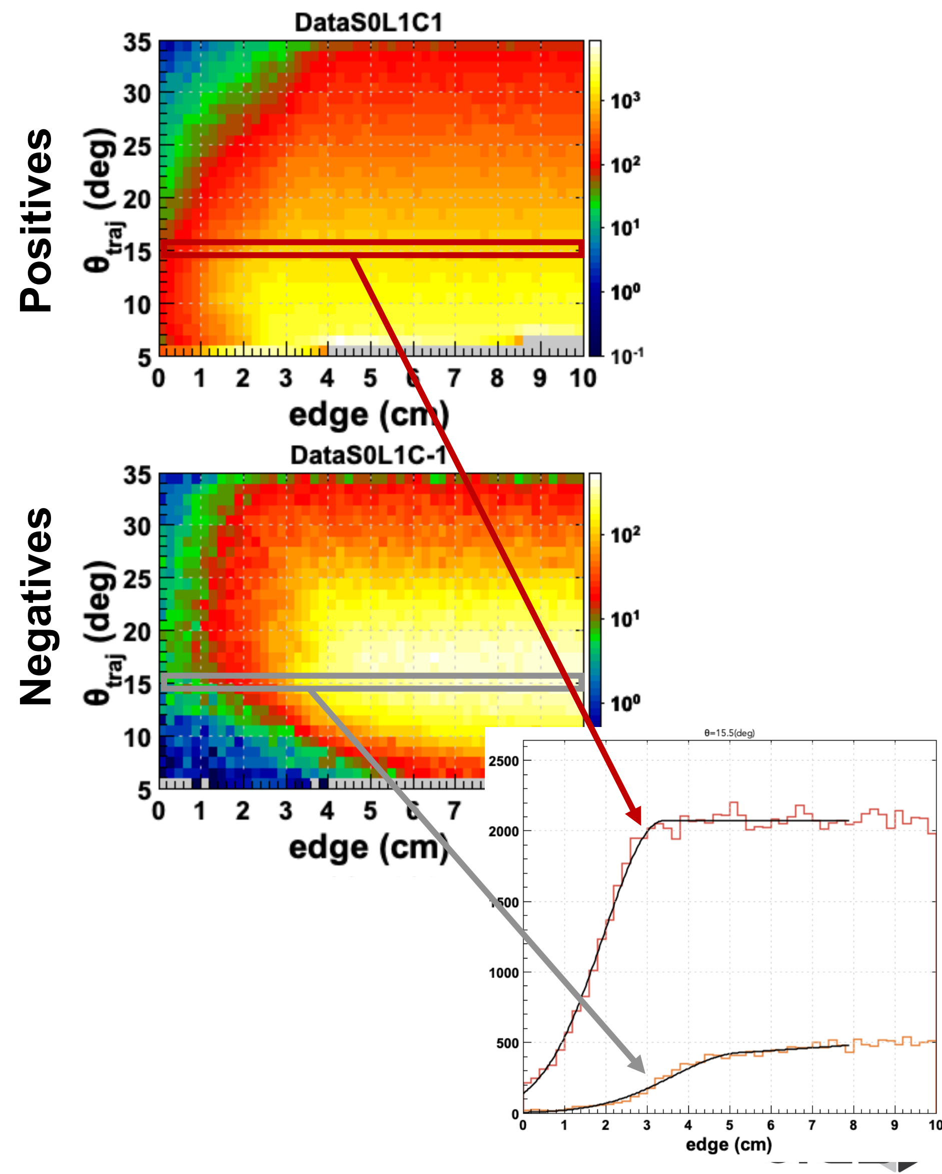
- DC cells in GEMC are rectangles instead of hexagons:
  - Affects resolution and efficiency
  - Fix in place, to be deployed
- First wire in the first two layers of superlayers 1, 2, 3, and 5 does not exist
- Phi coverage of R2 larger in simulation than in data
  - still being investigated
- Superlayer 5 and 6 wire positions do not account for ministagger of reference wire
  - 300 um shift
- (Not really a geometry issue) In R3, drift time is different in the left and right side of the cell because of the wire ministagger
  - Appears as an “effective”, larger ministagger size
- Wires position does not account for the size of the wire feed-throughs
  - Shifts in the radial direction and change of stereo angle (dependent on the wire length)
  - Fix in progress





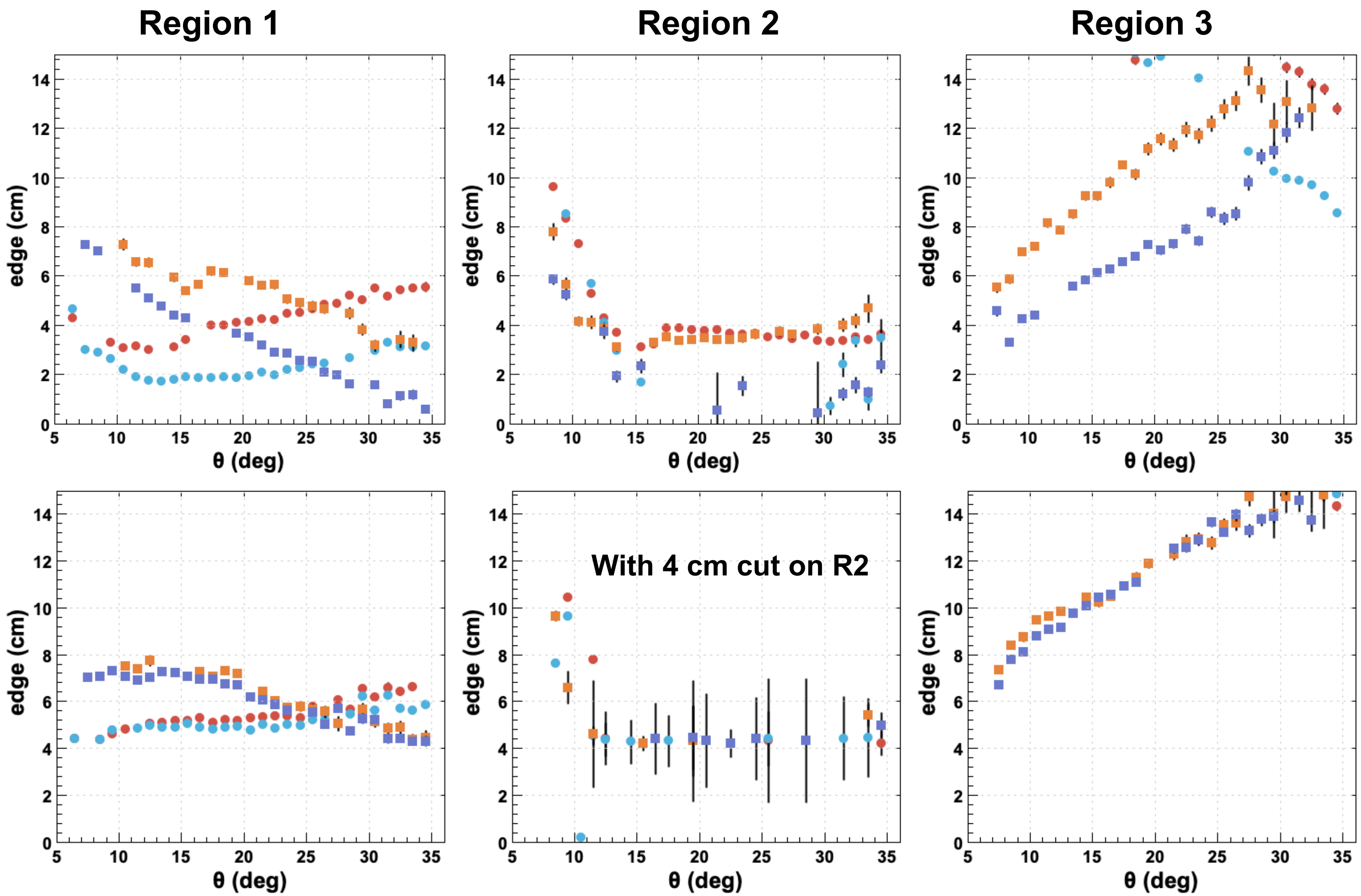
Fiducial cuts

- Determine the individual sub-detectors efficiency as a function of **REC::Traj.edge**, distance between the trajectory point and the detector edge
- Build the fiducial cut for a particle as the AND of the fiducial cuts of the detector involved in that particle reconstruction
  - Start with DC fiducial, studying charged hadrons in the FD since they depend only on DC and FTOF
  - Efficiency for each DC region inferred from the flatness of the particle yield at fixed theta vs phi
  - Verified agreement between sectors and particle type
  - Interdependency between regions complicates the interpretation initially
  - Found that R2 dominates over R1 and R3, and is responsible for the MC-data discrepancy
    - Under investigation
  - When resolved, final cleanup and testing still due



Fiducial cuts

Data: positives/negatives
MC: positives/negatives



Summary

- Many high-impact projects in progress
- Progress slower than expected because of unexpected findings
  - Focus shifted to addressing these findings
  - Once resolved, improvements are likely to be larger than originally anticipated for the given task
  - More soon…