CLAS12 Forward Tracking Status and Plans

T. Cao

Nov. 12, 2024

CLAS Collaboration Meeting

Jefferson Lab
Thomas Jefferson National Accelerator Facility

clas
CEBAF Large Acceptance Spectrometer

Forward Tracking and Efficiency

Recon. with Forward Tracking

AI-denoising

↓

Clustering

↓

6-cluster and 5-cluster combos predicted by AI

↓

pattern recognition

↓

HB Tracking for both 6-cluster and 5-cluster cases

↓

HB event reconstruction

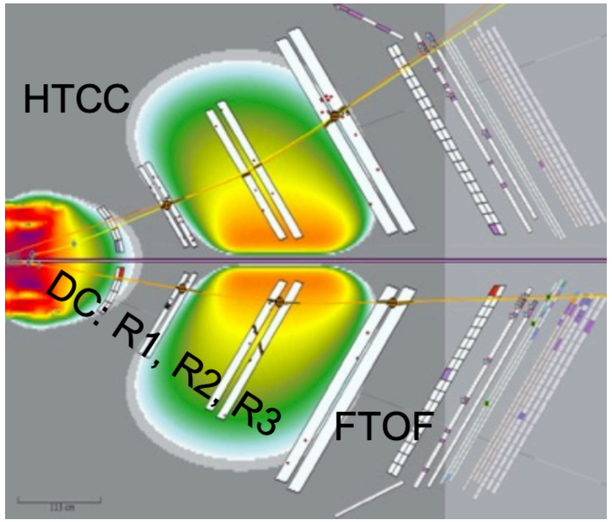
↓

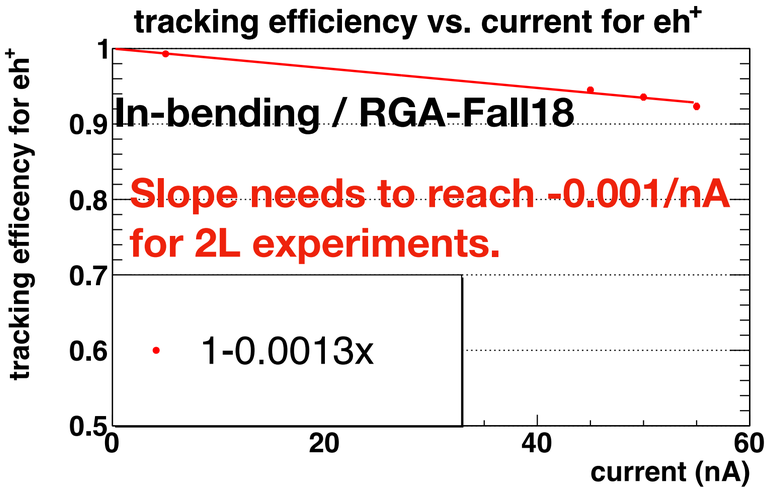
TB re-tracking with TB hit measurements

↓

TB event reconstruction

- To run future CLAS12 experiments with double luminosity (2L), forward tracking efficiency is required to be at least 85% at beam current of 150 nA.
- With AI-denoising for removal of DC noise hits, AI assistance for prediction of DC clusters, and recent new tracking with cooperation of Kalman Filter (KF) and Deterministic Annealing Filter (DAF), tracking efficiency has been significantly improved.
- However, tracking efficiency needs to be further improved to reach the goal of running at 2L.



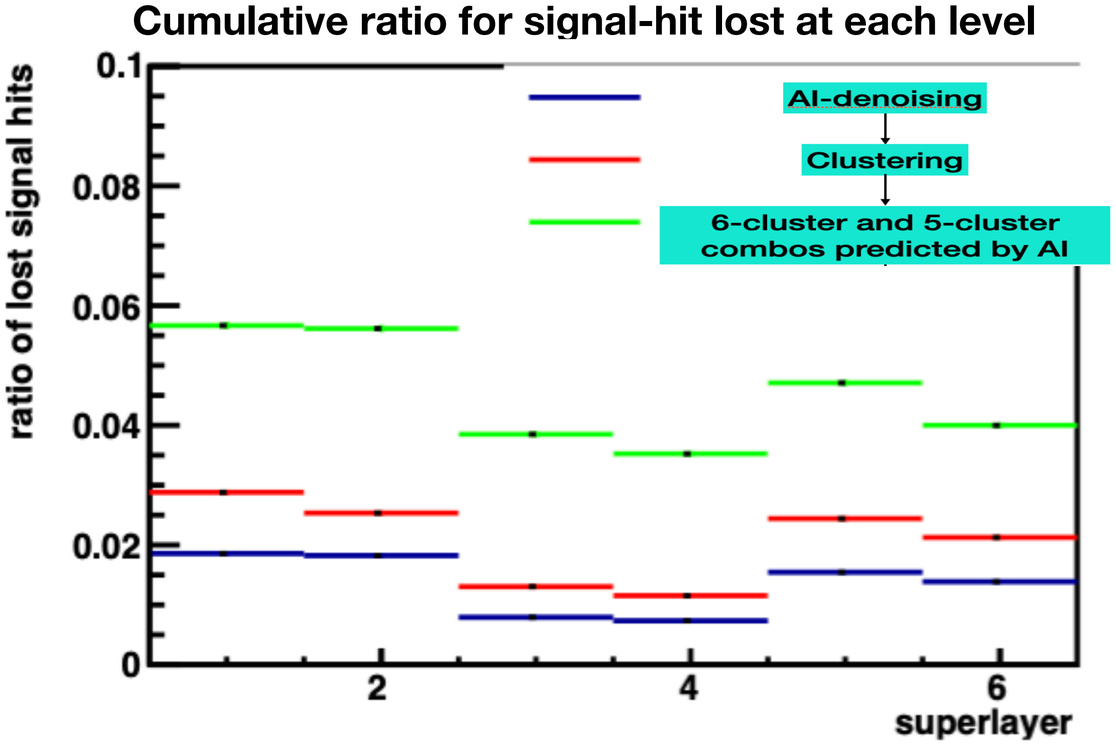


Outline

- Investigation of track lost
- Discussion about AI-denoising
- Updates of DC clustering
- Discussion about AI-assisted tracking
- Status and plans for DC-uRWell tracking
- Summary

Investigation of Track Lost

- An investigation is in process to comprehensively understand how noise hits affects tracking efficiency.
- The investigation will provide insights for improvement of tracking efficiency.
- MC sample with truth information is more suitable for the investigation. However, due to discrepancy between MC and data, the current study explores background effects through track-to-track comparison between 5nA and 5nA-50nABg RGA data.

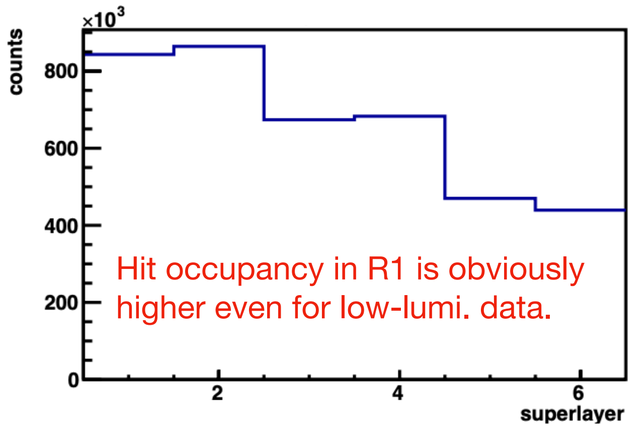


- Preliminary, the first three levels in reconstruction, which are the key for tracking efficiency, were investigated.
- The investigation tells that part of signal hits are lost at each level.
- More signal hits are lost in R1 due to higher noise hit occupancy, while less signal hits are lost in R2.

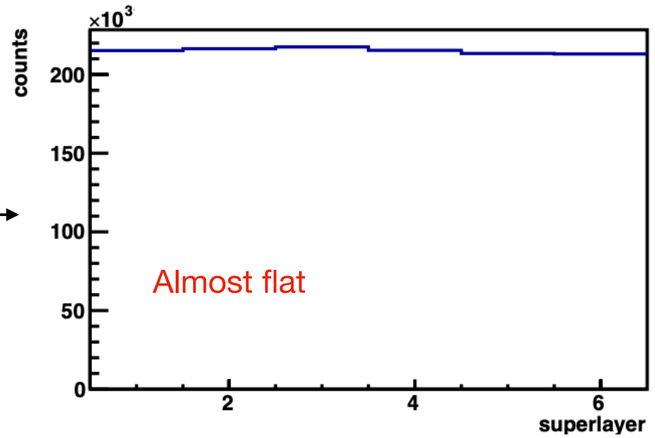
Discussion about AI-denoising

By checking remaining hits after AI-denoising

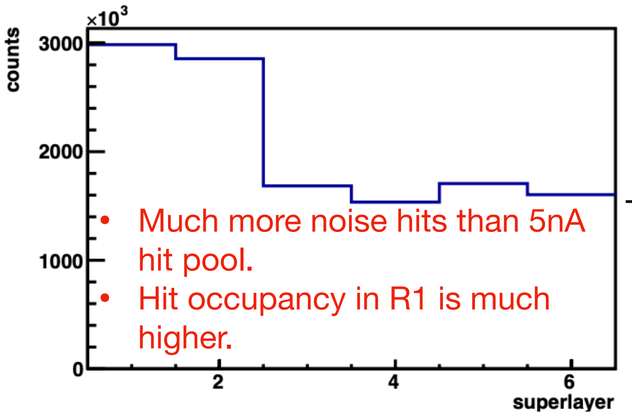
Remaining hits from 5nA hit pool in 5nA-50nABg data



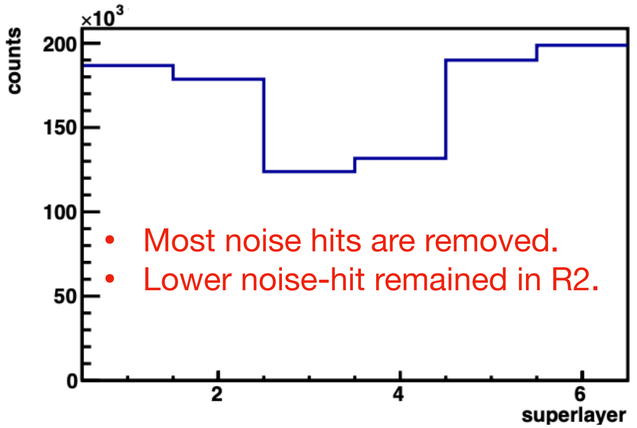
AI-denoising



Remaining hits from 50nABg hit pool in 5nA-50nABg data



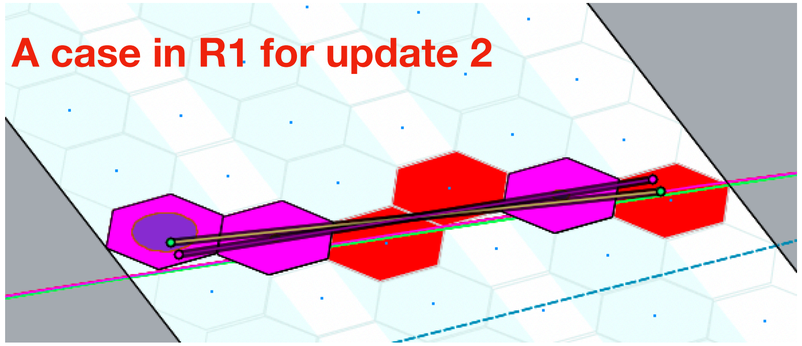
AI-denoising

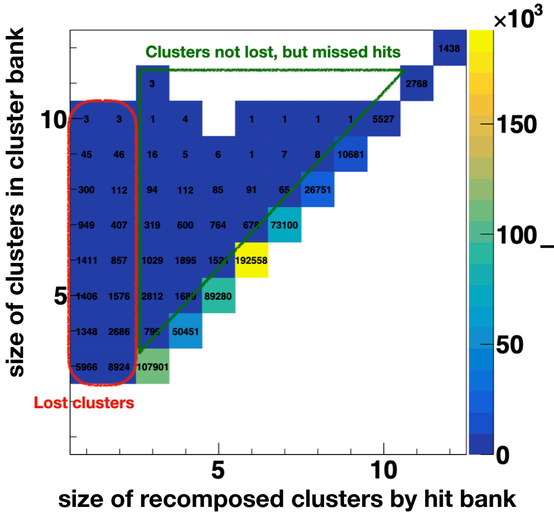


AI-denoising seems to work better for R2 with higher noise-hit suppression and lower signal-hit lost.
Should threshold in the model of AI-denoising be different for different regions?

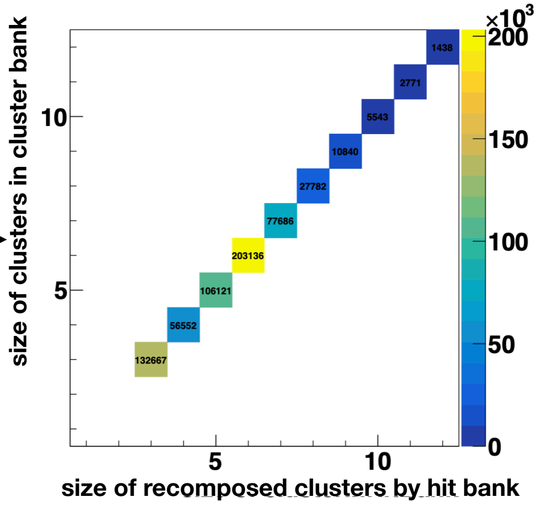
Updates of DC Clustering

- By investigation of DC clustering algorithms, some issues, leading to cluster lost, were found.
- Updates with issue fixing and algorithm optimization retrieve missing real clusters, while do not increase too much extra noise clusters.
- Main updates include:
  1. Fix a bug in the splitter for complicated hit clumps, and optimize the routine for choice of hit-overlapping clusters from the splitter.
  2. With consideration of signal-hit lost due to dead strips, AI-denoising, edge effect, etc, limit of cluster size for exceptional cases is loosed from 4 to 3.
  3. Fix an issue that a hit is shared by multiple clusters in clustering, but is only associated with one cluster when writing it into the hit bank.





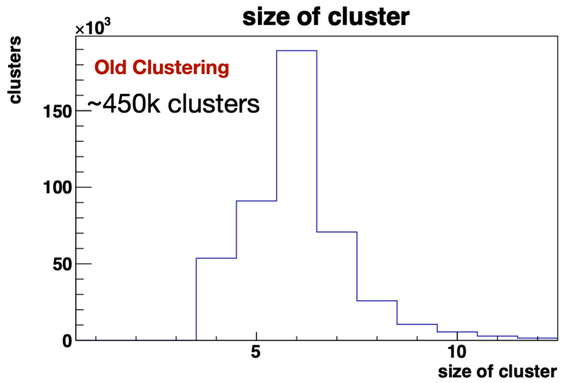
with update 3

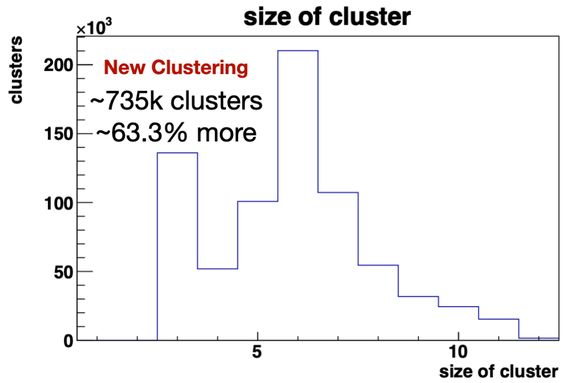


10k events from RGA Run 6666 (50 nA)

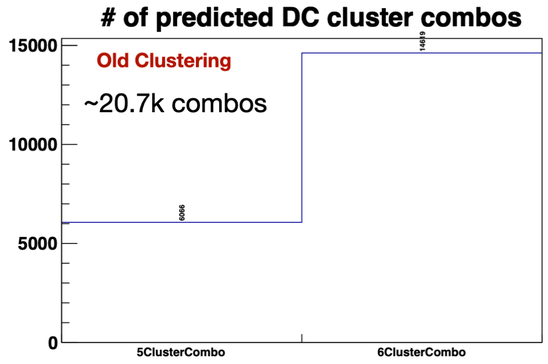
Effects of Clustering Updates

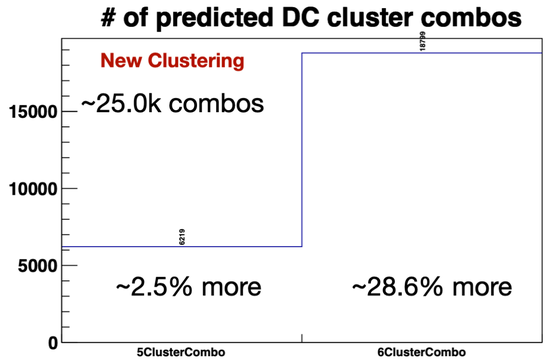
Clusters from clustering



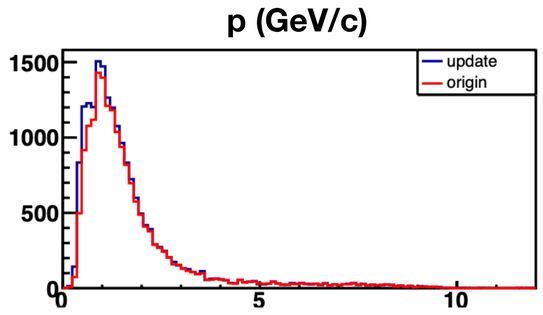


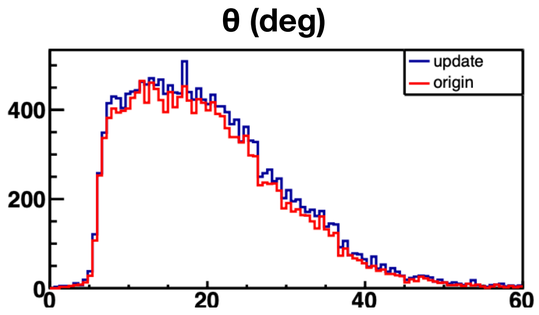
Cluster combos by AI prediction

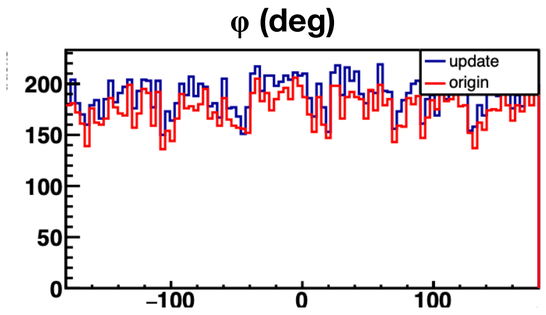




AI-assisted TB Tracks ~8.6% more







RGA-Fall18

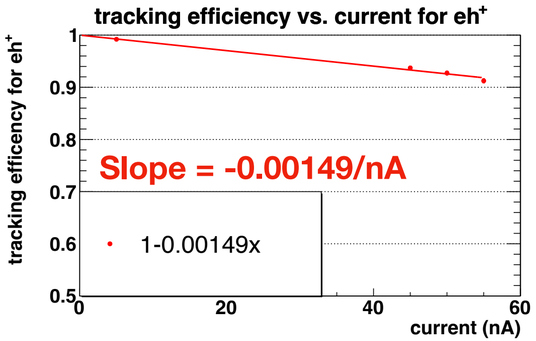
Tracking Efficiency Scan

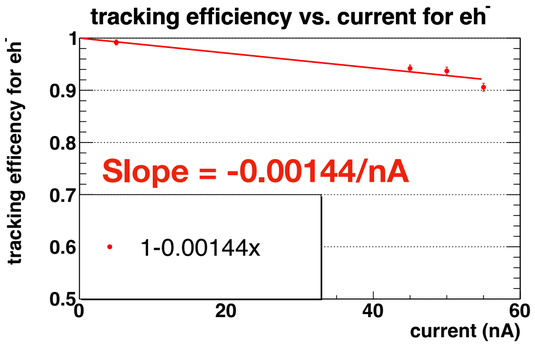
In-bending

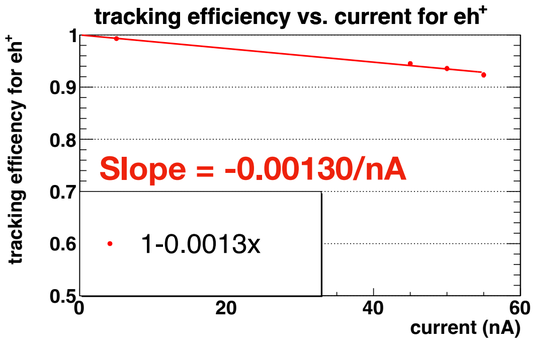
pass2 including AI-denosing and AI-assisted tracking

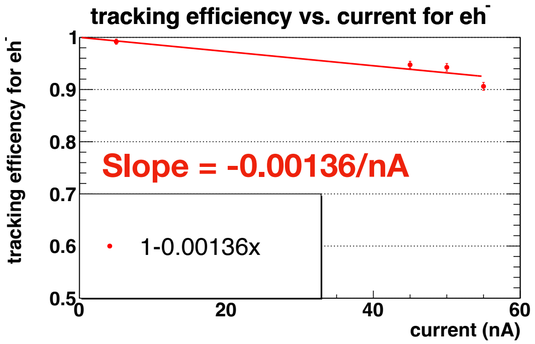
Update with new tracking

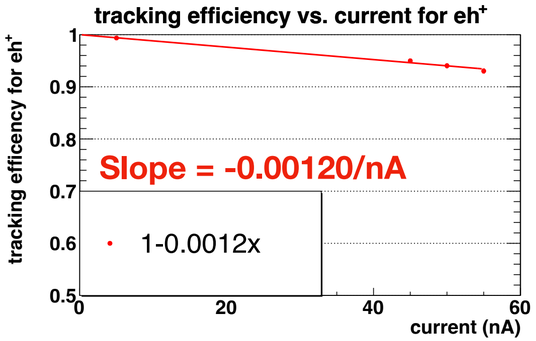
Further update with optimized DC clustering

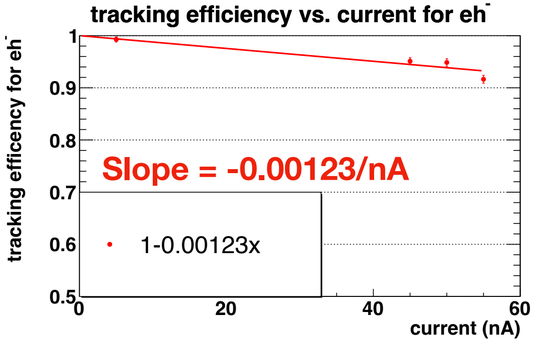












Need further improvement

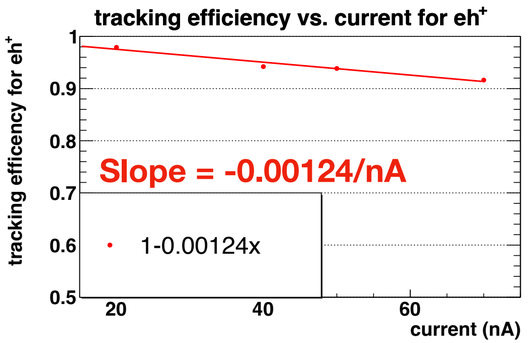
Tracking Efficiency Scan by RGA-Fall18

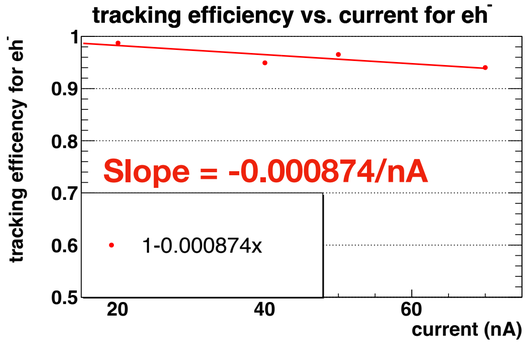
Out-bending

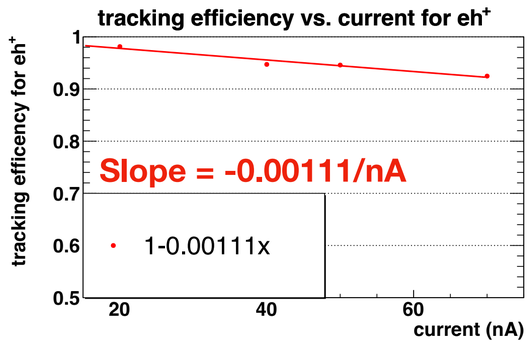
pass2 including AI-denosing and AI-assisted tracking

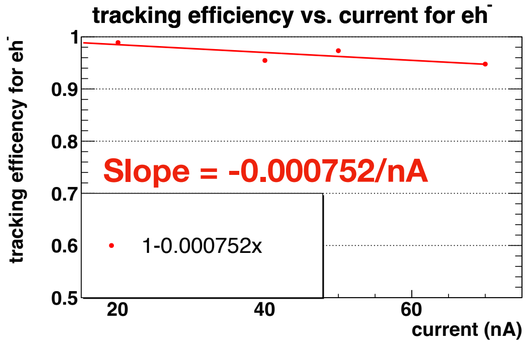
Update with new tracking

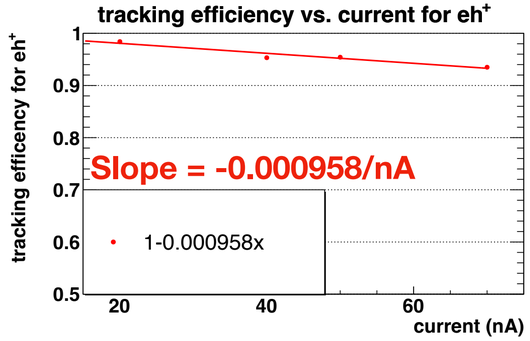
Further update with optimized DC clustering

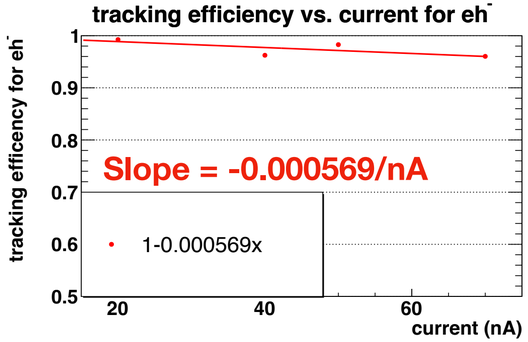












Reached the goal for 2L

Comparison between Conventional and AI-assisted Tracking

Conventional tracking

6-cluster tracking for almost all arbitrary 6 cluster combos with loose cut in the pattern recognition

↓

Make choice for cluster-overlapping track candidates based on χ^2/NDF

↓

5-cluster tracking with remaining clusters

↓

Make choice for cluster-overlapping track candidates based on χ^2/NDF

AI-assisted tracking

6-cluster-combo prediction

↓

Cut off combos with low probability and make choice for cluster-overlapping combos based on probability

↓

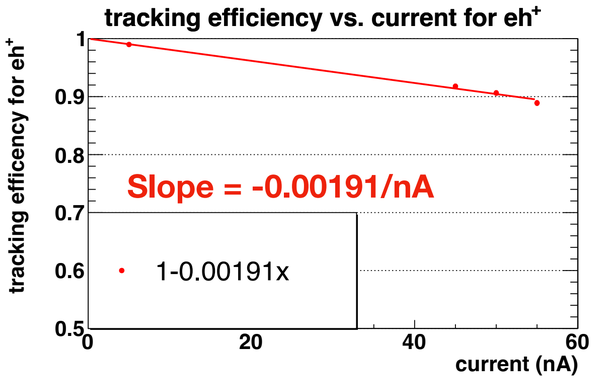
5-cluster-combo prediction with remaining clusters

↓

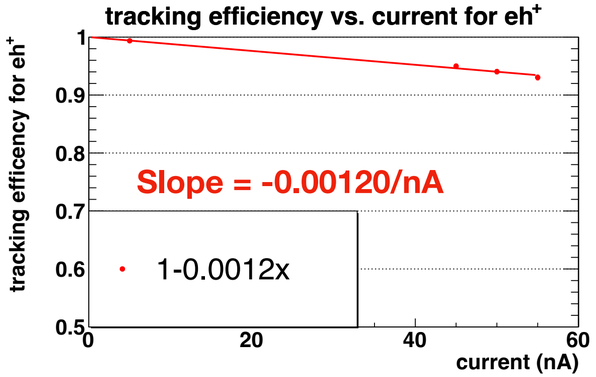
Cut off combos with low probability and make choice for cluster-overlapping combos based on probability

↓

Tracking for both 6-cluster and 5-cluster cases



In-bending / RGA-Fall18



Why AI-assisted Tracking Works Better?

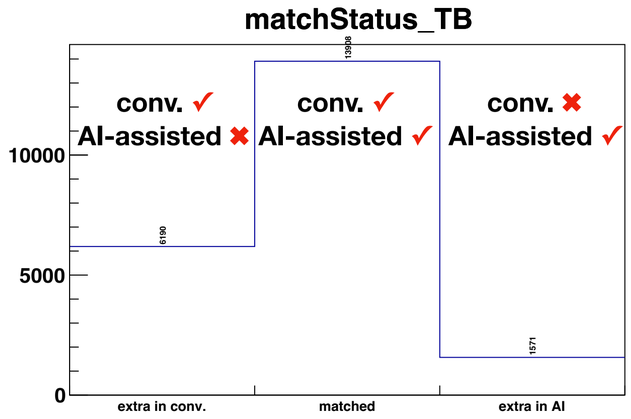
charge	type	conventional	ai	matched	predicted	gain	efficiency	inference
neg	0	1088248	1117574	1047081	1053197	1.0269	0.9622	0.9678
neg	6	1036171	1036276	1008670	1008931	1.0001	0.9735	0.9737
neg	5	51026	80394	37708	43284	1.5755	0.7390	0.8483

charge	type	conventional	ai	matched	predicted	gain	efficiency	inference
pos	0	2485735	2552343	2383956	2403644	1.0268	0.9591	0.9670
pos	6	2359951	2359283	2288125	2290857	0.9997	0.9696	0.9707
pos	5	123058	190902	94149	110224	1.5513	0.7651	0.8957

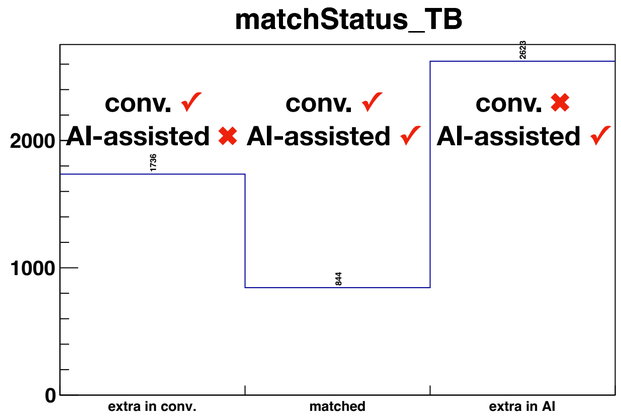
V_z is involved into routine for selection of overlapping track candidates for conventional tracking.

type	e	eh+	eh−	eh+/e	eh−/e
conventional	122537	94290	18898	0.7695	0.1542
ai	124459	97832	19797	0.7861	0.1591

6-cluster tracks

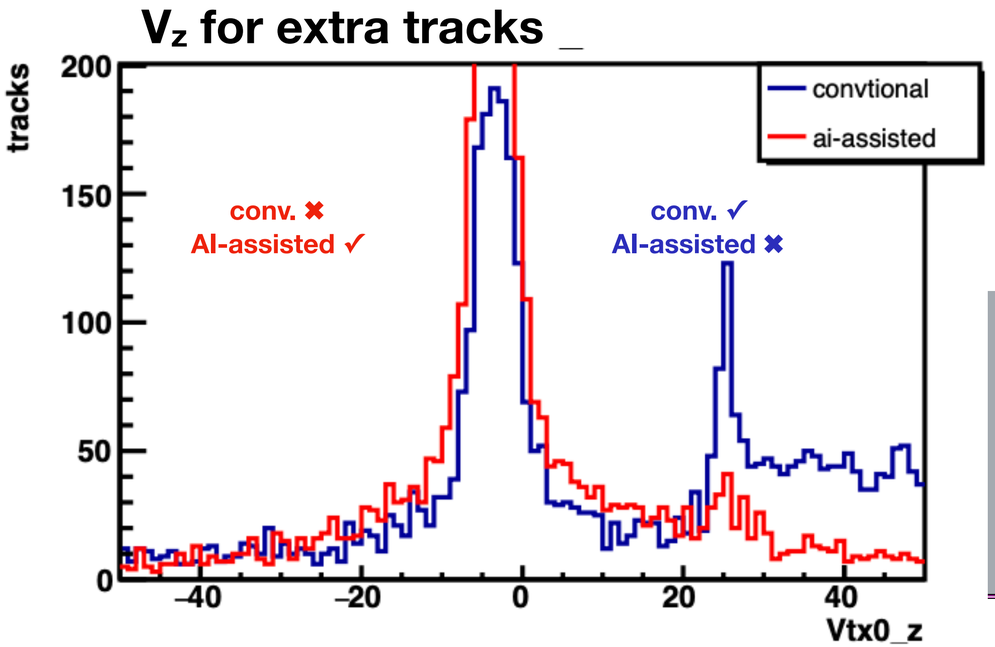


5-cluster tracks



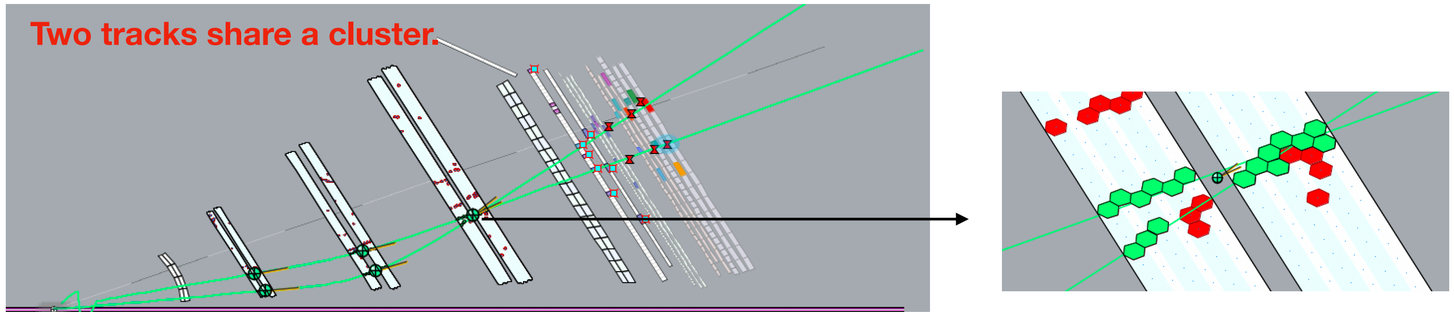
- For tracking efficiency, contribution by 6-cluster tracks are almost the same between AI-assisted and conventional tracking, but AI-assisted tracking extracts over 50% more 5-cluster available tracks.
- The reason is that AI model helps to cut off plenty of fake 6-cluster combos, while a lot of fake 6-cluster tracks must be kept for conventional tracking since no proper cuts could be applied as test.
- Much more fake 6-cluster tracks for conventional tracking cause that remaining clusters, which enter to 5-cluster tracking, are much less than AI-assisted tracking.

Track Lost and New AI Model



Two main reasons cause track lost:

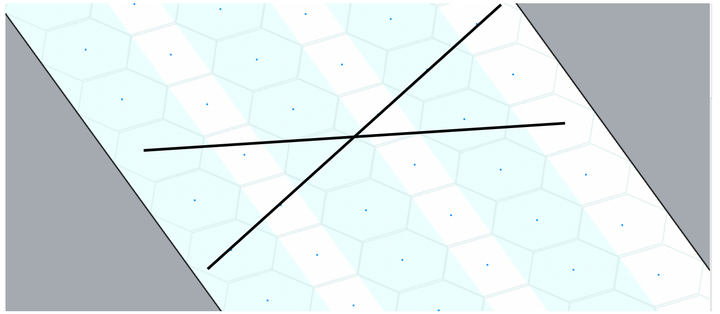
- For overlapping combos/track candidates, both AI-assisted and conventional tracking could make wrong choice.
- Tracks are not allowed to share clusters.



New AI model is in development by Gagik.

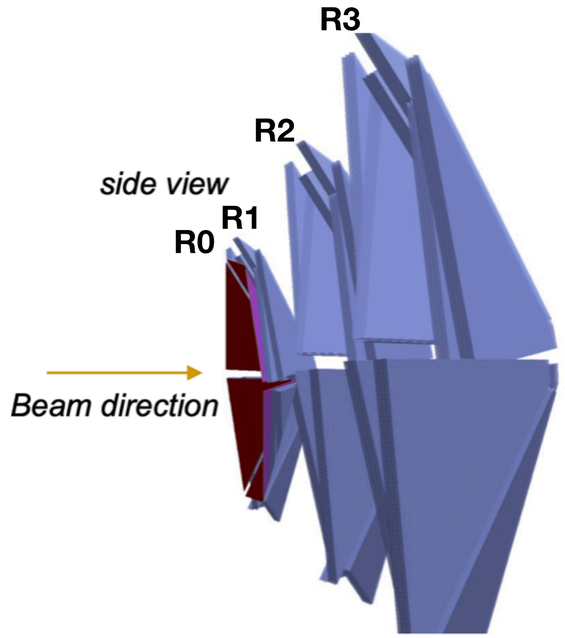
- Average wires of clusters are features of the old model, but average wires for clusters could be very close.
- New model adopts wire numbers at layers 1&6, calculated by function from linear fitting of clusters, as features.
- For overlapping combos, choice will be taken after tracking based on probability from AI and tracking results. Tracks will be allowed to share one and only one cluster.
- Still, cluster combos with low probability will be cut off for suppression of fake tracks.

Demon. for close average wires of two clusters



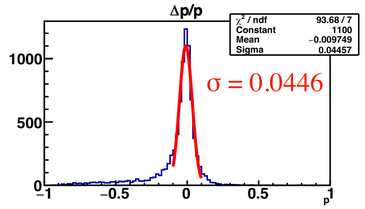
Status of DC-uRWell Tracking

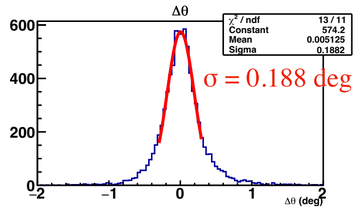
- DC-uRWell tracking package with application of KF and DAF has been developed with input of uRWell cluster measurements and DC hit measurements.
- Conventional tracking is in optimization, while AI-assisted tracking needs AI model for prediction of DC-uRWell cluster combos.
- Resolutions for both momentum and vertex are improved by test of single-electron MC.

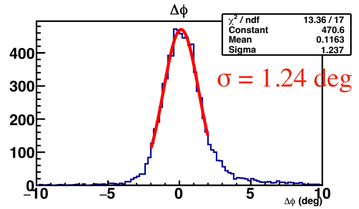


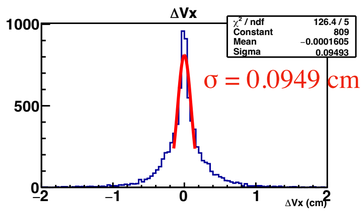
Δ = track - truth

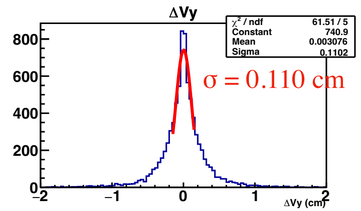
DC-only tracking

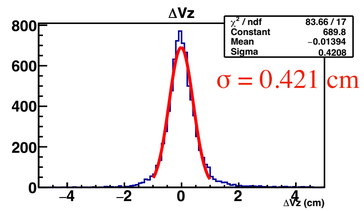




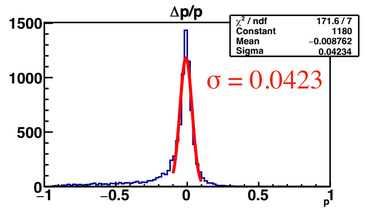


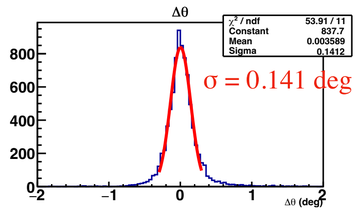


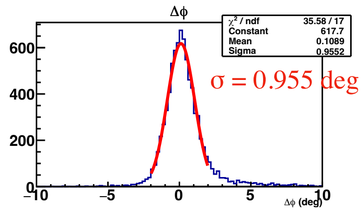


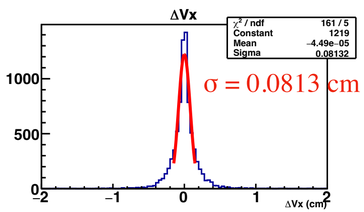


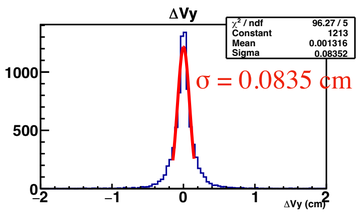
DC-uRWell tracking

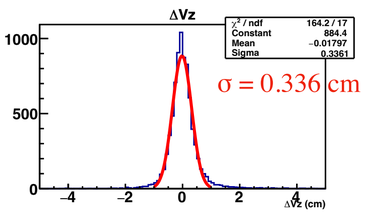












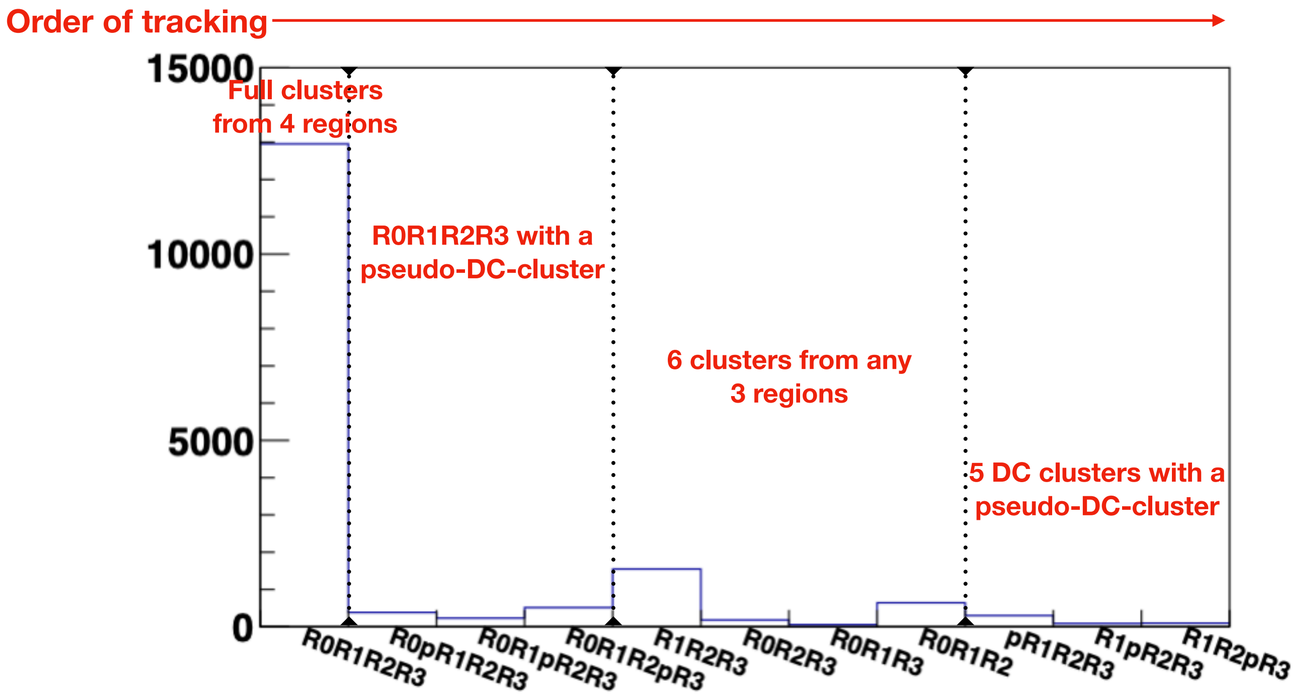
Resolution improvement

~5%↑ ~25%↑ ~23%↑ ~14%↑ ~24%↑ ~20%↑

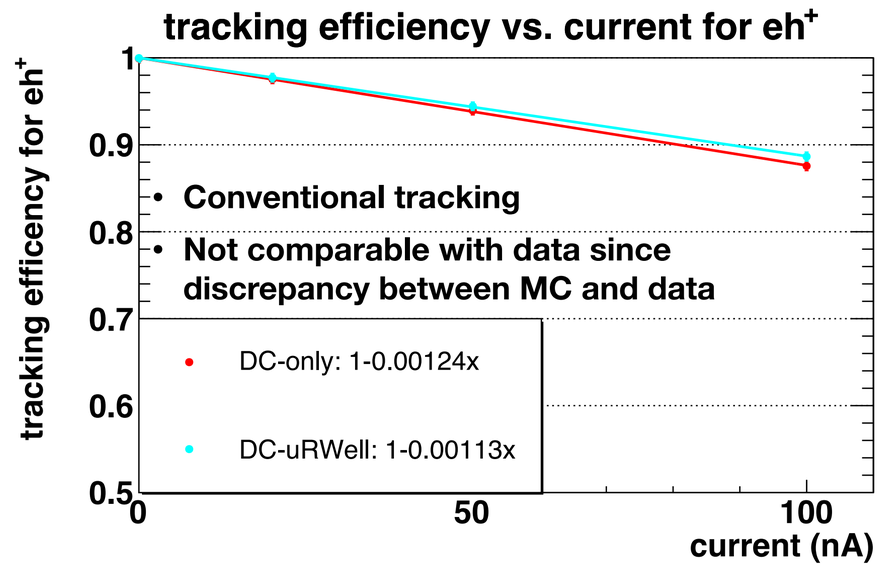
RGA-SIDIS MC

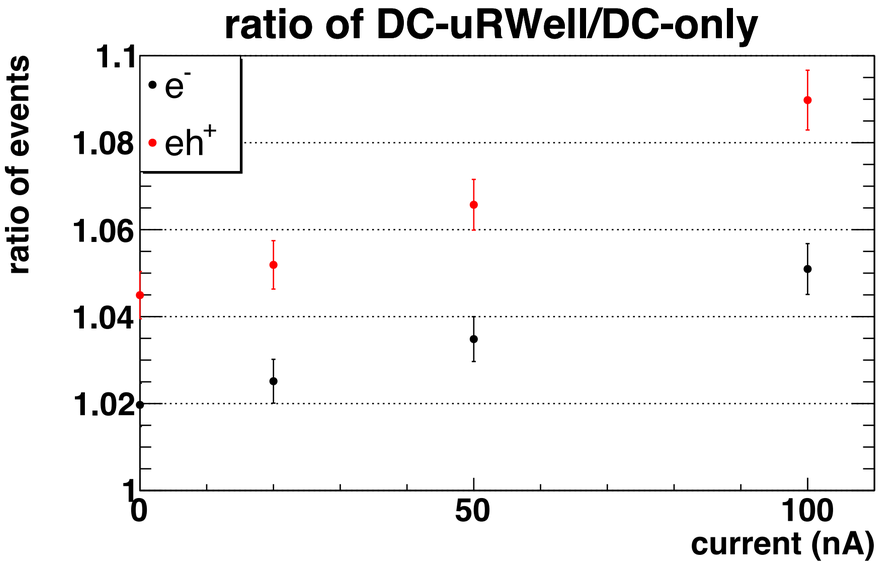
Status of DC-uRWell Tracking

- Due to addition of uRWell region (R0), tracking could be processed even if a region is lost.
- For tracking with 3 regions, currently it is only allowed that a pseudo-DC-cluster is estimated by 5 real DC clusters due to limits of algorithms for pseudo-cluster estimation and pattern recognition.
- Tracking is processed in order.



- Tracking efficiency is improved with addition of uRWell. Slope difference should be bigger once DC hit occupancy and other aspects for MC match better with data.

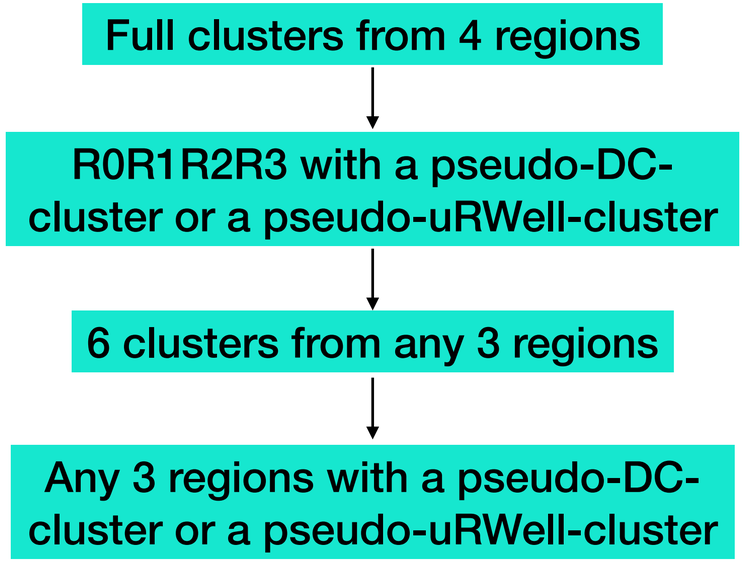




Plans for DC-uRWell Tracking

- Algorithm for estimation of pseudo-cluster is in update so that a pseudo-uRWell-cluster or a pseudo-DC-cluster could be estimated for any 3 region cases.
- Correspondingly, algorithm for the pattern recognition by at least 3 crosses will be updated so that it could work for any 3 region cases with a pseudo-uRWell-cluster or a pseudo-DC-cluster.
- Like DC-only tracking, remaining clusters from previous stage enter next stage. To properly remove fake tracks in each stage, AI model for prediction of uRWell-DC cluster combos for each stage is necessary.
- To further relief effects of high hit occupancy in R1 on tracking efficiency, uRWell measurements will join DC clustering in R1.
- Besides, it will be discussed if uRWell measurements could join AI-denoising to better suppress noise hits and reduce signal-hit lost, especially for R1.

Order in next tracking version



Summary

- With AI-denoising, AI-assistance for cluster-combo prediction, new tracking with KF and DAF, and updates of DC clustering, tracking efficiency has been significantly improved. As test by RGA-Fall18 data, slope of tracking efficiency is -0.00120/nA for eh^+ and -0.00123/nA for eh^- for in-bending, and -0.000958/nA for eh^+ and -0.000569/nA for eh^- for out-bending.

- However, tracking efficiency needs to be further improved to reach the goal for future 2L experiments: slope >= -0.001/nA.

- For DC-only tracking, new AI model is in development by Gagik, and a new routine will be explored for selection of overlapping track candidates based on probability from AI prediction and tracking results.

- Improvement for resolution and tracking efficiency has been observed with uRWell adding into tracking. To further relief effects of missing clusters and high hit occupancy in R1, some algorithms for DC reconstruction are in update with participation of uRWell measurements.

- AI model for prediction of uRWell-DC cluster combos is necessary for selection of overlapping track candidates and fake-track suppression at each stage of DC-uRWell tracking.

- AI-denoising with participation of uRWell will be discussed.

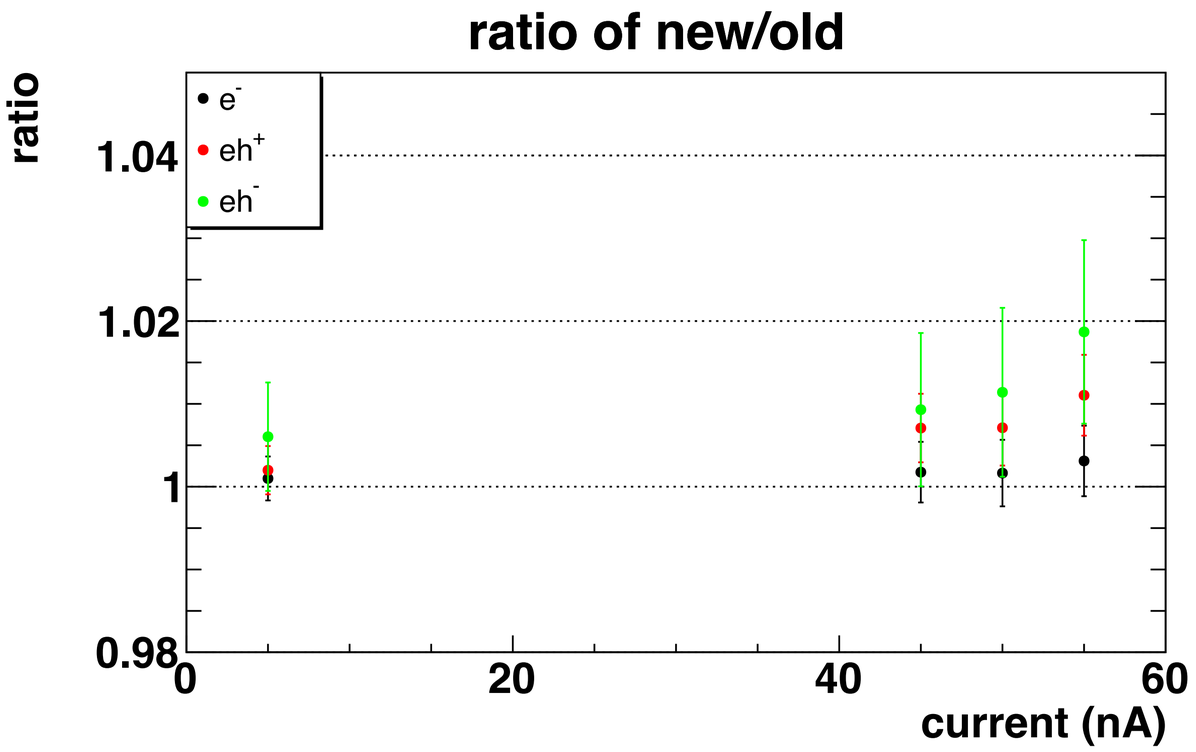
Backup

RGA-Fall18 AI-denoising AI-assisted tracking

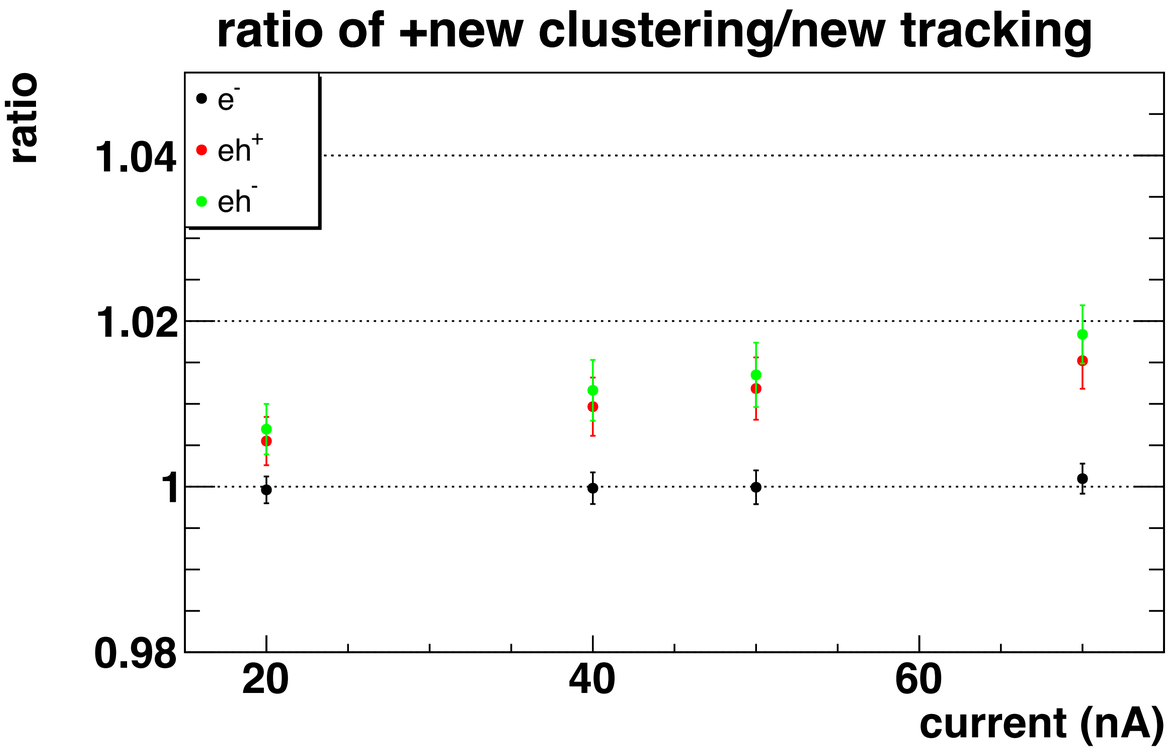
Ratio of Event Reconstruction

Ratio = coatjava v10.1.0 with new tracking / update-to-date coatjava with new clustering

in-bending



out-bending

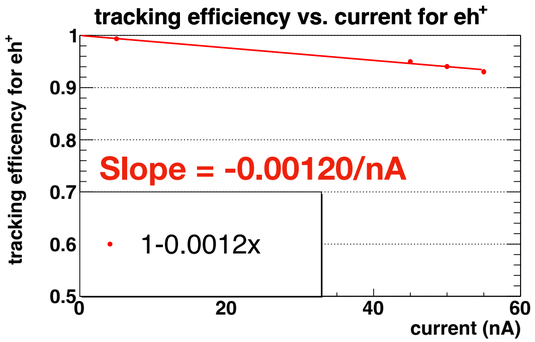


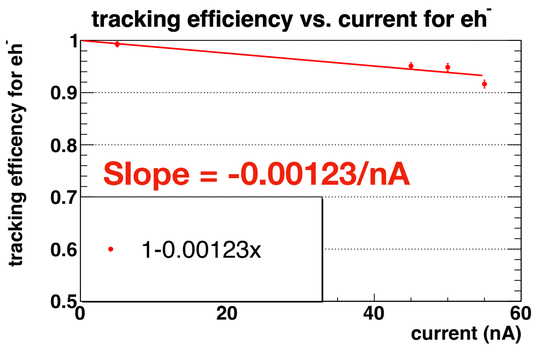
AI-assisted Tracking

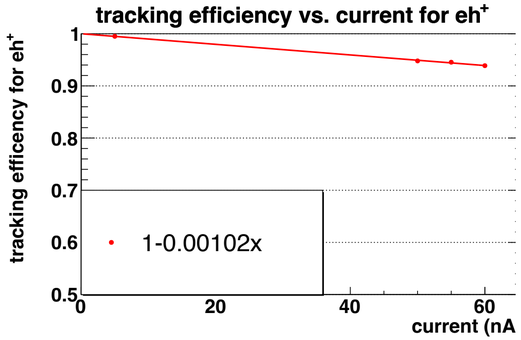
In-bending; AI-denoising

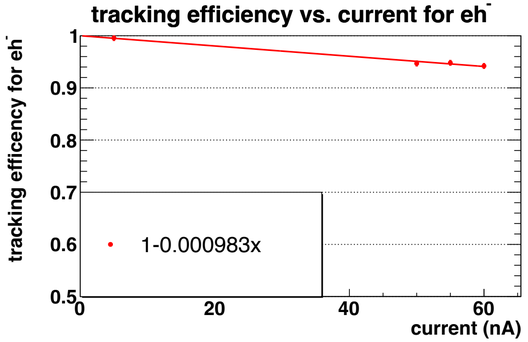
Comparison of Tracking Efficiency between Data and MC

RGA-Fall208 Data





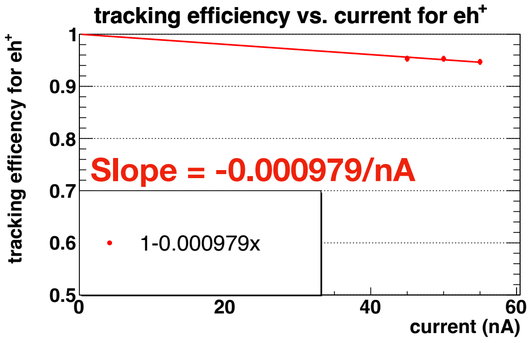


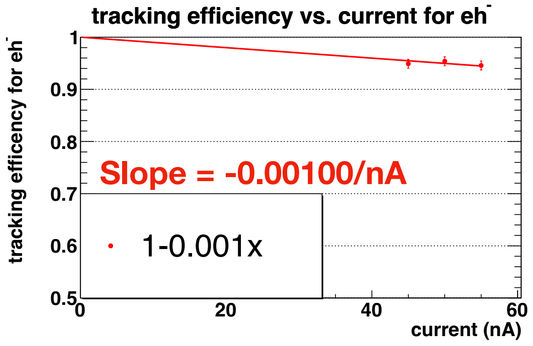


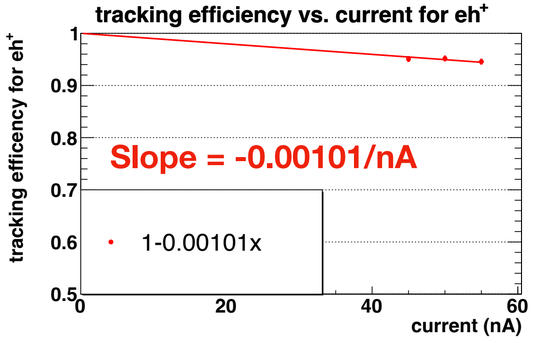
All-detector merging

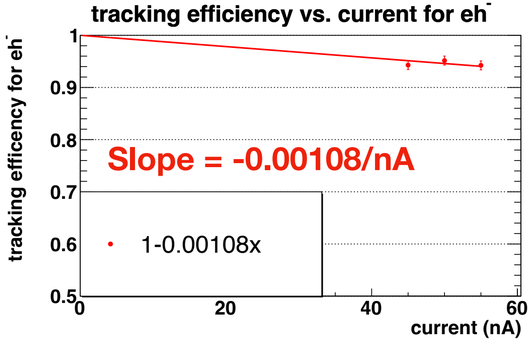
RGA-SIDIS MC + Bg from Data

DC-only merging



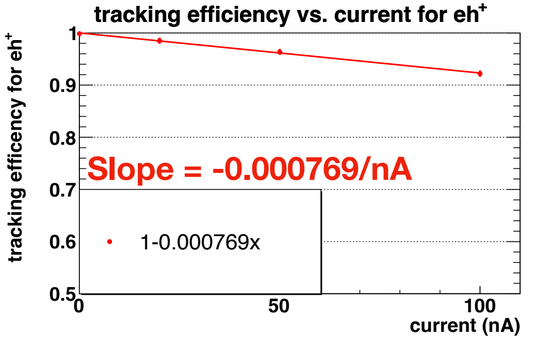


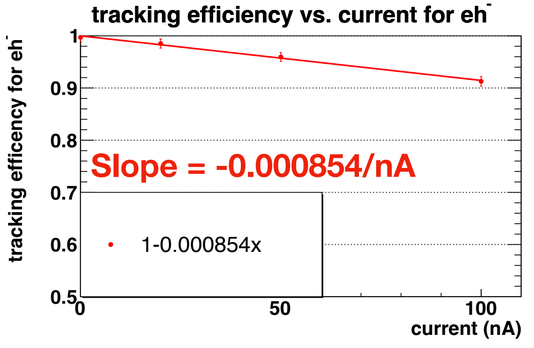


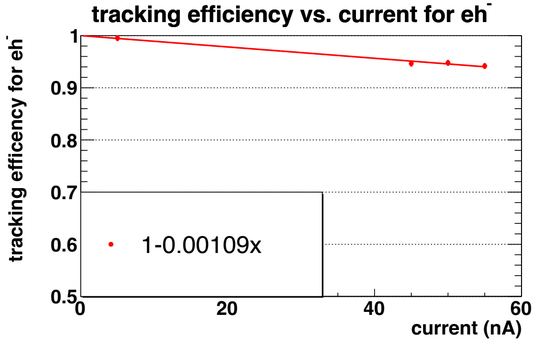


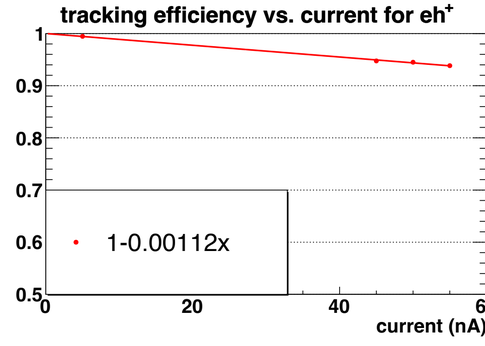
RGA-SIDIS MC + Bg MC

DC-only merging







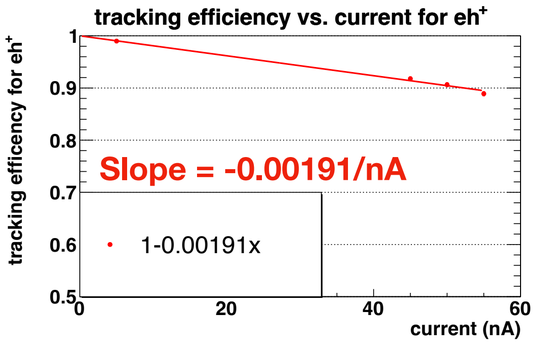


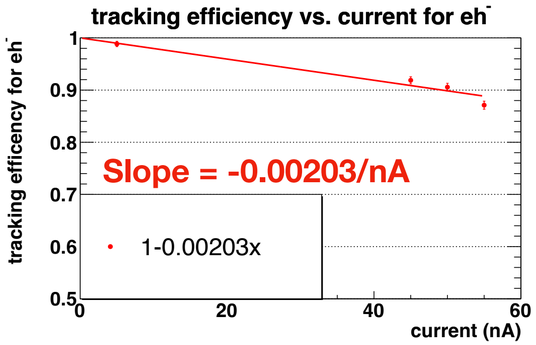
Conventional Tracking

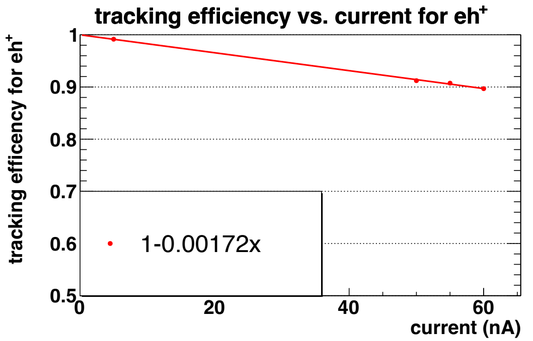
In-bending; AI-denoising

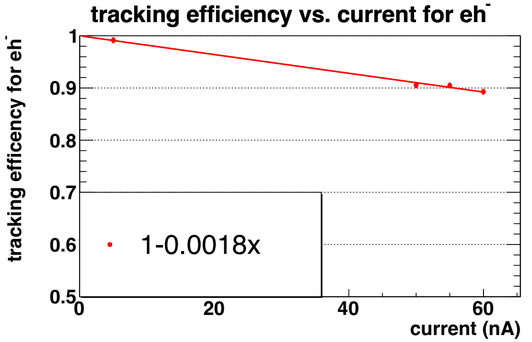
Comparison of Tracking Efficiency between Data and MC

RGA-Fall208 Data





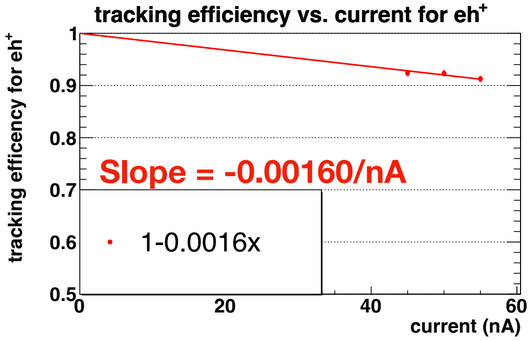


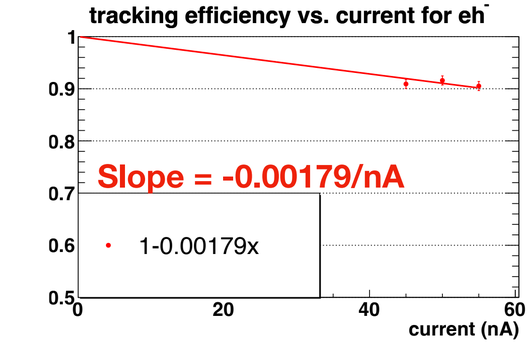


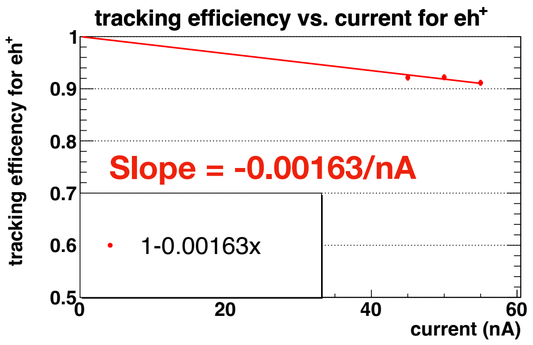
All-detector merging

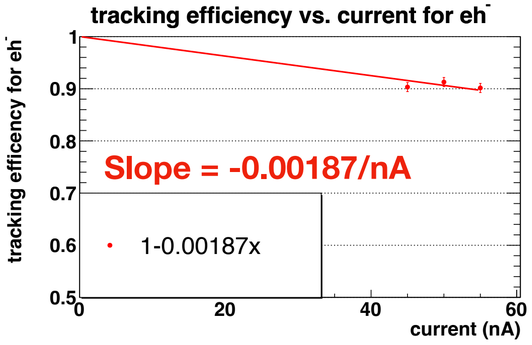
RGA-SIDIS MC + Bg from Data

DC-only merging



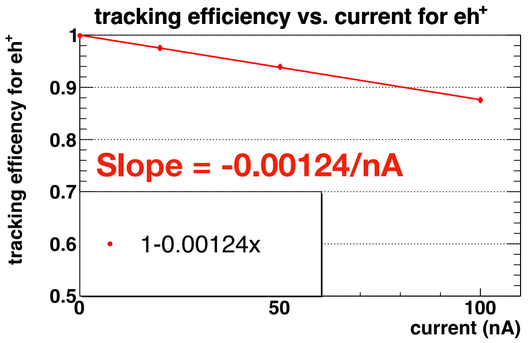


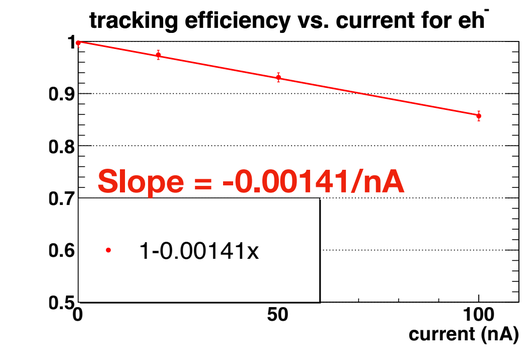


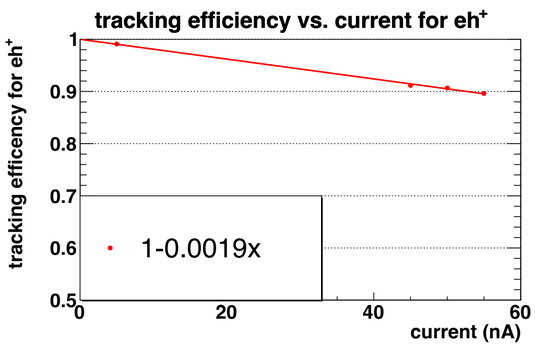


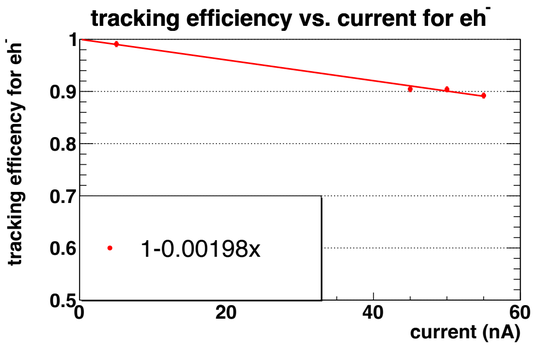
RGA-SIDIS MC + Bg MC

DC-only merging





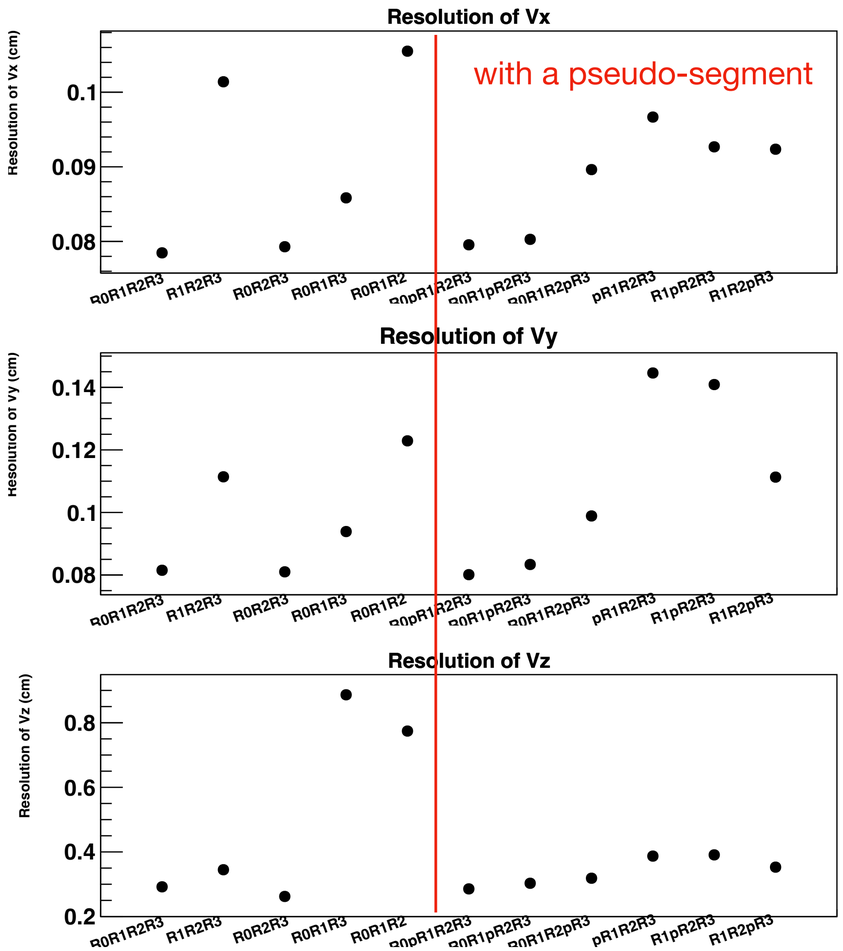


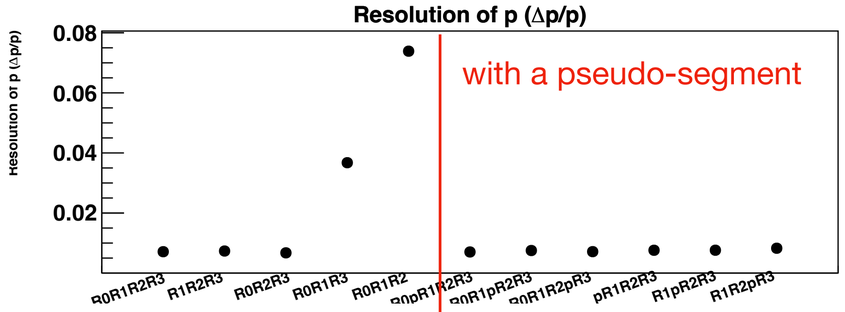


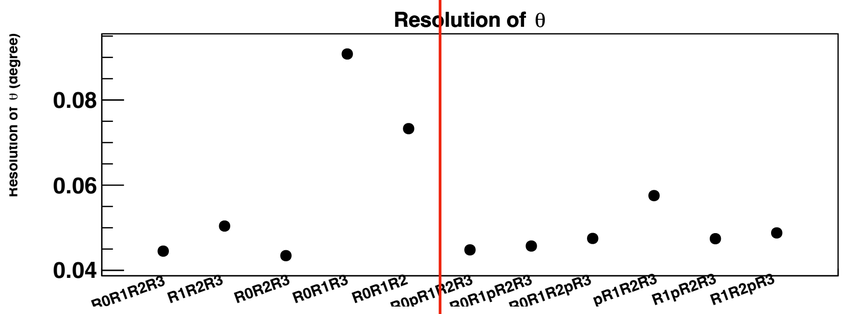
TB tracks

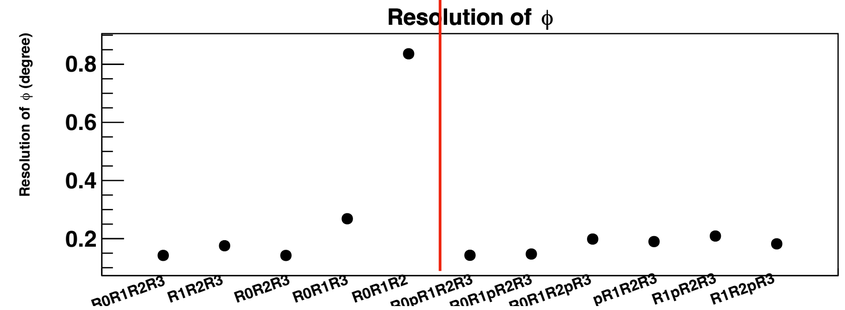
Tests

Comparison of resolution









- Overall, cases of R0R1R2R3, R1R2R3 and R0R2R3 works well. For cases of R0R1R3 and R0R1R2, tracking quality is worse.
- In other words, tracking quality is still good without R0 or R1, while it is not so good without R2 or R3.