Alma 9 Validation

Dark Photon Samples

February 6, 2025

Pawan Johnson

Prof. Carl Gwilliam

Dr. Monika Weilers

Prof. Monica D'Onofrio

University of Liverpool

Introduction

- Validate ALMA9 version of Calypso for the track variables
- Sinead already looked at single muon samples
- Ansh looked at the A' analysis cutflow
- We want to look at the "two track reconstruction" as a function of separation between them

Data Description

- We will to look at Dark-Photon decays to electron pairs
- Data samples used are
  /eos/experiment/faser/data0/sim/mc24/foresee/1100{33,38,51}/
  - 110033 : Mass = 10 MeV, epsilon = 1E-5
  - 110038 : Mass = 100 MeV, epsilon = 1E-5
  - 110051 : Mass = 10 MeV, epsilon = 1E-4
- ALMA 9 samples : ./phy/s0008-dev/
- CENTOS 7 samples: ./phy/s0008-r0019/
- Chaining them together for better statistics [total 60k events]
- Can separate based on mass/couplings if interested

Overview of Validation

Objective: To quantify the Efficiency of two track reconstruction as a function of separation between tracks

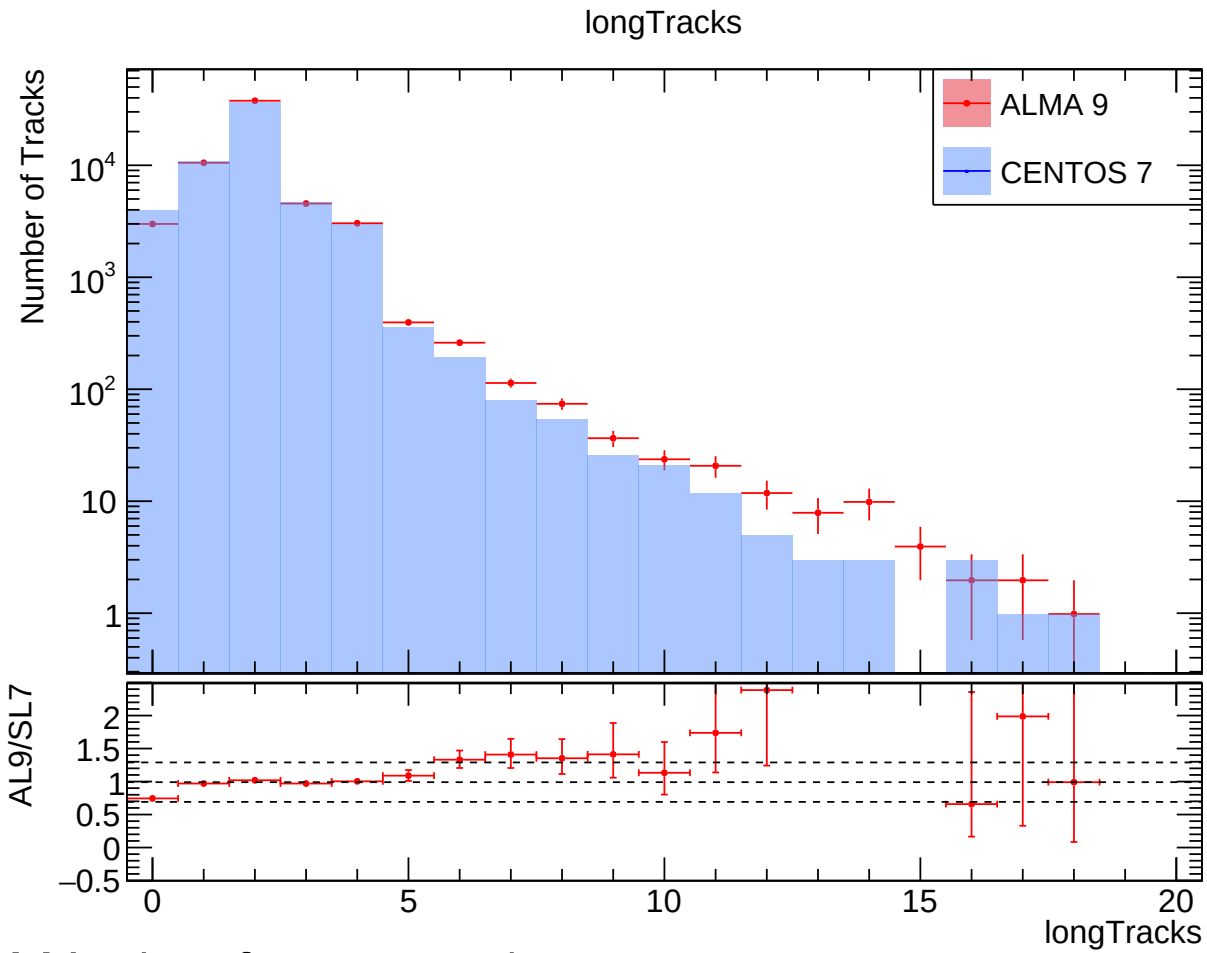
- Perform an initial assessment of the Track Parameters
- Quantify the separation between tracks
- Evaluate generic track reconstruction performance
- Define a metric for "Reconstruction Efficiency"

Distribution of Track Parameters

Basic Plot Parameters to assess

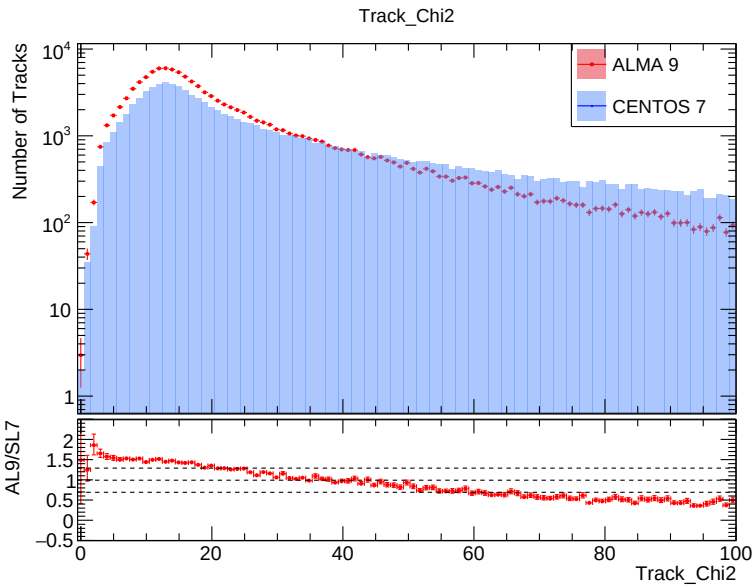
- longTracks
- Track Chi2
- Track Chi2perDoF
- Track nDoF
- Track charge
- Track nLayers

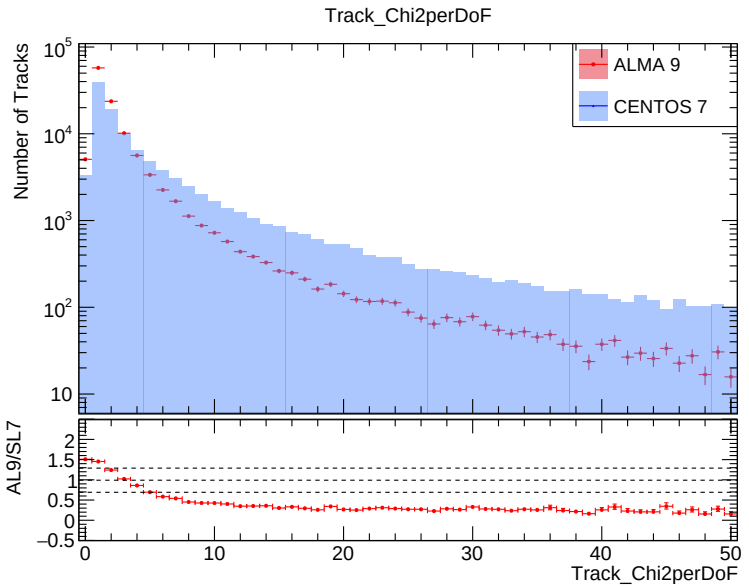
Distribution of longTracks



- ALMA9 has fewer 0-track events;
- Also reconstructs more events with more than 5 tracks.
- Total number of longTracks: CENTOS7: 115,206, AL9: 118,491
- A 2.8% increase in longTracks in AL9.

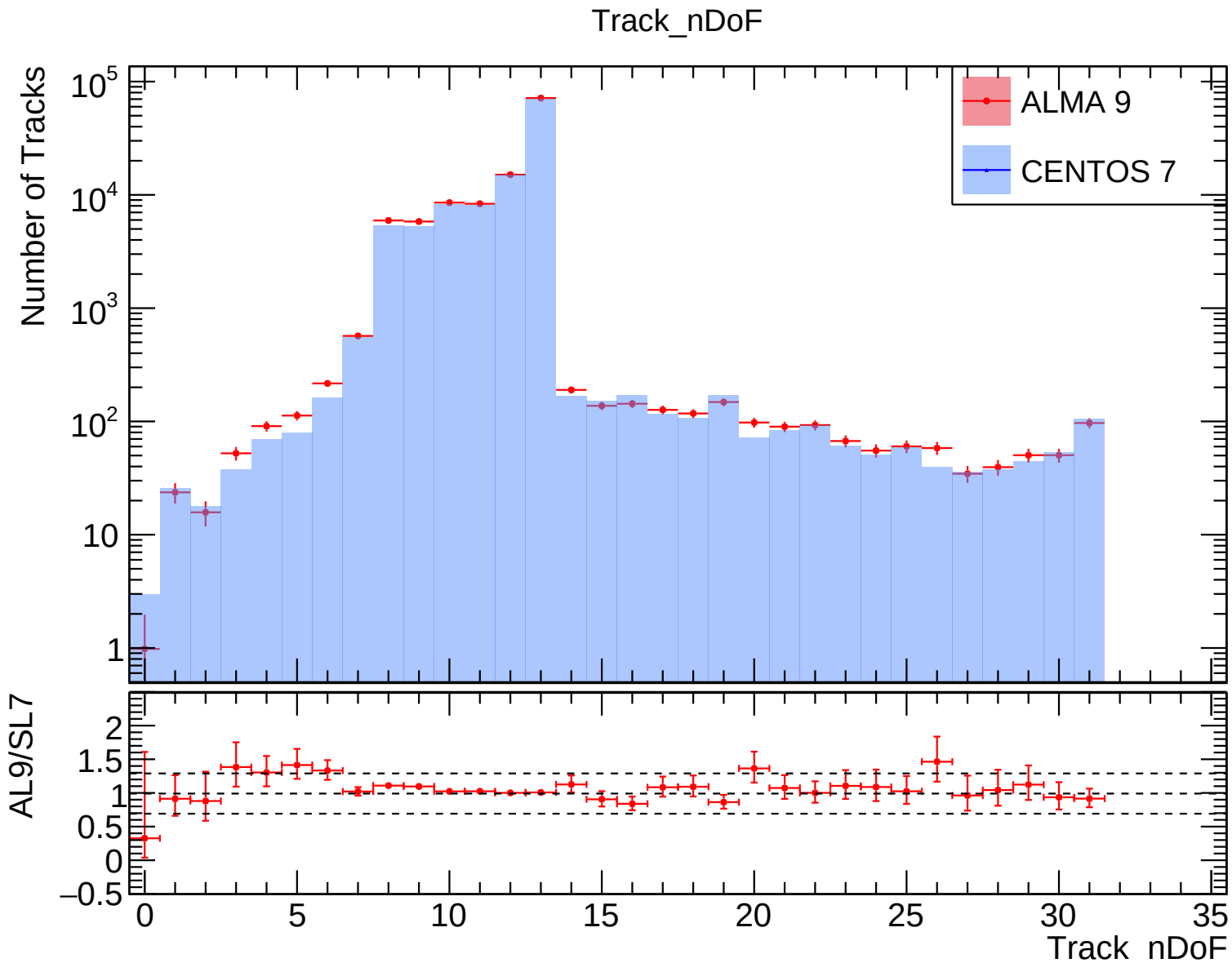
Distribution of TrackChi2





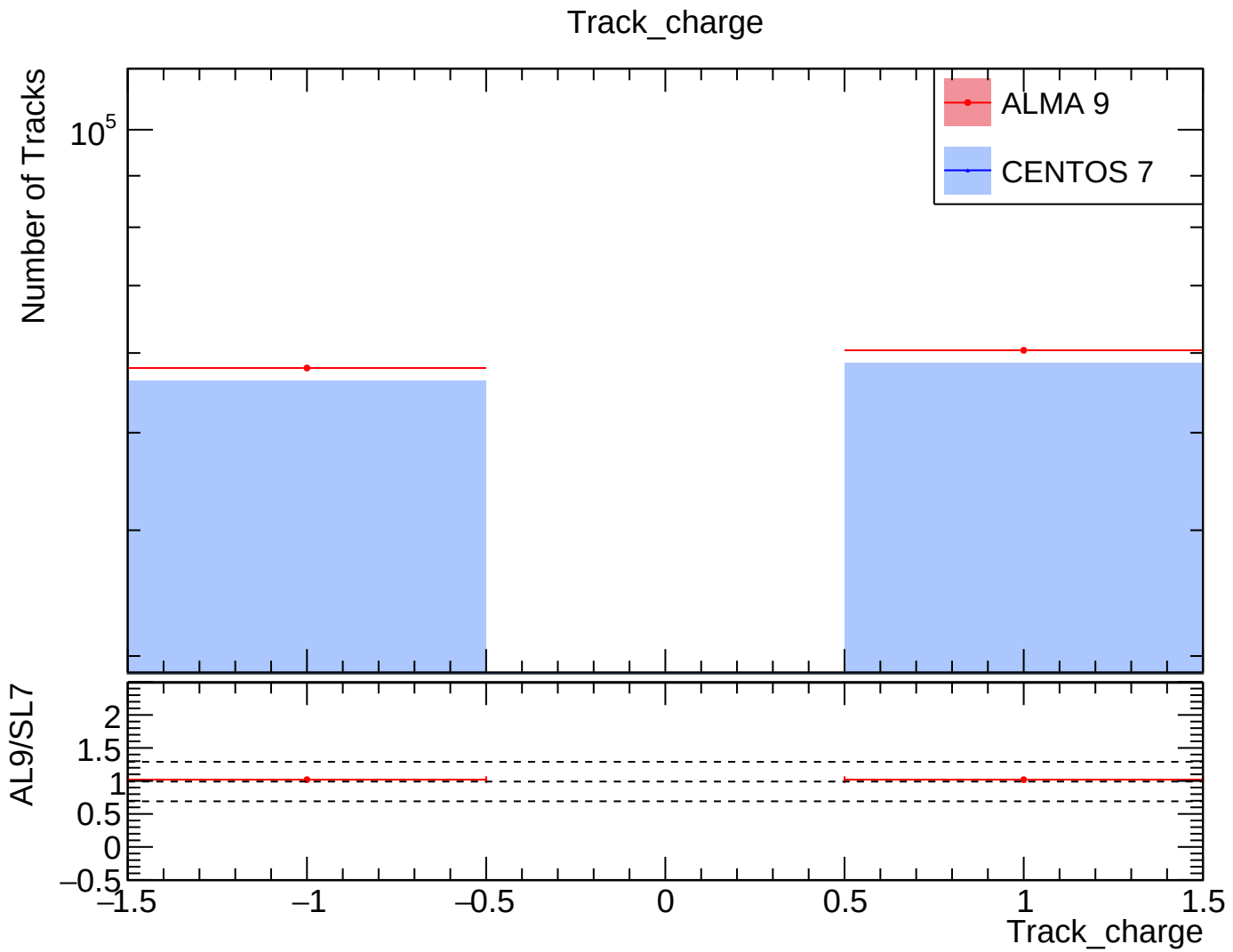
- Displays the greatest improvement in ALMA9.
- Overall Tracks have lower Chi2/DoF in ALMA9.

Distribution of TrackNDoF



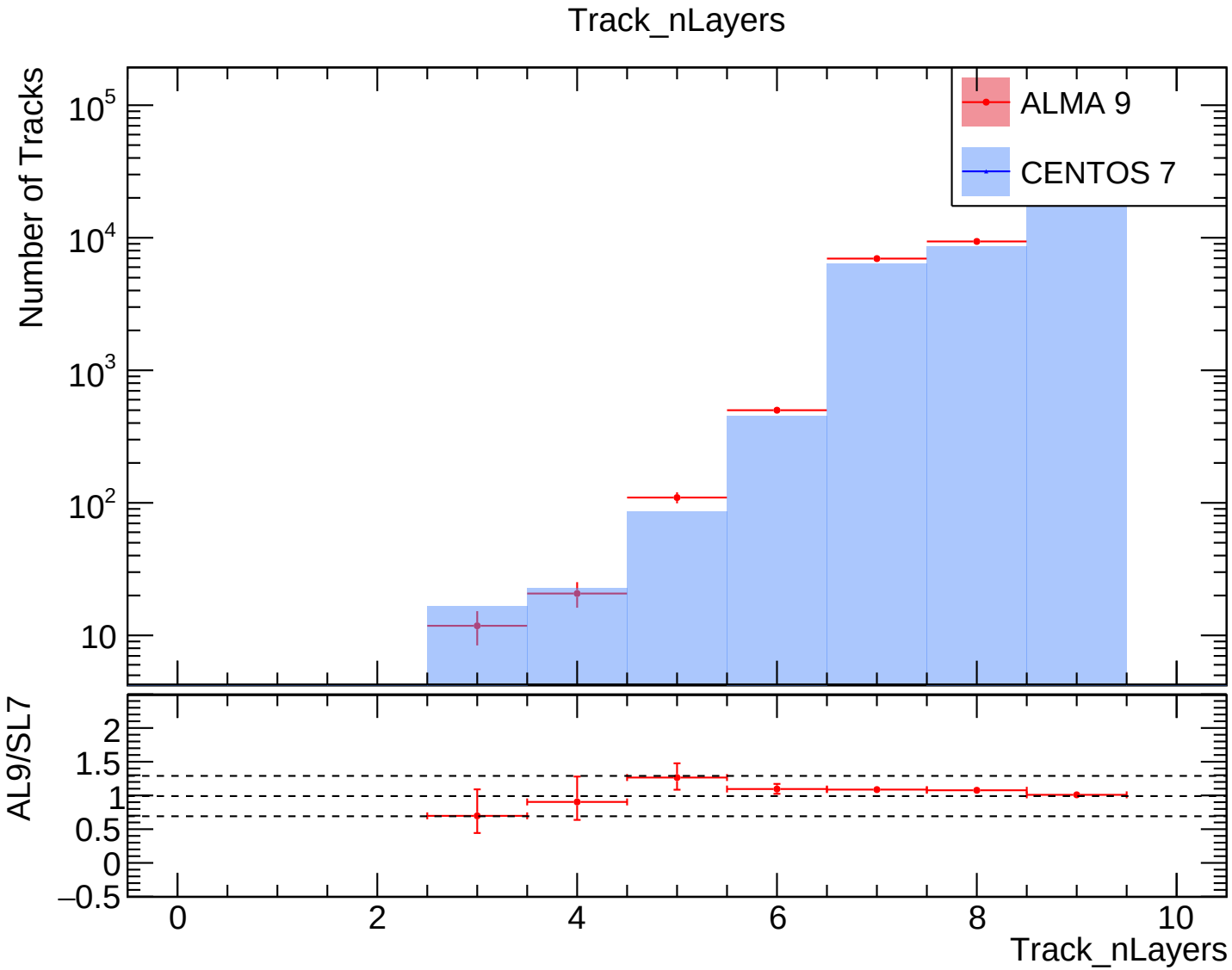
- Generally strong agreement, except for bins 3-5.

Distribution of Track Charge



- The ratio is 1.028 [same factor of increase seen in longTracks]
- Ratio of positive to negative tracks is 1.04 in both. ChargeMISID?

Distribution of Track nLayers



- Similar agreement within statistical uncertainties.

Quantifying Separation

Possible Track Separation Variables

- $\Delta R_{0,1}$: Separation between the electron and positron at the first/last tracking station in the x-y plane
- $\theta_{0,1}$: Angle between the line connection decay vertex to the two tracks at the first/last tracking station
- $\Delta X_{0,1}$: Same as above but only in x direction
- $\Delta Y_{0,1}$: Same as above but only in y direction
- $\Delta R_P = \sqrt{\Delta\eta^2 + \Delta\phi^2}$: Momentum space separation between electron and positron

Notes:

- Particle predominantly separated in the y-direction due to magnetic field
- DeltaX looks symmetric but separation here is much lower.[DeltaR $\approx$ DeltaR]
- DeltaRP has a relatively "flat" distribution and is not a good separation to quantify reconstruction
- We shall use DeltaR and Theta as the primary separation variables

Track Separation Calculations

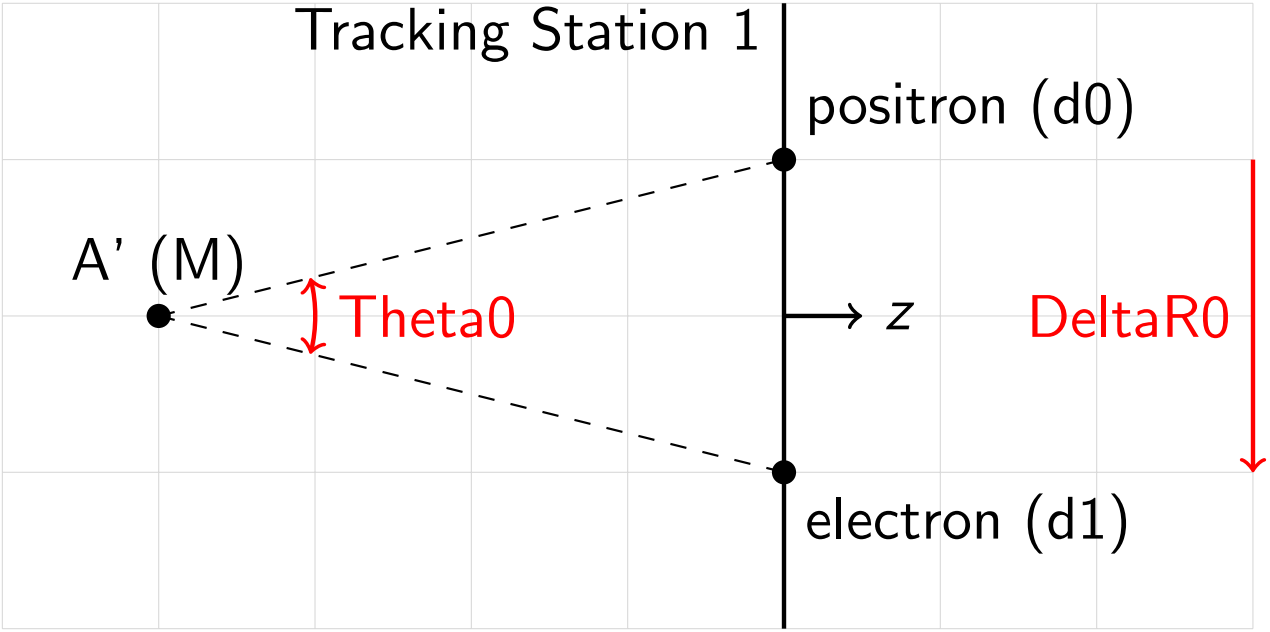


Figure: Angle and separation between the particles defined

Calculation of Separation Variables

- Separation variables are calculated using MC-truth data.
- This ensures consistency across both ALMA9 and CENTOS7.
- The separation varaiabes used are :
  - truthd0_x, truthd0_y, truthd1_x, truthd1_y, truthd0_z, truthd1_z
  - They are vectors containing the "truth positions" of d0 ($e+$) and d1 ($e-$):
  - at the vertex, the first, second, and third tracking stations.

Distribution of DeltaR

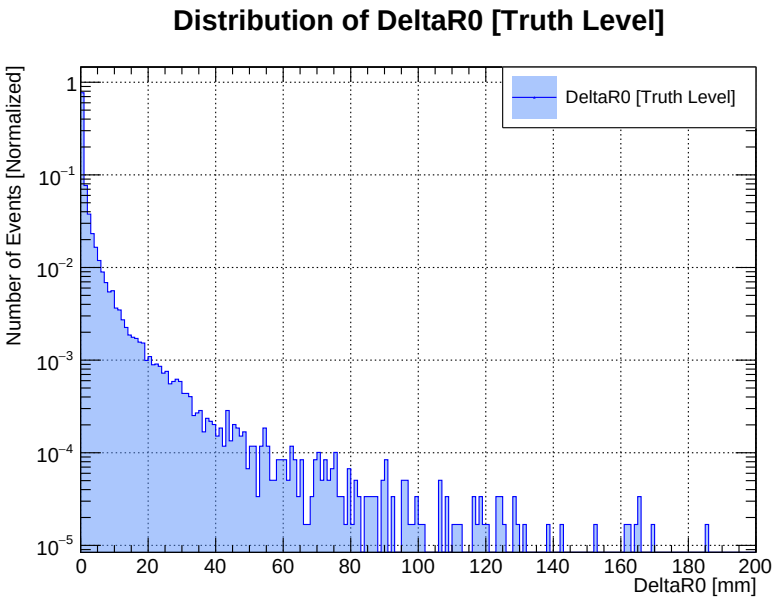


Figure: Distribution of DeltaR0 [DeltaR at Station 1]

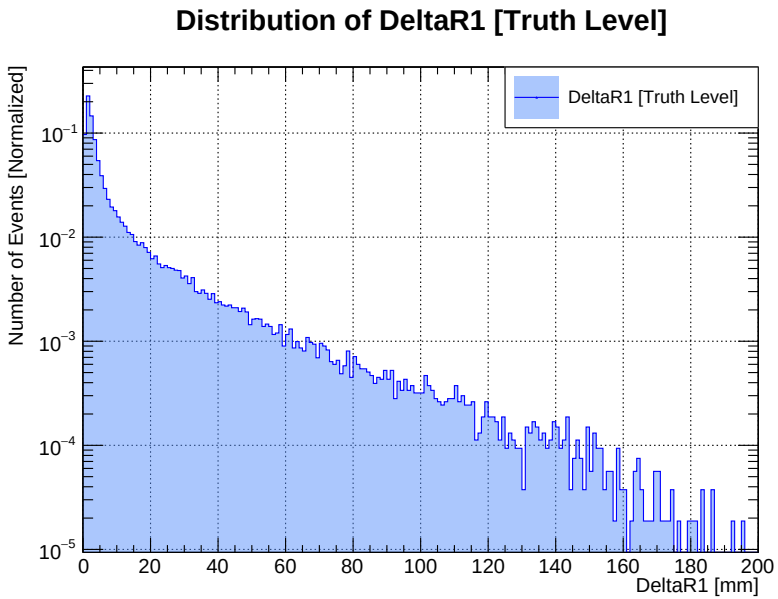
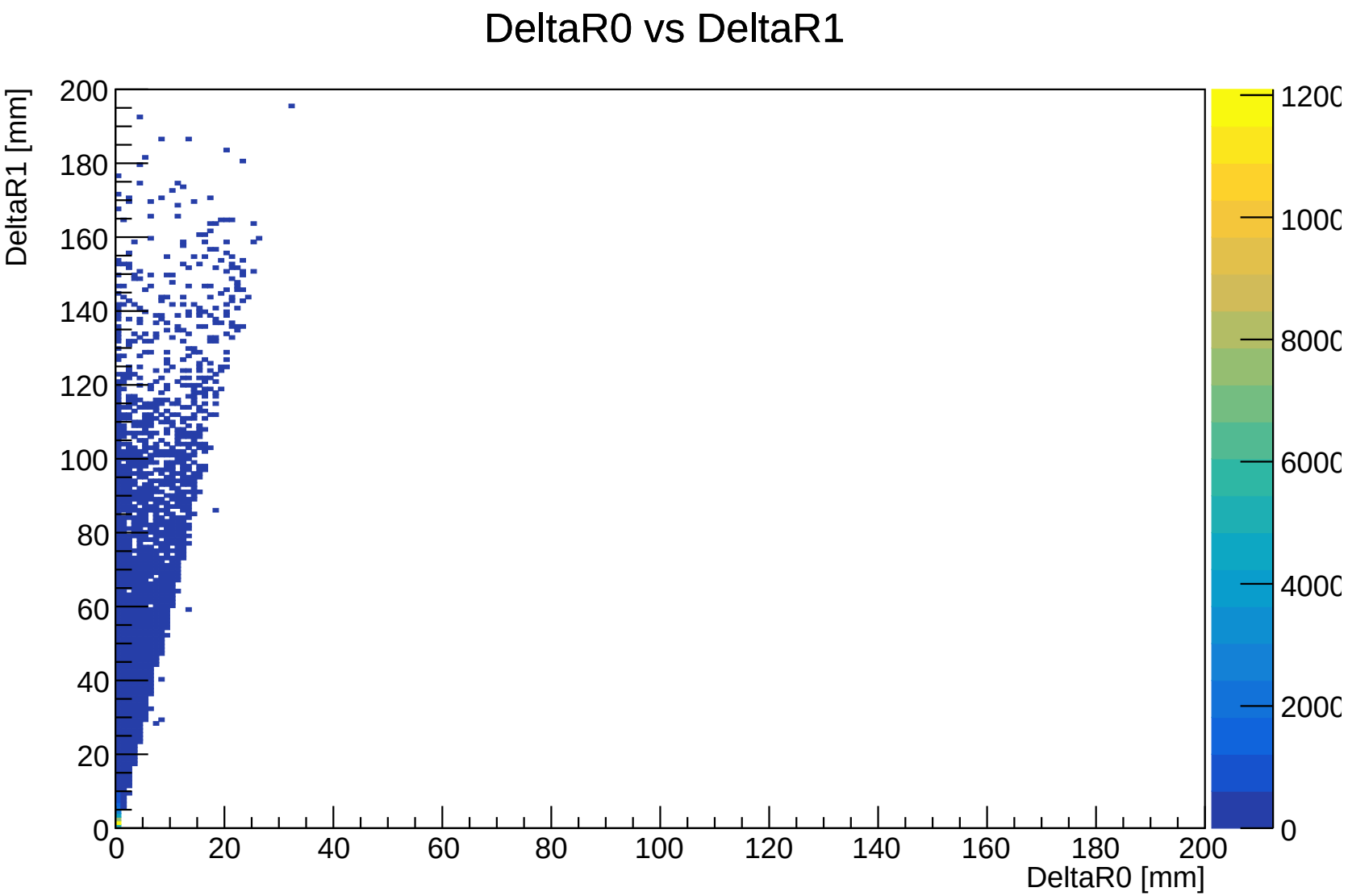


Figure: Distribution of DeltaR1 [DeltaR at Station 3]

- NEvents decay with increasing separation as expected
- Separations increase at the last tracking station due to the magnetic field
- Large separations at Station1 [>30 mm] aren't reconstructed at Station3

Transfer Plot between DeltaR0 and DeltaR1



Distribution of Theta

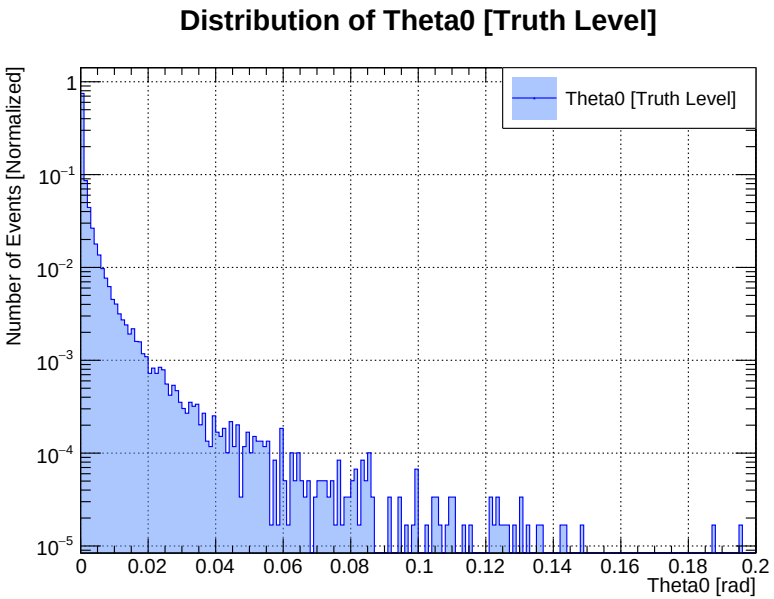


Figure: Distribution of Theta0 [Theta at Station 1]

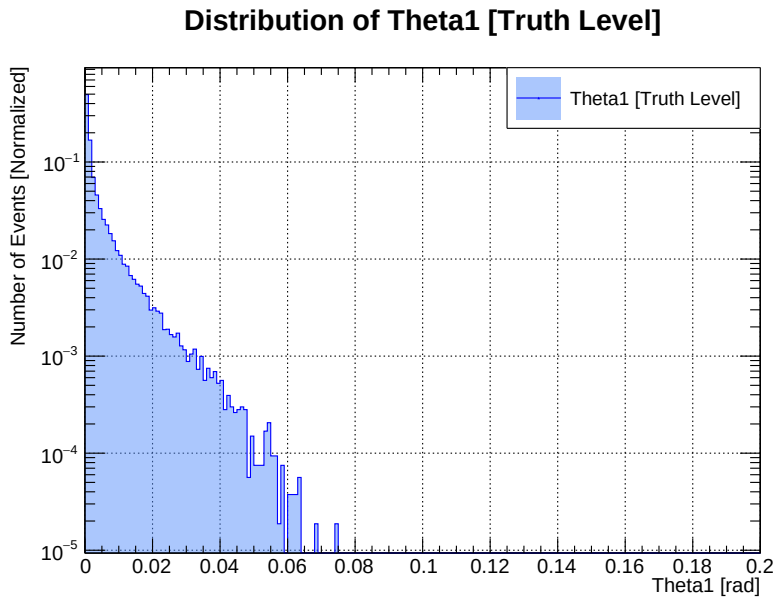
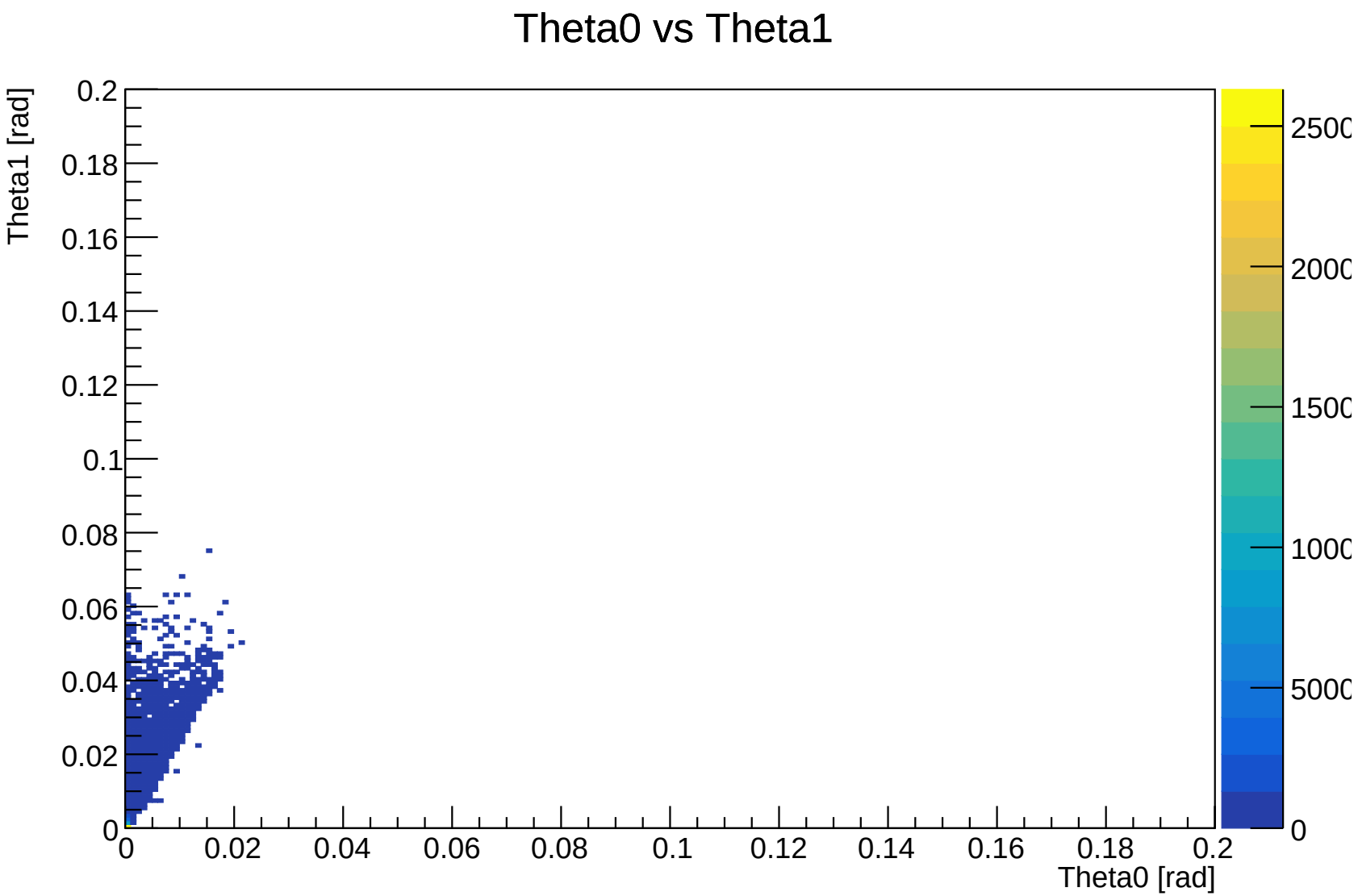


Figure: Distribution of Theta1 [Theta at Station 3]

- As previously seen, separations increases at the last tracking station.
- Large separations at Station1 [$>$0.03 rad] aren't reconstructed at Station3
- Minimum angle for which a track can still go through the tracking spectrometer 0.083 rad [arctan $\frac{0.2m}{2.4m}$, without accounting for the magnetic field]

Transfer Plot between Theta0 and Theta1



Previous Efficiency Studies : Overlay

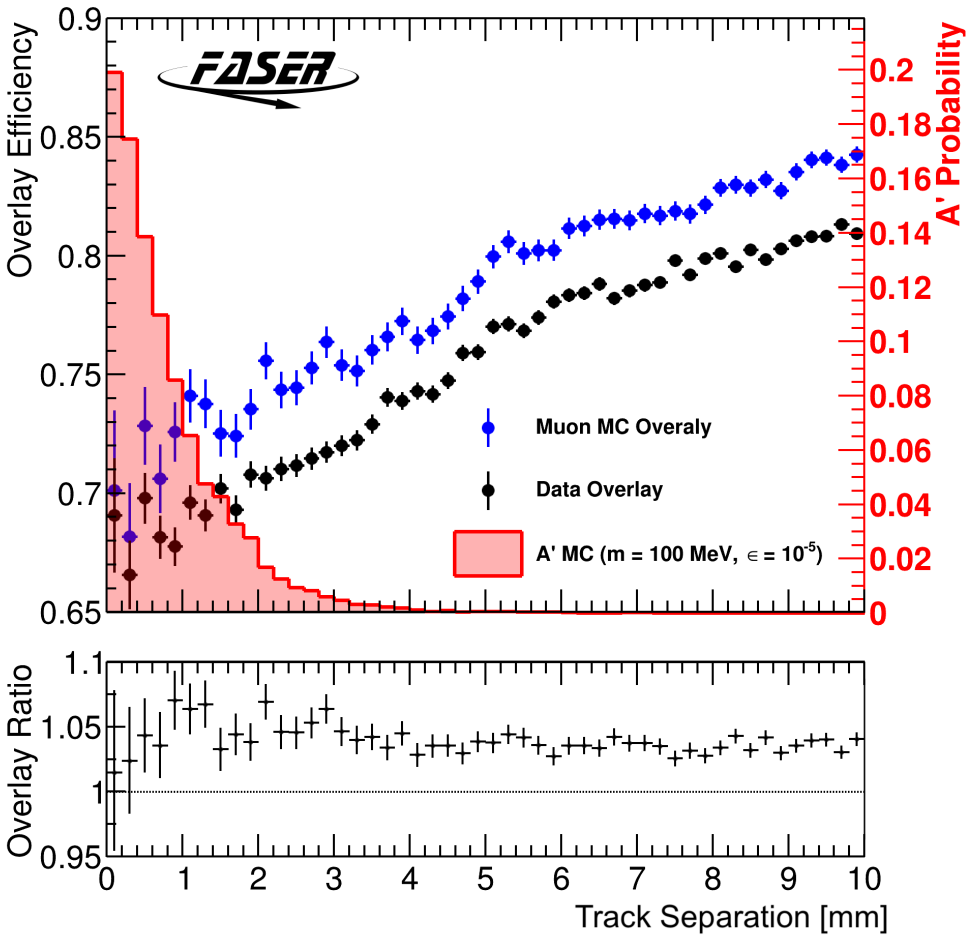


Figure: Overlay plot from Search for dark photons with the FASER detector at the LHC
The separation variable used is the "distance between the two tracks at their first measurements"

Previous Efficiency Studies : A' Tracking CutFlow

Selection	ALMA9				CENTOS7				ΔEff.
	Pass	All	Eff.	Cum. Eff.	Pass	All	Eff.	Cum. Eff.	
≥1 LongTracks	56989	60000	94.98	94.98	56002	60000	93.34	93.34	1.64
≥2 LongTracks	46416	56989	81.45	77.36	45210	56002	80.73	75.35	0.72
=2 LongTracks	37807	46416	81.45	63.01	36746	45210	81.28	61.24	0.17
Opposite Charge	32427	37807	85.77	54.04	30375	36746	82.66	50.62	**3.11**
MaxRadius < 100	31489	32427	97.11	52.48	29520	30375	97.19	49.20	-0.08
goodTrack Cuts									
≥ 7 Layers	31435	31489	99.83	52.39	29472	29520	99.84	49.12	-0.01
χ^2/DoF < 25	31121	31435	99.00	51.87	27710	29472	94.02	46.18	**4.98**
≥ 7 DoF	31115	31121	99.98	51.86	27706	27710	99.99	46.18	-0.01

Table: Comparison of efficiency and cumulative efficiency for ALMA9 and CENTOS7.
Note: The Cutflow is at an Event Level (not track level), thus the conditions have to met by all tracks in the event.

- Highest improvement in goodTrack Cut of $\chi^2/\text{DoF} < 25$
- Better ChargeID in ALMA9?
- We want to take a look at it differentially

Definition of Efficiency

General Idea

- One Track $\Longrightarrow$ NOT reconstructed
- Two Tracks with opposite charges $\Longrightarrow$ Reconstructed
- More than two track $\Longrightarrow$ Complicated

Potential Efficiency Metrics

- Number of Events with ≥ 2 longTracks [good proxy]
- MC Based Effi. [Using reconstructed-truth level data]

Definition of Fiducial

Before we define the efficiency we must account for the detector acceptance by ensuring the particle to be Fiducial.

Fiducial Criteria Based on Truth-Level Data

- truthd0_r [{1,2,3}] < 100
- truthd1_r [{1,2,3}] < 100
- truthd0_pz > 20 GeV
- truthd1_pz > 20 GeV

A truth-level prescription is preferred because the reconstruction-based approach depends on reconstruction performance, whereas "fiducial" should ideally be a function of the detector alone.

Note: truthd0_r $= \sqrt{\text{truthd0\_x}^2 + \text{truthd0\_y}^2}$, and similarly for d1.

Fiducial CutFlow

Selection Criteria	Pass	All	Eff (%)	Cum. Eff (%)
truthd_st1_r $<$ 100	59228	60000	98.71	98.71
truthd_st2_r $<$ 100	56549	59228	95.48	94.25
truthd_st3_r $<$ 100	52640	56549	93.09	87.73
truthd_pz $>$ 20 GeV	50444	52640	95.83	84.07

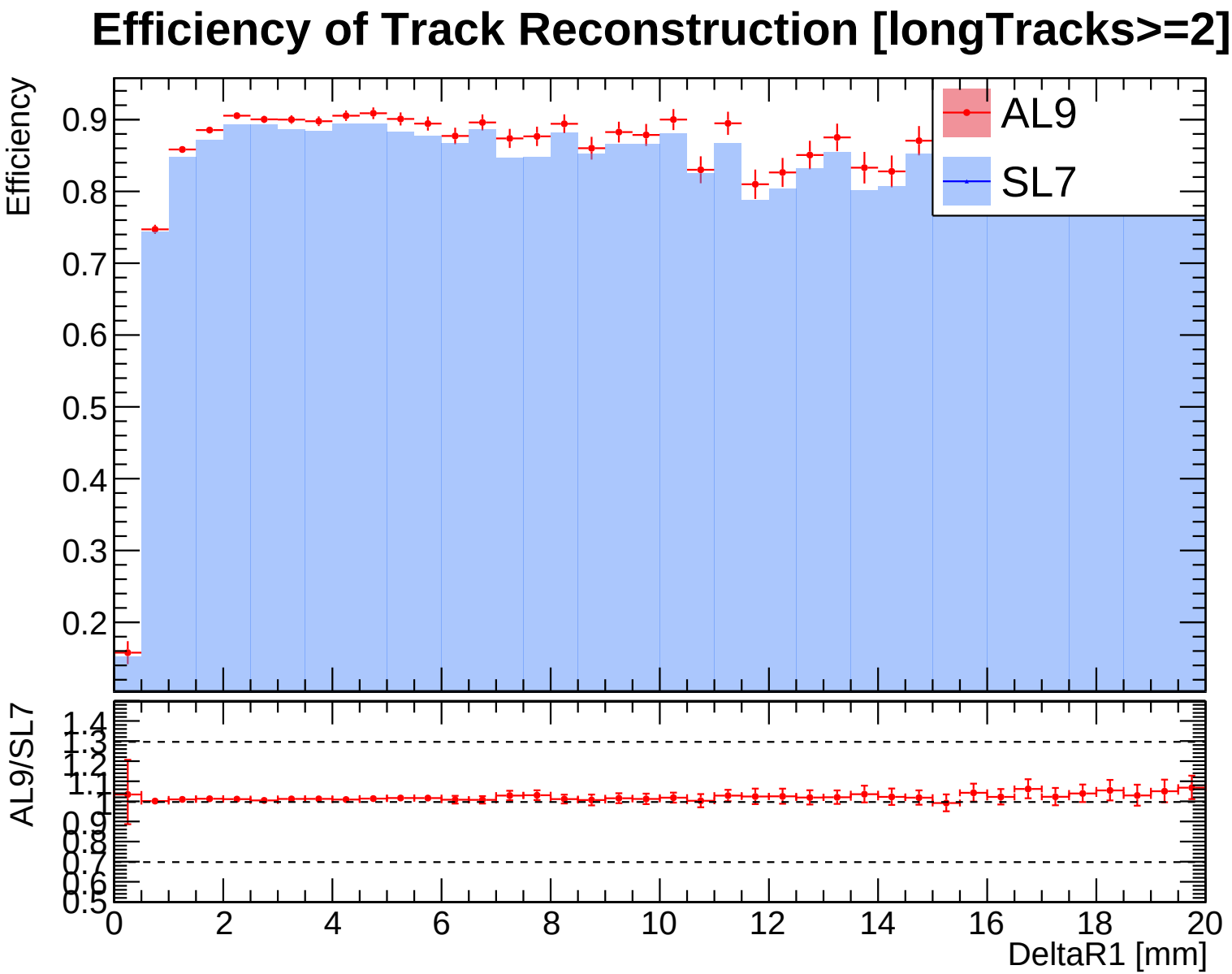
Table: Efficiencies and cumulative efficiencies for truth-level selection steps. [same for ALMA9/CENTOS7]

Track Reconstruction Efficiency Defined

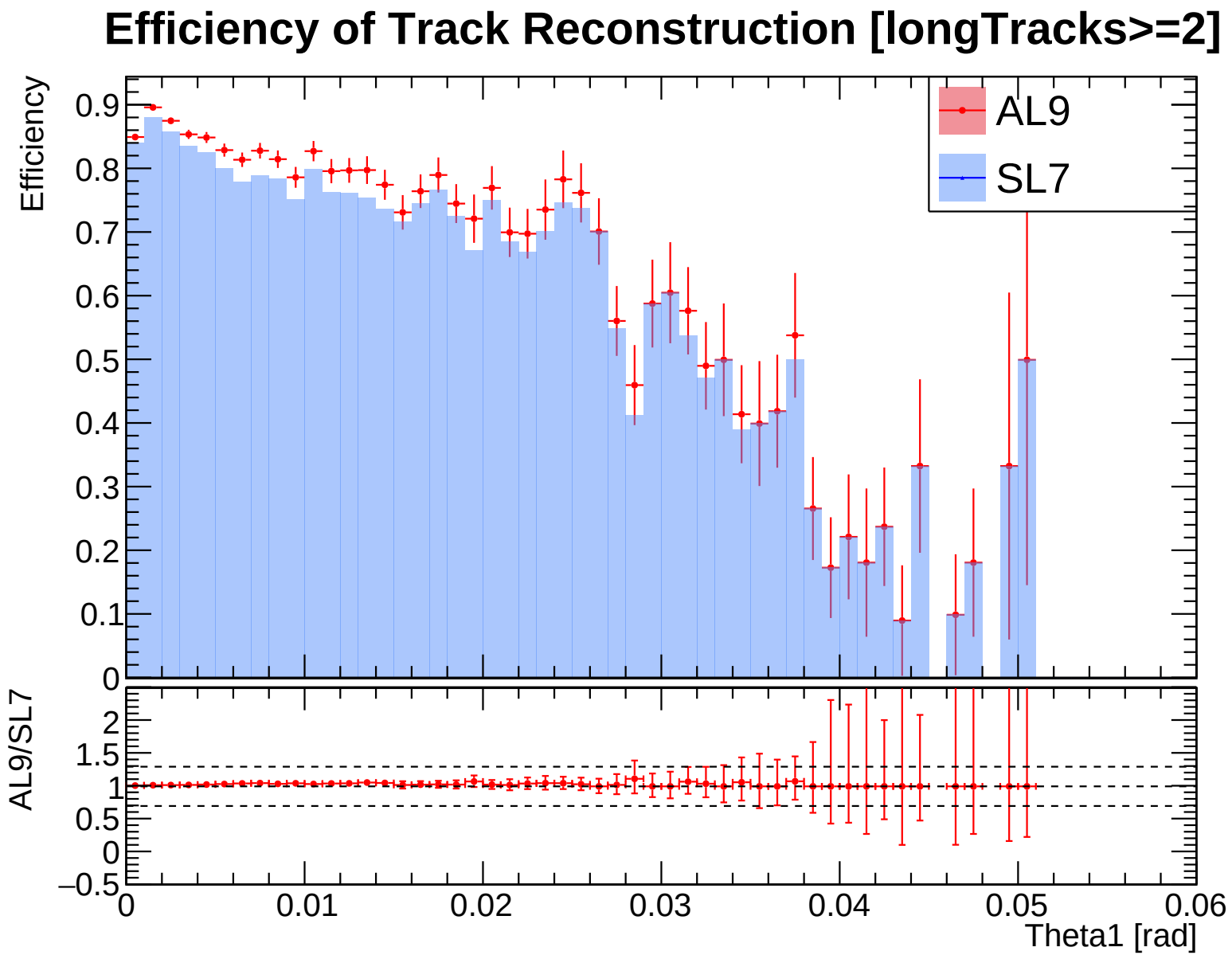
- Remove acceptance based on fiducial cuts at truth level
- Define Efficiency as the fraction of events with $\geq$2 reconstructed longTracks divided by the total number of events which is same as $\geq$0 longTracks. [Given both are fiducial]

$$\text{Efficiency} = \frac{\text{NEvents}(\geq 2\text{longTracks} \mid \text{fiducial})}{\text{NEvents}(\geq 0\text{longTracks} \mid \text{fiducial})}$$

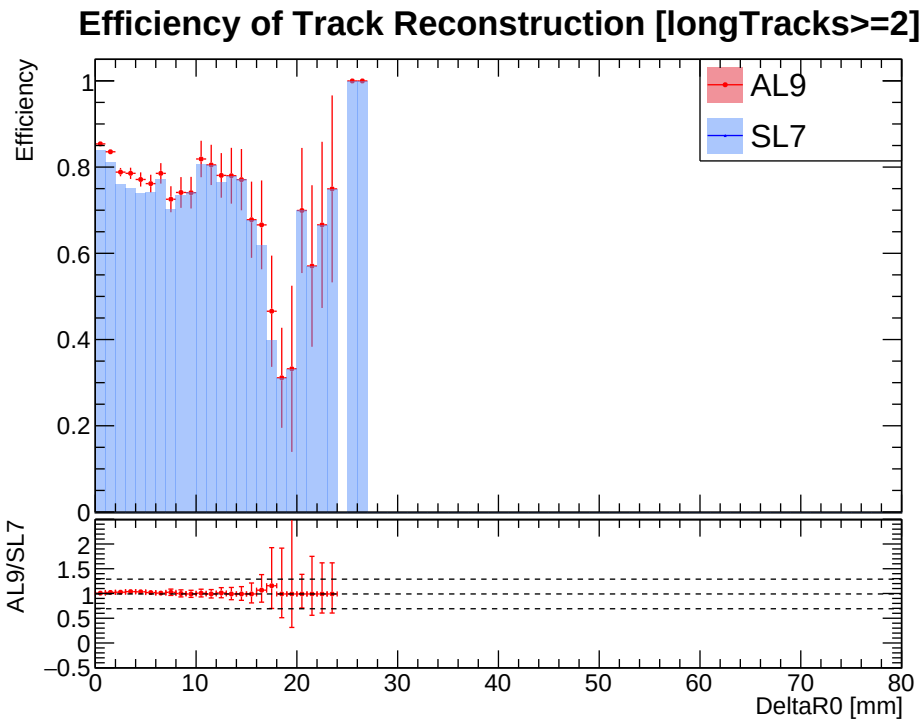
$\geq$ 2 Track Efficiency as a function of DeltaR1

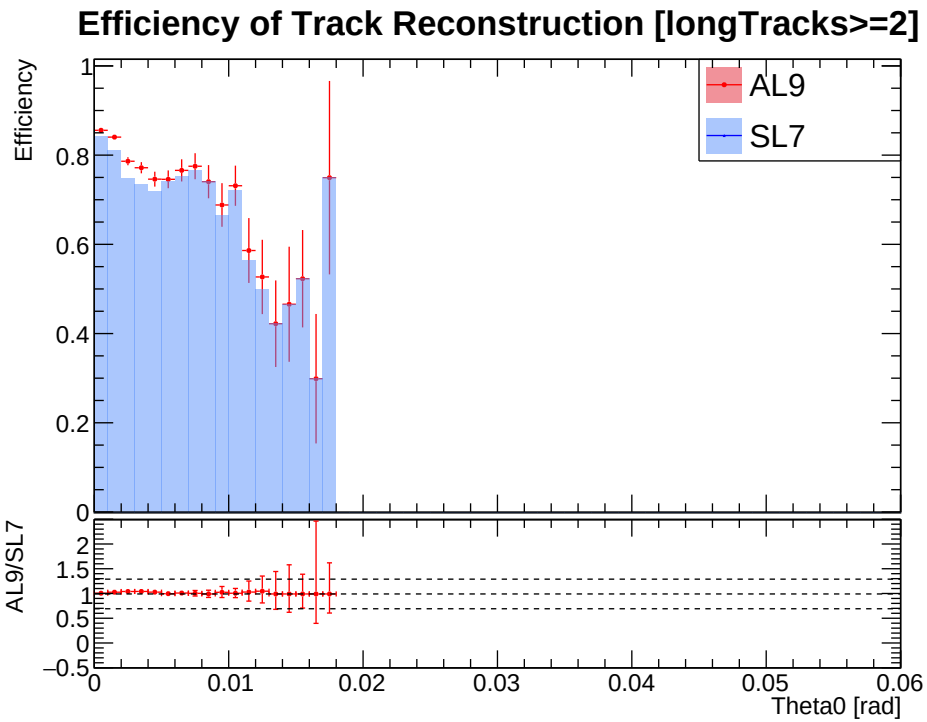


$\geq$ 2 Track Efficiency as a function of Theta1



Efficiencies as a Function of Separation @ Station 1





Comments on $\geq$2 Track Efficiency

- Good agreement between ALMA9 and CENTOS7
- Low statistics at large separation
- The reconstruction is independent of momentum-separation (DeltaRP)
- Efficiencies at Station1 are confusing?

Efficiency Metric based on Reconstructed Truth

We are interested only in the two primary tracks from e^+e^-. Thus we can use the reconstructed truth variables to check if the underlying truth particles are reconstructed.

- For acceptance: Truth Position of $e^+e^- < 100$
- **Identify the two primary tracks**
  - t_barcode = 2,3 $\implies$ primary tracks
  - Also require t_barcode_parent = 1 $\implies$ from DarkPhoton
- TLDR; Ran the analysis with the above definition, but the results are the exact same as ≥ 2 case
- So if event reconstructed has more than 2 longTracks, the two primary tracks are reconstructed, according to above definition.

Efficiency Metric based on Reconstructed Truth

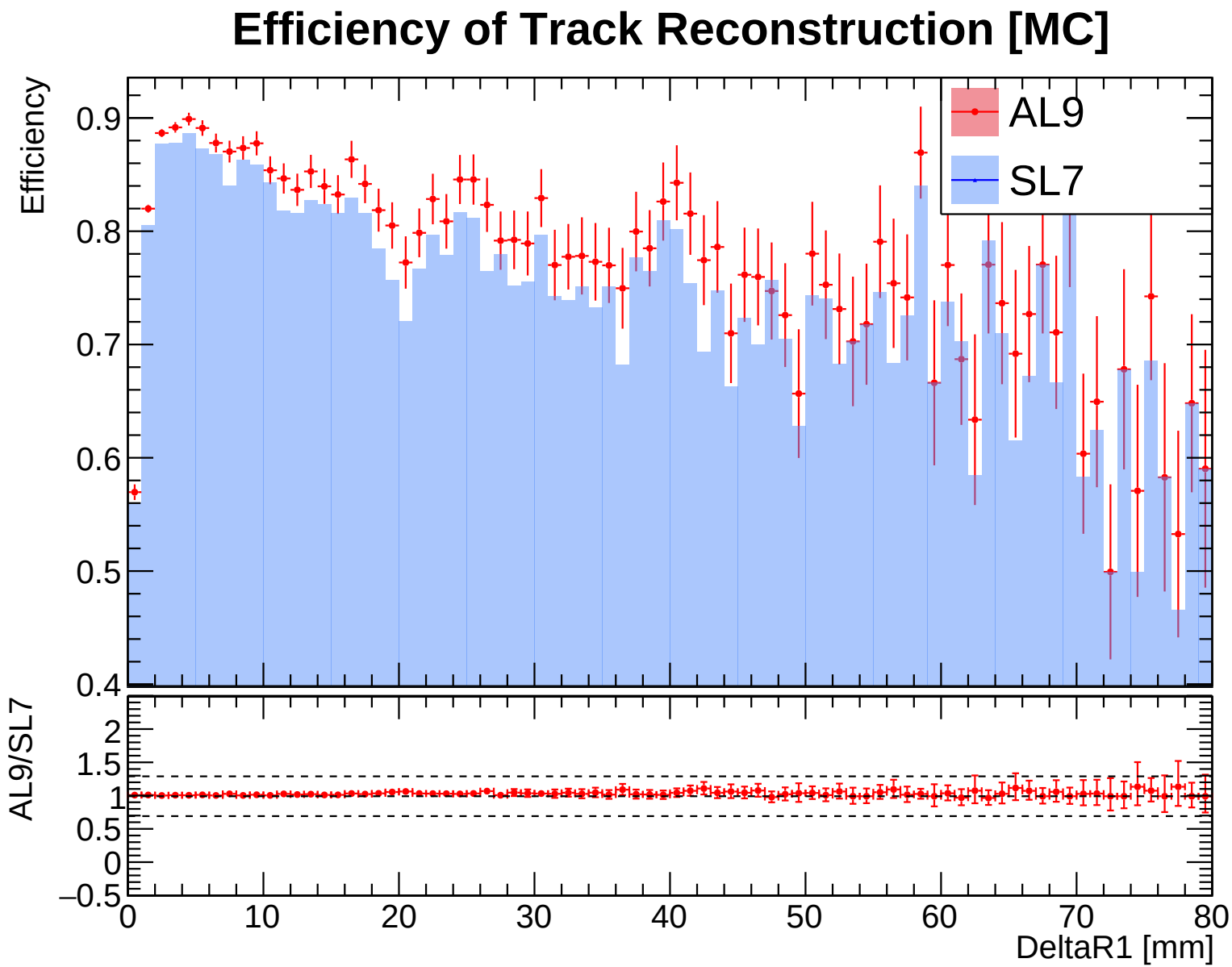


Figure: Efficiency defined using reconstructed truth variables is same as the $>$ 2 longTracks definition

Conclusions

- The “efficiencies” generally agrees between ALMA9 and CENTOS7
- Need a better metric for efficiency accounting for track quality parameter like Track Chi2
- Possible duplications in tracks
- The efficiencies at Station 1 are confusing, needs to be better understood