MATHUSLA Simulation Report

Single track study

Tom Ren for simulation working group, 2023-04-26

Tracker algorithm

The Tracker consists of three steps

1. **Pattern recognition**
2. **Track parameter estimation**
3. **Vertex parameter estimation**

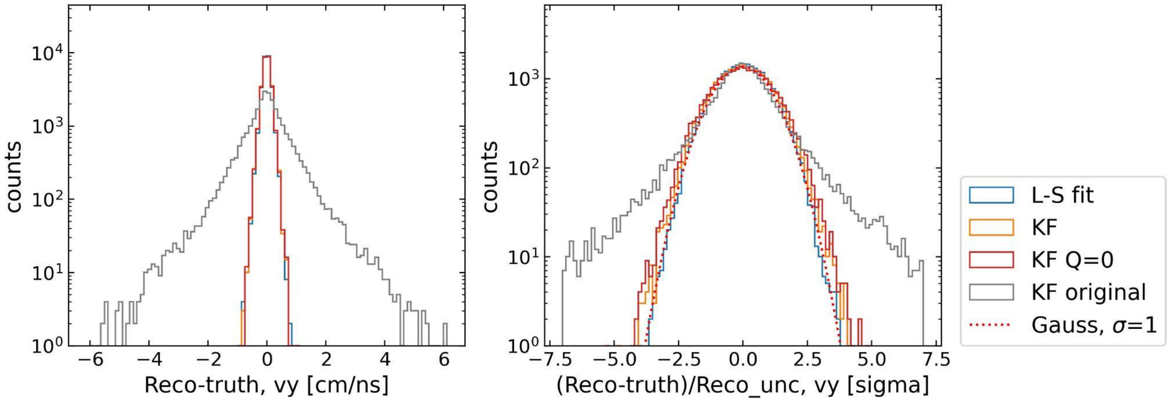
Kalman filter (KF)

The performance has been greatly improved after fixing bugs and tuning parameters in the KF.

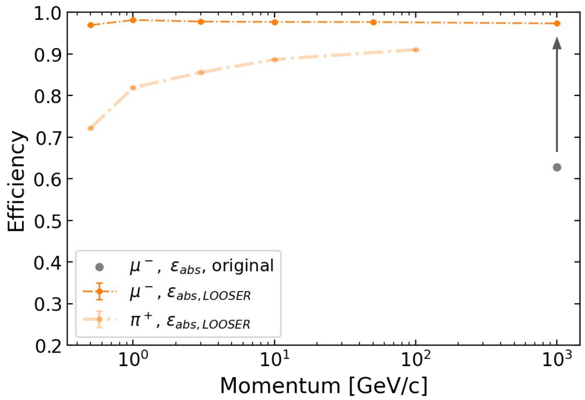
This study focuses on evaluating the performance of <u>track parameter estimation</u> on single-track event, where pattern recognition is maximally simple, limited to helping reconstruct low momentum tracks with some secondary particle production.

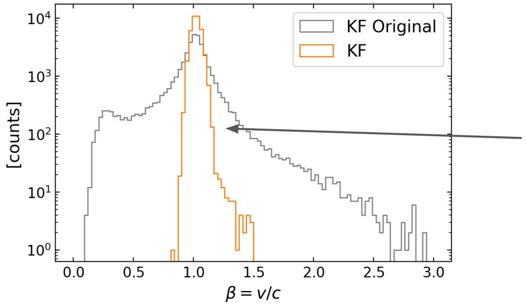
Progress we made

1) Parameter resolution



2) Efficiency





We no longer need the fixed beta assumption because of improved resolution.

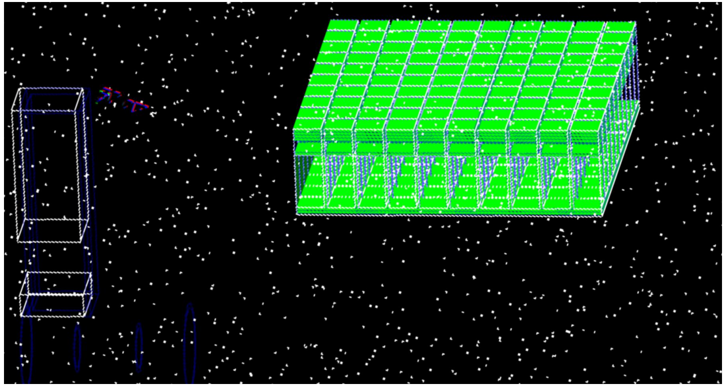
Simulation

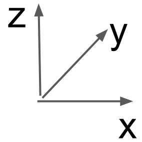
GEANT4: Full geometry with 10x10 modules.

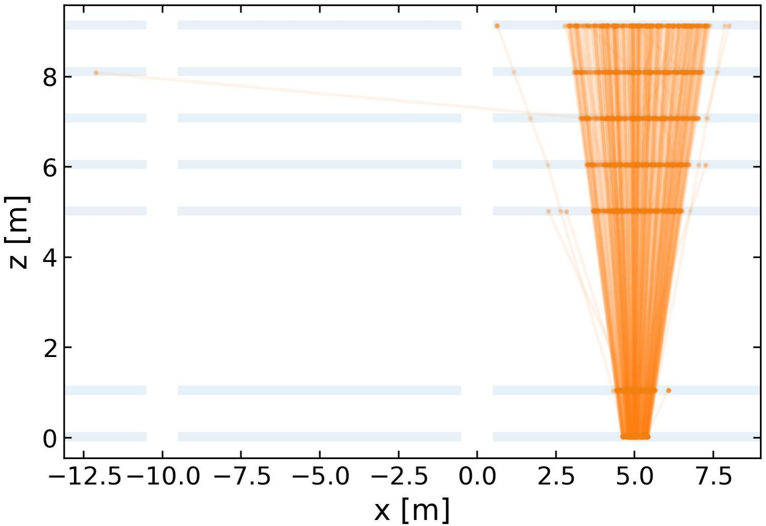
Module: 9m*9m*(3 floor + 2 bottom + **5 top**) layers

Particle gun with origin below the bottom layer

- Single muon events 0.5-1000 GeV/c
- Single pion events 0.5-100 GeV/c







Example of 300 single muon events. Blue shadowed region are locations of detector layers.

Angles are constrained in a small range vertically so the event is contained within a module.

Digitization

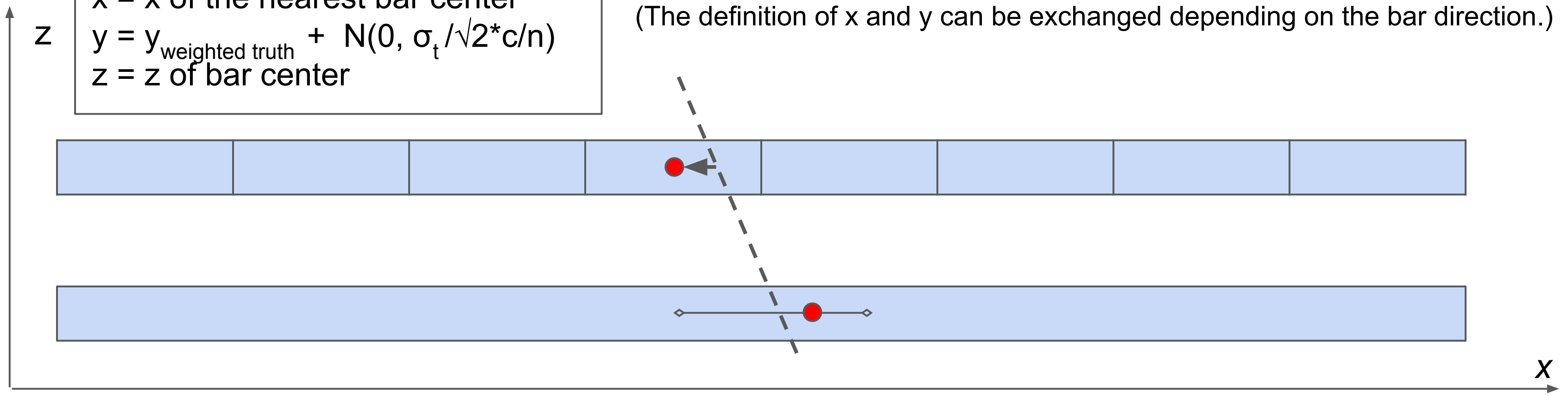
Digitization is the step to convert simulation truth to measurable quantities: (x, y, z, t)

1) **Grouping the truth:** All simulation hits that are in the same bar and happened within 20 ns are combined into one measured hit with energy weighted position and time. No noise hits added, no detector inefficiency.

2) **Modeling timing and spatial resolution**: The time and position along bar (either x or y, depending on the direction) are smeared with detector resolution (1 ns, 14 cm).

$t = t_{\text{weighted truth}} + N(0, \sigma_t)$
x = x of the nearest bar center
$y = y_{\text{weighted truth}} + N(0, \sigma_t / \sqrt{2} * c/n)$
z = z of bar center

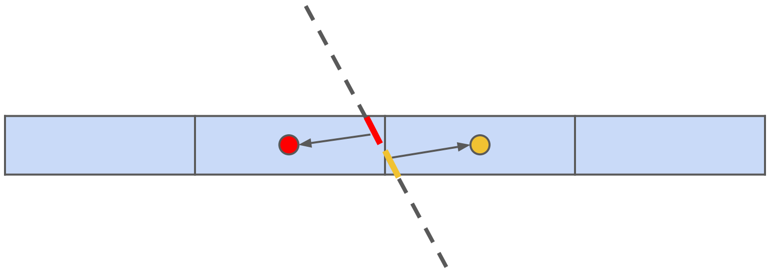
(The definition of x and y can be exchanged depending on the bar direction.)



Digitization

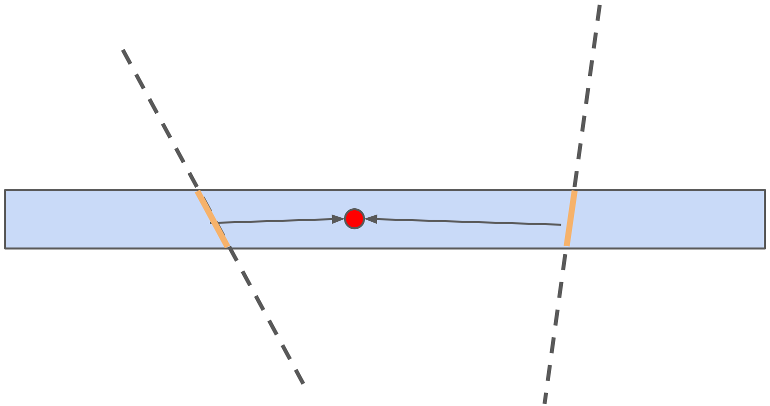
Special cases

1) Energy sharing: if the track hit two adjacent bars and deposit energy that is above threshold in each bar, then **two** digitized hits are generated, one in each bar.



(Pattern recognition will then only pick one hit per layer to assign to a track)

2) Hits clustering: if two tracks hit the same bar and are within 20 ns, **only one** digitized hit is generated.



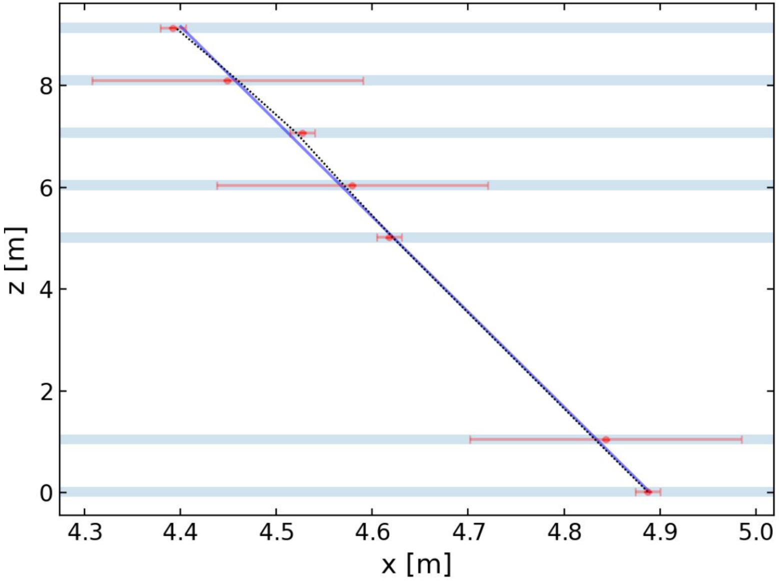
x

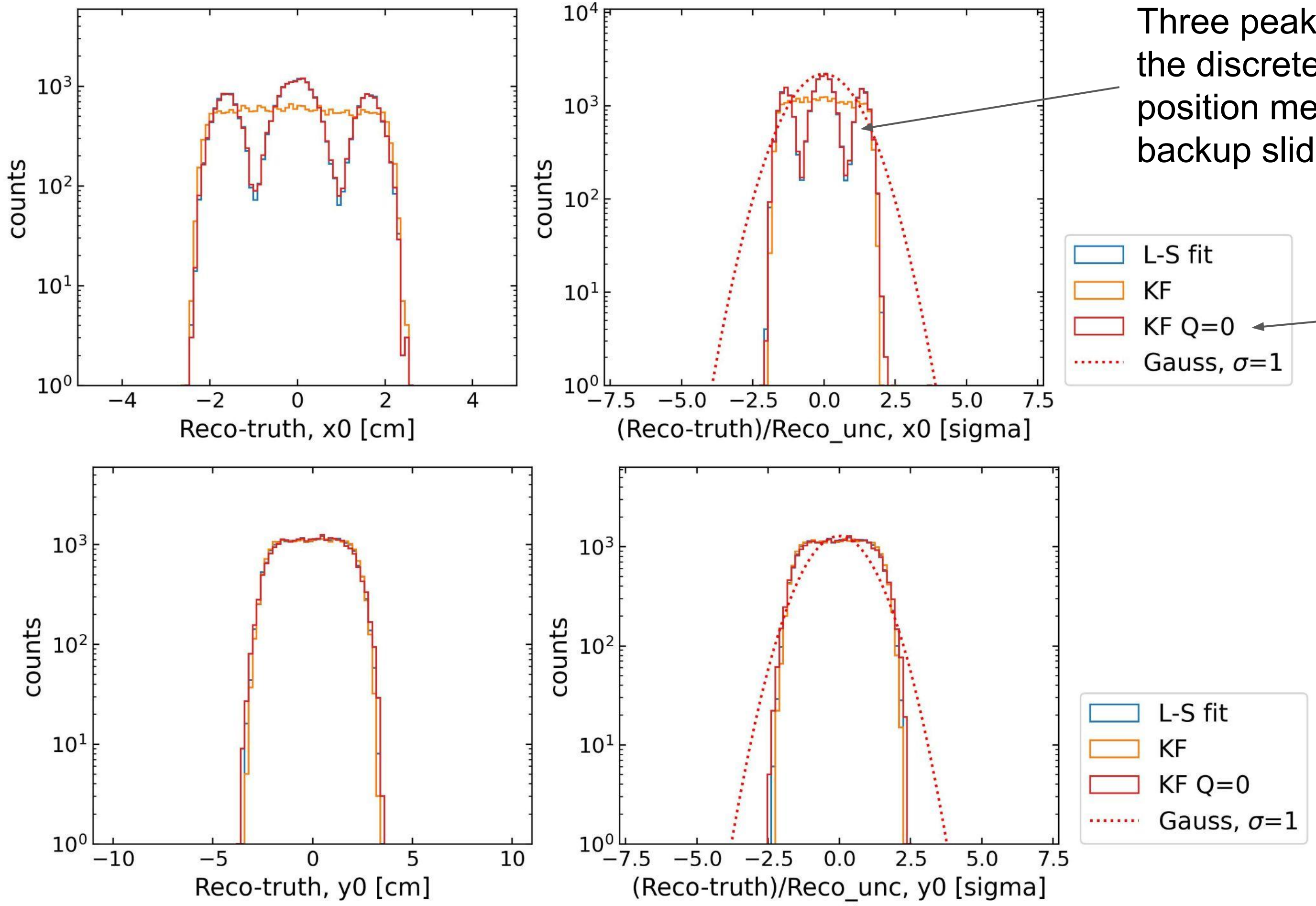
Understanding straight lines and KF

- **Event:** we use the simplest case to understand how "ideal" events will look like in our detector:

  1 TeV/c muon with no secondary particles

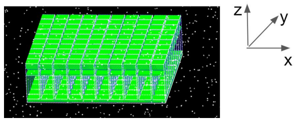
- **Methods:** we use two other methods for comparison
  a. **LS fit,** since LS is well understood and gives optimal parameter estimation in the no-multiple scattering limit
  b. **KF with zero multiple scattering** (Q=0), which is expected to be a sub-optimal version of the LS

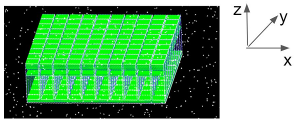


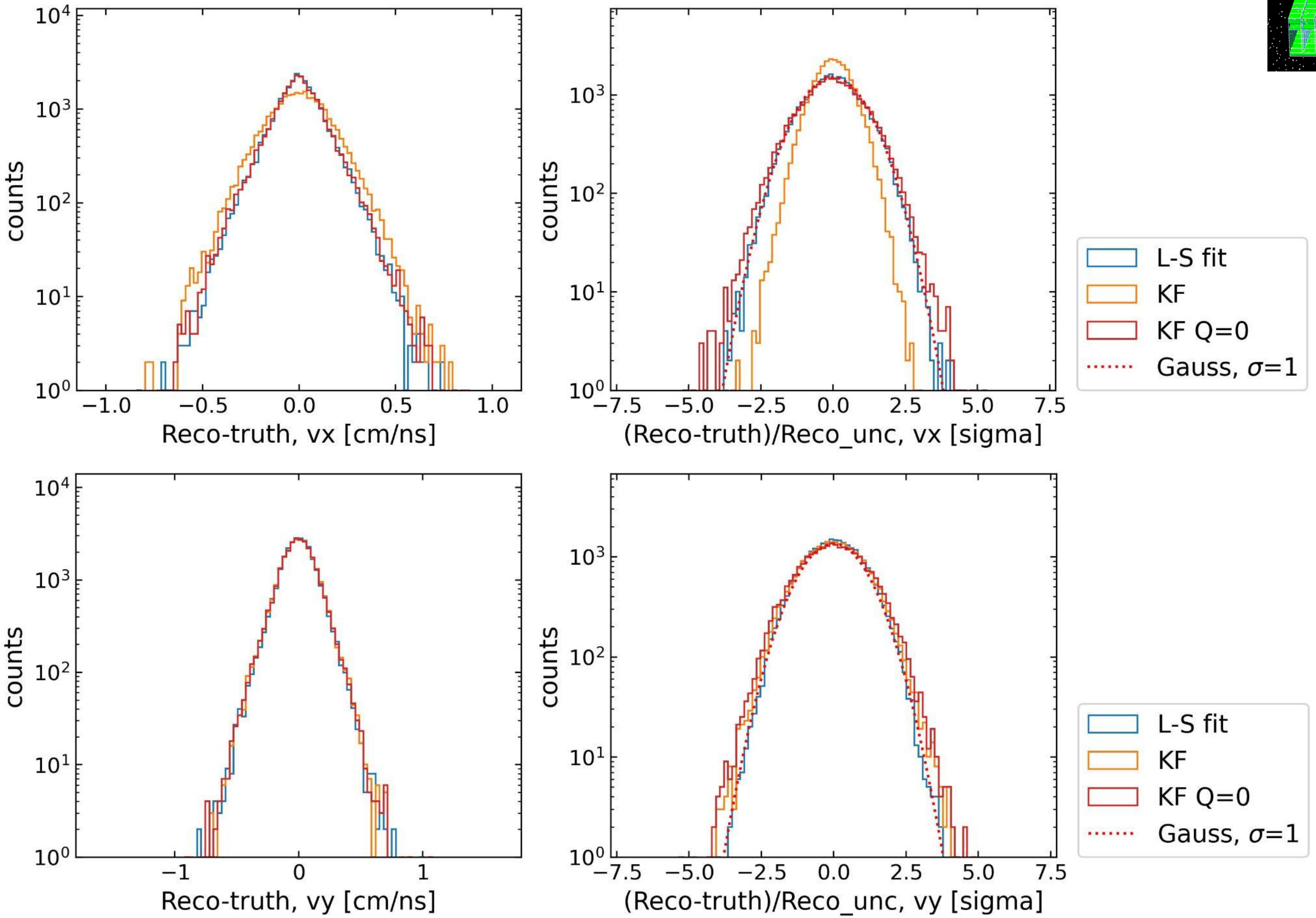


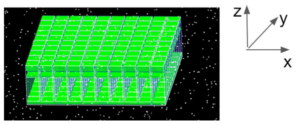
Three peaks are caused by the discrete nature of the position measurement (see backup slides for details)

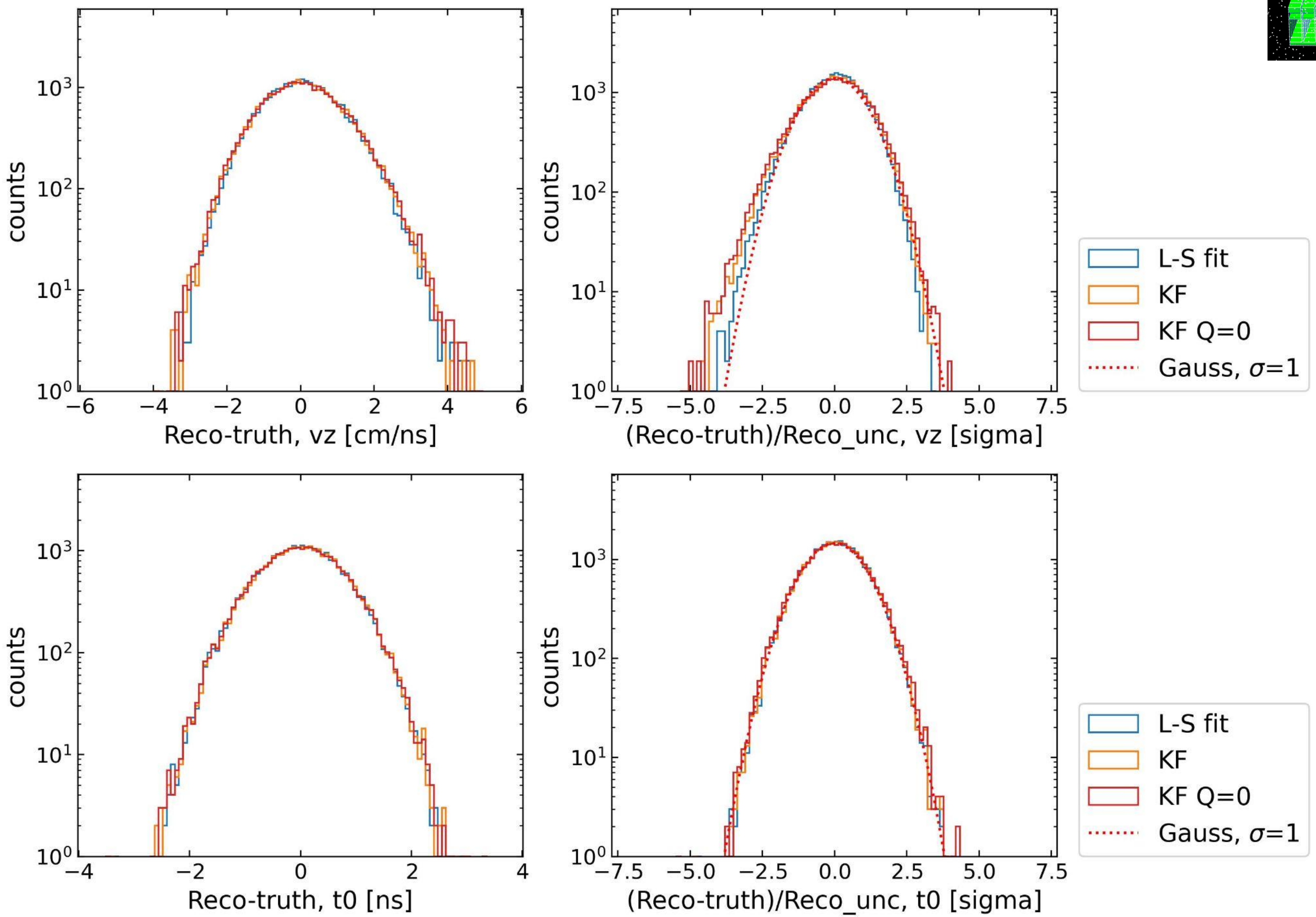
KF with no multiple scattering




z
y
x


counts
Reco-truth, vx [cm/ns]
(Reco-truth)/Reco_unc, vx [sigma]
Reco-truth, vy [cm/ns]
(Reco-truth)/Reco_unc, vy [sigma]
L-S fit
KF
KF Q=0
Gauss, σ=1





counts
Reco-truth, vz [cm/ns]
(Reco-truth)/Reco_unc, vz [sigma]
Reco-truth, t0 [ns]
(Reco-truth)/Reco_unc, t0 [sigma]
L-S fit
KF
KF Q=0
Gauss, σ=1

1. **Both KF results are very close to LS** (as expected)
2. **KF with no multiple scattering has better resolution than the default KF** (as expected)

The following behaviours are not actually problem, just a feature of our discrete sensor:

- The three peaks in the x0 distribution
- The pull distribution is not Gauss(0,1) for x0, y0

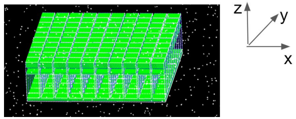
KF is able to achieve similar performance as LS at high energy limit

Tracker performance on single muon/pion

A. Parameter resolution

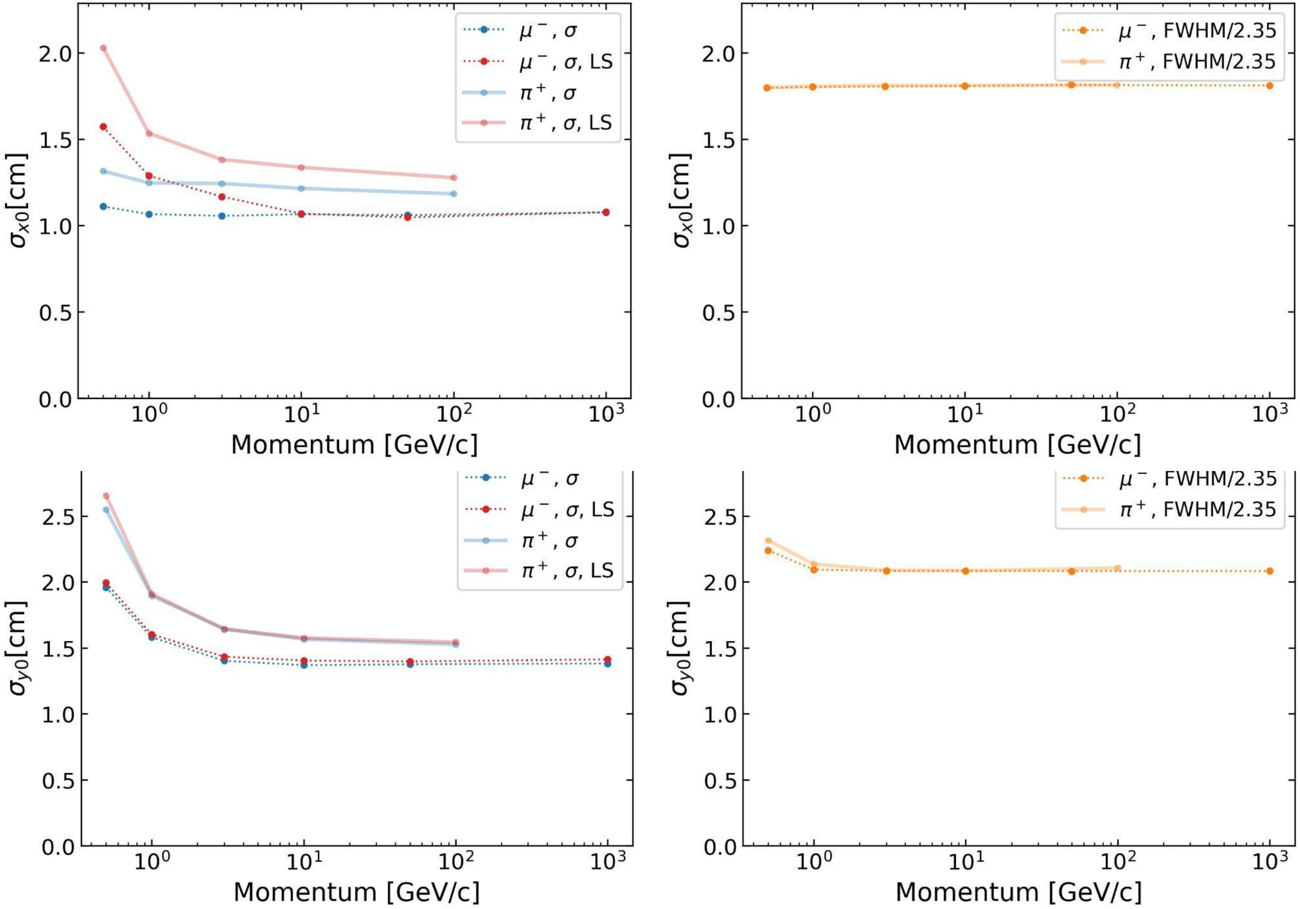
B. Efficiency

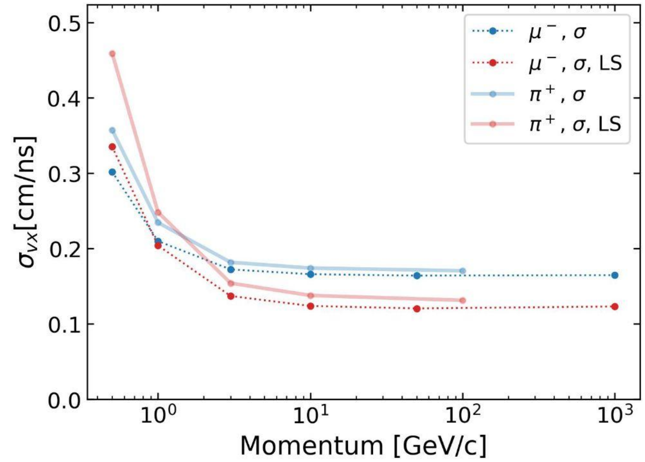
A. Parameter resolution

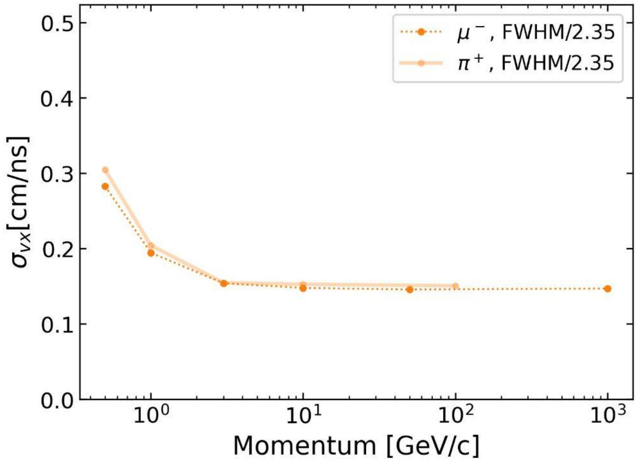


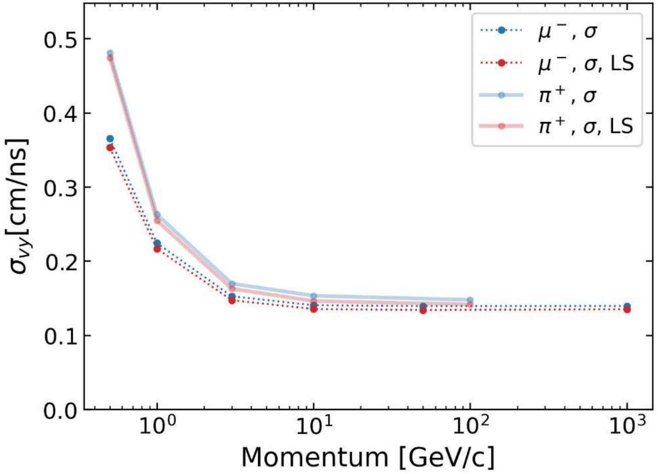
We report both σ and FWHM

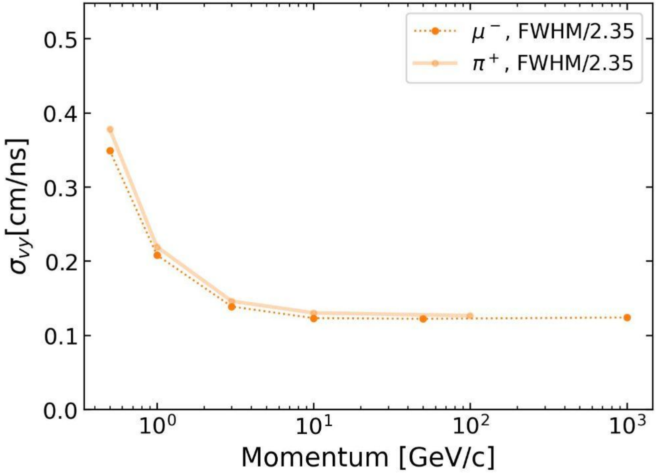
The σ of LS fit is added for comparison. The LS is slightly worse on x0 at lower energy, but better on vz.

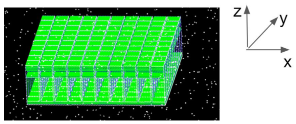


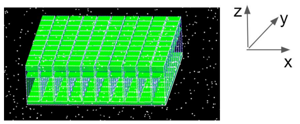

μ^-, σ
μ^-, σ, LS
π^+, σ
π^+, σ, LS
σ_{vx}[cm/ns]
Momentum [GeV/c]

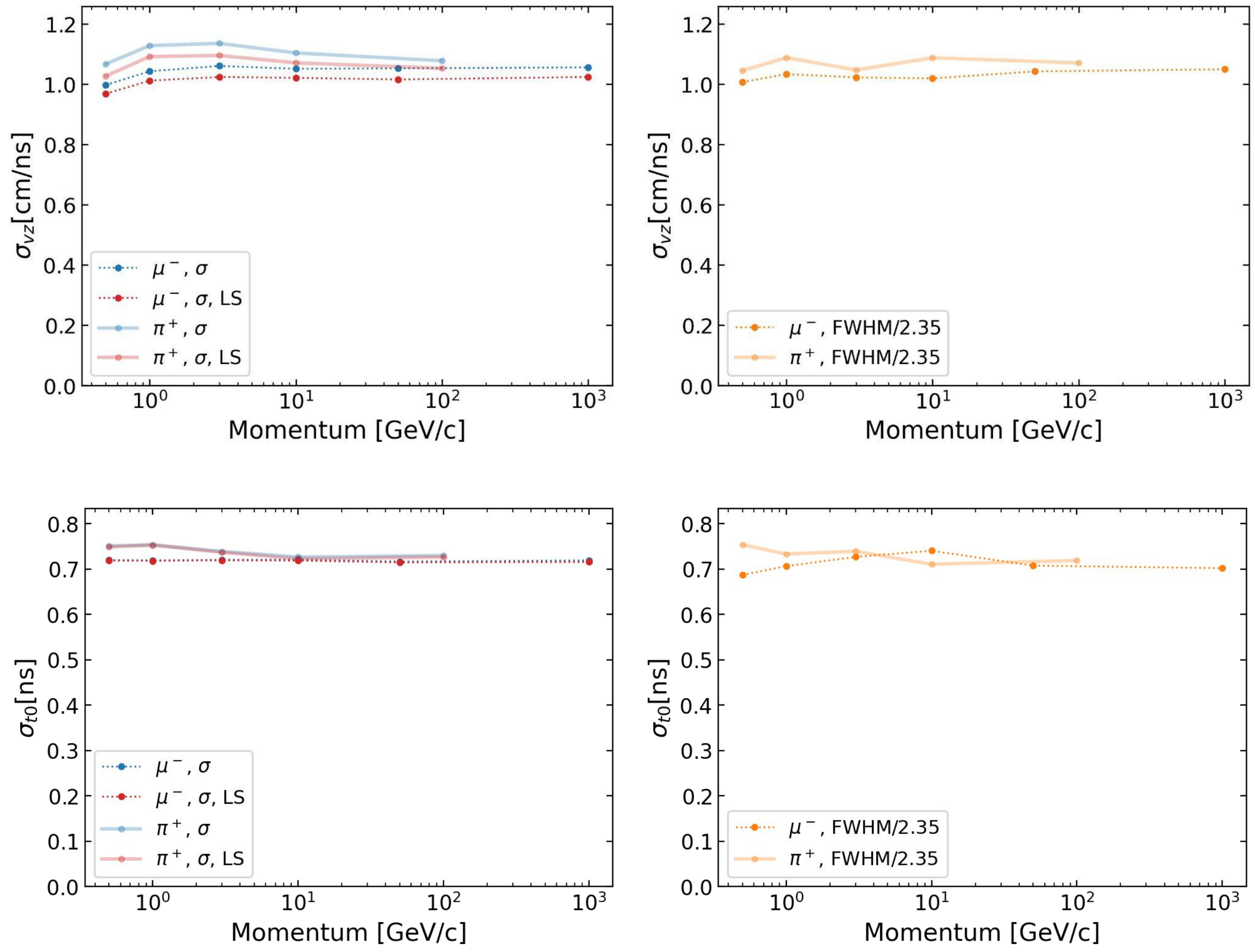

μ^-, FWHM/2.35
π^+, FWHM/2.35
σ_{vx}[cm/ns]
Momentum [GeV/c]


μ^-, σ
μ^-, σ, LS
π^+, σ
π^+, σ, LS
σ_{vy}[cm/ns]
Momentum [GeV/c]

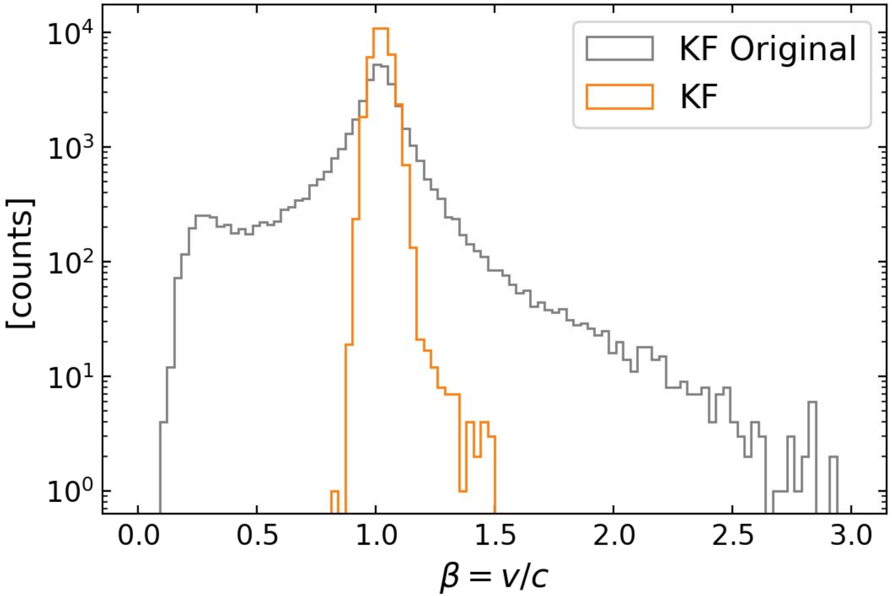

μ^-, FWHM/2.35
π^+, FWHM/2.35
σ_{vy}[cm/ns]
Momentum [GeV/c]


z
y
x


z
y
x


σ_{vz}[cm/ns]
Momentum [GeV/c]
μ^-, σ
μ^-, σ, LS
π^+, σ
π^+, σ, LS
μ^-, FWHM/2.35
π^+, FWHM/2.35
σ_{t0}[ns]
Momentum [GeV/c]
μ^-, σ
μ^-, σ, LS
π^+, σ
π^+, σ, LS
μ^-, FWHM/2.35
π^+, FWHM/2.35

With the improved resolution, we no longer need to constrain beta during the fit.



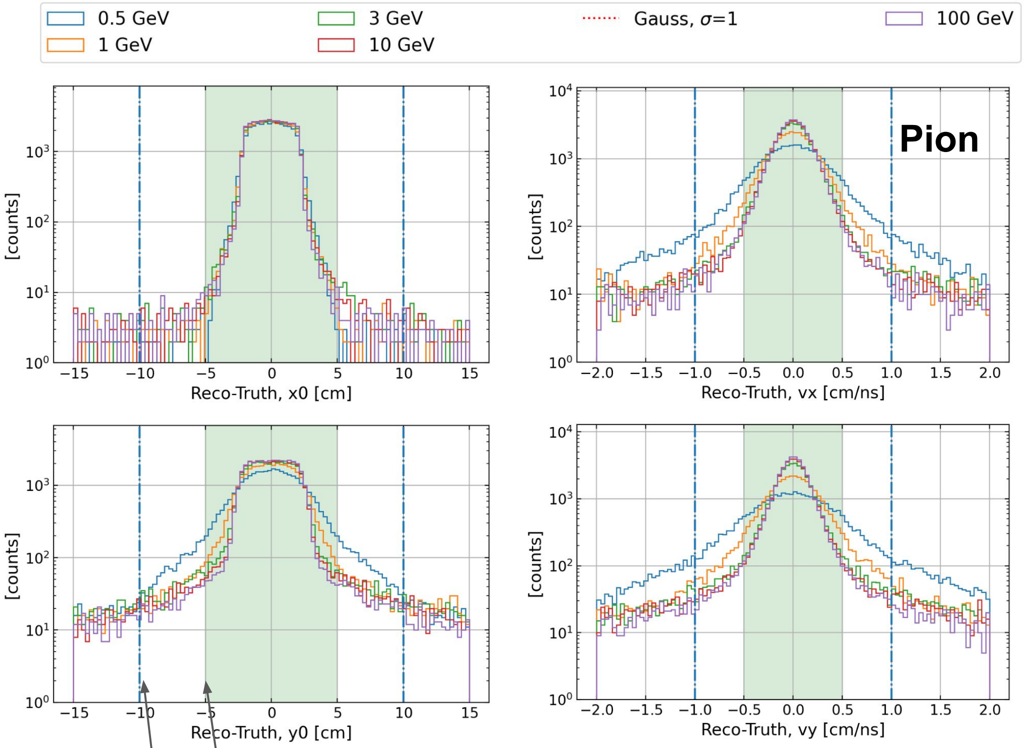
B. Efficiency

We define absolute efficiency ε_{abs} as the ratio of events with a correct track parameter estimation to all the events.

This efficiency reflects the total fraction of good events that we can get.

$$\epsilon_{abs} = \frac{\text{N(correctly estimated)}}{\text{N(all)}}$$

We define a *tight* cut and a *looser* cut to select "correctly estimated" events:



- Tight: |Δx|, |Δy| < 5 cm, |Δvx|, |Δvy| < 0.5 cm/ns,
- Looser: |Δx|, |Δy| <10 cm, |Δvx|, |Δvy| < 1 cm/ns

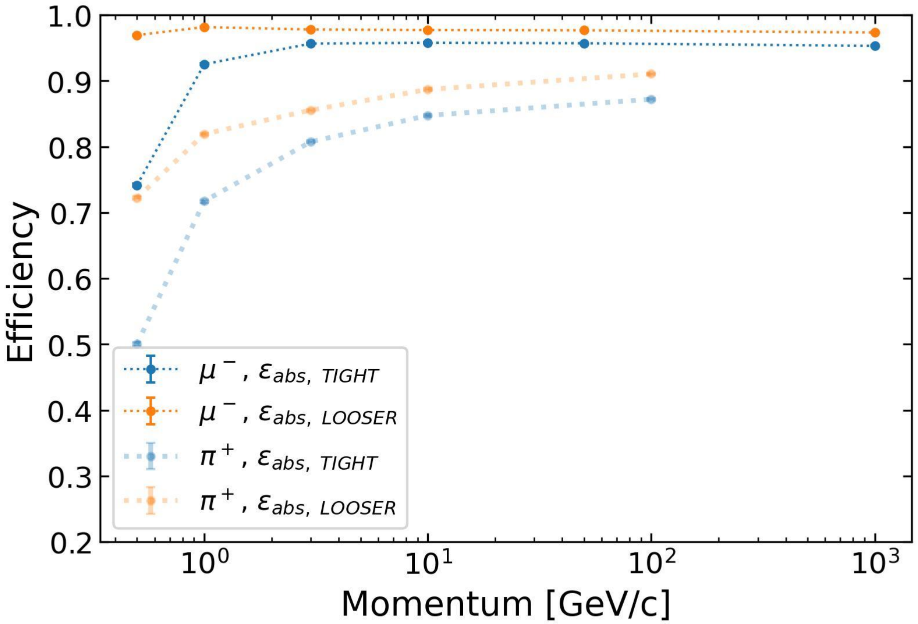
B. Efficiency

We define absolute efficiency ε_{abs} as the ratio of events with a correct track parameter estimation to all the events.

This efficiency reflects the total fraction of good events that we can get.

$$\epsilon_{abs} = \frac{\mathrm{N(correctly\ estimated)}}{\mathrm{N(all)}}$$

We define a *tight* cut and a *looser* cut to select “correctly estimated” events:



- Tight: |Δx|, |Δy| < 5 cm, |Δvx|, |Δvy| < 0.5 cm/ns,
- Looser: |Δx|, |Δy| <10 cm, |Δvx|, |Δvy| < 1 cm/ns

There are three different contributions to the loss of efficiency:

1) Physics: the particle needs to hit enough layers. We define particles with 4 or more hits as reconstructible

$$k_{reconstructible} = \frac{\text{N(reconstructible)}}{\text{N(all)}}$$

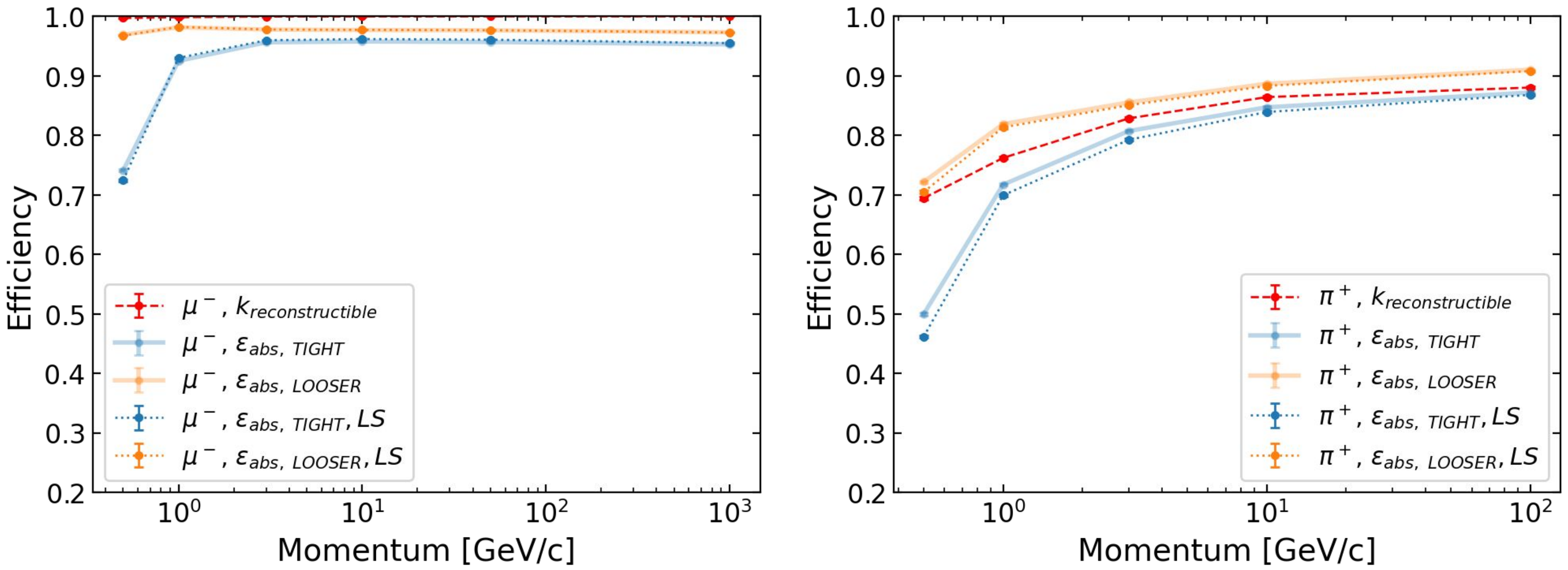
2) Pattern recognition: for reconstructible events, the pattern recognition has probability to fail (correct means >=4 hits from truth and 0 hits from not truth)

$$\epsilon_{recog} = \frac{\text{N(reconstructible \& correctly recognized)}}{\text{N(reconstructible)}}$$

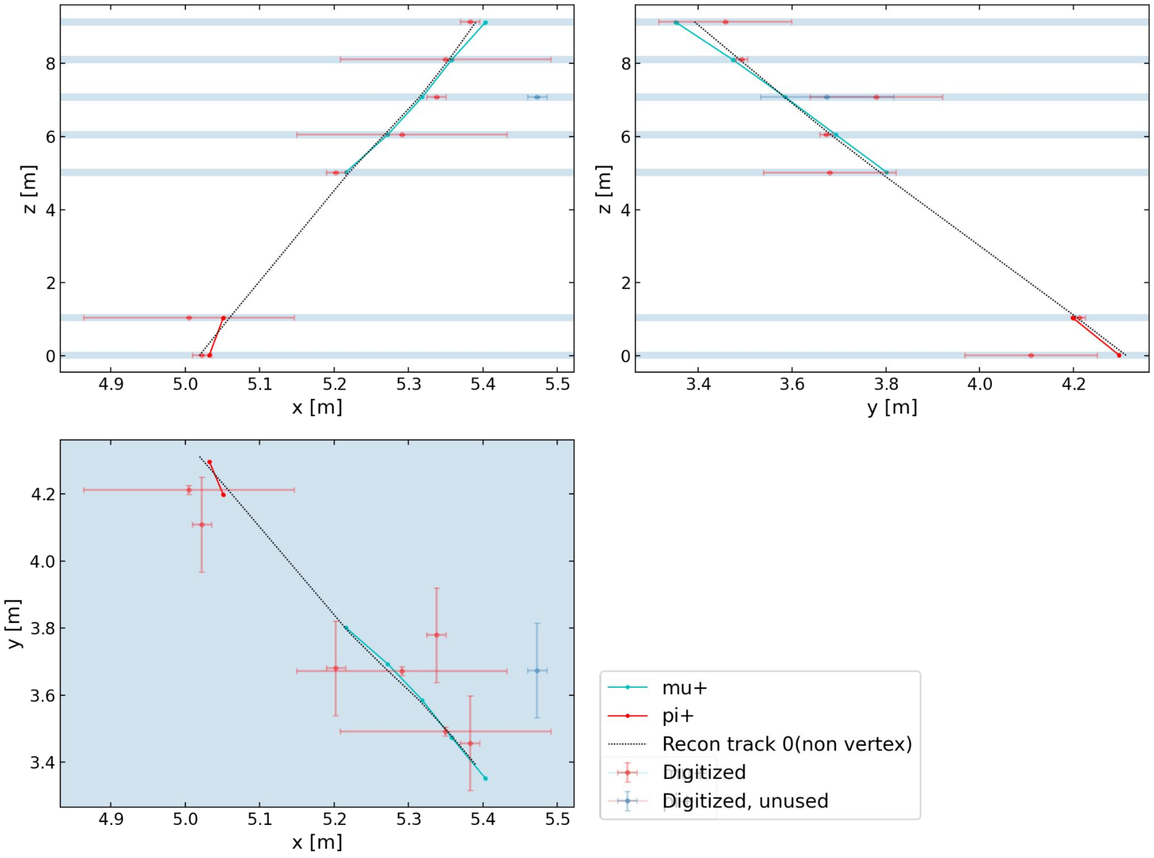
3) Parameter estimation: for events that are correctly identified by pattern recognition, we sometimes do not get track parameters that are consistent with the truth (using the same “LOOSER” cut).

$$\epsilon_{param} = \frac{\text{N(correctly recognized \& correctly estimated)}}{\text{N(reconstructible \& correctly recognized)}}$$

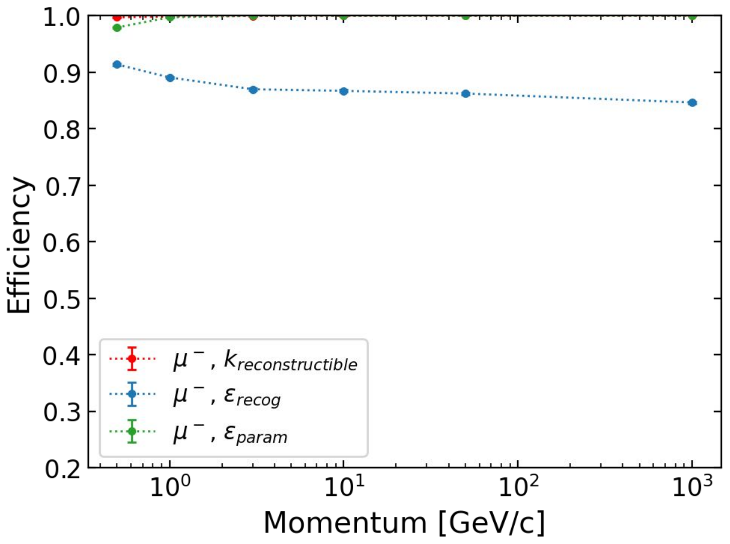
- $k_{reconstructible}$ **is roughly an upper bound for** ε_{abs}**.** With the LOOSER cut, ε_{abs} is close to / have surpassed the $k_{reconstructible}$ (see next slide)
- **We also evaluated the efficiency using LS** (using the same pattern recognition, difference is only in resolution)

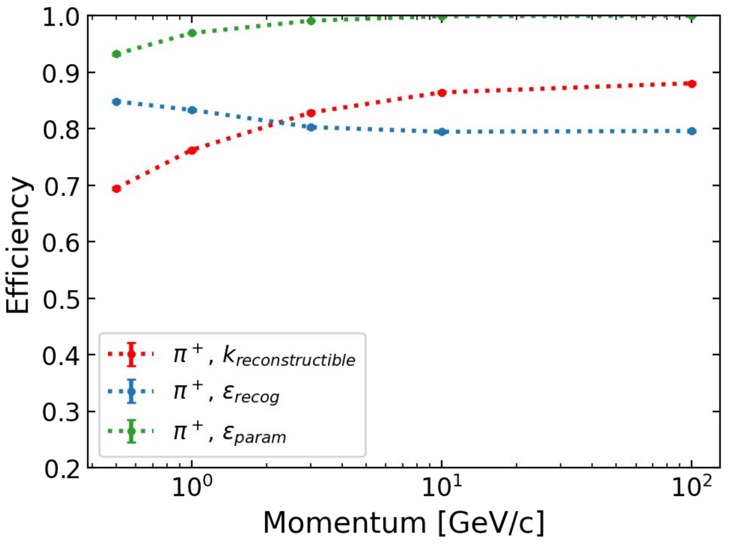


Example of a "non-reconstructible" pion event (<4 hits) that get correct track parameters because of the help of secondary particles (in this case, muon from pion decay).



Apart from $k_{reconstructible}$, the only other main effect is that for ~10% of tracks, you don't assign all the correct hits (ε_{recog}). ε_{param} is 'close to perfect' once previous two conditions are satisfied



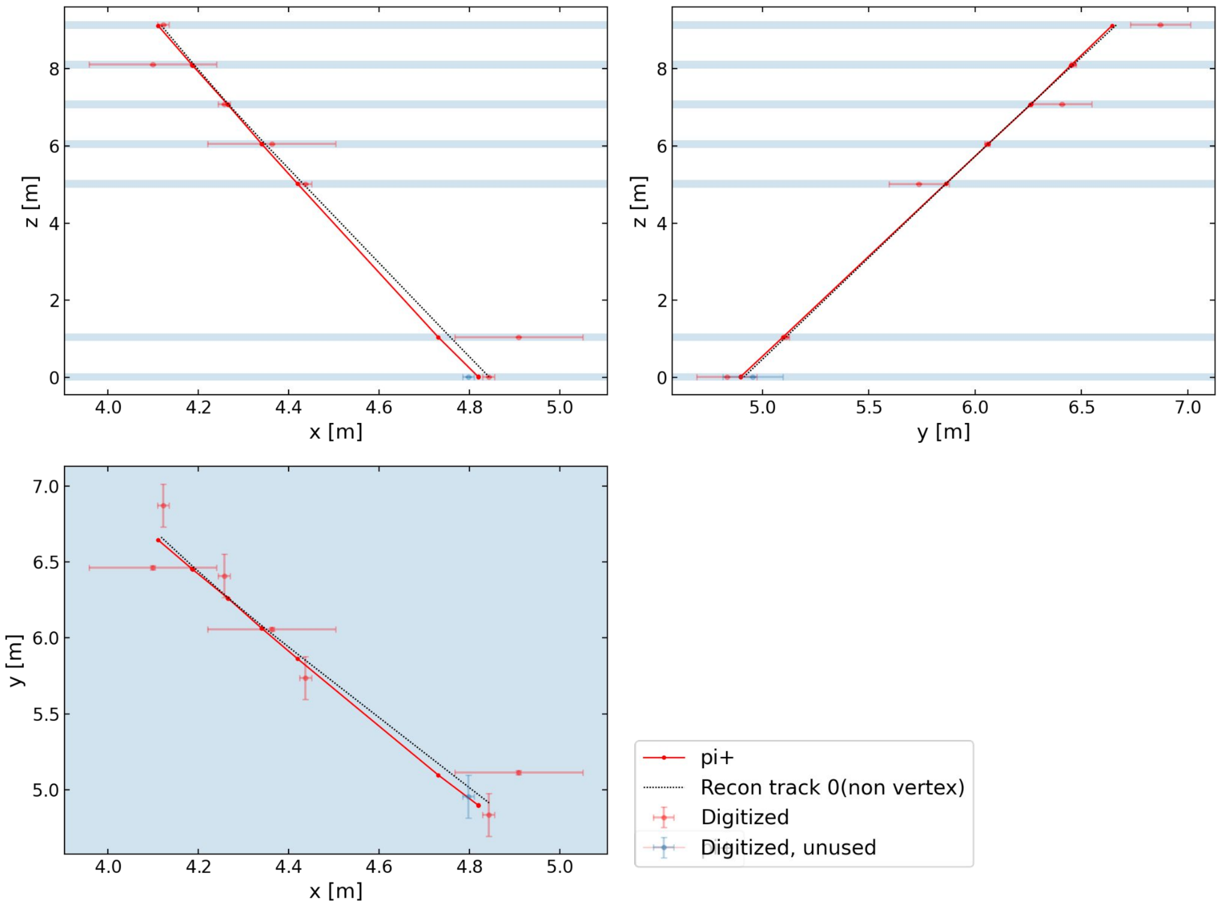


$k_{reconstructible} * \varepsilon_{recog} * \varepsilon_{param}$ does NOT equal ε_{abs} !

- Some non-reconstructible (hits<4) events are correctly reconstructed because of secondary particles (see previous slide for example)
- Some events that have a few hits from secondary particles (decay or material interaction) are correctly reconstructed (see next slide for example)

Example of a pion track with a wrongly assigned hit.

It is not “correctly recognized” but has good reconstruction.



Conclusion

The tracker performance on single tracks has been greatly improved:

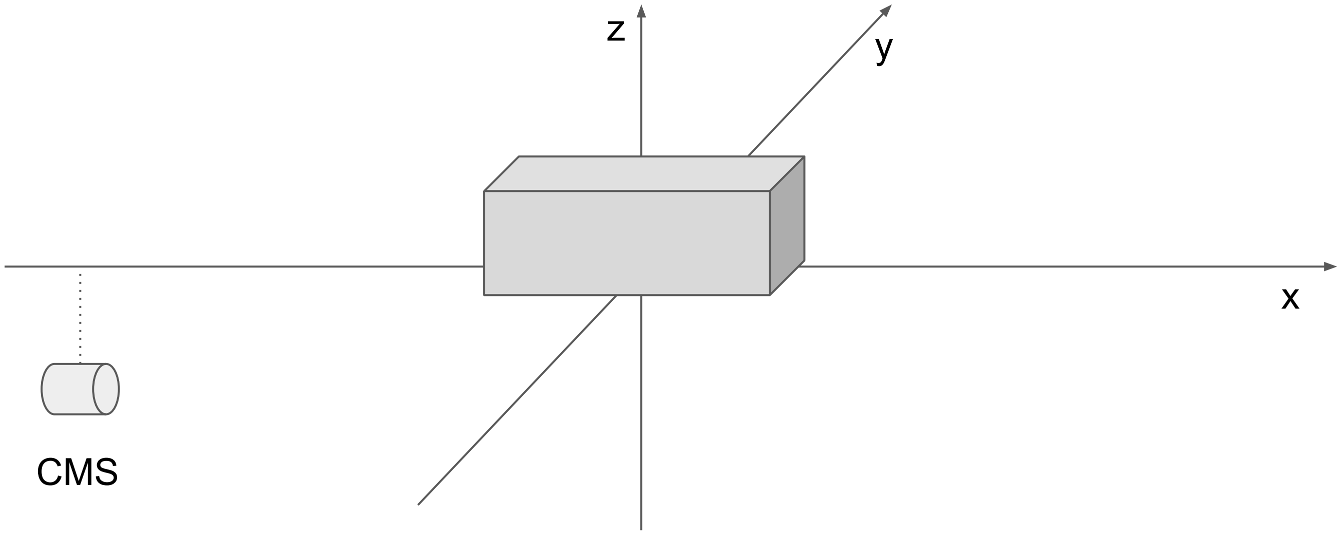
- Parameter resolution: KF is better than LS at low energy, and close to LS at high energy.
- Muon efficiency: 63% → 96% (loose cut)

Further studies to finalize the 1-track analysis:

- Change the geometry to 6 tracking layers to match CDR
- Include lower energies down to 0.1 GeV and add electrons
- Evaluate tracker performance at larger angle/with missing hits
- (Optional) Improve the pattern recognition performance

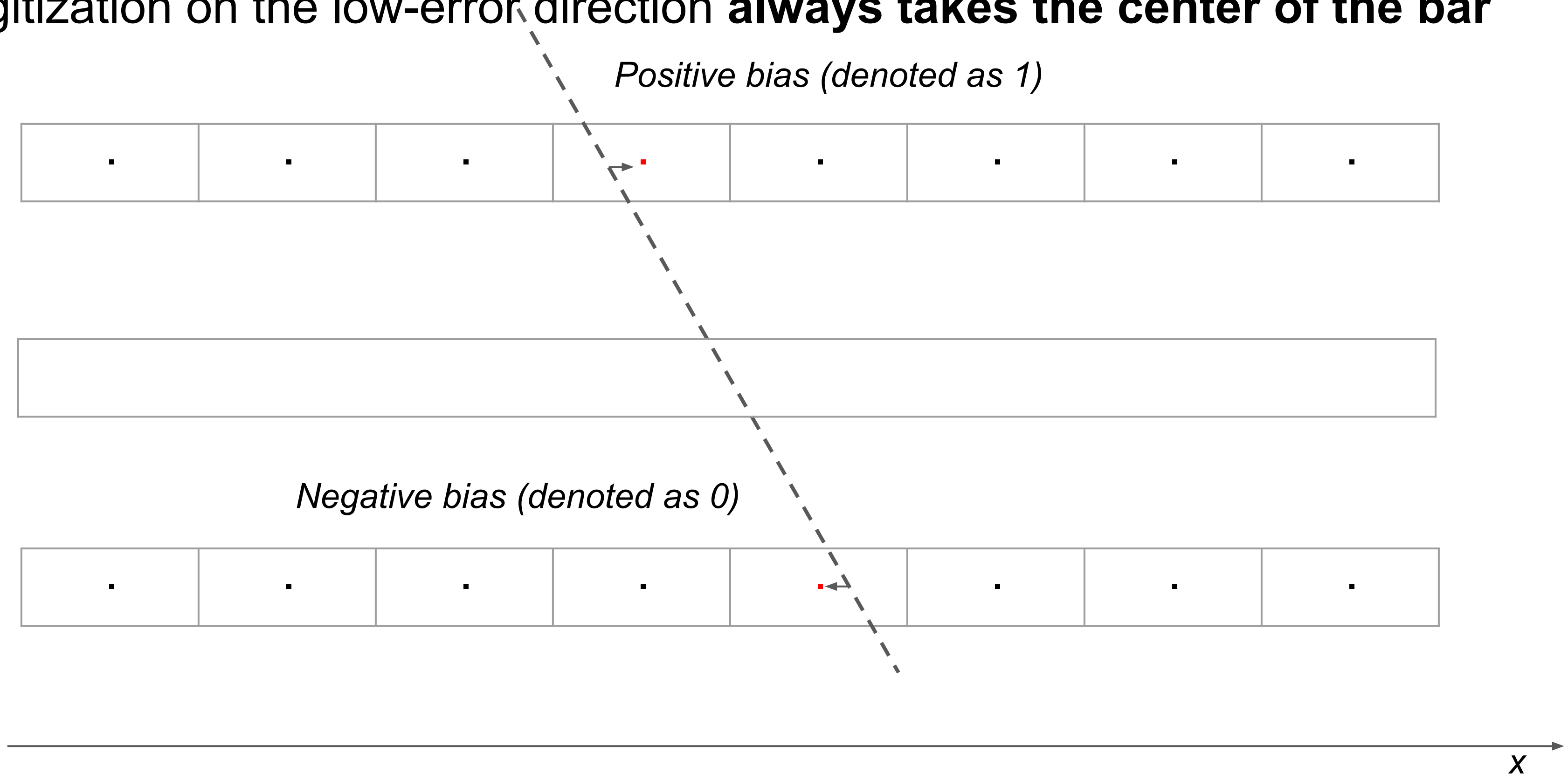
Backup slides

Coordinate



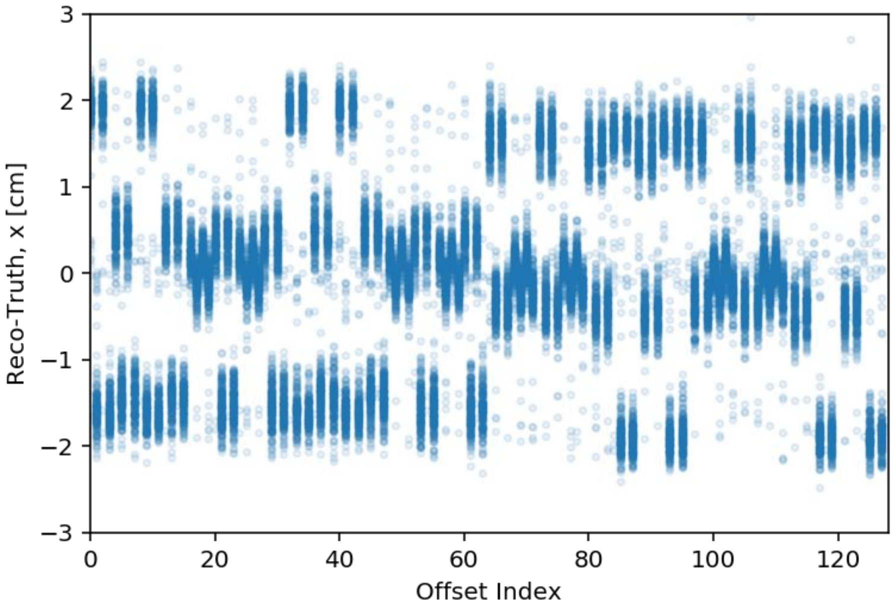
Origin of the 3-peak structure

Digitization on the low-error direction **always takes the center of the bar**



Combine the bias direction of all 7 layers as a binary number

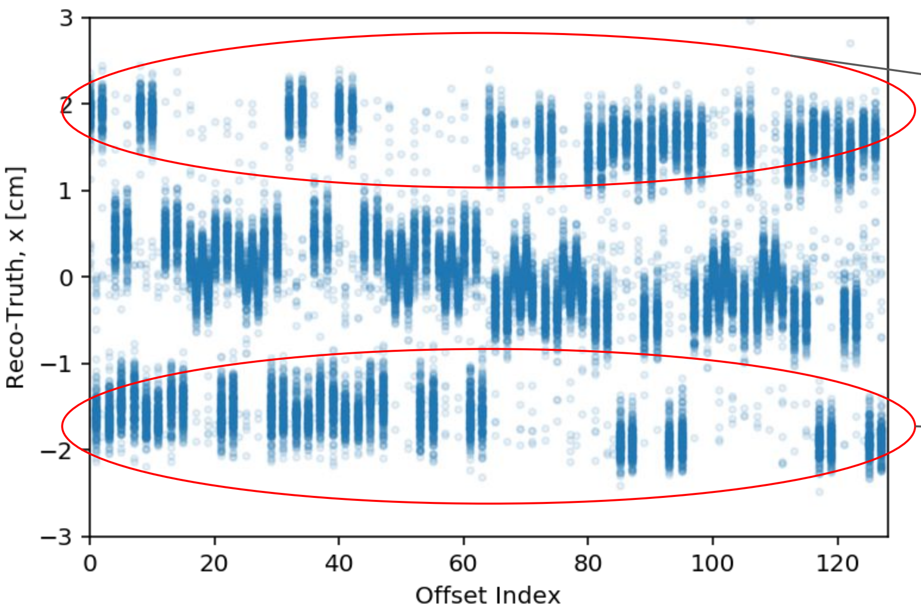
→ Each situation correspond to a unique number ranging from 0-127, referenced as “offset index”



	Layer	Bias direction
Top layer (MSB)	7	0
	6	1
	5	0
	4	0
	3	0
Bottom layer (LSB)	2	0
	1	1

0b0100001=33

Combine the bias direction of all 7 layers as a binary number



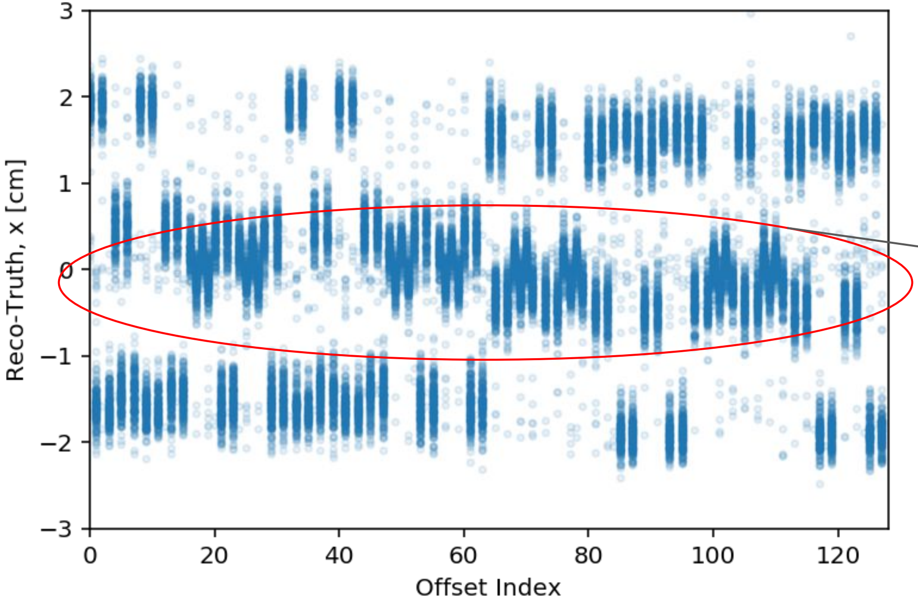
The 32 situations with a positive bias have indices of:

0b1x1x1x**1**
0b0x1x1x**1**
0b0x0x1x**1**
0b0x0x0x**1**

x takes 1 or 0. And the 32 situations with a negative bias have indices of:

0b0x0x0x**0**
0b1x0x0x**0**
0b1x1x0x**0**
0b1x1x1x**0**

Combine the bias direction of all 7 layers as a binary number

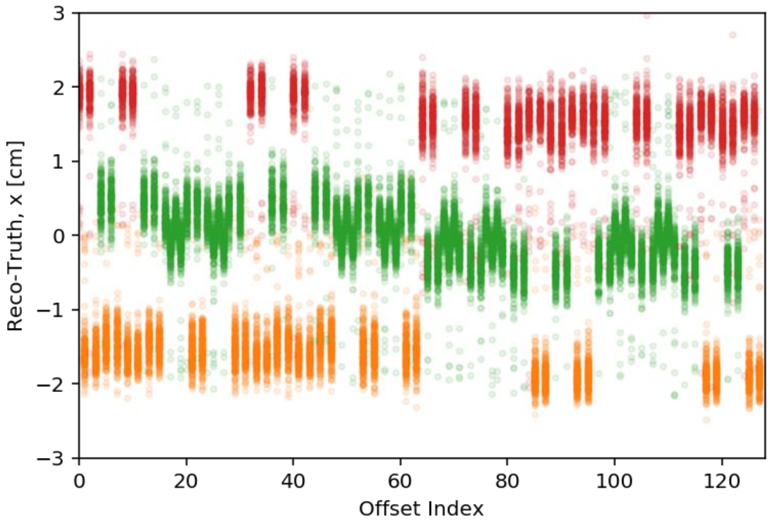


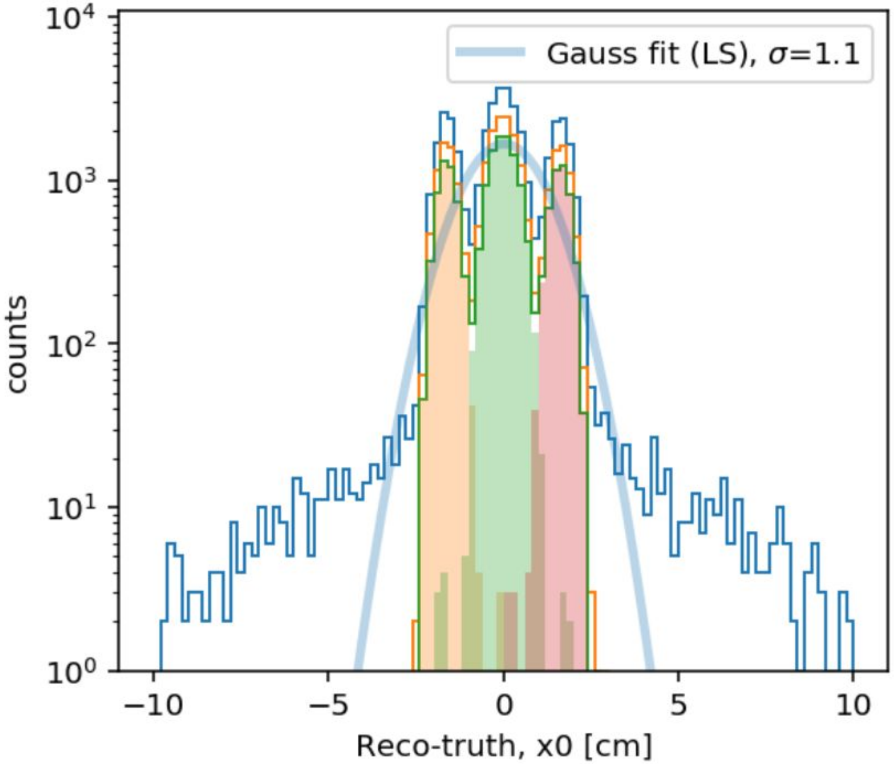
The 64 situations with a ~0 bias have indices of:

0b1x0x1x**1**
0b1x1x0x**1**
0b1x0x0x**1**
0b0x1x0x**1**

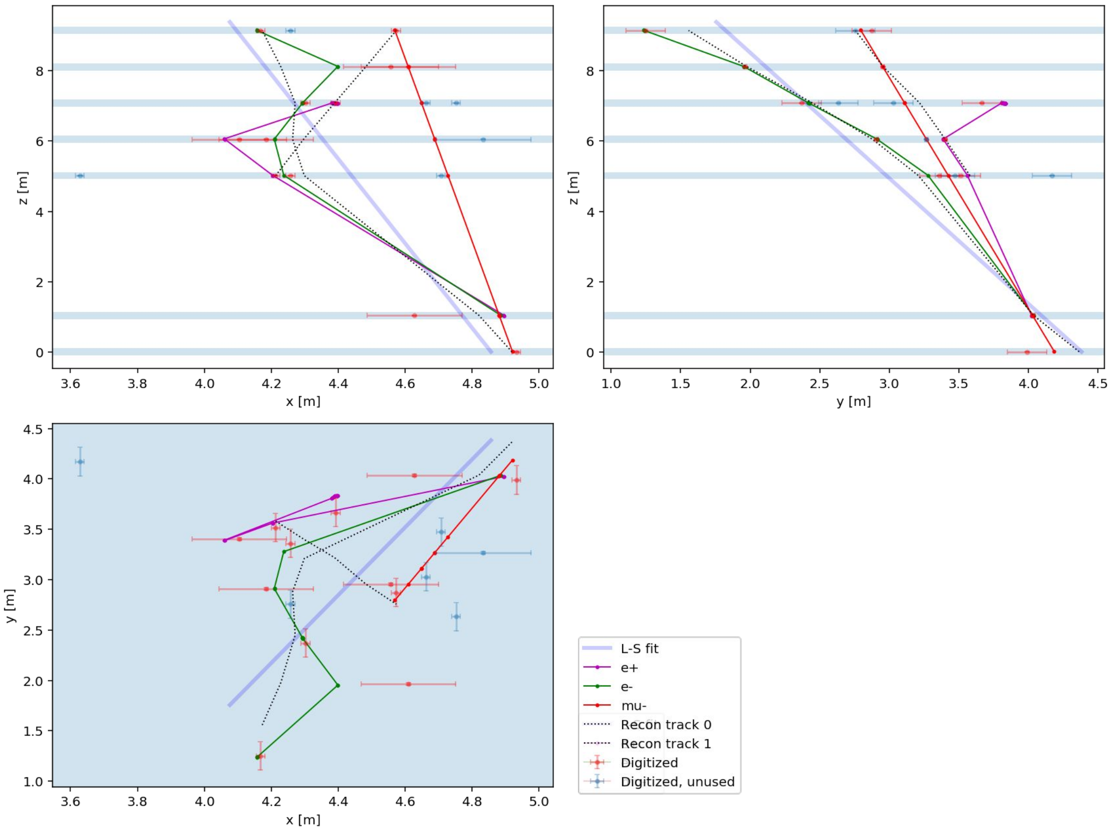
0b0x1x0x**0**
0b0x0x1x**0**
0b0x1x1x**0**
0b1x0x1x**0**

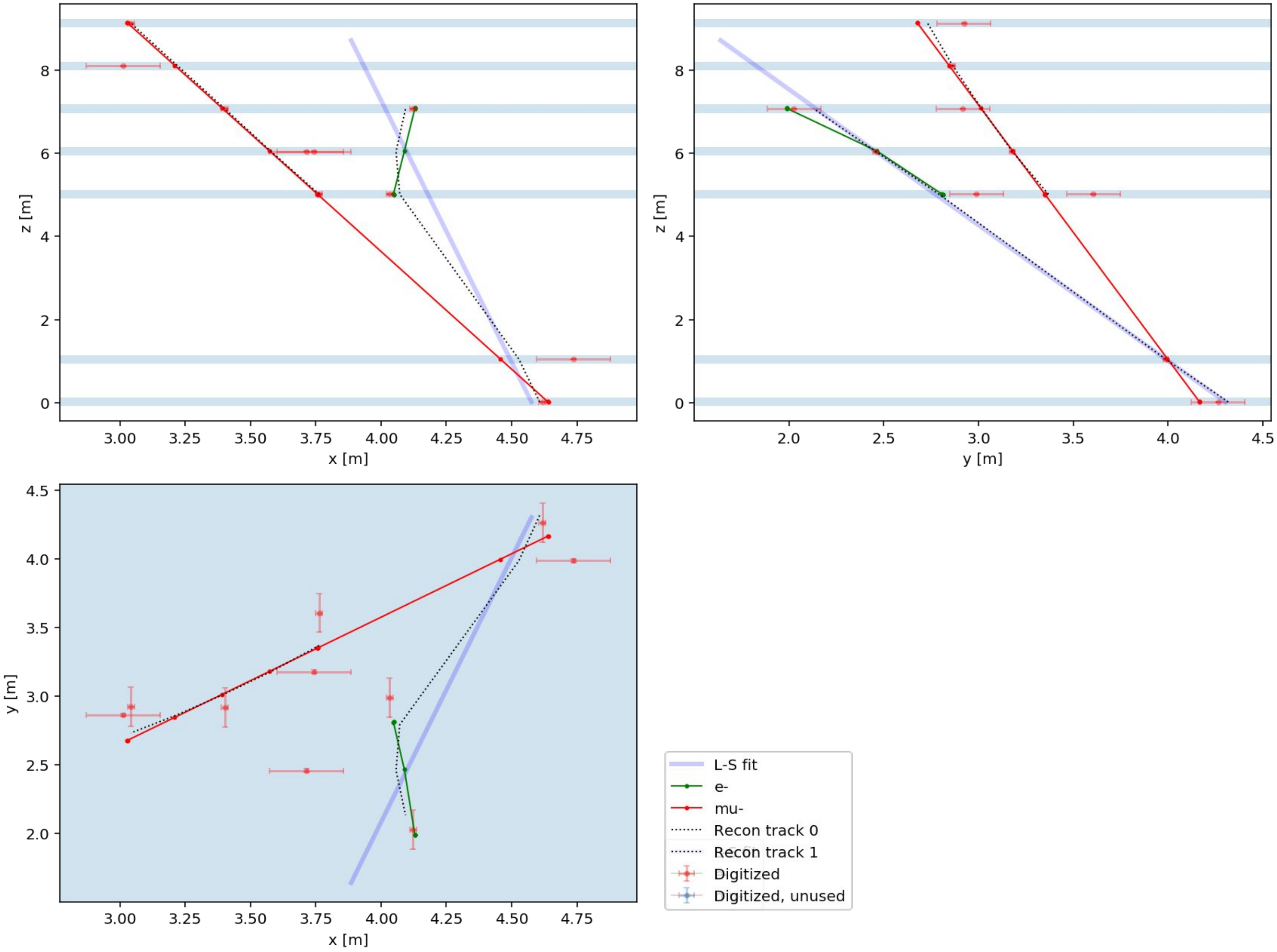
Applying a selection on the offset indices:



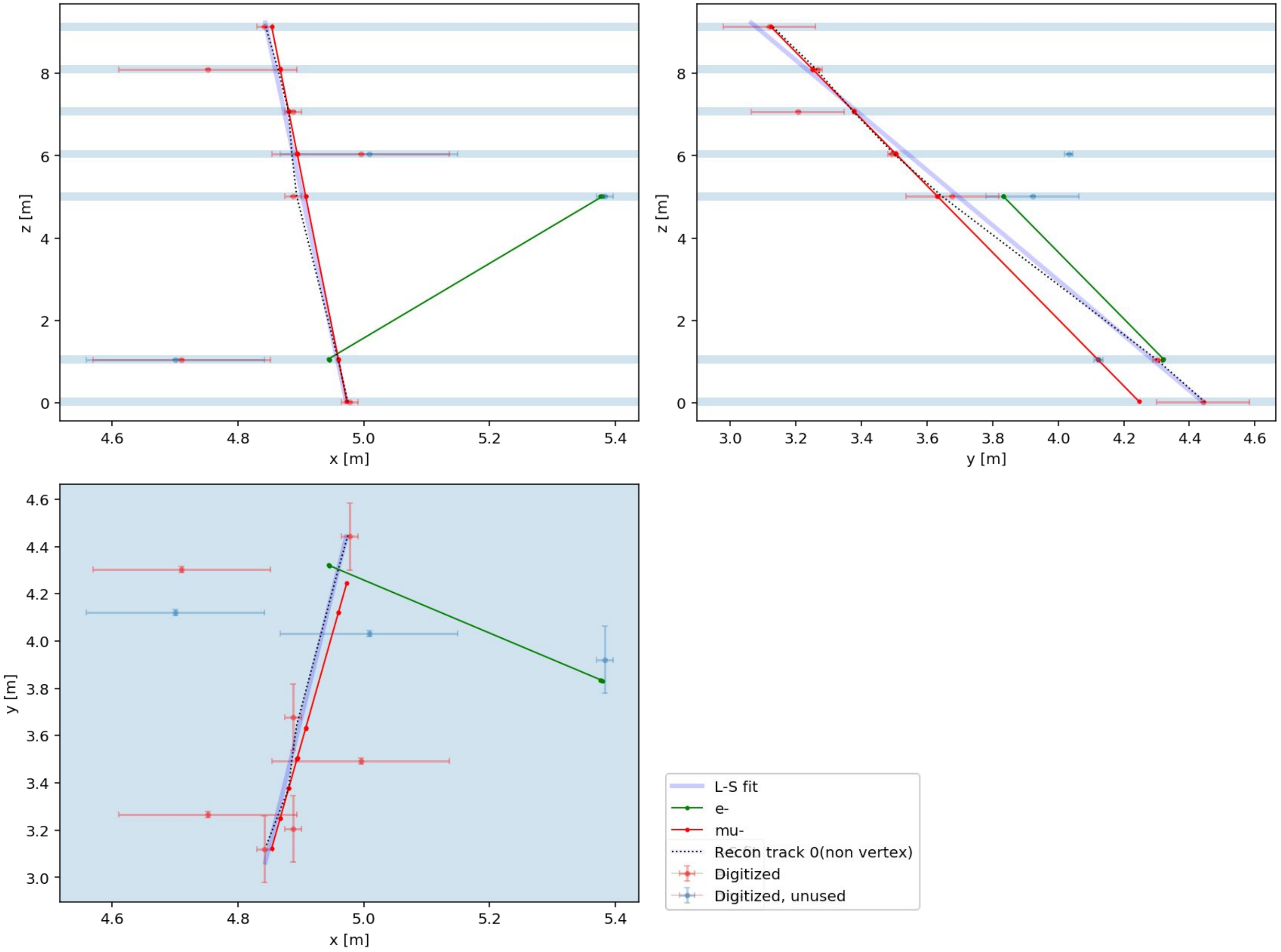


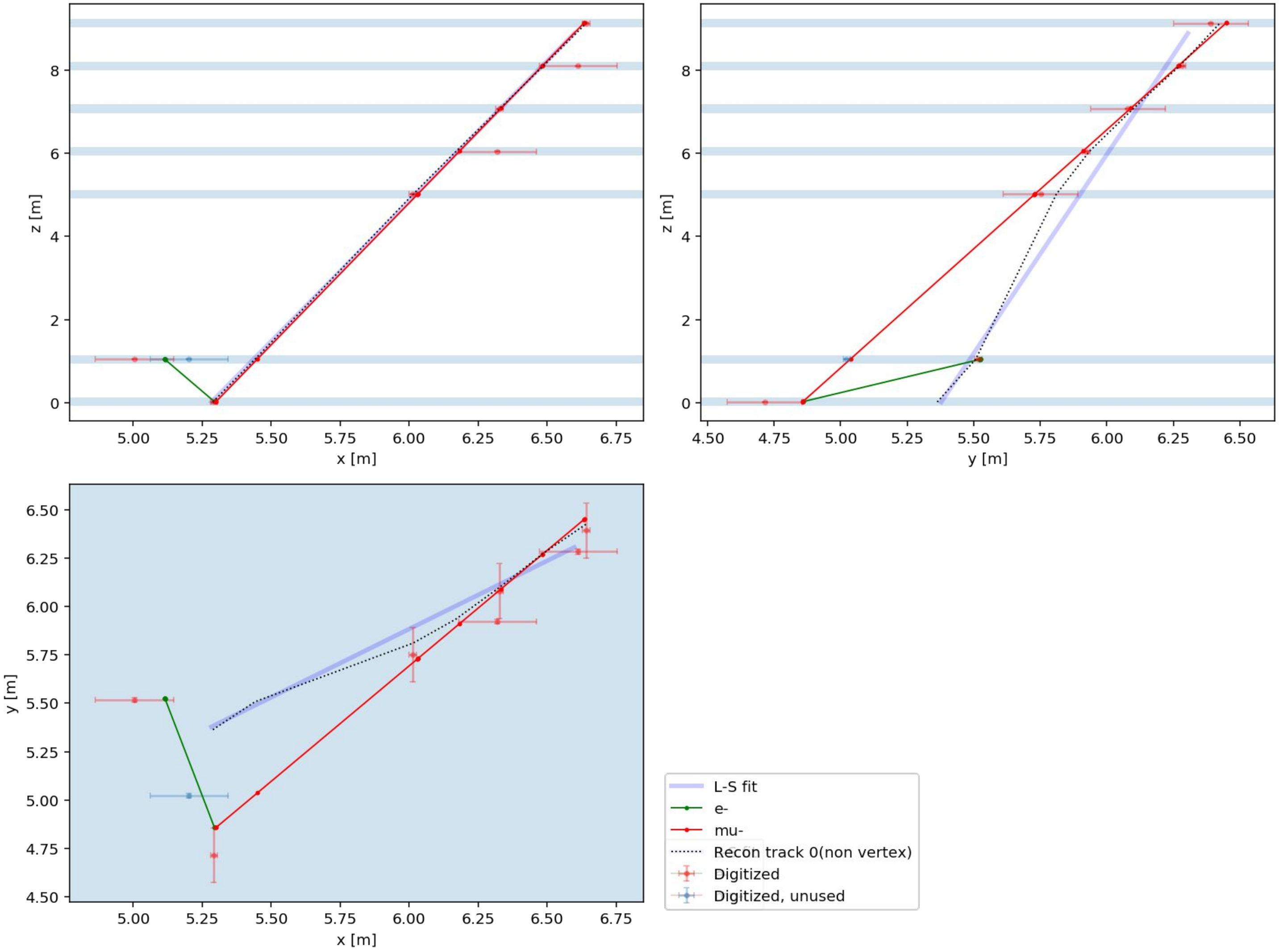
Example of muon events that failed pattern recognition





z [m]
x [m]
y [m]
L-S fit
e-
mu-
Recon track 0
Recon track 1
Digitized
Digitized, unused


z [m]
x [m]
y [m]
4.6
4.8
5.0
5.2
5.4
3.0
3.2
3.4
3.6
3.8
4.0
4.2
4.4
4.6
0
2
4
6
8
L-S fit
e-
mu-
Recon track 0(non vertex)
Digitized
Digitized, unused


z [m]
x [m]
z [m]
y [m]
y [m]
x [m]
L-S fit
e-
mu-
Recon track 0(non vertex)
Digitized
Digitized, unused