Tracking Fit Studies

Angelo Santos - 2020/06/24

SPRACE

Recent Results

Track Reconstruction

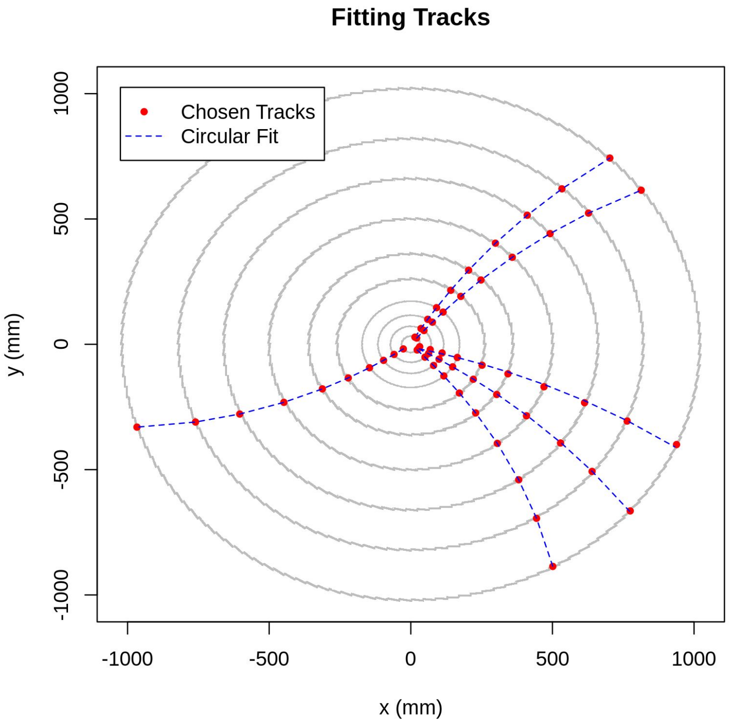
- Dataset (30k tracks)
  - $p_T > 1$ GeV/c
  - $|\eta| < 0.5$
  - no ϕ cut
  - 10 hits per track
- Machine Learning Method
  - BDT-GBM ((x,y,z) coordinates)
  - 5th hit from previous 4 ones
  - Samples
    - Training: 21k tracks (70%)
    - Optimization: 3k tracks (10%)
    - Application: 6k tracks (20%)

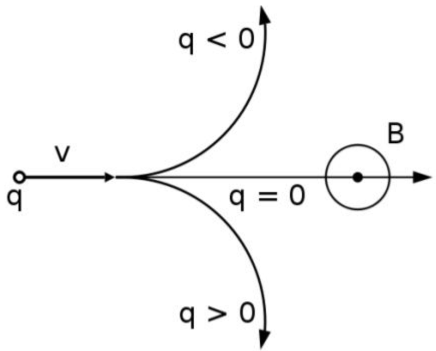
Discussing about Failures

- Maybe failures are not that bad
- Failures to find a perfect track (according the Kaggle challenge)
  - does not mean a failure to find the particle features
  - and may be only due hits (from different tracks) very close to each other
- We can check it
  - comparing pT values between the dataset and a circle fit on the track

Update Dataset	Layers	5	6	7	8	9	10	Whole Track
Yes	Success	5026 (84%)	4815 (80%)	4626 (77%)	4272 (71%)	3793 (63%)	3552 (59%)	2411 (40%)
	Failure	974 (16%)	1185 (20%)	1374 (23%)	1728 (29%)	2207 (37%)	2448 (41%)	3589 (60%)
No	Success	5125 (85%)	5084 (85%)	5017 (84%)	4820 (80%)	4445 (74%)	4426 (74%)	3280 (55%)
	Failure	875 (15%)	916 (15%)	983 (16%)	1180 (20%)	1555 (26%)	1574 (26%)	2720 (45%)

Tracking Fit





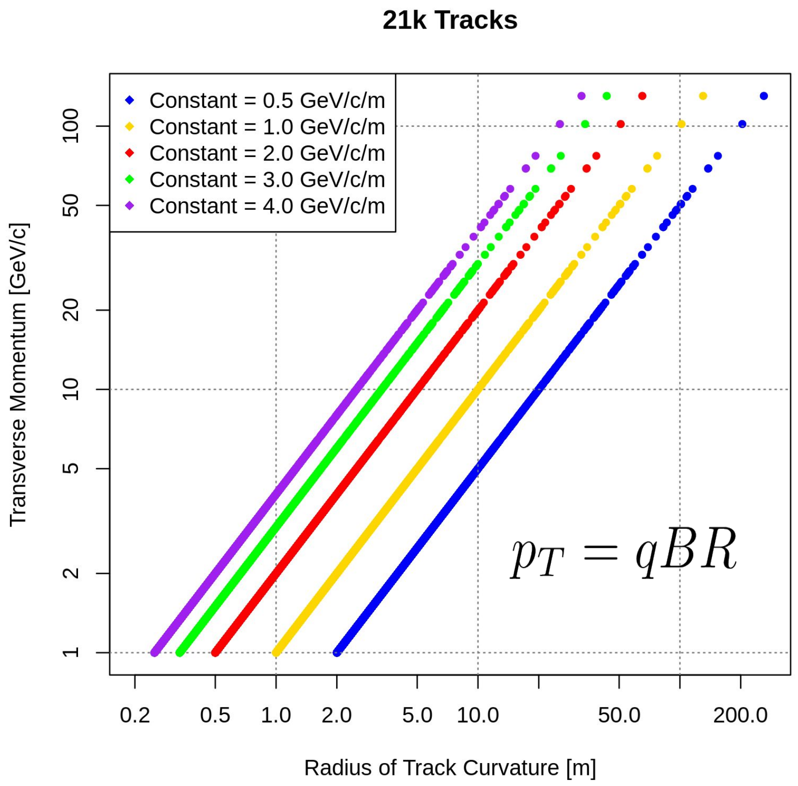
$$p(x,y) = \sqrt{p_x^2 + p_y^2} \equiv p_T$$

$$p_T = qBR$$

$$\frac{p_T}{R} = q.B = constant$$

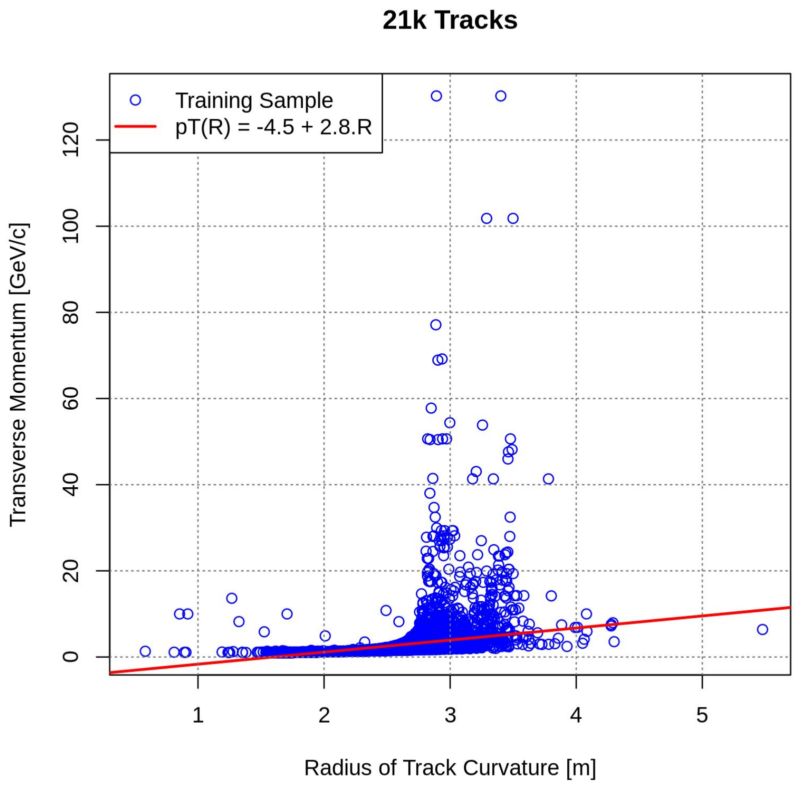
Constant value would come from a fit in pT vs. R

Phenomenology Expectations



- What would we expect to get from?
  - A line consistently increasing
  - A smooth line
  - Highest radius less than 300 m
  - The higher pT, less points

Unexpected Result from a Circular Fit



- But real life is hard
- Two different scenarios
  - Points
    - consistently increasing
    - in a smooth line
  - Points
    - insensitive to the formers
    - without any structure
- Highest radius
  - lower than 6 m
- Constant = 2.8 GeV/c/m
  - Highest expected radius: 40 m

Unexpected Tracking Fit: Hypophysis

- Failure to fit tracks
  - R library (circular) may not be good enough
  - Would be the best scenario (unfortunately)

- Presence of outliers
  - But are not expected to be so much

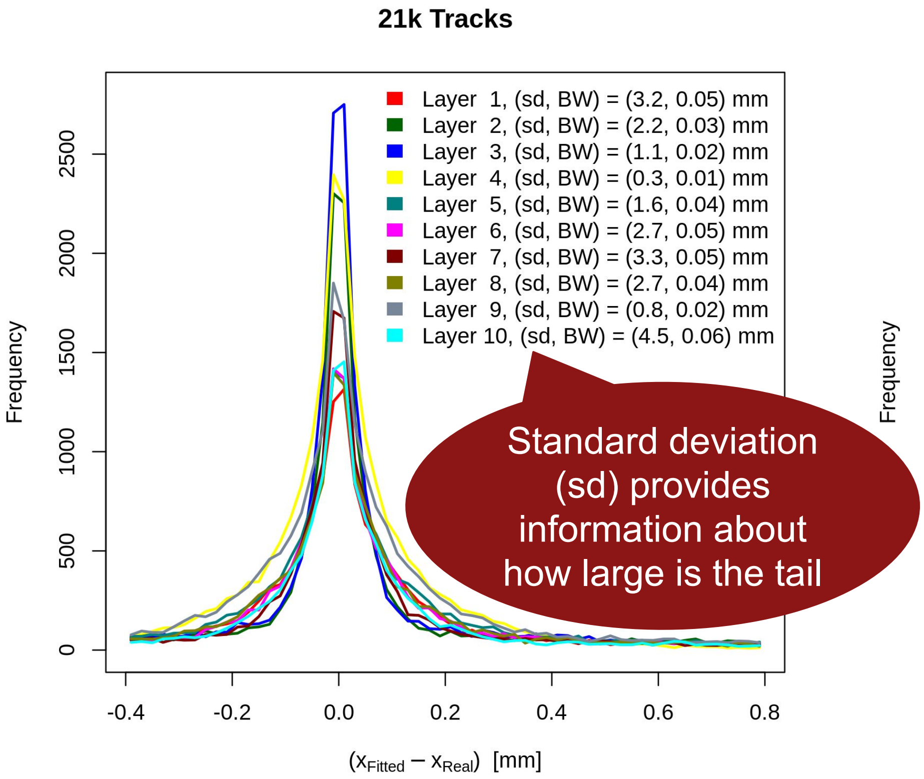
- The constant is not a constant
  - It is a residual assumption
  - But must to be considered if the two first hypophysis are not confirmed

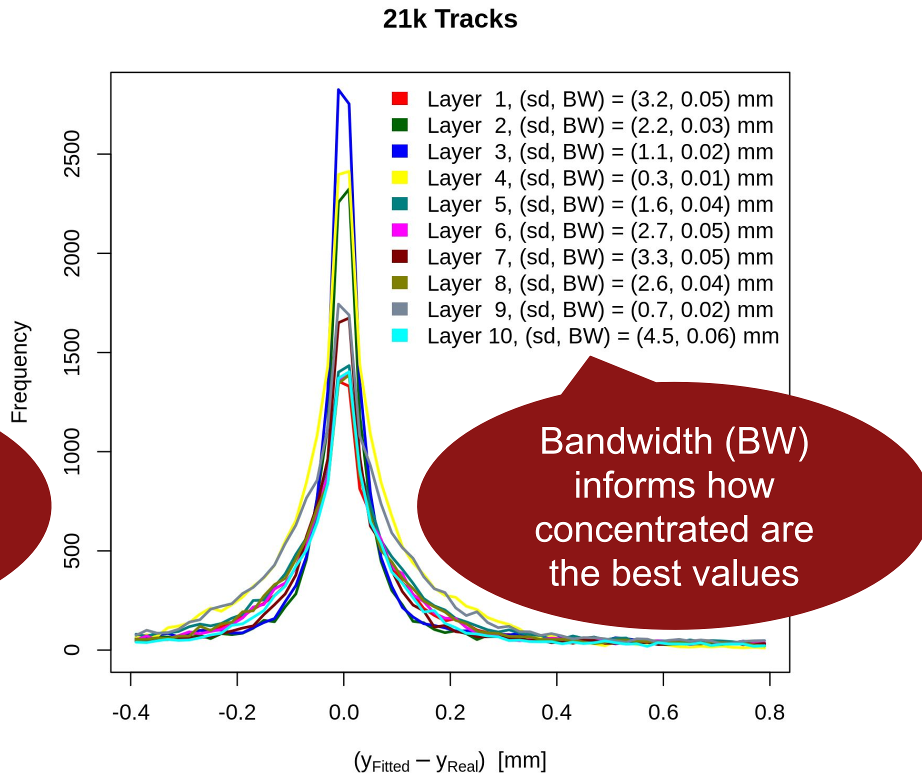
Checking Fit

```
# R
> library(circular)
```

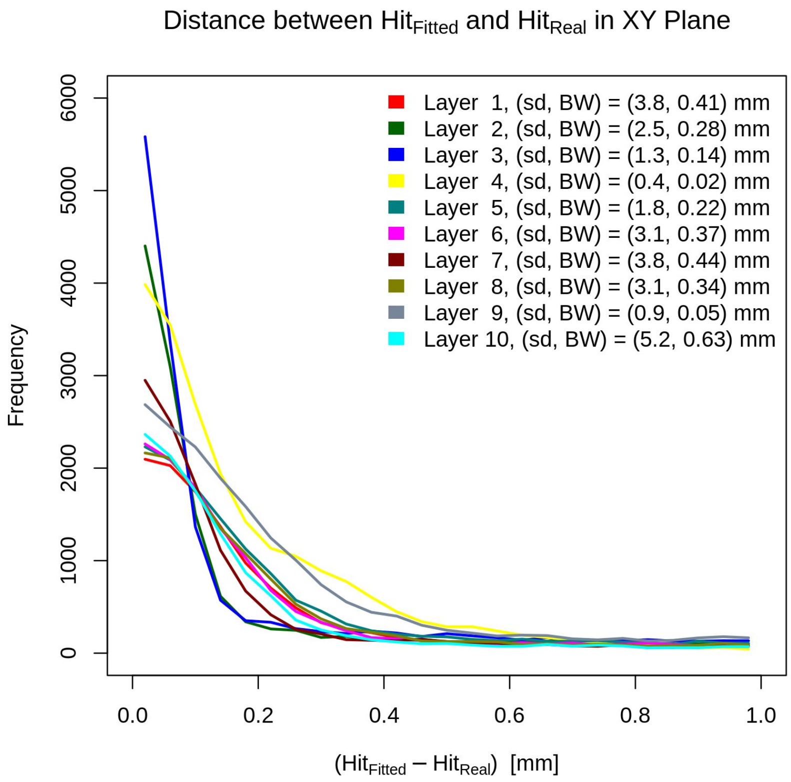
- It is mandatory to analyze how far/close
    - points from tracking fit
    - are from real hits

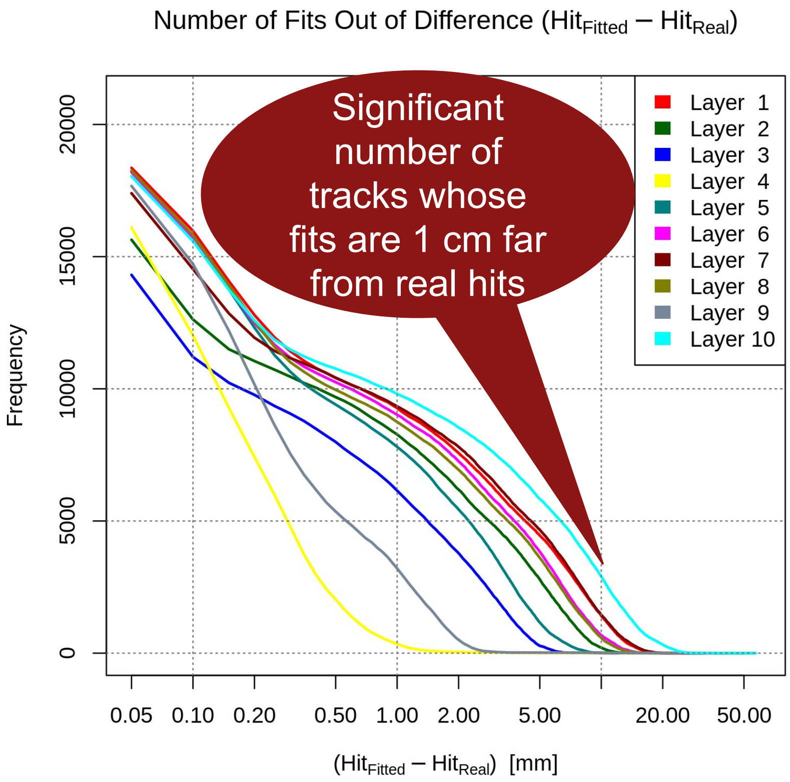
X-Y: Distance between Fitted and Real Points



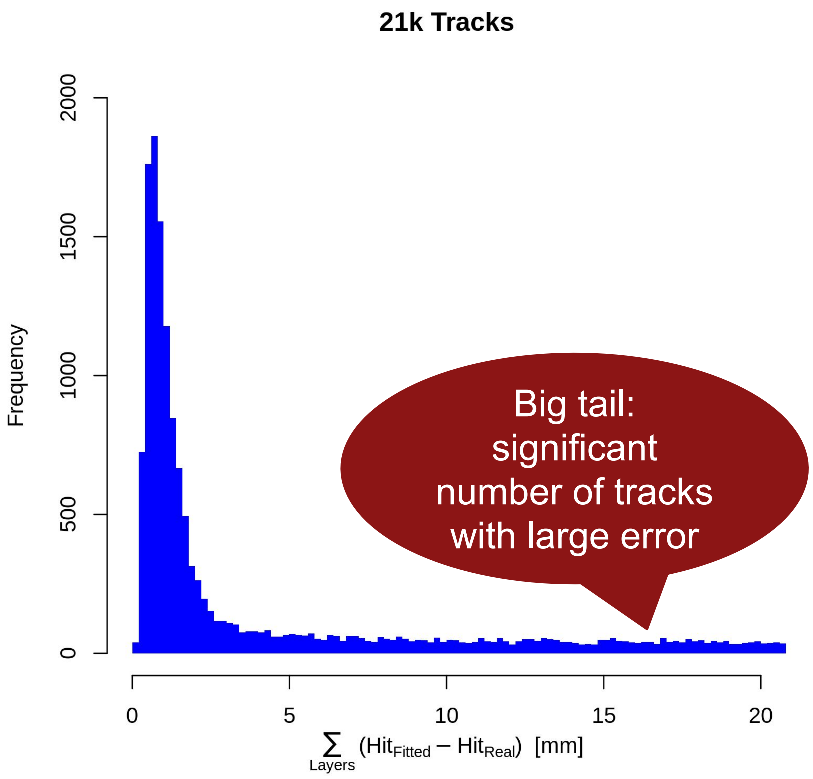


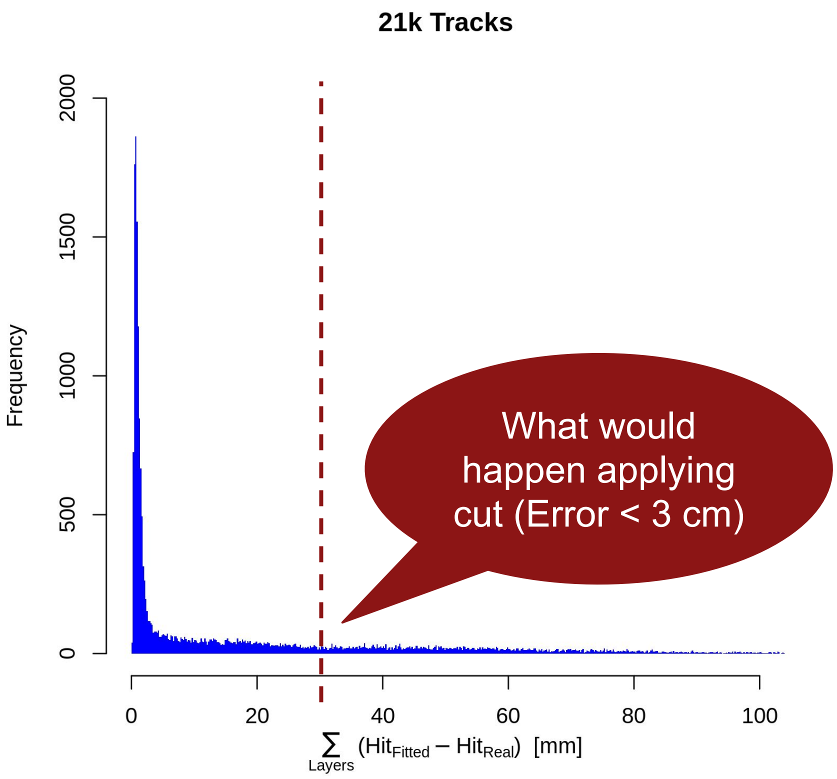
2D Distance between Fitted and Real Points

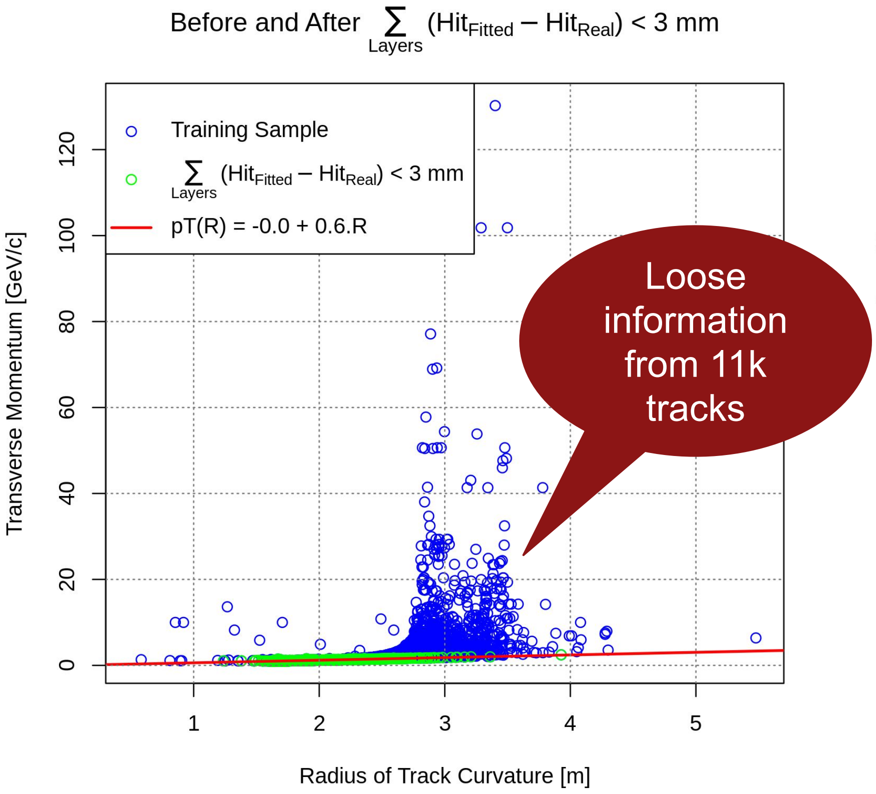




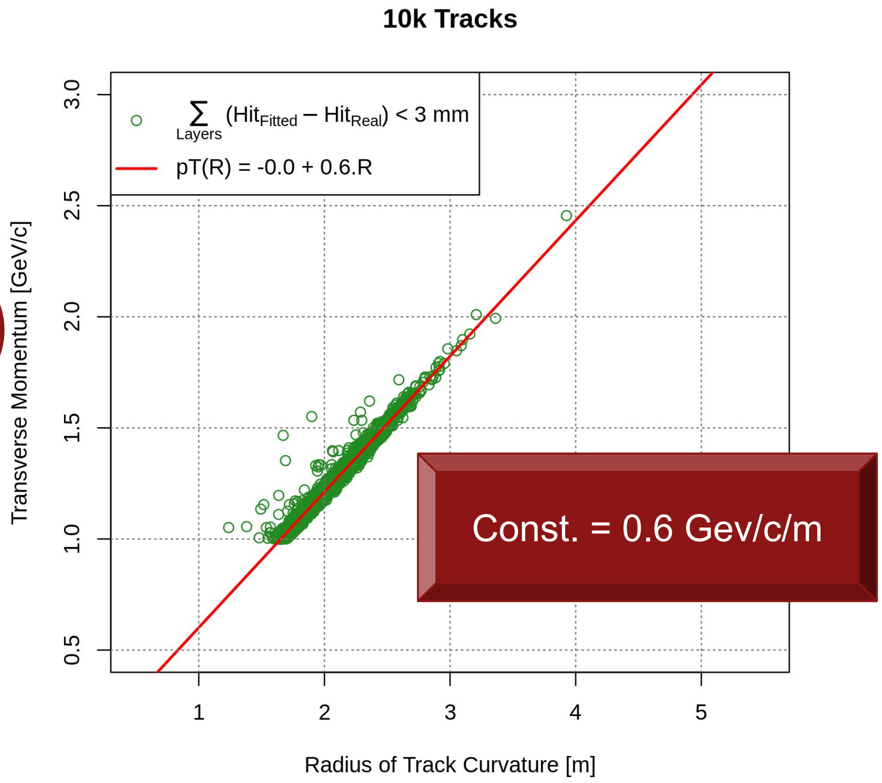
Summing Errors over All Hits in Track





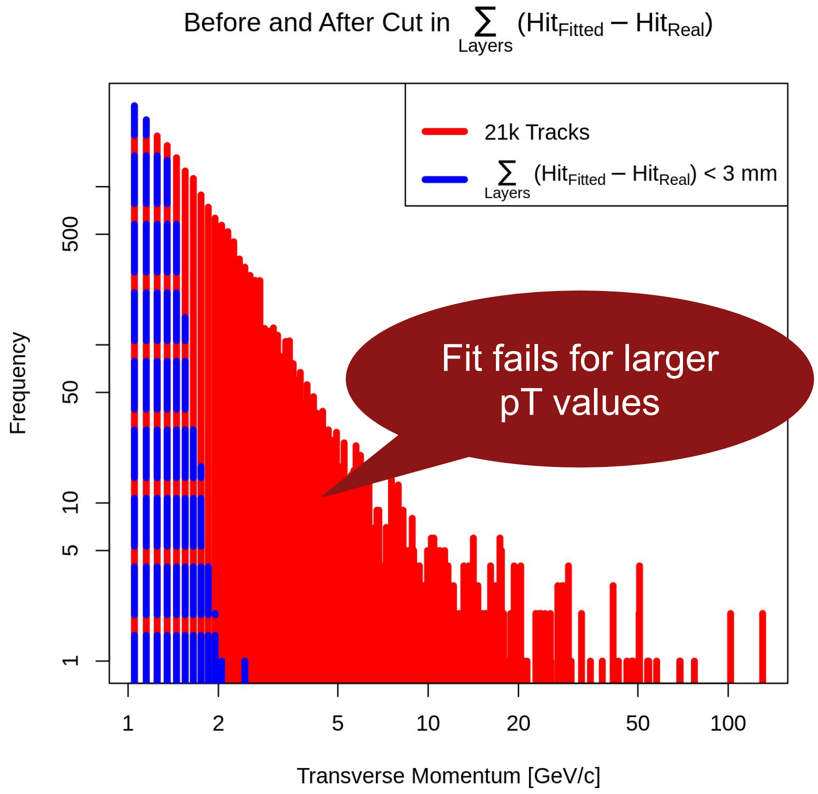


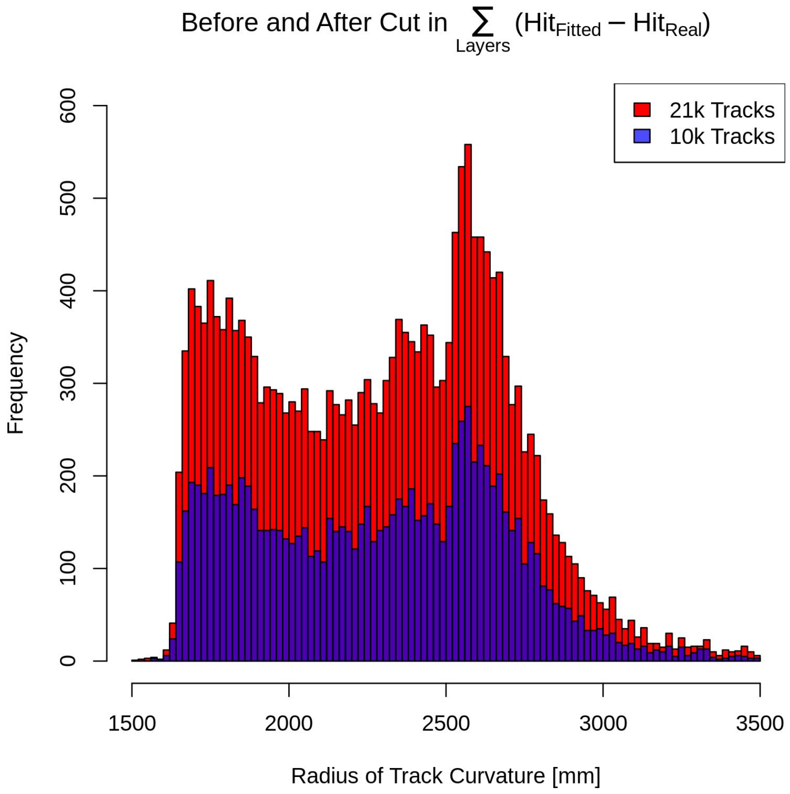
Before and After $\sum_{Layers} (Hit_{Fitted} - Hit_{Real}) < 3$ mm
Training Sample
$\sum_{Layers} (Hit_{Fitted} - Hit_{Real}) < 3$ mm
pT(R) = -0.0 + 0.6.R
Loose information from 11k tracks
Transverse Momentum [GeV/c]
Radius of Track Curvature [m]
0
20
40
60
80
100
120
1
2
3
4
5



10k Tracks
$\sum_{Layers} (Hit_{Fitted} - Hit_{Real}) < 3$ mm
pT(R) = -0.0 + 0.6.R
Const. = 0.6 Gev/c/m
Transverse Momentum [GeV/c]
Radius of Track Curvature [m]
0.5
1.0
1.5
2.0
2.5
3.0
1
2
3
4
5

Cutting Out Cases with Larger Fit Error

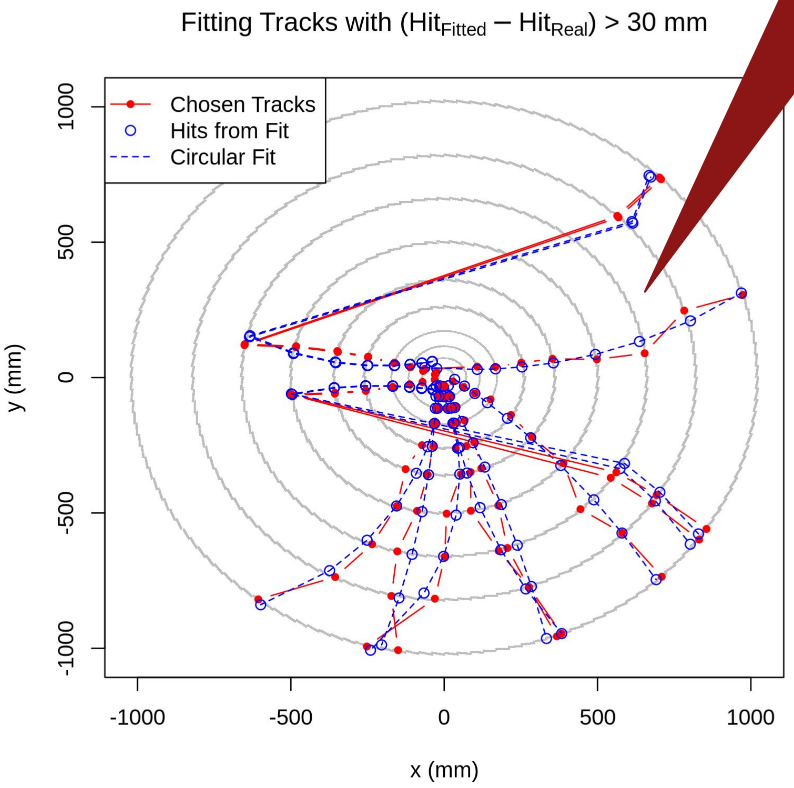


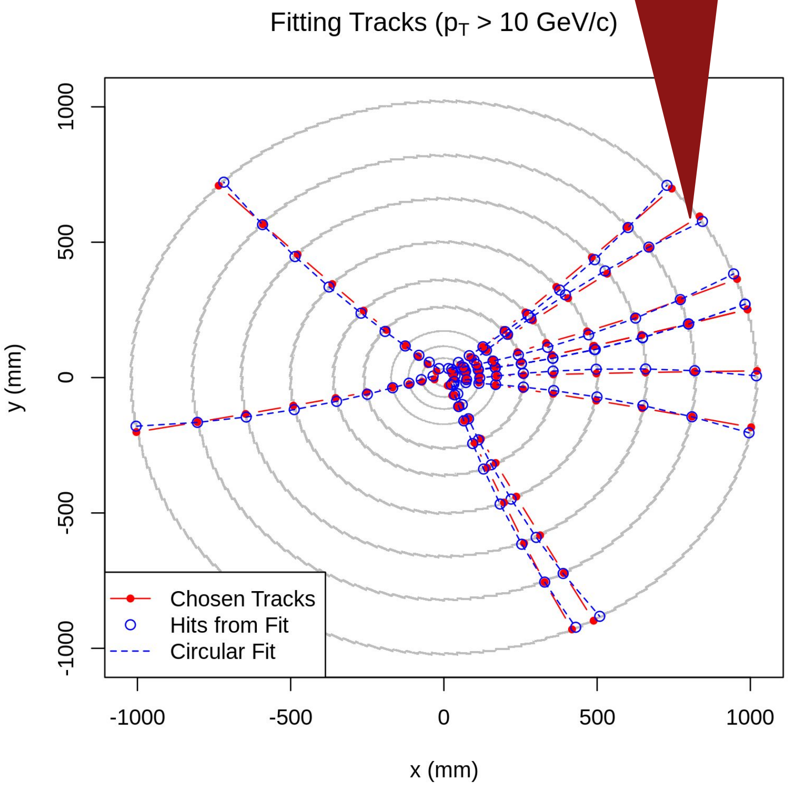


Fit Failures

Yes!
There are some outliers

Fit fails for larger pT values: larger track curvature

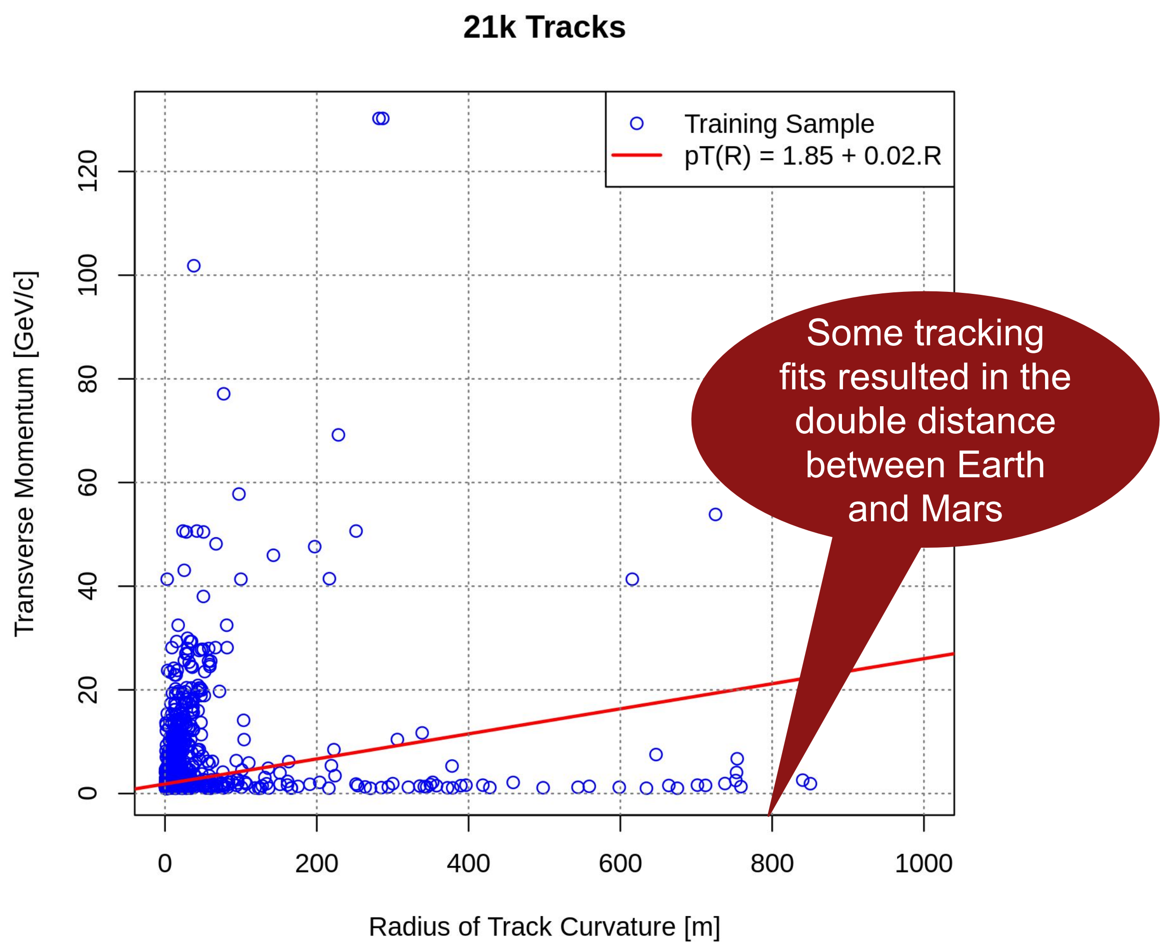




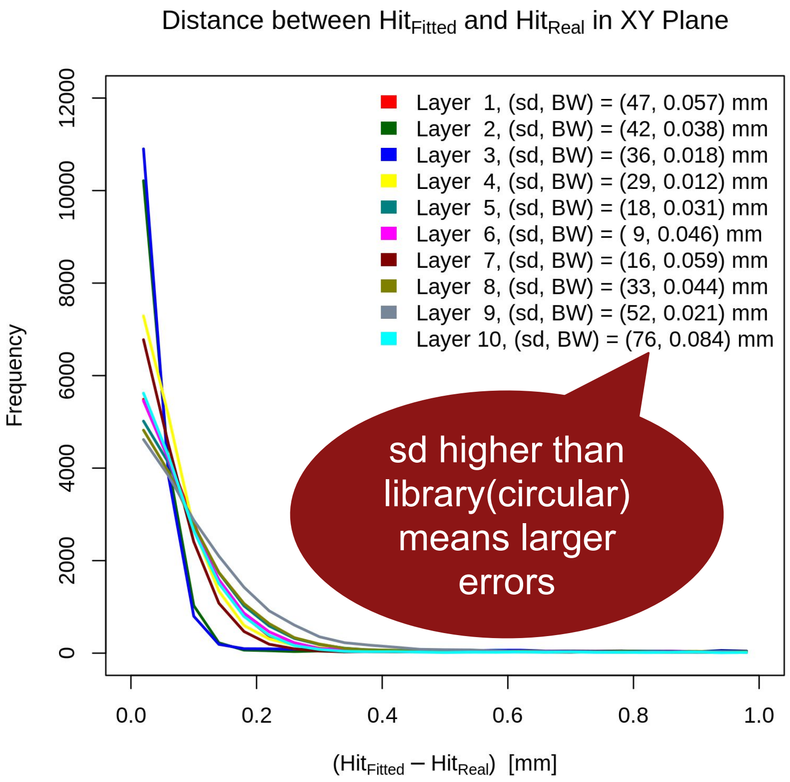
Trying with Other Library

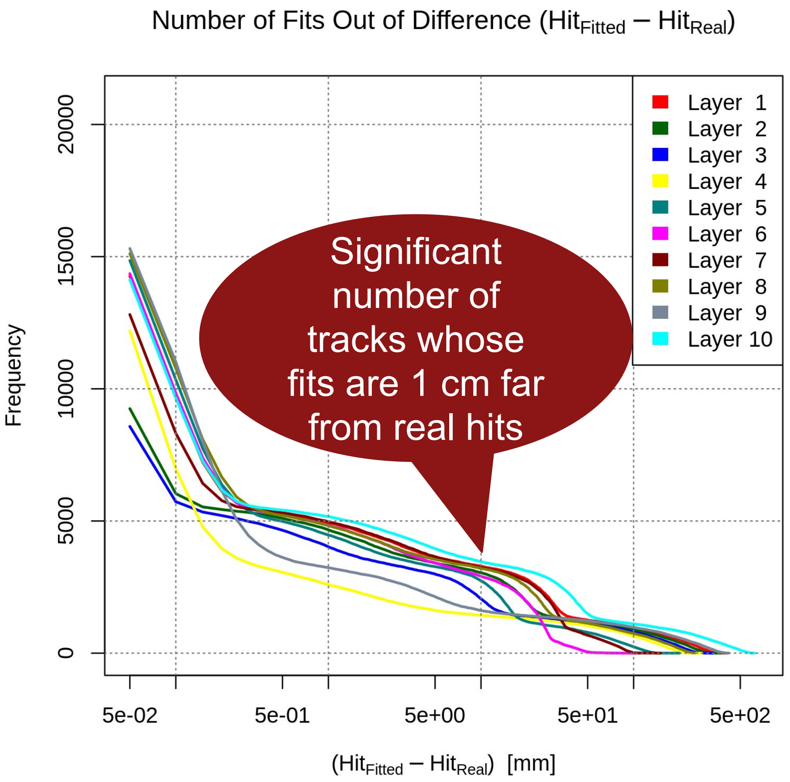
R

> library(pracma)



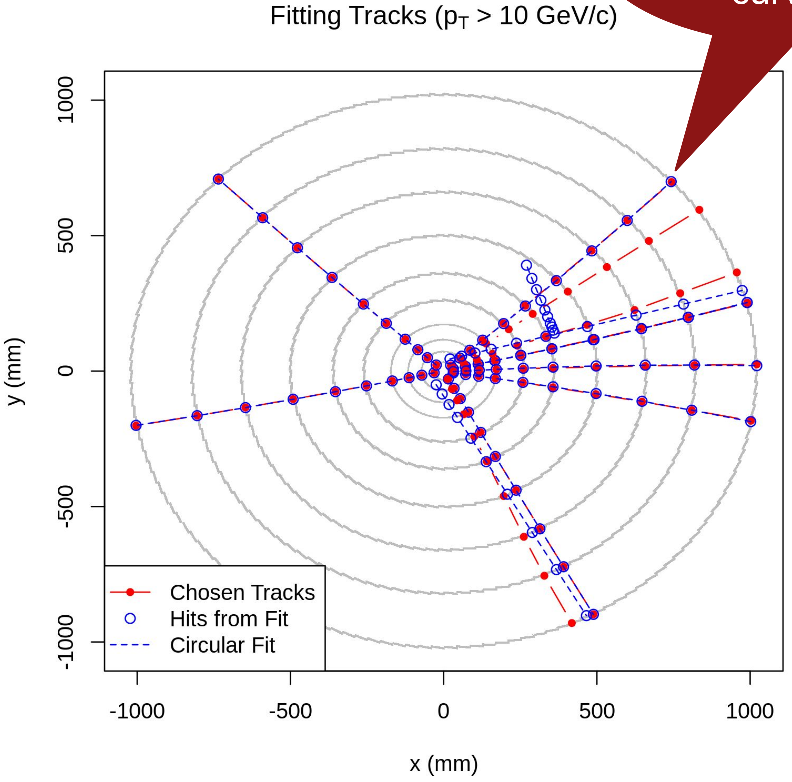
2D Sum of Errors



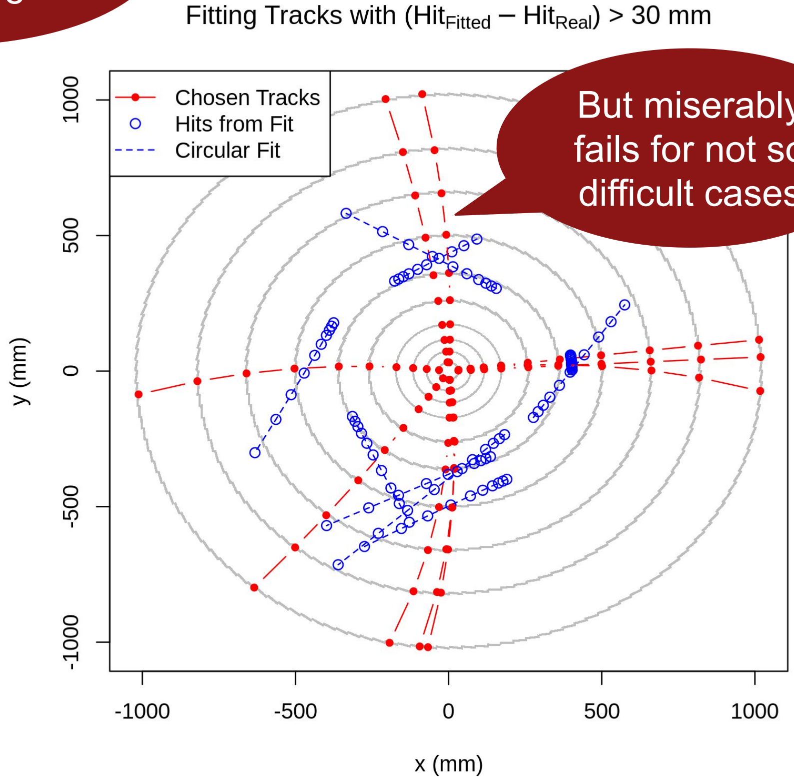


Checking Fits

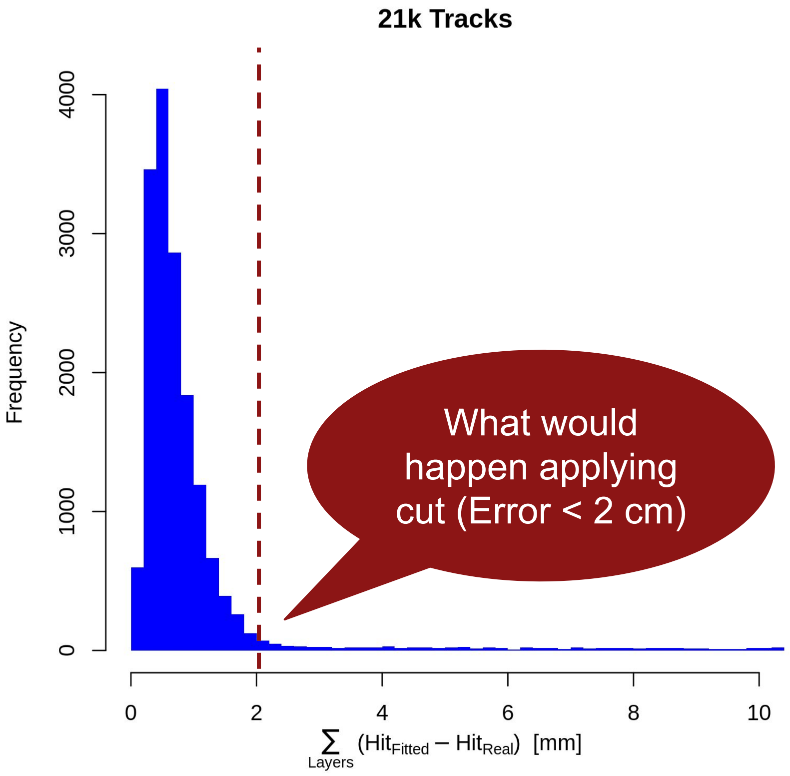
Works fine for larger tracking curvature

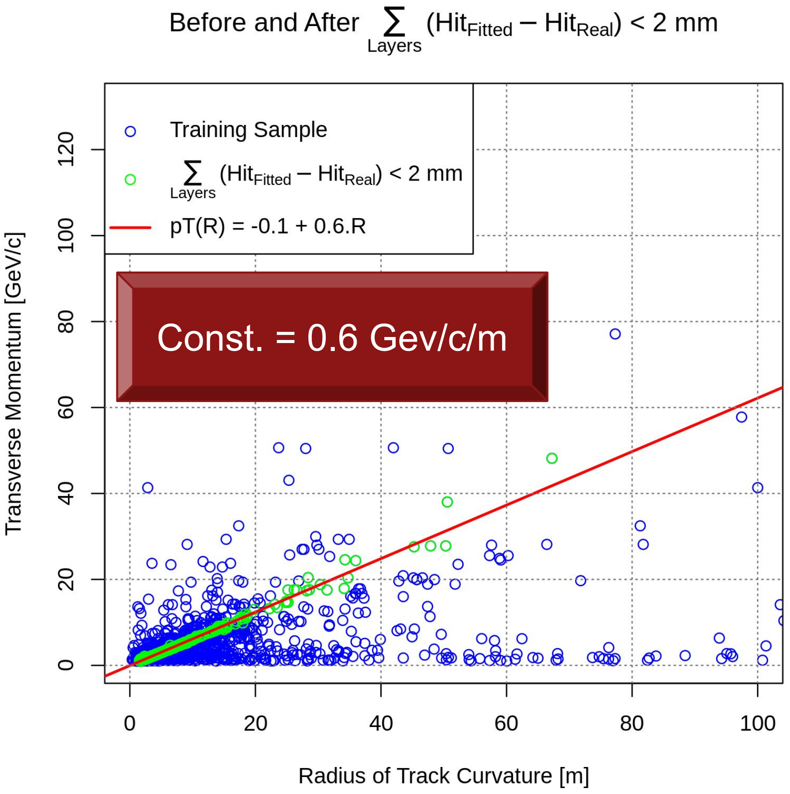


But miserably fails for not so difficult cases

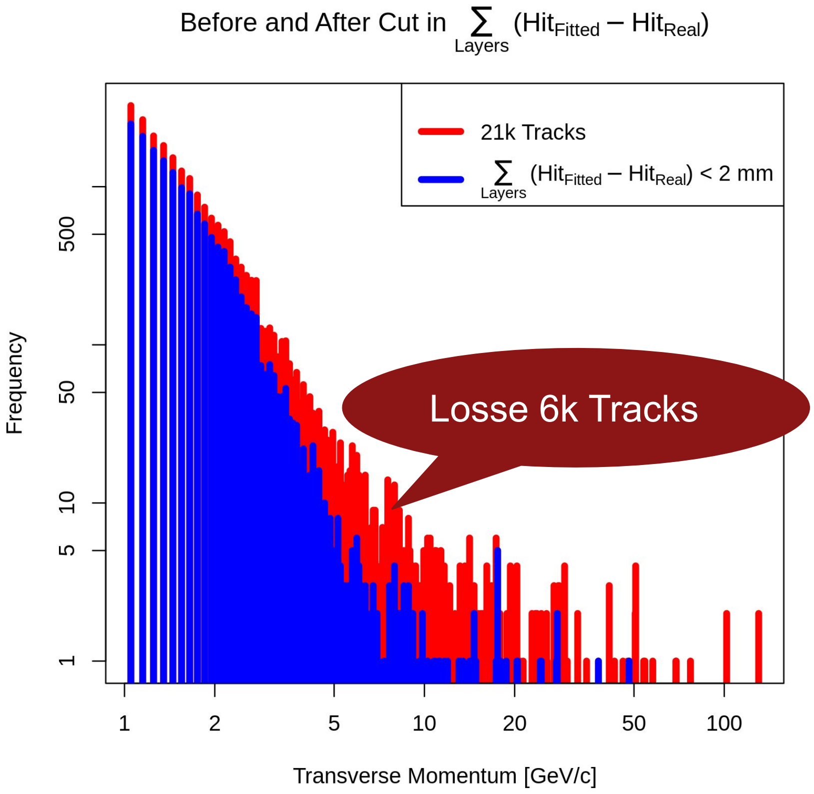


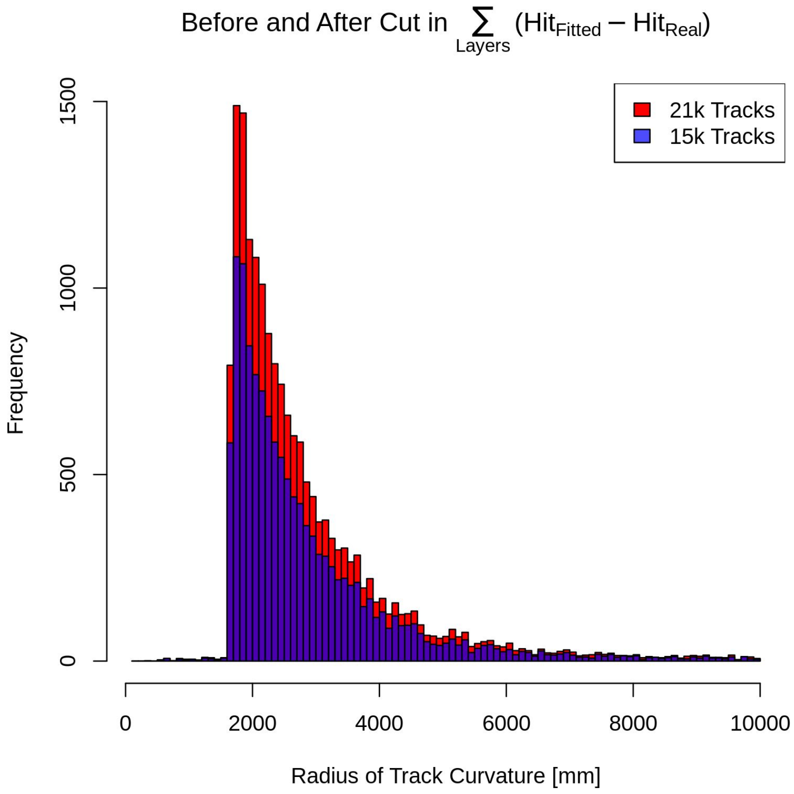
Summing Errors over All Hits in Track





Cutting Out Cases with Larger Fit Error





Solution to Computing Constant

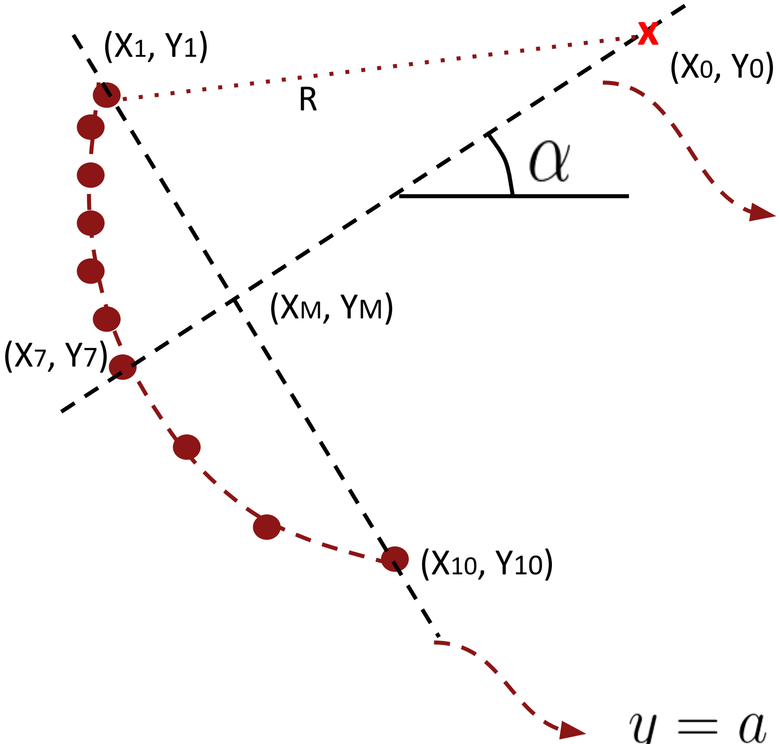
- Library “pracma” is very simple
  - Does not allow to set initial (r, x0, y0)
- Library “circular” is more complete
  - Allow to inform to set initial (r, x0, y0)
  - Providing triplet may be helpful

$$R_i = \frac{p_T}{k_i}, k_i = \underbrace{\{0.2, 0.4, 0.6, \ldots, 4.0\}}_{\text{20 different constant values}} \times 10^{-3} \frac{GeV/c}{mm}$$

20 different constant values

- However
  - Radius is computed from pT
  - What about circumference center (x0, y0)?

Computing Center of the Circumference



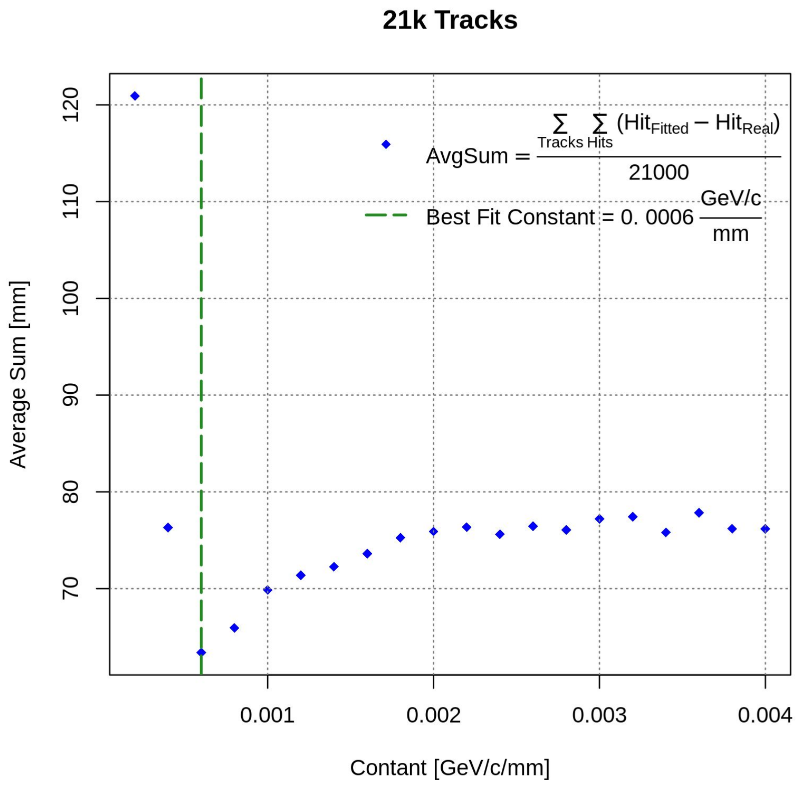
$$x_0 = x_7 \pm R.cos(\alpha)$$
$$y_0 = y_7 \pm R.cos(\alpha)$$

$$y' = a' + b'.x$$

$$b' = arctan(-(\frac{x_{10} - x_1}{y_{10} - y_1}))$$

$$y = a + b.x \quad b = arctan(\frac{y_{10} - y_1}{x_{10} - x_1})$$

Find Best Constant Value

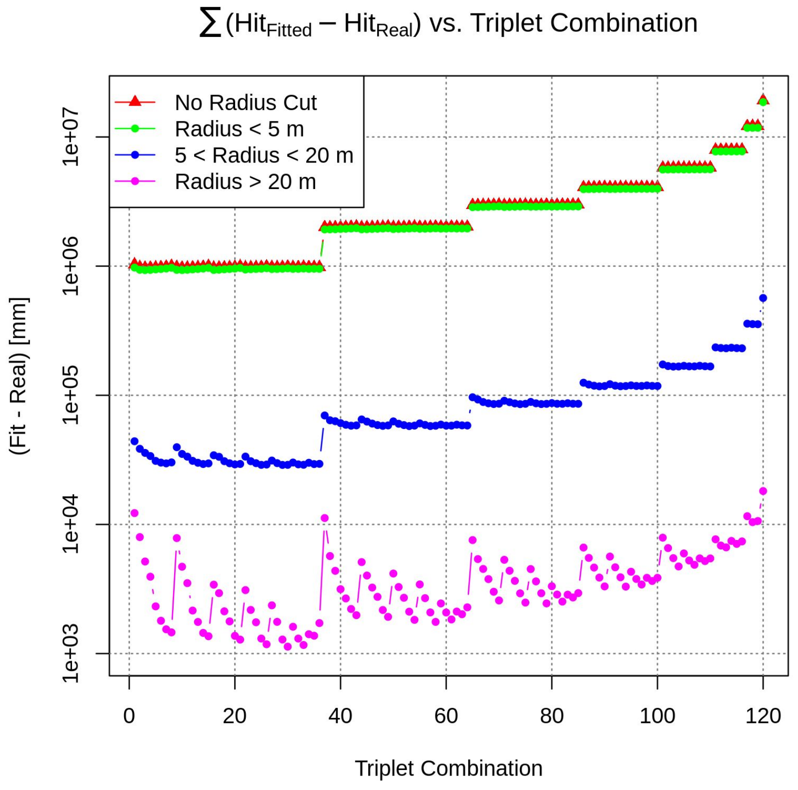


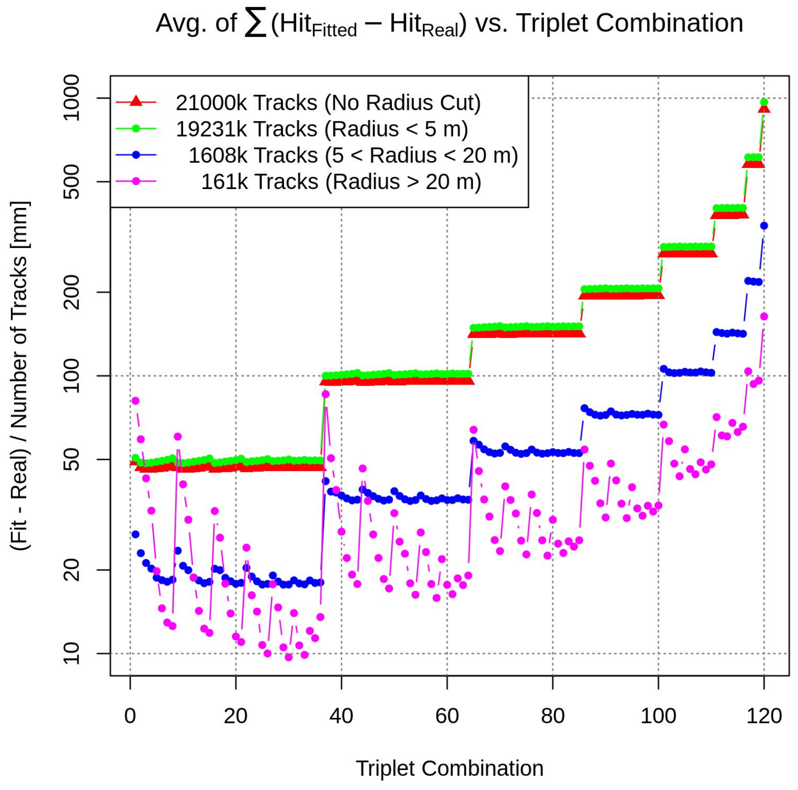
- q.B = 0.6 GeV/c/m
- For High Energy Physics
  - $e = \sqrt{(4\pi\alpha)}.\sqrt{(\hbar c)}$
    $= 0.30282212088\sqrt{(\hbar c)}$

My solution for Tracking Fit

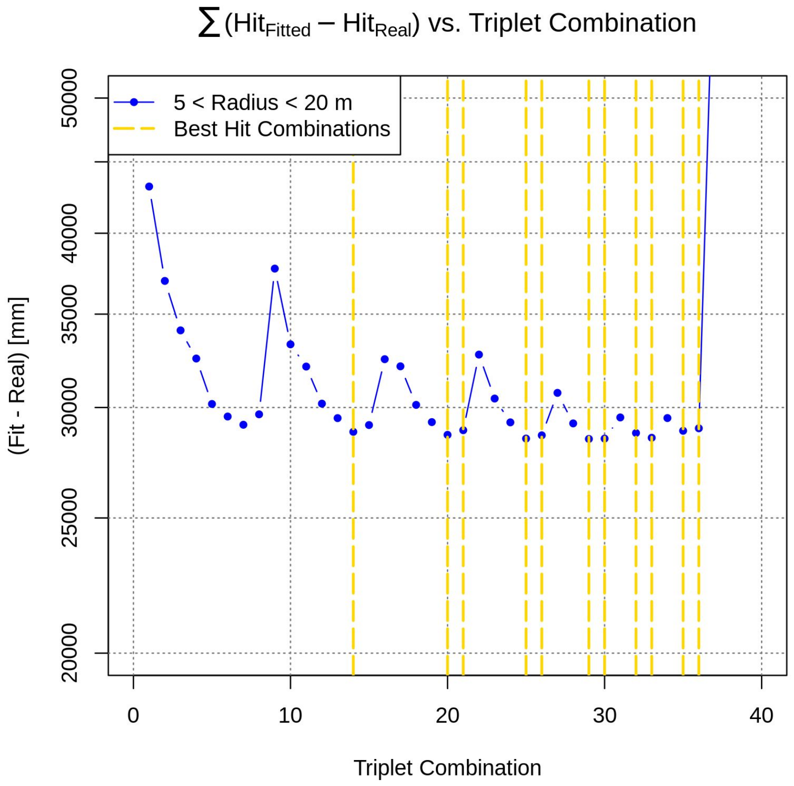
-

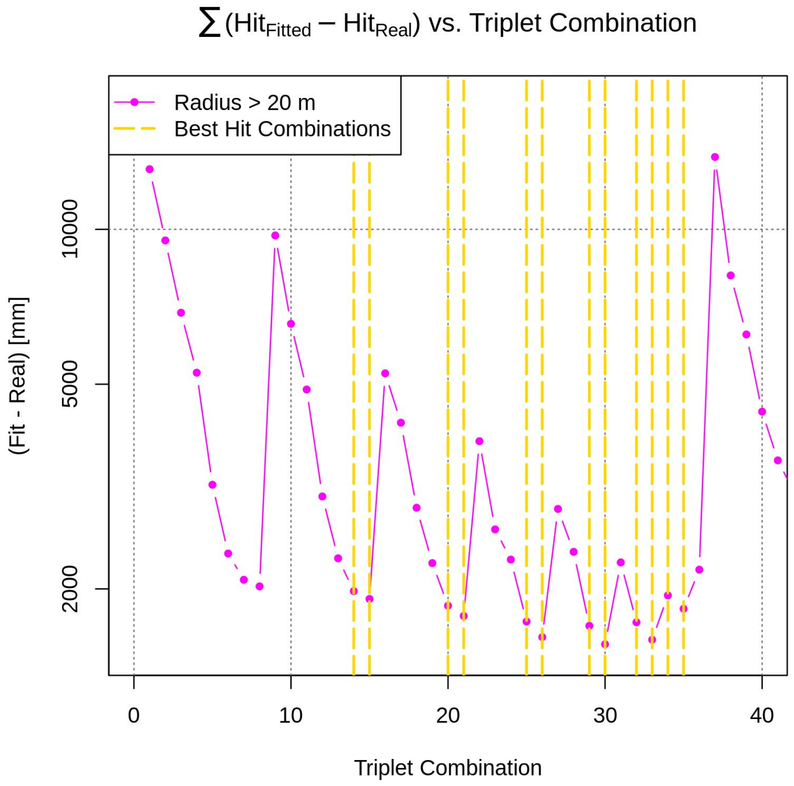
Find Best Triplet Combinations (H_1, H_2, H_3)



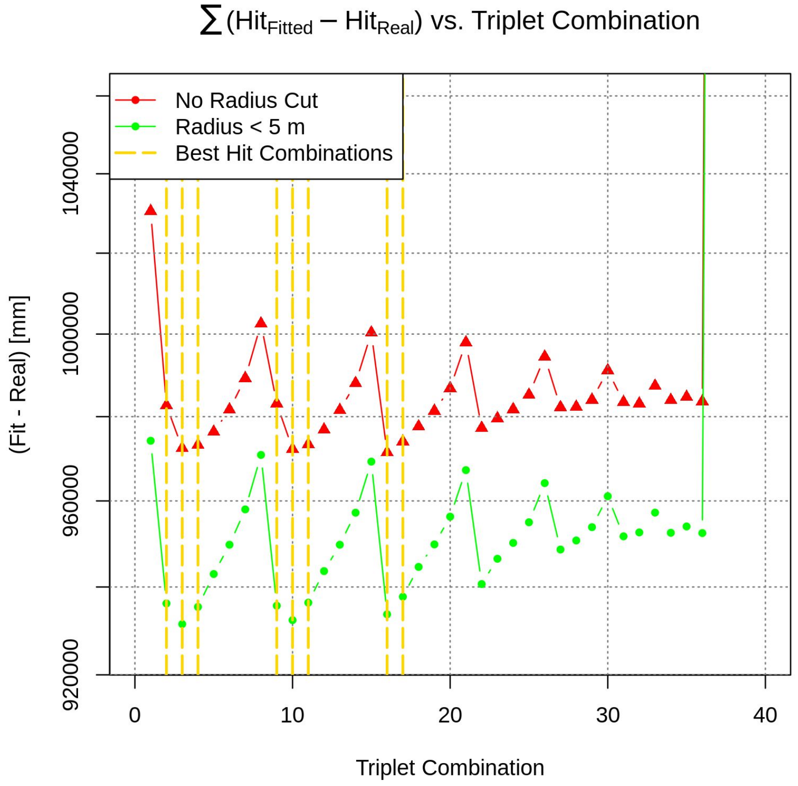


Best Triplet Combinations: Looking Closer (I)



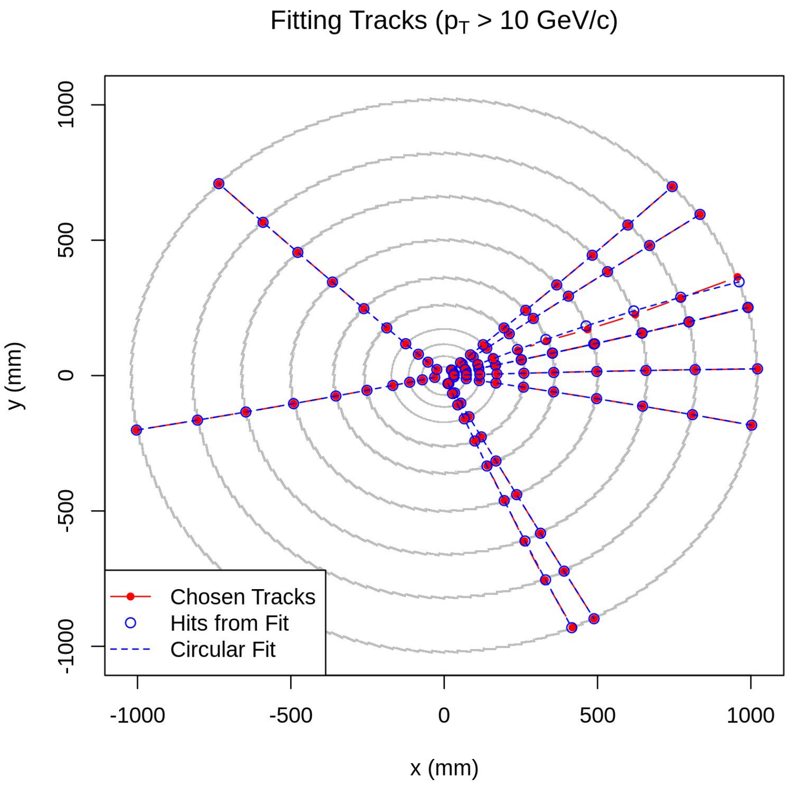


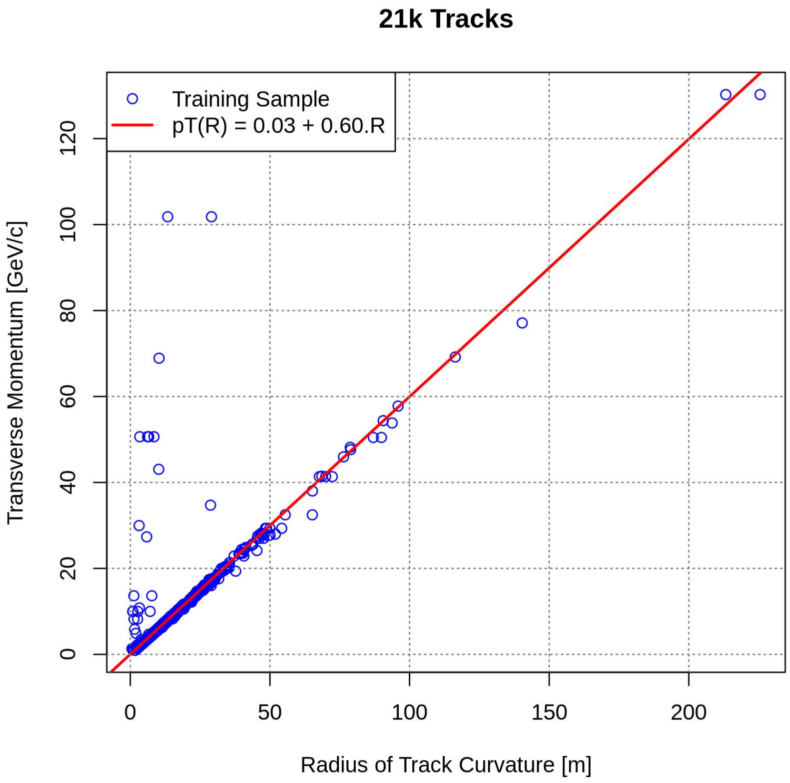
Best Triplet Combination: Looking Closer (II)



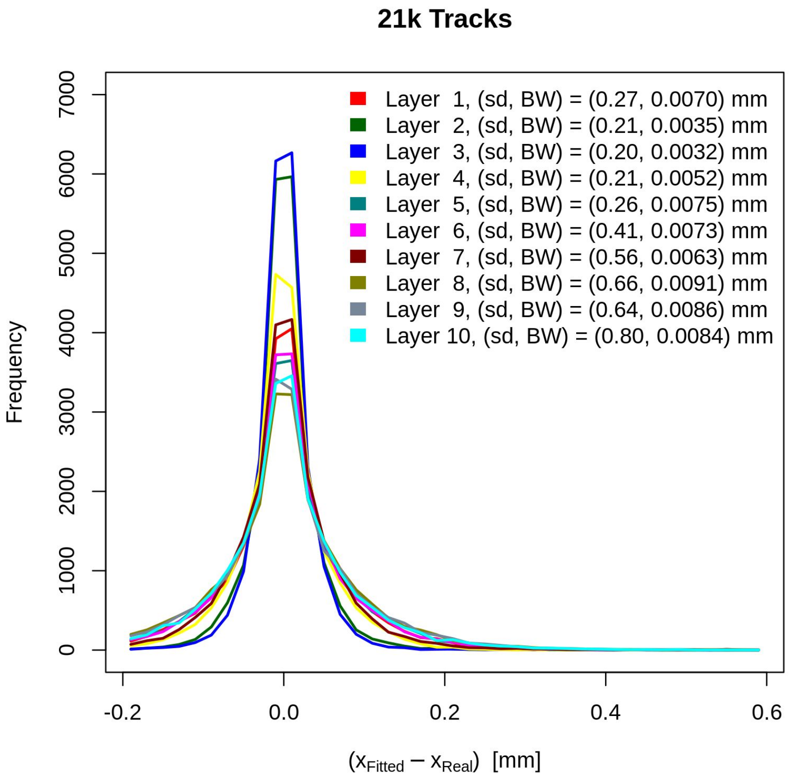
- The idea is to get mean values of (r, x0, y0) using the best triplet combinations
- And use them as input for library “circular”

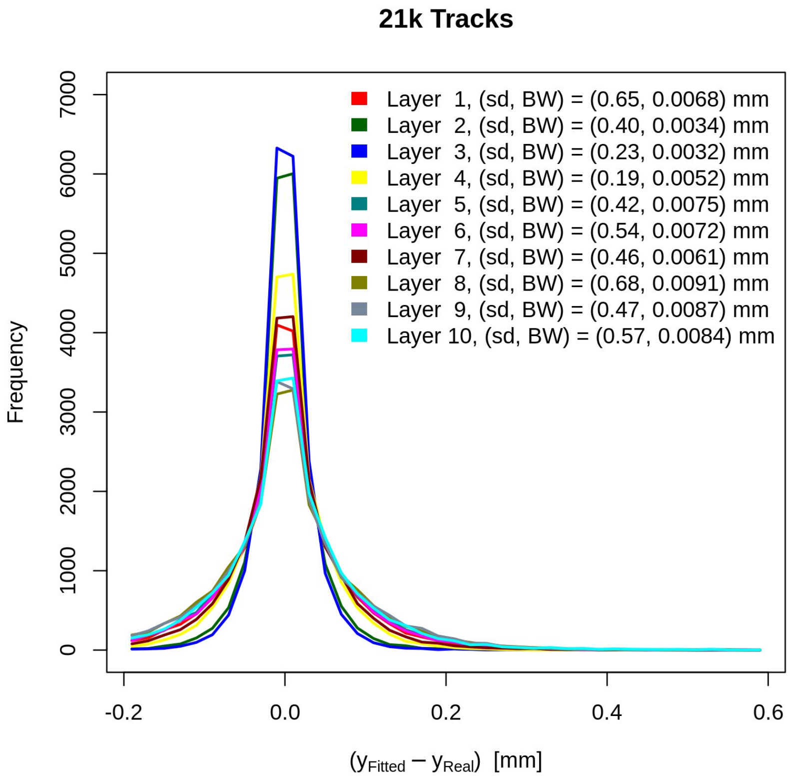
My Solution (I)



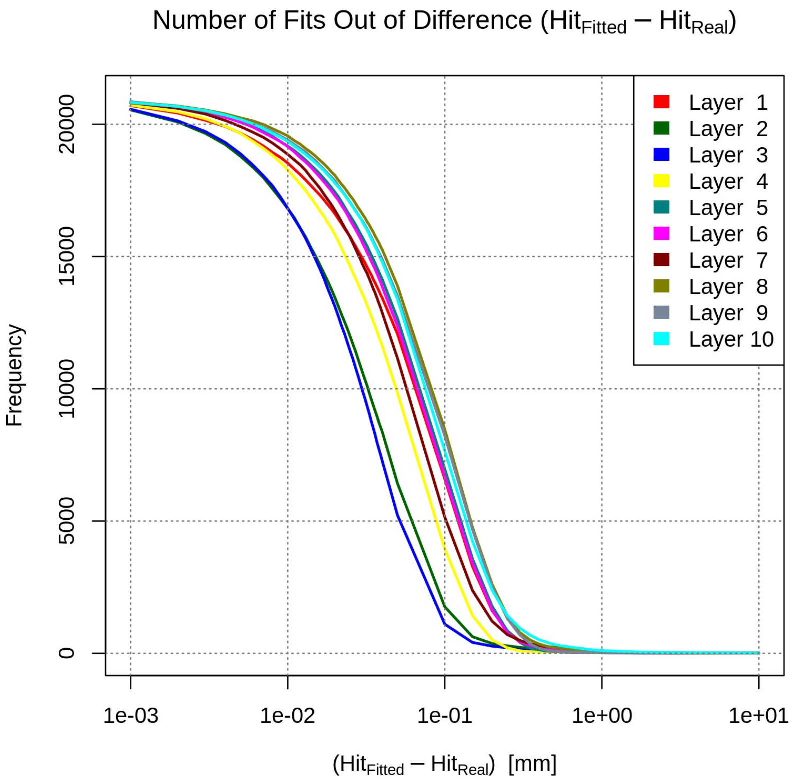


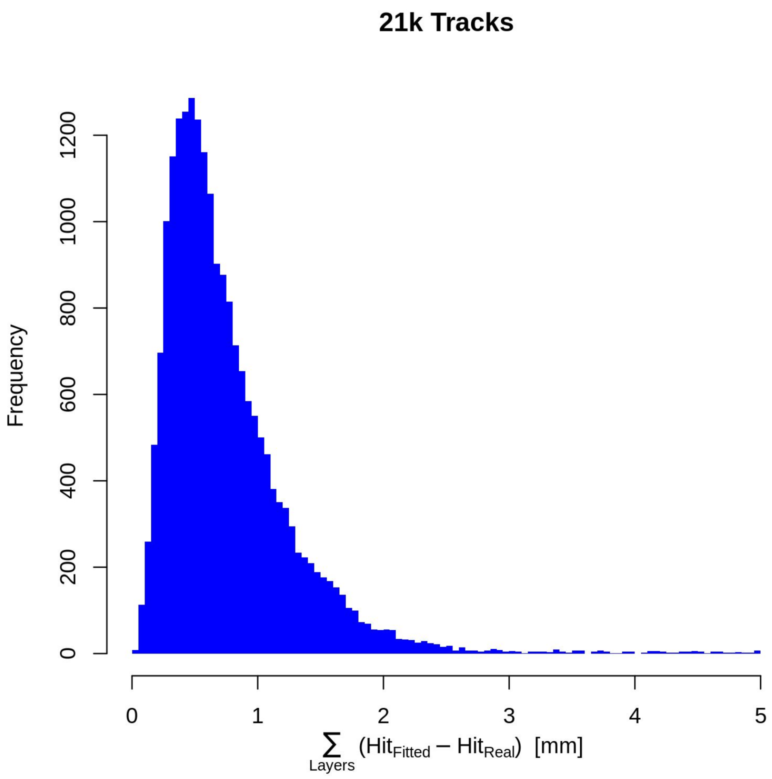
My Solution (II)



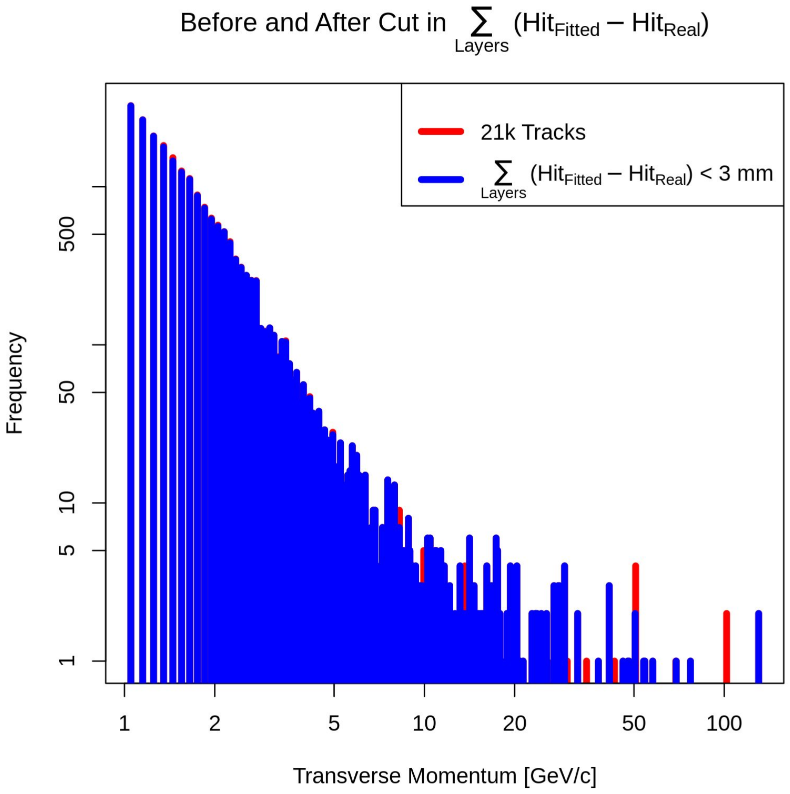


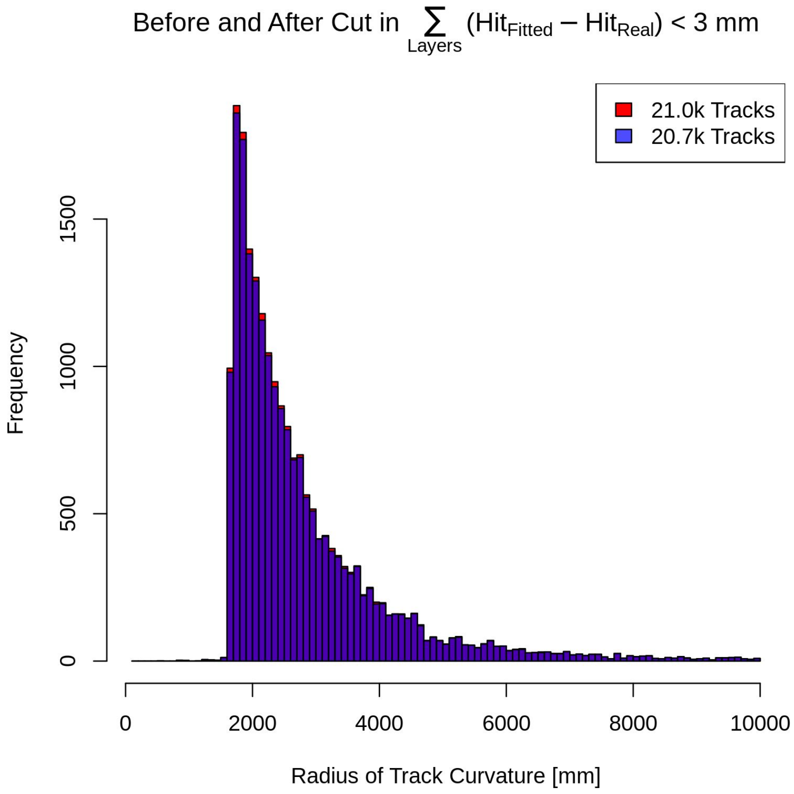
My Solution (III)



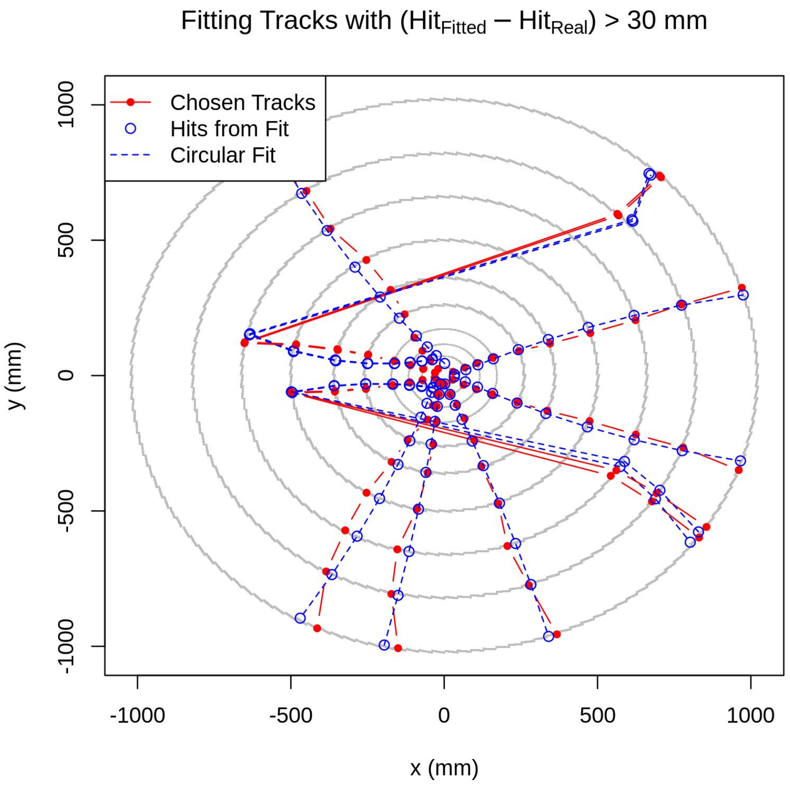


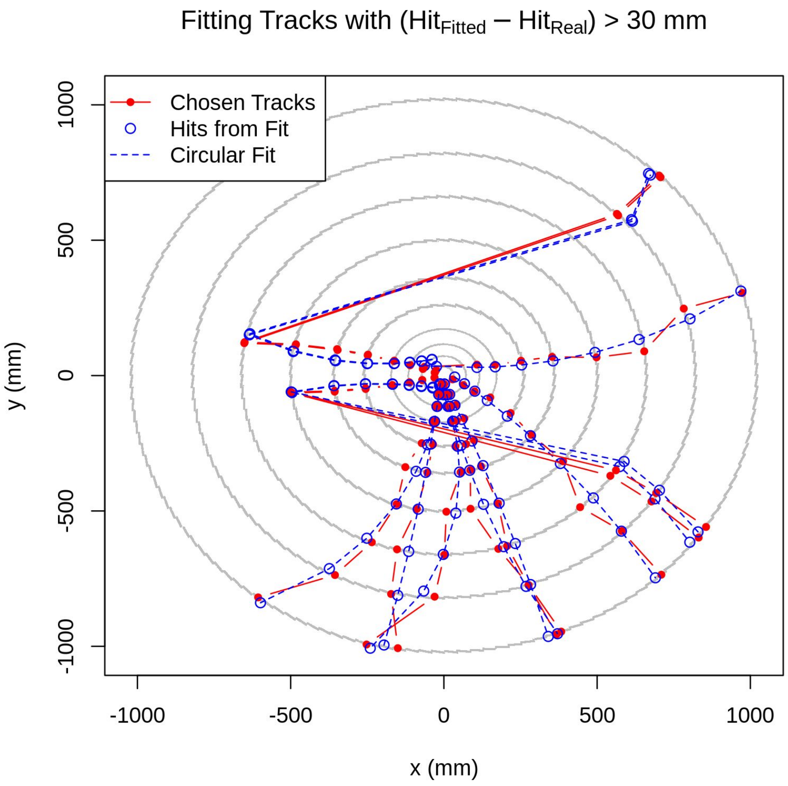
My Solution (IV)





My Solution (V)





To-Do List

- Use Optimization sample (3k tracks) to validate the fit
- Compare pT values between
    - values from dataset and values from fits on the reconstructed tracks