Adding Accurate Hit Errors to A Pixel Telescope

DESY Summer Student Programme, 2014

Mohan Li

Tsinghua University, China

Supervisor

Phillip George Hamnett

11th September 2014

Abstract

The pixel telescope, which has the corresponding EUTelescope analysis framework, is a high precision tracking device used for testing future detector components for particle physics experiments. To exhibit precise tracks, accurate hit errors are important. This report presents the principles and processes of using clustering information to add accurate hit errors to the analysis framework. The result shows improvement of the track precision in low and middle threshold, but deterioration in high threshold. Besides, it also validates the original guess of hit error as 0.005mm. More researches are needed to understand the relationship between the track error and the threshold.

Contents

1 Introduction

1.1 The pixel telescope

The pixel telescope is a high precision tracking device that can be operated inside a solenoidal magnetic field of up to 1.2 T. It is made up of readout board, cooling pipes, six MIMOSA26 sensor planes and so on. Each plane contains 576×1152 pixels and each pixel is $18.4 \times 18.4\mu m$, and the whole active area of one sensor is $10.6 \times 21.2mm^2$. A general purpose cooling, positioning and readout infrastructure is available. The structure of the pixel telescope is shown in Fig. 1.

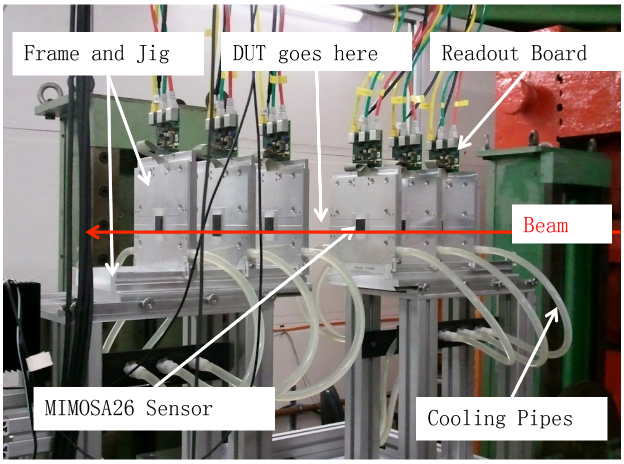


Figure 1: The structure of telescope, which is made up of six sensors.

The EUTelescope analysis framework is the software that analyzes the raw data taken by the pixel telescope. There are five steps in the analysis process, as follows.

- Converter, which converts raw data to .slcio format.
- Clustering, which searches clusters in the six sensors.
- Hitmaker, which calculates the hit position and it's hit error for each cluster.
- Alignment, which aligns the six sensors according to their relative positions.
- Fitter, which fits the track of the particle.

1.2 Motivation

Because of it's high precision, the pixel telescope can be used for testing future detector components for high energy physics. As a result, improving the track precision is of vital importance. The track error has a close relationship with the hit error, and the improvement of the hit error will help to reduce the number of fake tracks and find more true tracks, as well as improve track precision. Currently the hit error is guessed as 0.005mm, and improving the knowledge of the hit error will be the focus of the analysis.

The track error can be explained by Fig. 2. A particle flying through the six sensors will leave information about hit position and it's hit error on each sensor, and the EUTelescope framework can use such position information to fit a track of the particle. Where the track crosses the sensor is called the track hit position. There is always a distance between the real hit position and the track hit position, and such a distance is called the residual. The RMS of the histogram that is filled with the residual is the track error.

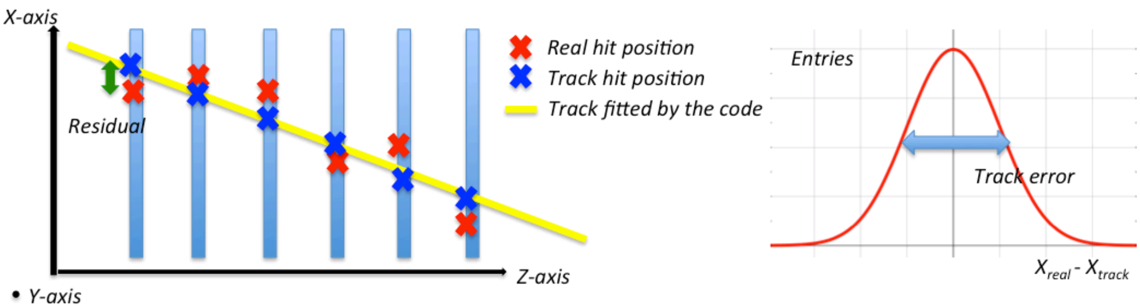


Figure 2: Explanation of the track error. Residual means the distance between real hit position and track hit position, and the RMS of the histogram that is filled by residual is the track error.

1.3 The clustering information

When a particle flies through one sensor, electrons and electron holes will be ionized in the sensitive volume. As there exists voltage between the anode and the cathode, electrons will fly to the anode where pixels exist. A pixel is a independent unit with micro-circuits that can collect and count electrons. If the electrons that one pixel collects are more than the threshold, which is set before the experiment, the pixel will be activated.

The process of an electron flying to a pixel will be effected by many factors, such as electromagnetic repulsion between electrons. As a result, the process of the electron's drift is more or less random in the direction to the anode. Even though a particle just hits one particular pixel, the neighboring pixels will also have the possibility to collect electrons and to be activated. A cluster is made up of all the connected and activated pixels, and clustering information refers to the size and shape of a cluster. Clustering information is very important because it is essential for fitting the track. The micro-structure of sensors and the collecting process is shown in Fig. 3.

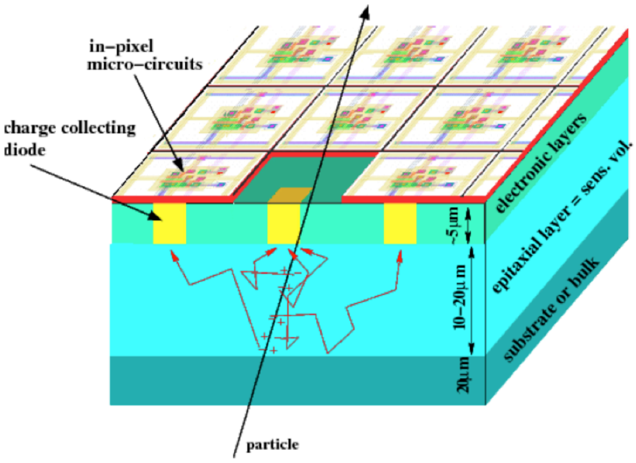


Figure 3: When a particle flies through the sensor, electrons are collected and counted by each pixel.

1.4 The center hit position assumption

However, to fit a track, the code needs the hit position and the hit error rather than the clustering information. As there is no way to know where exactly the real hit position is, it is always assumed that the real hit position is in the center of a cluster, which is called the center hit position assumption.

Fig 4 shows a cluster made of four pixels. The blue point refers to possible real hit positions, which can be anywhere in a cluster, and even outside. However, for example, if the real hit position is in the top left in the pixel 1, the pixel 5, 6, and 7 will have a larger probability to be activated, rather than the pixel 2, 3, 4. In other word, the most probable hit position is the center of a cluster. Currently, the EUTelescope analysis framework will use the algorithm "center of gravity" to calculate the center, taking the size and shape into consideration.

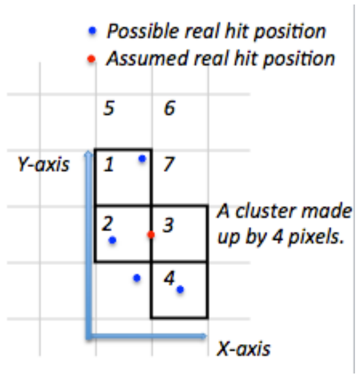


Figure 4: Explanation of center hit position assumption. Real hit position can be anywhere in a cluster, but the most probable hit position is in the center of a cluster.

2 Experiment set up

2.1 Get the hit error

Sensor2 is chosen to do analysis because it has the best resolution among all 6 sensors. The hit error in sensor2 is accessed in the residual histogram and drawn in a curve as vertical axis, and the horizontal axis is the threshold. Minimum biased data is test first to check the understanding of how the hit error depends on the threshold, and then unbiased data. Minimum biased data means the threshold of other 5 sensors also change with that of sensor2, and unbiased means the threshold of other sensors are fixed.

In fact, the track error is used to replace hit error for analysis. The hit error means the distance between the real hit position and the center of a cluster, which is no way to know because the real hit position is unknown. The track error means the distance between track hit position and the center of a cluster. Because track hit position is very close to the center of a cluster, and because of center hit position assumption, track hit position is close to real hit position. So the track error can be used to replace the hit error.

2.2 Fit the curve

After getting the curve, the next step is to fit it, i.e. to find function $\sigma_{hiterror} = f(threshold)$. Each curve is fitted by four functions: Explinear, Shah, Polynomial3, Belehradek. Unfortunately, there is no physical reason for choosing these functions. Instead, they are chosen just by their shapes. This is because the threshold is set before the experiment, so it does not depend on any other variables, which means that no connection between the threshold and the hit error can be found. The four fit functions are listed below, in which x is the argument, y is the dependent variable and other letters are fitting parameters.

- Explinear: $y = a \times e^{-\frac{x-e}{b}} + c + d \times (x - e)$
- Shah: $y = a + b \times (x - e) + c \times d^(x - e)$
- Polynomial3: $y = a + b \times x + c \times x^2 + d \times x^3$
- Belehradek: $y = a \times (x - b)^c$

Though there are many fit results, and definitely using one specific fit function for one specific curve is more precise than just applying one same fit function to all situations, the most general fit function, which is valued by χ^2, is chosen because the purpose is not just to improve the track precision of the specific data, but to use the same fit function found to other datas. Then the general fit function $\sigma_{hiterror} = f(threshold)$ is put back to the EUTelescope analysis framework and the same data is rerun for checking the influence of the change of the hit error on the track error.

2.3 Check the outcome

There are two methods to check whether the track error improves or not. The first and direct one is to check the RMS of the residual histogram. The second method is called in-pixel track hit position, which means moving all the track hit positions from the same kind of cluster into one pixel. If the in-pixel track hit position gets more concentrated, the track precision is improved.

In one run, there are a lot of clusters, as well as track hit positions in sensor2. Picking up all the size 1 cluster, for example, and move all of their hits into one pixel, without changing the relative position of a hit in one pixel, there should be a concentrated area in the center of the pixel. It is the same principle for shape2, 3 and 4, as shown in Fig 5.

The key to draw the in-pixel track hit position is to find the cluster that one hit belongs to. As there is no direct connection between the cluster and the track hit, the solution is to search for a track hit around the center of a cluster. One track hit is considered to belong to one cluster if it is close enough to the center. The principle is explained in Fig 6.

3 Result and analysis

Fig 7 shows how the hit error depends on the threshold from minimum biased data, whose threshold in all the sensors changes from 3 to 12. It clearly exhibits the pattern

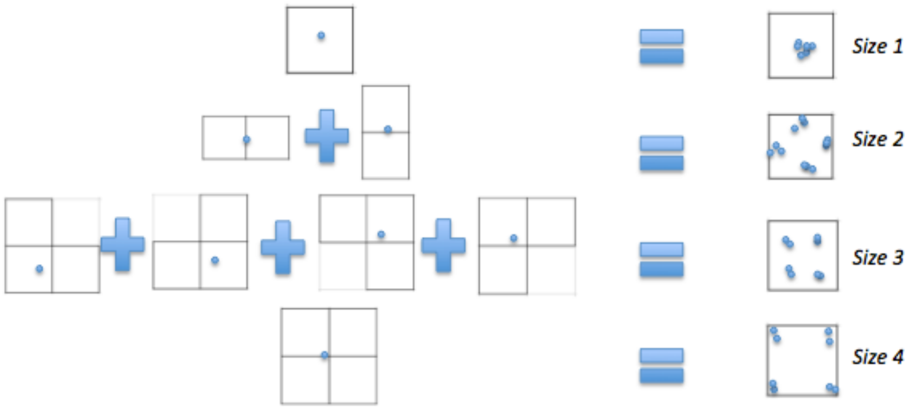


Figure 5: Explanation of what is in-pixel track hit position. Blue point refers to the track hit position, which is always around the center of the cluster. If accumulating hits from the same cluster into one pixel, there will be concentrated areas.

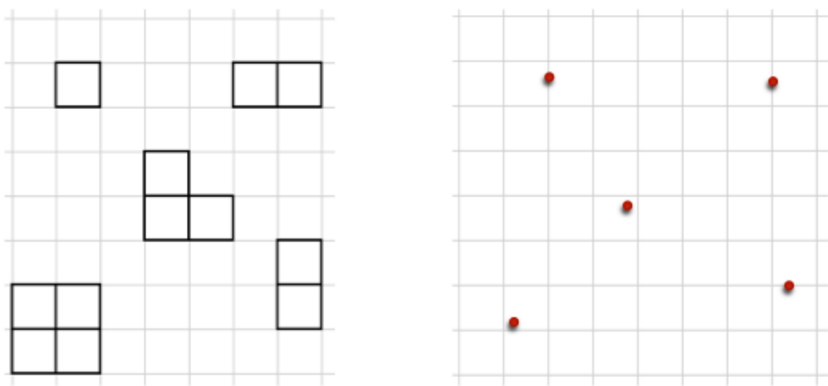
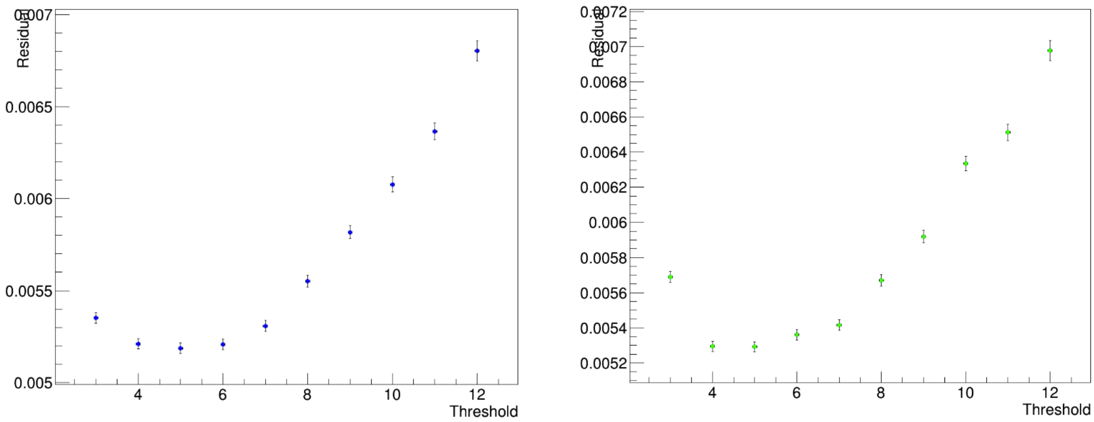


Figure 6: Explanation of how in-pixel track hit position works. Track hit is searched around the center of a cluster. This schematic assumes no noise, which means no superabundant cluster, and every track can be fit successfully, which means no scarcity of the track hit.

that the hit error goes down first and then goes up with the increasing of the threshold.

Fig 8 shows the same curve from unbiased data, whose threshold in sensor2 changes from 4 to 10, while the threshold of other 5 sensors are fixes as 6, 7, 8, 9 and 10. The pattern becomes obscure though tendency is clear in unbiased data because the number of event in one run in unbiased data is much smaller than that in minimum biased data, so the statistic regularity is not so clear.There are some differences in x and y axis, which can be caused by difference of beam width and alignment.



(a) x axis from minimum biased data.

(b) y axis from minimum biased data.

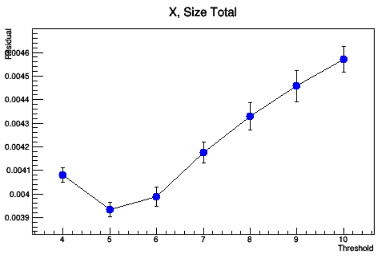
Figure 7: The dependent relationship of the hit error on the threshold in x and y axis from minimum biased data.

Basically, the hit error is effected by the threshold, cluster size and cluster shape. If the cluster keeps unchanged, the probability of real hit position on the edge of a cluster will reduce with the increasing of the threshold. This is because higher threshold means more electrons are needed to activate a pixel, and the real hit position must be around the center of a cluster so that it can provide enough electrons to all the pixels. In the specific example in Fig 9, it is very hard for a particle hitting the top left corner of the left pixel to activate the right pixel if the threshold is high.

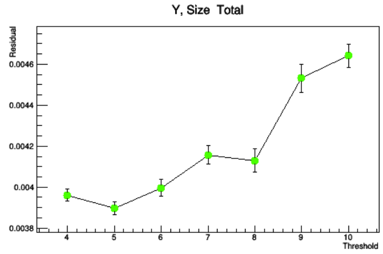
However, the cluster size and shape will also be effected by the threshold, which makes things a bit more complicated. Fig 10 shows the possible change of a cluster created by the same particle with the change of the threshold. Suppose the same particle always hits the top left corner of one pixel. When the threshold is low, the particle can activate 3 pixels. With the increasing of the threshold, it can only activate 2 pixels, and then 1 pixel. At the same time, the distance between real hit position and the center of a cluster becomes larger, which means the track precision gets worse.

Taken all the factors into consideration, the going down part of the pattern corresponds to the situation where clusters do not change a lot while the hit error decreases in one cluster, and the going up part corresponds to the situation where clusters already change a lot with the increasing of threshold.

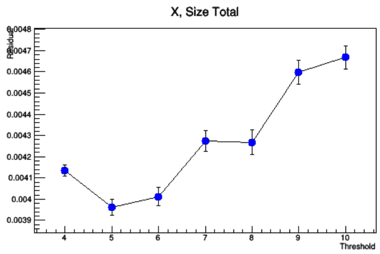
The χ^2 of the 4 functions in fitting minimum biased data is shown in Tab. 1. Here we can see the Explinear and Shas are better than Polynomial3 and Belehradek. Then Explinear and Shah are used to fit unbiased data, and the χ^2 is shown in Tab. 2. Here '-' means the fitting fails, and 'threshold' is the threshold in other 5 sensors.



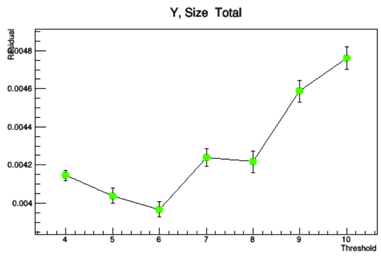
(a) x axis from unbiased data. The threshold of other five sensors is 6.



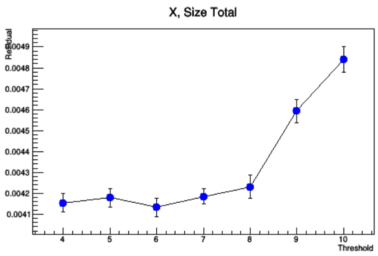
(b) y axis from unbiased data. The threshold of other five sensors is 6.



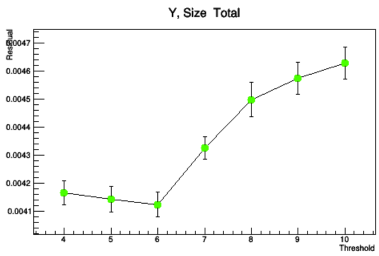
(c) x axis from unbiased data. The threshold of other five sensors is 7.



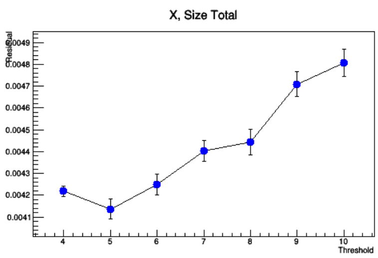
(d) y axis from unbiased data. The threshold of other five sensors is 7.



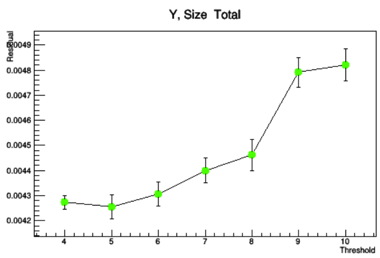
(e) x axis from unbiased data. The threshold of other five sensors is 8.



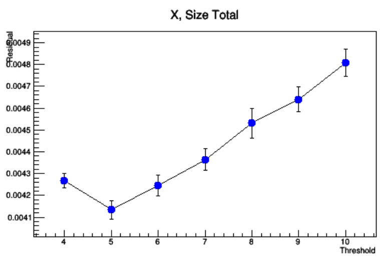
(f) y axis from unbiased data. The threshold of other five sensors is 8.



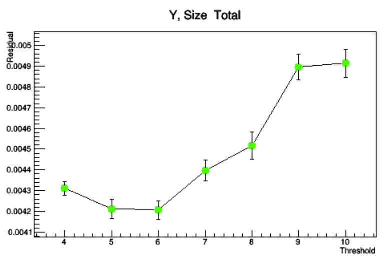
(g) x axis from unbiased data. The threshold of other five sensors is 9.



(h) y axis from unbiased data. The threshold of other five sensors is 9.



(i) x axis from unbiased data. The threshold of other five sensors is 10.



(j) y axis from unbiased data. The threshold of other five sensors is 10.

Figure 8: The dependent relationship of the hit error on the threshold in x and y axis from unbiased data.

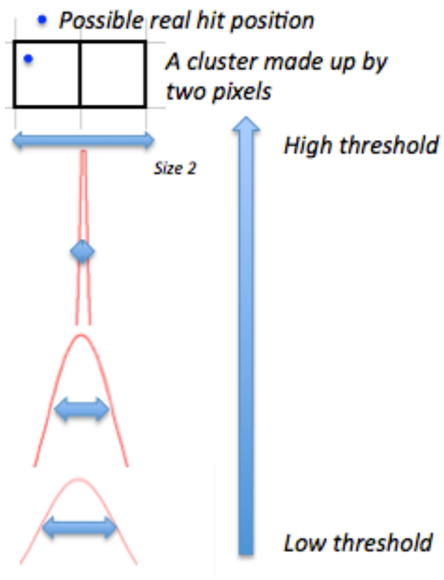


Figure 9: Within the same cluster, the hit error will reduce with the increasing of the threshold.

Axis	Explinear	Shah	Polynomial3	Belehradek
x	0.798258	0.798258	1958.83	58.6415
y	3.57374	3.57374	1383.91	86.8563

Table 1: The χ^2 of 4 fit functions in x or y axis of minimum biased data.

Finally, the fit function of explinear in threshold 9 in x axis is chosen because of it's best χ^2, which is 1.31284. The function is $y = 5.54 \times 10^{-4} \times e^{-\frac{x-3.59}{5.79\times10^{-1}}} + 3.89 \times 10^{-3} + 1.44 \times 10^{-4} \times (x - 3.59)$. Fig 11 is the fitting result.

After putting the fit function mentioned above back into the EUTelescope analysis framework and rerunning the unbiased data, track error are checked by comparing the RMS of residual histogram. Fig. 12 shows the residual histograms of threshold 4 in x axis before and after the change of the hit error as a example. The red line refers to residual histogram before change, and blue line refers to that after change. The 'threshold' means the threshold in other 5 sensors, and the threshold of sensor2 is kept 9. it shows that the blue line gets narrower and higher, which means the improvement of the track precision.

All the other change of the residual histograms look like this example, and the RMSs are shown in Tab. 3. Here 'threshold' means the threshold in other 5 sensors, because the track error is associated with all the 6 sensors. The threshold of sensor2 is kept 9.

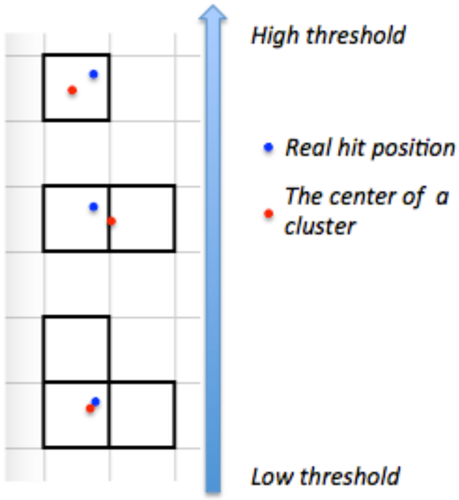


Figure 10: Within the same cluster, the hit error will reduce with the increasing of the threshold.

Axis	Function	Threshold 6	Threshold 7	Threshold 8	Threshold 9	Threshold 10
x	Explinear	0.567786	-	3.80126	1.31284	0.120467
y	Explinear	4.47215	-	4.51743	2.27658	3.2177
x	Shah	-	-	3.80126	-	0.120467
y	Shah	-	-	4.51743	-	3.2177

Table 2: The χ^2 of Explinear and Shah in fitting unbiased data.

When threshold is low, there is improvement in track precision, and in some cases the improvement can be as large as 6%. However, in high threshold, the track precision gets worse.

As mentioned above, another method to check the track error is the in-pixel track hit position. The in-pixel track hit position before and after change, whose threshold in sensor2 is 4, and threshold in other 5 sensors is 9, is shown in Fig. 13. The center of the concentrated area should be in the center of a pixel, but misalignment can cause a shift. As shown in shape 1, both of their concentrations are not in the center. But it is confusing why their concentrations are different, because the in-pixel track position will not effect the alignment.

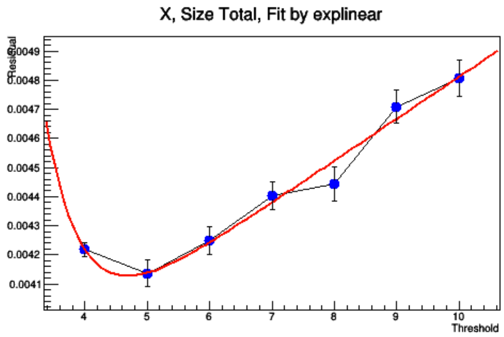


Figure 11: The ultimately chosen fitting is Explinear in threshold 9 in x axis.

4 Conclusion

In this summer program, accurate hit error is found by the fit function that has the best χ^2 in fitting the dependent relationship of the hit error on the threshold. The accurate hit error is added into the EUTelescope framework analysis to check whether it will help to improve the track precision. The result shows that in low and middle threshold, the track precision will improve. However, the track precision gets worse in high threshold. Even though there are changes, they are so small that the original of the hit error as 0.005mm is thought accurate. More research is needed to understand the relationship between the track error and the threshold.

5 Acknowledge

The author acknowledges the guidance and attendance of supervisor Phillip George Hamnett, who is the reference of the article. The author also sincerely thanks for all the support of Deutsches Elektronen-Synchrotron (DESY).

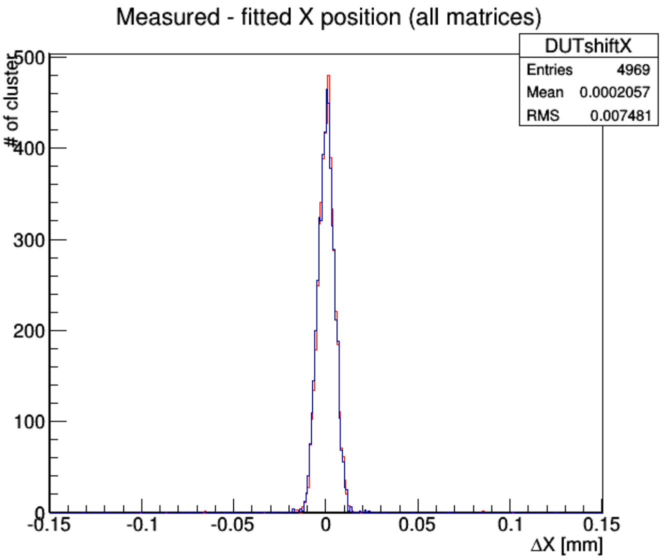
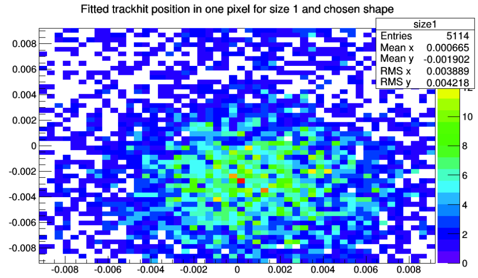


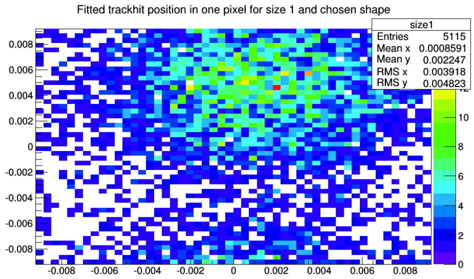
Figure 12: The residual histogram before and after change of threshold 4 in x axis as an example. Red line represents before change, an d blue line represents after change.

Threshold	Axis	Original error	Current error	Percent change
4	x	0.007481	0.006978	-6.72370004
4	y	0.00835	0.008075	-3.293413174
5	x	0.006719	0.006704	-0.223247507
5	y	0.006157	0.00617	0.21114179
6	x	0.00553	0.00551	-0.361663653
6	y	0.005221	0.005204	-0.325608121
7	x	0.005822	0.00554	-4.843696324
7	y	0.005125	0.005125	0
8	x	0.006266	0.006244	-0.351101181
8	y	0.005592	0.005587	-0.089413448
9	x	0.005406	0.005394	-0.221975583
9	y	0.005432	0.005466	0.625920471
10	x	0.006083	0.006096	0.21371034
10	y	0.005739	0.005749	0.174246384

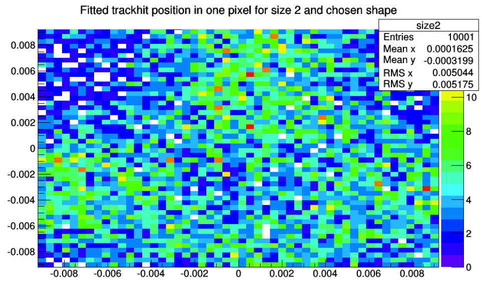
Table 3: The track error before and after using the fit function to calculate the hit error.



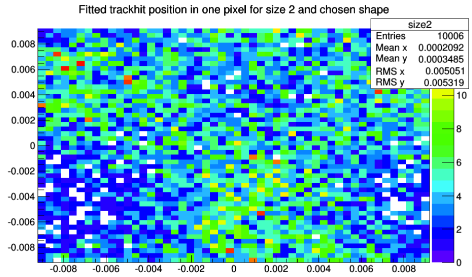
(a) Before change in shape 1.



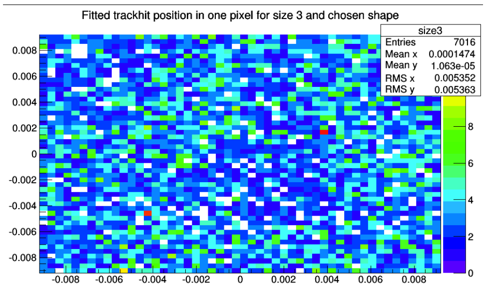
(b) After change in shape 1.



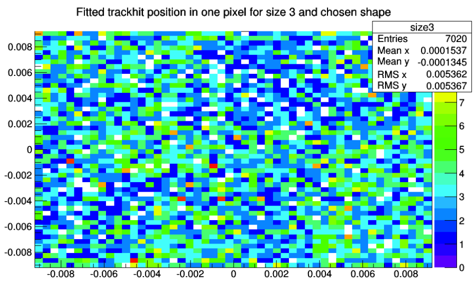
(c) Before change in shape 2.



(d) After change in shape 2.



(e) Before change in shape 3.



(f) After change in shape 3.

Figure 13: The in-pixel track hit position before and after change, whose threshold in sensor2 is 4, and threshold in other 5 sensors is 9.