Resonance search analysis of 2016 HPS spring run data.

Bump hunt folks

December 13, 2019

Contents

Introduction

The Heavy Photon Search (HPS) experiment has capability to search for a so called heavy photon (A') with two complementary methods.

1 Data set

Describe the data, beam energy, beam current, target runs, etc.

2 Event Selections

This section describes all the cuts that are applied to get the final vertex candidate distribution. The main goal of event selection cuts is to maximize signal sensitivity.

In this analysis only events with "Pair1" trigger (see [1] for the description of HPS triggers) are used.

2.1 Cluster timing cuts

The readout window of ECal FADC data is 200 ns. Clusters coming from the physics events, that generated the trigger, are located in a narrow time range (few ns width because of the trigger jitter) in the readout window around t = 56 ns. In Fig.1 shown "time vs Energy" distributions of ECal

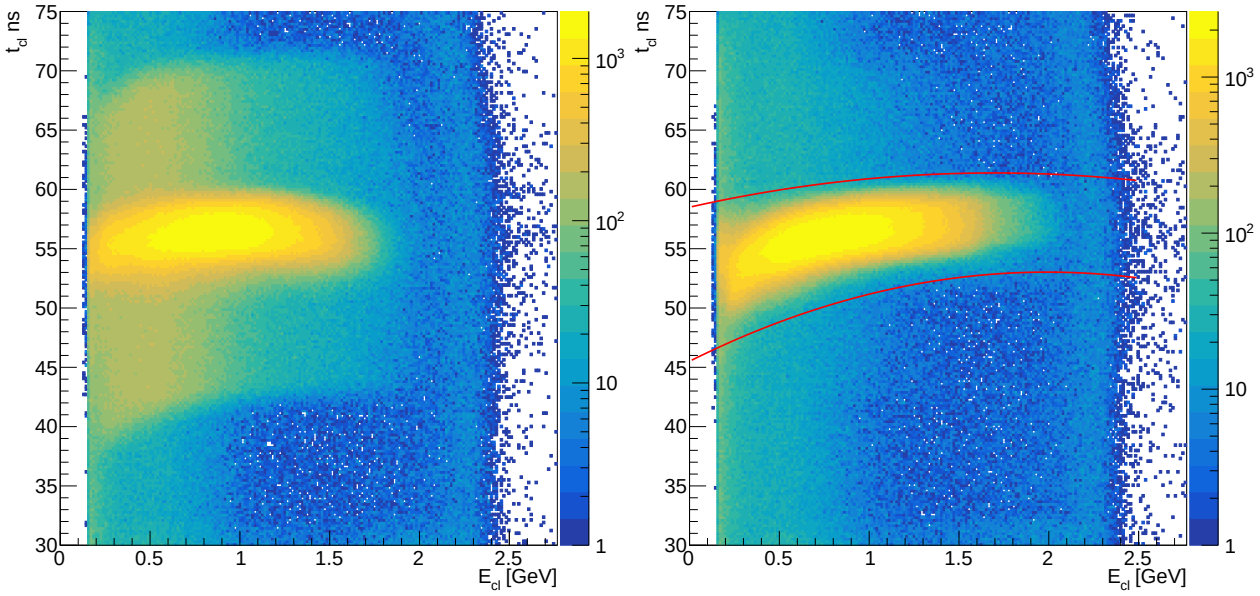


Figure 1: "time vs Energy" distributions of ECal clusters in the Top (Left) and Bottom (Right) half. Red curves in the in the right plot indicate cuts that are applied to clusters in the bottom half for the initial cluster selection. See the text for the description of the difference between left and right plots.

clusters in the Top (Left) and Bottom (Right) halves. The bulge of events in the right plot are clusters that generated the trigger, and also trigger time is defined by these clusters. The noticeable energy

dependence is due to the so called "time walk Corrections" [2]. During initial event selection only clusters that are inside the outlined red curves are used, since the rest are accidentals that didn't come from the beam bunch generating the trigger. One can notice that for the clusters in the top half, in addition to the central bulge, there is an extra occupancy of events in region (40 ns $< t_{cl} <$ 70 ns). This is because the coincidence time between clusters in the "Pair" trigger was 12 ns [1], and the trigger time is determined by the bottom cluster. Unlike to clusters in the bottom half, in the initial event selection, we have not cut on time of the top cluster, but rater we have applied cut on the cluster time difference between top and bottom clusters.

2.1.1 Ad hoc ECal time corrections

The next step is to cut pairs of top-bottom clusters that are far from each other in terms of time. During the analysis is was found that ECal cluster times can be improved, in particular the dashed

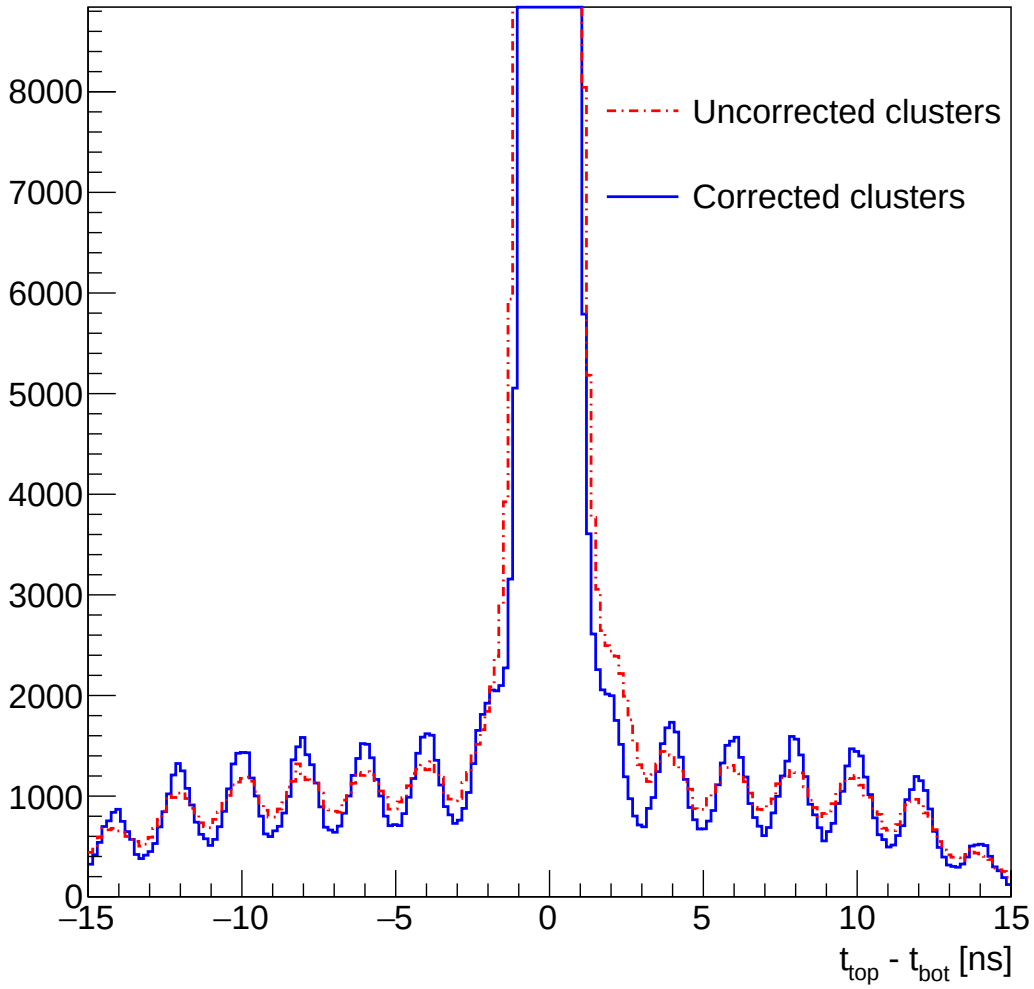


Figure 2: Time difference between top and bottom clusters at high E_{sum} region. Dashed red lines show uncorrected clusters, and the blue curve show corrected clusters.

red histogram in Fig.2 shows the time difference between top and bottom clusters[1] at high E_{sum} region

[1]For the sake of better visualization, the plot doesn't fully show the entire central peak.

(1.9 GeV $<$ E_{Sum} $<$ 2.4 GeV). As one can see there is a bump at around 2 ns, while the at -2 ns there is no clear bump. This suggests that time offsets of some crystals might be wrong. To check

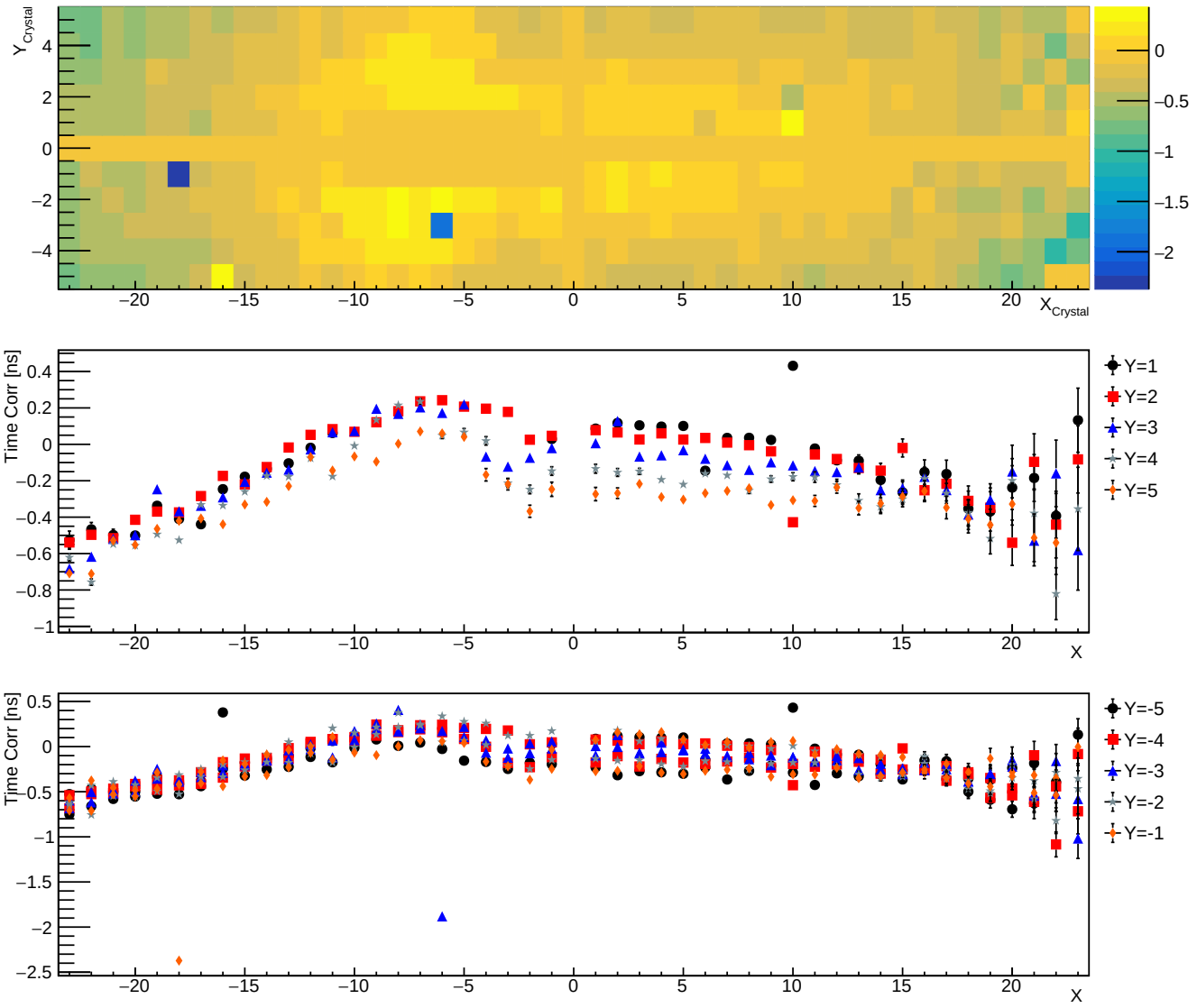


Figure 3: Time Corrections for each crystal.

this, for each crystal the time difference between that crystal and it's pair (in opposite half) crystal is constructed. The Top 2D plot of Fig. 3 shows mean values of each of crystals. The middle and the bottom plots show same mean values as a function of crystal X index for Top and Bottom crystal respectively. Different markers show different rows. As one can see crystals (-18, -1) and (-6, -3) are shifted from their immediate neighbors by about 2 ns. There are some other crystals which are shifted significantly too (but less than 2ns). These include for example crystals (-16, -5), (10, -5), (10, 1) and (10, 2). In addition to this, we see that there is a general crystal X index (and slightly Y index) dependence too. In reality crystal X index is correlated to the charged particle energy too, and the original dependence might be not on X but on energy. Studying it is out of the scope of this note, and here for each crystal we have corrected the time, by subtracting these calculated mean values from the reconstructed cluster time. After the correction the cluster time difference is depicted by blue solid line in Fig.2. One can see that the excess of events at 2 ns disappeared. Dips and peaks between bumps

indicating difference beam bunches also got sharper, which is an indication of an improvement of the cluster time resolution.

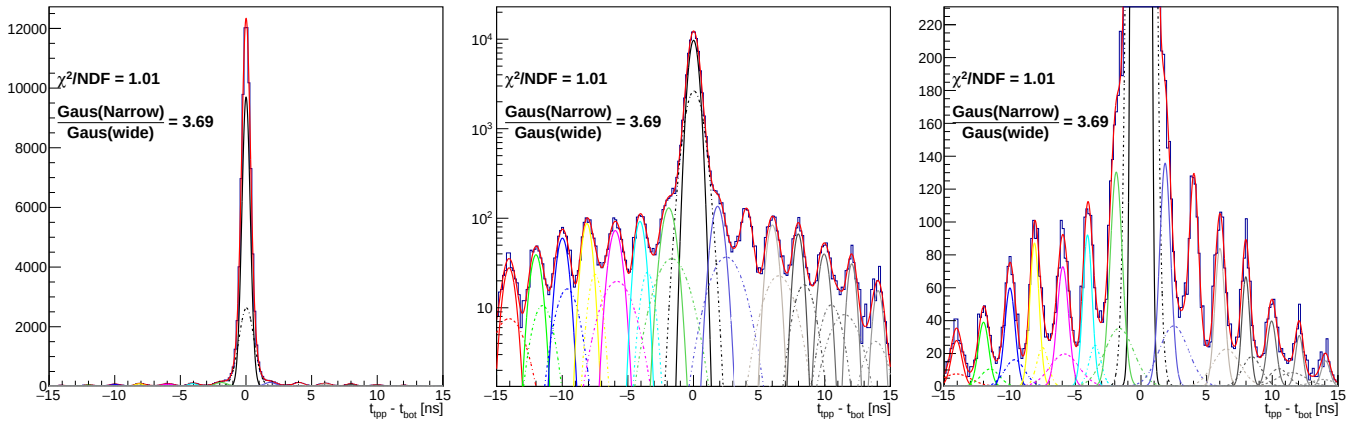
2.1.2 Fitting Cluster time difference

After correction of individual cluster times, the Top-Bottom cluster time difference was fitted with a following function:

$$F = \sum_{i=0}^{N_{peak}} a_i \cdot \left(\text{Gaus}(x - \mu_i^1, \sigma_i^1) + b \cdot \text{Gaus}(x - \mu_i^2, \sigma_i^2)\right) \quad (1)$$

where N_{peak} is the number of peaks. Each peak is described by the sum of two Gaussian functions $\text{Gaus}(x - \mu_i^1, \sigma_i^1)$ and $\text{Gaus}(x - \mu_i^2, \sigma_i^2)$ with their amplitude ratio "b'. The parameter "b" is the same for all peaks. In the fit, free parameters are $a_i,\ \mu_i^1,\ \sigma_i^1,\ \mu_i^2,\ \sigma_i^2,$ b.

The fit result is shown in Fig.4. Different peak components of the function are depicted by different



(a) (b) (c)

Figure 4: Fit of the Top and Bottom cluster time difference. Left: linear scale, Middle: Log scale and right: Linear scale but includes only low magnitude peaks. Main Gaussian functions are represented by solid lines, and the secondary Gaussian (with wider width and lower magnitude) are represented by dashed lines.

color. The main Gaussian function of each peak is represented by a solid line, while the secondary Gaussian is represented with a dashed line. One can see that this function fits the distribution reasonably well.

Then in order to determine the optimal cut on the cluster time difference, we will use the value, which maximizes the $\frac{S}{\sqrt{S+Bgr}}$ ratio, where "S" is the signal (in our case the central peak), and "S + Bgr" is the signal plus Background (the total fit function). The $\frac{S}{\sqrt{S+Bgr}}$ ratio as a function of cluster time difference cut is shown in Fig.5, where the maximum value at $\Delta t < 1.43$ ns is indicated by a vertical dashed line.

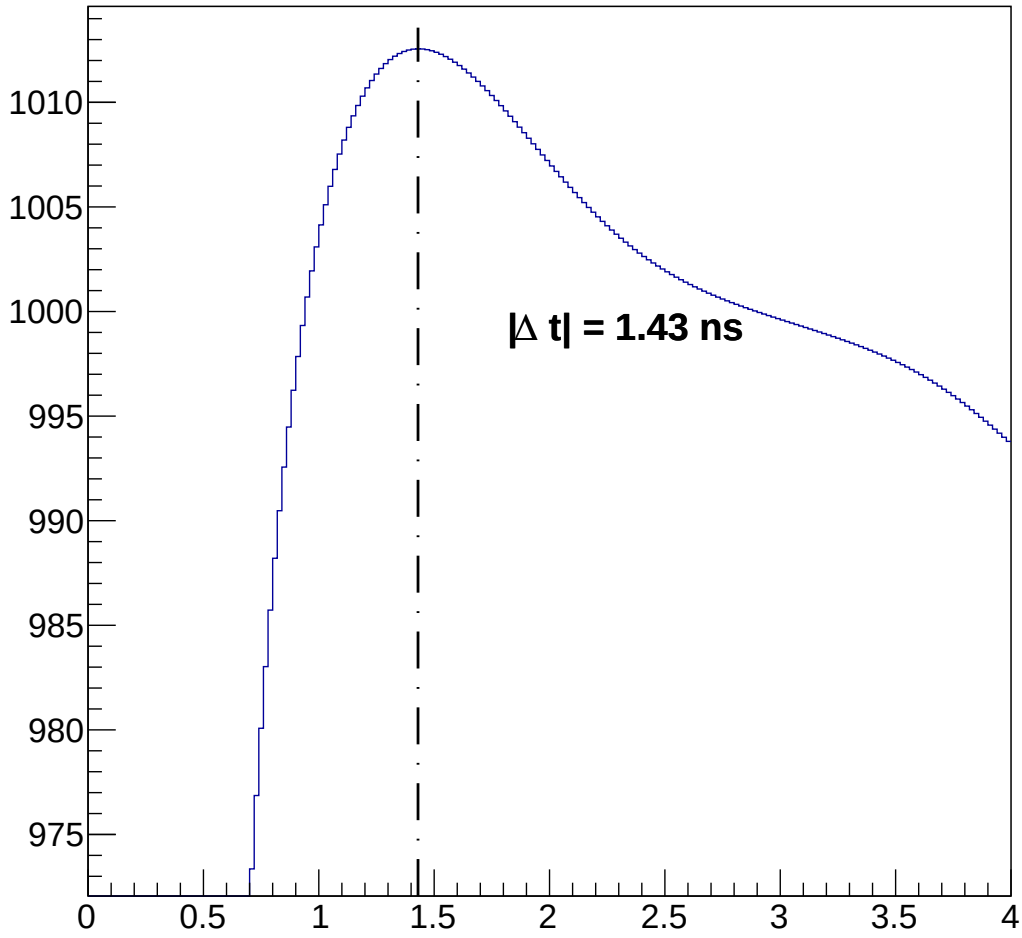


Figure 5: The $\frac{S}{\sqrt{S+Bgr}}$ ratio as a function of cluster time difference cut. The Dashed line indicates the maximum of the function.

2.2 Two dimensinal cuts

Some of event selections cuts described below are two dimensinal cuts, i.e. the cut value depends on the value of another variable. In most of cases two dimensional cuts are implemented as a function of particle's momentum.

In general, to study the distribution of a given variable for a "signal like" particle, the rest of event selection cuts are applied, to make as clean as possible signal. The only exception is the two cluster time difference cut, which is described in section 2.1). Applying the rest of cuts except the one under the investigation, will ensure the accidental background is minimal (negligible), and the resulting distribution will represent actuall signal (the e^-, e^+, (X) final state). In most of cases the distribution is not Gaussian, even when it represent a small momentum bin. In such cases, the conventional $\pm 3\sigma$ cuts will not be will not keep 99.7% but rather might cut more events. As an example in Fig.6 shown a toy distribution which is not a pure Gaussian, but rather has a tail on the left side. The Gaussian fit is shown on top if the histogram and $\pm 3\sigma$ limits are shown by vertical red lines. One can see that -3σ limit will cut several % of events rather than 0.3%. Instead it was decided to choose left and right cuts

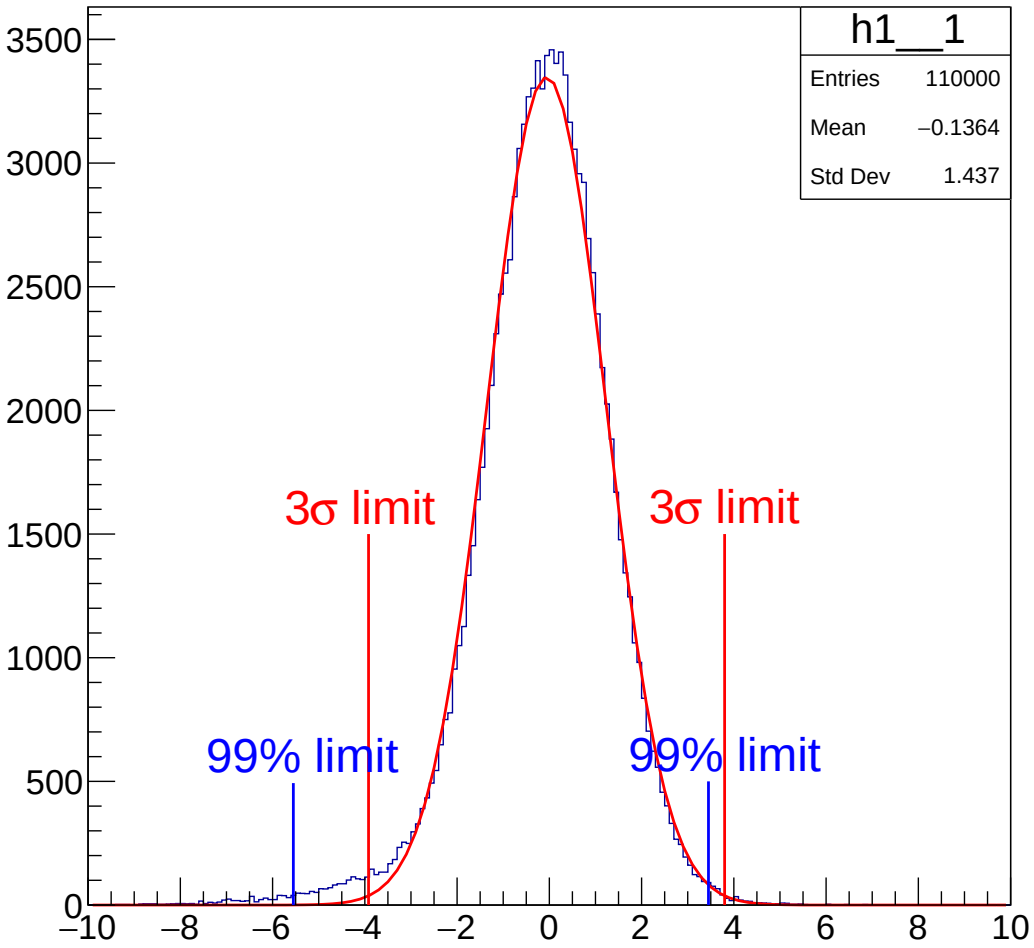


Figure 6: Illustration of 3σ cut limits vs 99% cut limits on a toy distribution.

limits such that will keep 99% of the signal and will throw 0.5% of signal events from each side. In this particular case 99% cut limits are shown by blue vertical lines.

There are some special cases in this algorithm, which are explained below.

1. The number of events in the one dimensional projected histogram is too low. In this case, when then number of events is below 45, then the entire momentum bin is considered as out of acceptance. Note: the data sample that we use, has such statistics, that two dimensional original histograms that we use have about 200K or more events, and binning is chosen such, that the total number of events that are in the "$N < 45$" category one dimensional histograms are significantly less than 1% of the original two dimensional histogram.

2. The number of events in the one dimensional histogram is more than 45, however it is not high enough (order of thousands). In this case the 0.5% events that should be cut out from each side of the distribution is very small number e.g. 1, 2, 3 or even 0, if the number of events is in between 45 and 200.

So in order to apply some cuts rather than no-cut or very loose cut, the 99% requirement is released for histograms having less than 500 events ($N < 500$). In general cut limits for different statistic cases are summarized in table 1.

# of events	Cut limit
N > 500	99%
200 < N < 500	98%
100 < N < 200	96%
45 < N < 100	95%
N < 45	0%

Table 1: Cut limits for different statistic scenarios.

2.3 Track-Cluster Matching

The offline reconstruction code forms particles by matching tracks and clusters to each other, by utilizing spatial coordinate and time differences between tracks and clusters. In the offline reconstruction the matching is quite loose ("Better to keep junk, rather than throwing a good particle"). In this section spatial and time matching cuts are described. Both, time and position resolution of e^- and e^+ clusters depend on particle momentum. The precision of the track projected coordinate at the ECal face does depend on the track momentum too. Because of these reasons we studied track-cluster matching as a function of momentum.

2.3.1 time matching

In addition to the momentum dependence we noticed also slight difference between top and bottom sectors, therefore two separate cuts are developed for each detector half. In Fig.7 shown Cluster-Track time difference as a function of particle momentum. Left plot represents particles in the bottom half, and the right plot represents particles in the top half of the detector. The Cl-Trk time difference plot should be replaced, i.e. The original 2D distribution here doesn't correspond to the "n-1" cut.

2.4 Track quality cuts

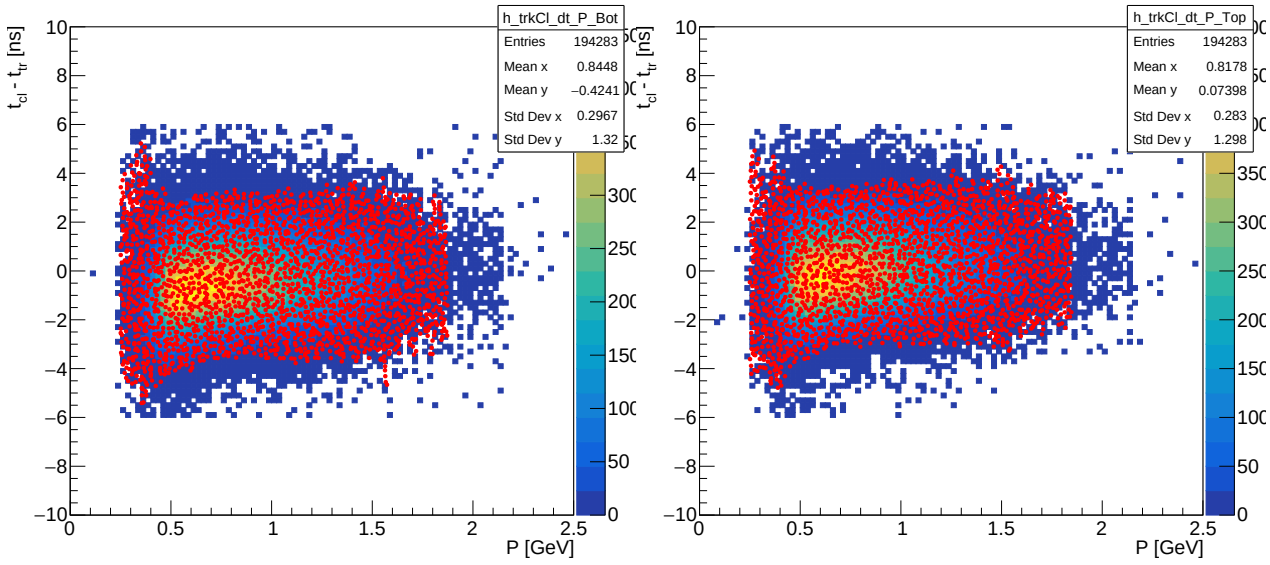
The point of track quality cuts is to maximize the Figure Of Merit (FOM). It is natural to think, if the track quality (in terms of χ^2 per degrees of freedom) is poor, then that could result in worse mass resolution, and consequently will have a negative impact on the experiment reach.

To maximize the the reach, we should maximize the FOM, which is

$$\text{FOM} \sim \frac{\sqrt{N_{Tot}}}{\sigma_m} \tag{2}$$

Where $\sqrt{N_{Tot}}$ is the number of events in the given mass bin, and σ_m is the mass resolution for the given mass (see appendix A for more details about eq. 2).

We have used the Møller process to estimate the impact of the track quality on the mass resolution. The square root of the center of mass energy in the Møller process, is fixed for a given beam energy,



(a) (b)

Figure 7: Cluster-Track time difference as a function of particle momentum. Left plot represents particles in the bottom half, and the right plot represents particles in the top half of the detector. The area marked by Red dots represents the acceptance region.

and is equal (neglecting electron mass square terms):

$$M_{ee}^{\text{c.m.}} = \sqrt{2 \cdot m_e \cdot E_b} \tag{3}$$

The square root of the center of mass energy is also equal to the invariant mass of final state electrons in the Møller process. Hence we will use the Møller process to estimate the effect of track quality on the mass resolution. As a measure of quality of track pairs, the combined $\chi^2_{\text{Sum}}/\text{NDF}_{\text{Sum}}$ was used, which is defined as:

$$\chi^2_{\text{Sum}}/\text{NDF}_{\text{Sum}} = \frac{\chi^2_{\text{Bot}} + \chi^2_{\text{Top}}}{2\left(\text{N}^{\text{hits}}_{\text{Bot}} + \text{N}^{\text{hits}}_{\text{Top}}\right) - 10} \tag{4}$$

Here $\chi^2_{\text{Bot(Top)}}$ is the χ^2 for bottom (top) track, and $\text{N}^{\text{hits}}_{\text{Bot(Top)}}$ is the number of 3d hits for the bottom (top) track. In the denominator 10 is the total number of constraints (5 from each track).

2.4.1 Selection of Møller events

Important: We need MC data to justify some of cuts. Some of them are natural, however they should be confirmed by MC.

As a starting point, we have used so called "Møller candidate events" defined by the so called "MOUSE" cuts [4]. Those are events which contain at least one negative track in each detector half. The magnitude of their momentum sum also should be within 20% of beam energy:

$$0.8E_{\text{b}} < P_{\text{Møeller}} \equiv |\vec{P}_{\text{Bot}} + \vec{P}_{\text{Top}}| < 1.2E_{\text{b}} \tag{5}$$

Note: here there is no requirement for a track to be associated with a cluster. In the Møller selection of 2015 run, it was required both electrons to have a cluster, however about $\times 2$ higher beam energy of

2016 run boosts electrons more, and it is almost impossible to get both electrons to hit the Calorimeter (The ECal hole significantly reduces Møller acceptance), however there is non zero acceptance in the tracker for Møller electrons. In Møller event selection, it is not critical to keep track on how much signal and background each cut throws, but rather it is important to have the final sample as clean as possible (event if some cuts might be tight). During the analysis the following cuts were used:

- P_{sum} momentum sum of final state electrons. Møller kinematics requires this to be equal to the beam energy. In the analysis we will chose a region where the P_{sum} is in the vicinity of beam energy.

- Δt_{tr} time difference between top and bottom tracks. Cutting Δt_{tr} around 0, will suppress accidentals coming from different beam bunches.

- Track-Cluster matching. In 2016 kinematics One of electrons misses the calorimeter, and, in the offline reconstruction, if both tracks are associated with a cluster, then it is very likely one of them is accidental or track and cluster are produced from different particles, and it will will yield a bad matching χ^2 value (large value). To suppress accidentals we will require one of tracks to have a good matching χ^2 (small value), and the other to have a poor χ^2 (large value).

- Cut on P_{diff}: Momentum difference between top and bottom electrons. This is cut is also based on the specific acceptance of Møller events. and certain momentum configuration enter into the detector acceptance. This also will help to suppress acceptance.

- dϕ: Azimuthal angular difference of two final state electrons wrt beam direction at the target. In the Møller kinematics dϕ should be 180°. This cut however is not used. It cuts almost nothing, when the rest of cuts is applied.

Similarly to the rest of event selection cuts for the trident final state, here for Møller event selection also, in order to understand where to put cuts on a given variable, cuts are applied to the rest of variables mentioned above, then the distribution of the given variable is studied. In Fig.8 shown Time difference between top and bottom tracks (left) and Momentum difference (P_{Diff}) vs Momentum sum (P_{Sum}) of Top and bottom tracks (right). Red lines indicate cut region for corresponding variables **Again, we need proper MC justify these cuts** .

2.5 WAB Suppression cuts

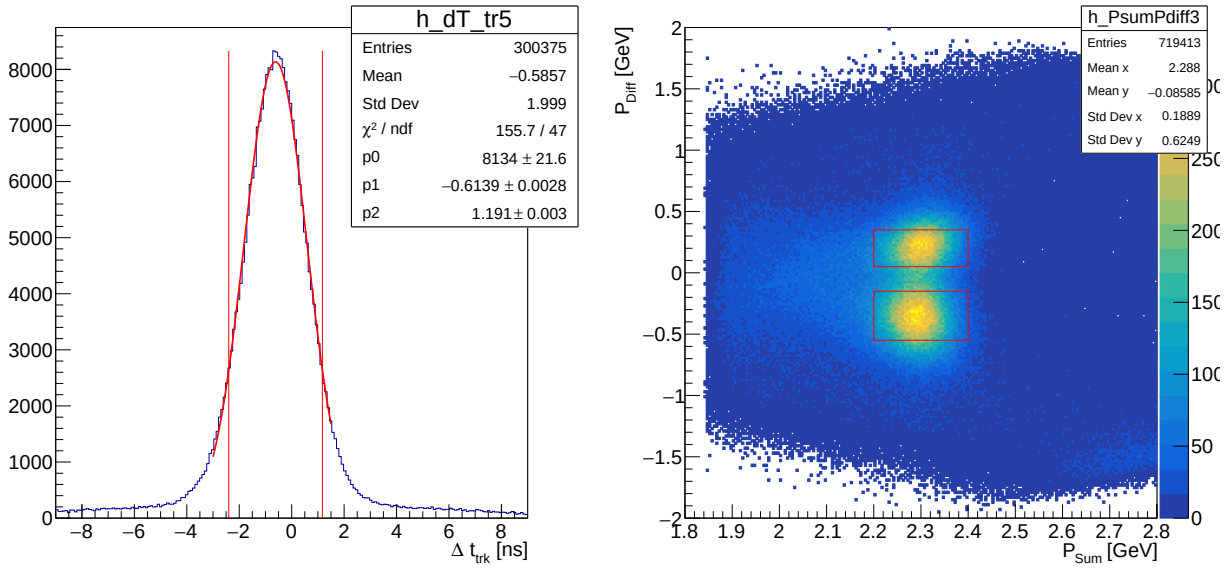
Describe L1 and d_0 cuts.
Mention why L1 is important, addition of suppressing WABs, it also significantly improves the mass resolution.

3 Parametrization of Mass resolution.

4 Bump hunt analysis

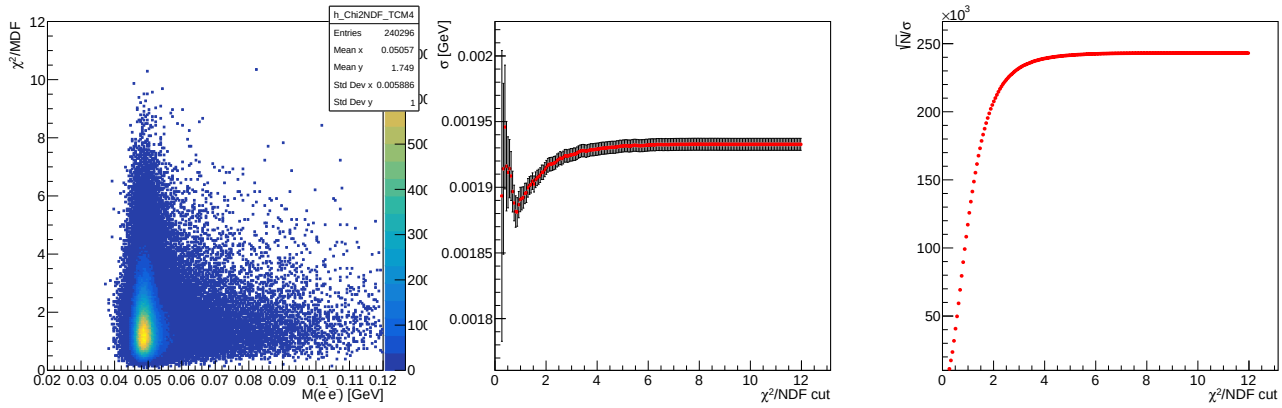
5 Study of systematics

Here goes studies on systematics.



(a) (b)

Figure 8: Left: Time difference between top and bottom tracks. Right: Momentum difference vs Momentum sum of Top and bottom tracks. Red lines represent cut limits. Each of this distributions are made when cuts were applied to the rest of Møller selection variables.



(a) (b) (c)

Appendices

A Figure of Merit in terms of Mass resolution

In general, the sensitivity for a signal (in our case a dark photon A') which is expressed in a form of a peak over a continuous background, is proportional to the number of signal events $N_{A'}$, and inversely proportional to the statistical uncertainty σ_{stat} of the distribution under the peak. So the figure of merit

is expressed as:

$$\text{FOM} = \frac{N_{A'}}{\sigma_{\text{stat}}} \tag{6}$$

The $\sigma_{\text{stat}} = \sqrt{N_{\text{Tot}}}$, and N_{Tot} is the total measured number of events in the given mass bin.

For a given A' mass, the expected number of dark photons, $N_{A'}$ events in the given mass bin can be expressed in terms of number of expected Radiative trident events N_{Rad} using the eq.(19) of [3]:

$$N_{A'} = \left(\frac{3\pi\epsilon^2}{2N_f\alpha}\right)\left(\frac{m_{A'}}{\delta m}\right) \cdot N_{\text{Rad}} = \left(\frac{3\pi\epsilon^2}{2N_f\alpha}\right)\left(\frac{m_{A'}}{\delta m}\right) \cdot N_{\text{Tot}} \cdot f_{\text{Rad}} \tag{7}$$

Here f_{Rad} is the radiative fraction, and δm is the width of the mass bin which is proportional to the mass resolution $\sim \sigma_m$ (look [3] for the description of the rest of variables). Using eq.7 for $N_{A'}$, $\sqrt{N_{\text{Tot}}}$ for σ_{stat}, and σ_m for δm, we can express FOM as

$$\text{FOM} \sim \frac{\sqrt{N_{\text{Tot}}}}{\sigma_{mass}} \tag{8}$$

References

[1] Kyle McCarty, Valery Kubarovsky and Benjamin Raydo, "Description and Tuning of the HPS Trigger", HPS Note 2018-002

[2] Holly Szumila-Vance, "HPS Ecal Timing Calibration for the Spring 2015 Engineering Run", HPS Note 2015-011.

[3] J. D. Bjorken, R. Essig, P. Schuster and N. Toro, "New Fixed-Target Experiments to Search for Dark Gauge Forces", Phys.Rev. D80 (2009) 075018, arXiv:0906.0580

[4] Miriam Diamond's talk at analysis group meeting On 16-Oct-2018, (password protected) https://confluence.slac.stanford.edu/display/hpsg/October+16%2C+2018