Resonance search analysis of 2016 HPS spring run data.

Bump hunt folks

February 15, 2020

Contents

1 Introduction

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

2 The final state of interest

Describe the Rad. Trident, BH, and converted WABs.

3 Data set

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

4 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.

4.1 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 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.

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

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

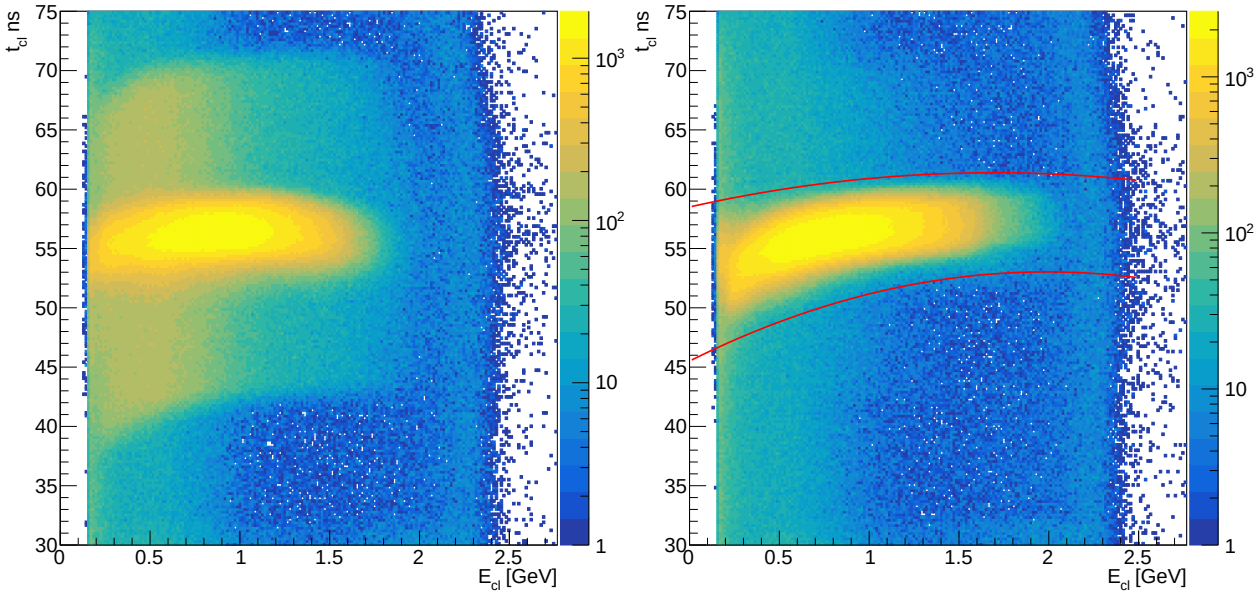


Figure 1: "time vs Energy" distributions of ECal clusters in the Top (Left) and Bottom (Right) half. Red curves in the in the right plot indicate cuts that are applied to clusters in the bottom half for the initial cluster selection. See the text for the description of the difference between left and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

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.

4.1.2 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, μ_i^1, σ_i^1, μ_i^2, σ_i^2, b.

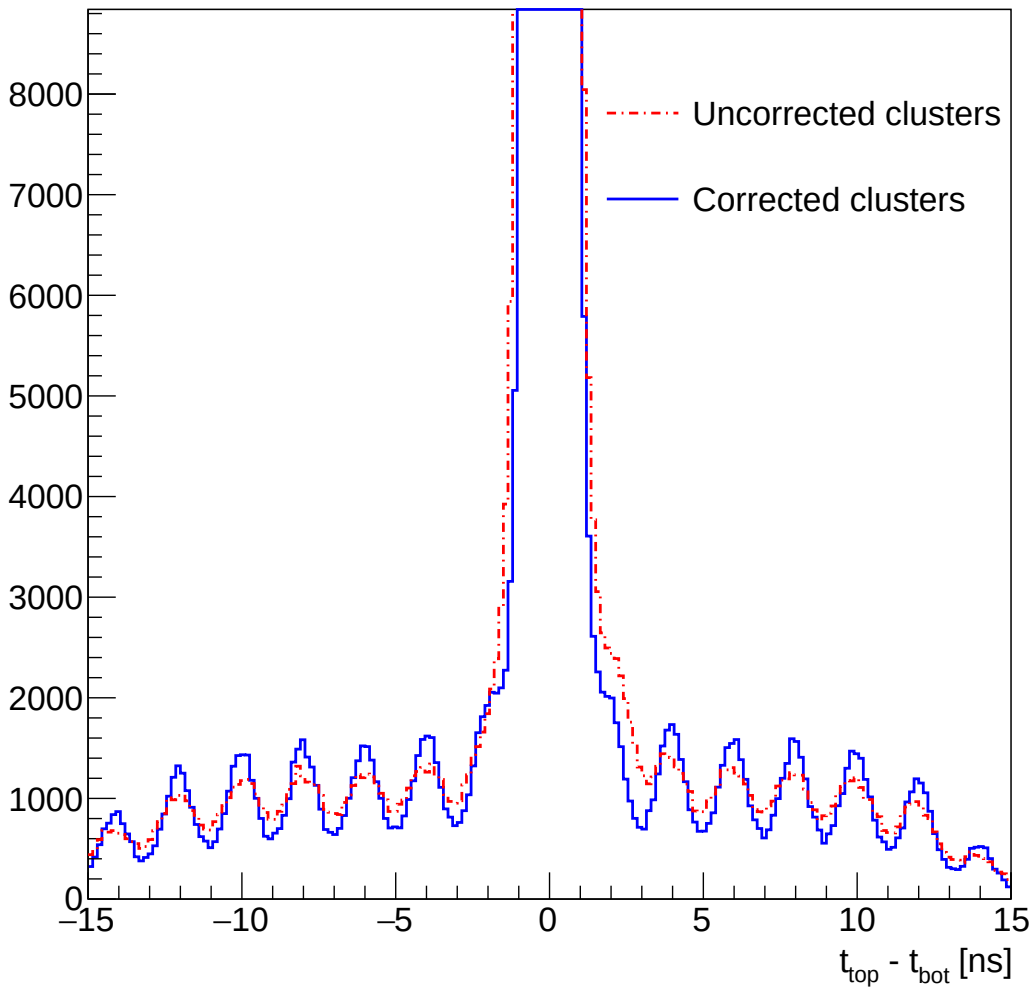


Figure 2: Time difference between top and bottom clusters at high E_{sum} region. Dashed red lines show uncorrected clusters, and the blue curve show corrected clusters.

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

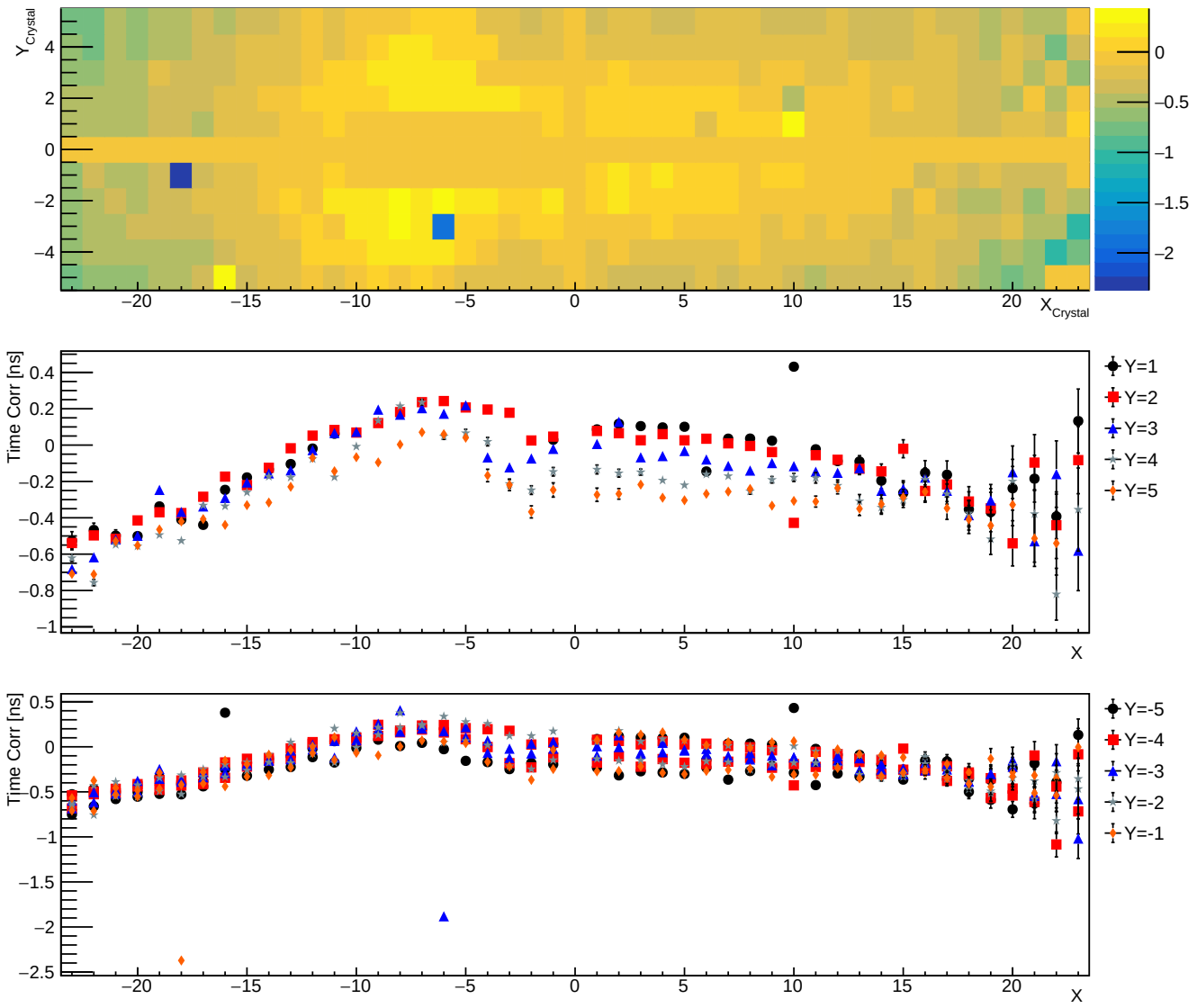
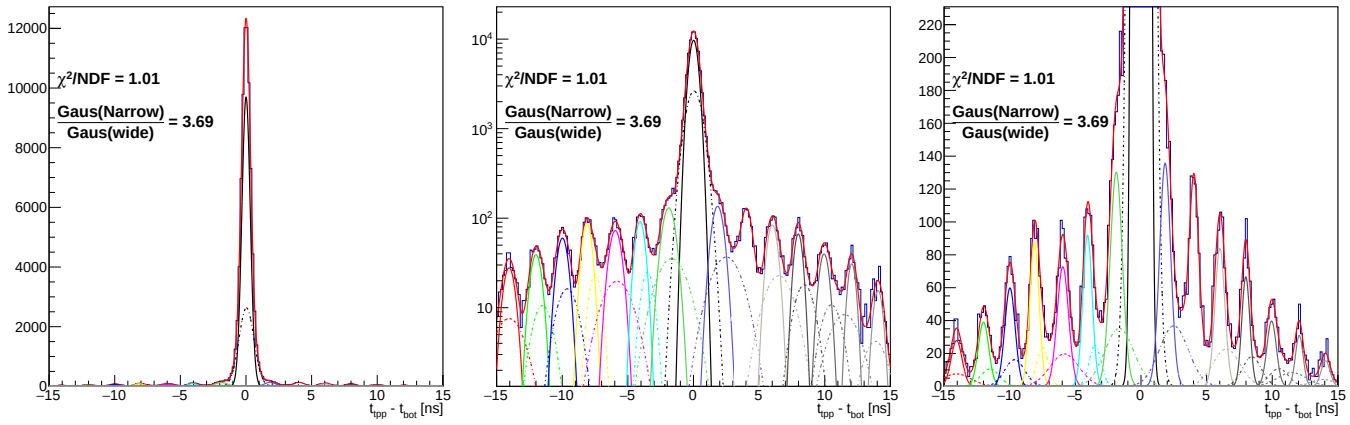


Figure 3: Time Corrections for each crystal.



(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.

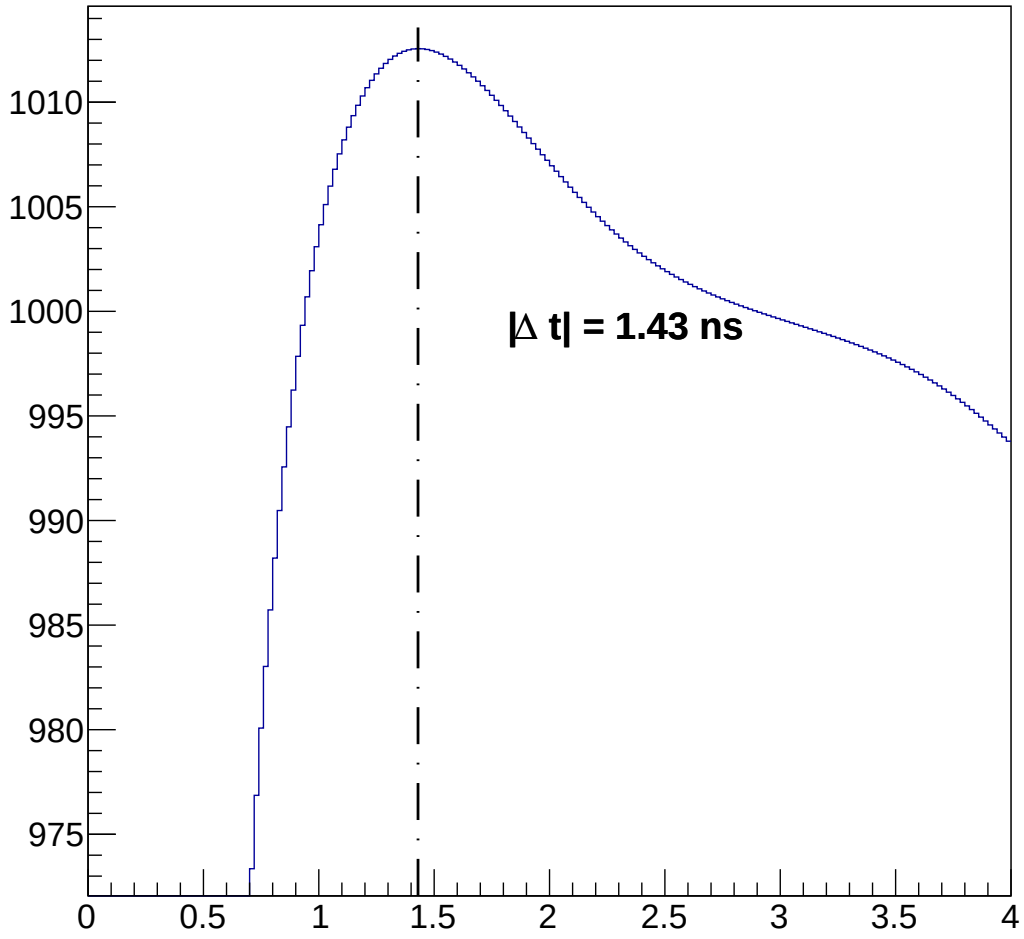


Figure 5: The $\frac{S}{\sqrt{S+Bgr}}$ ratio as a function of cluster time difference cut. The Dashed line indicates the maximum of the function.

4.2 Anti-FEE cut

The purpose of this cut is to eliminate elastically scattered electrons. In principle if we don't put any cut, then an elastically scattered electron (aka FEE) can be in accidental coincidence with a positron in the opposite half, and that electron-positron pair would look like a trident candidate. In Fig.6

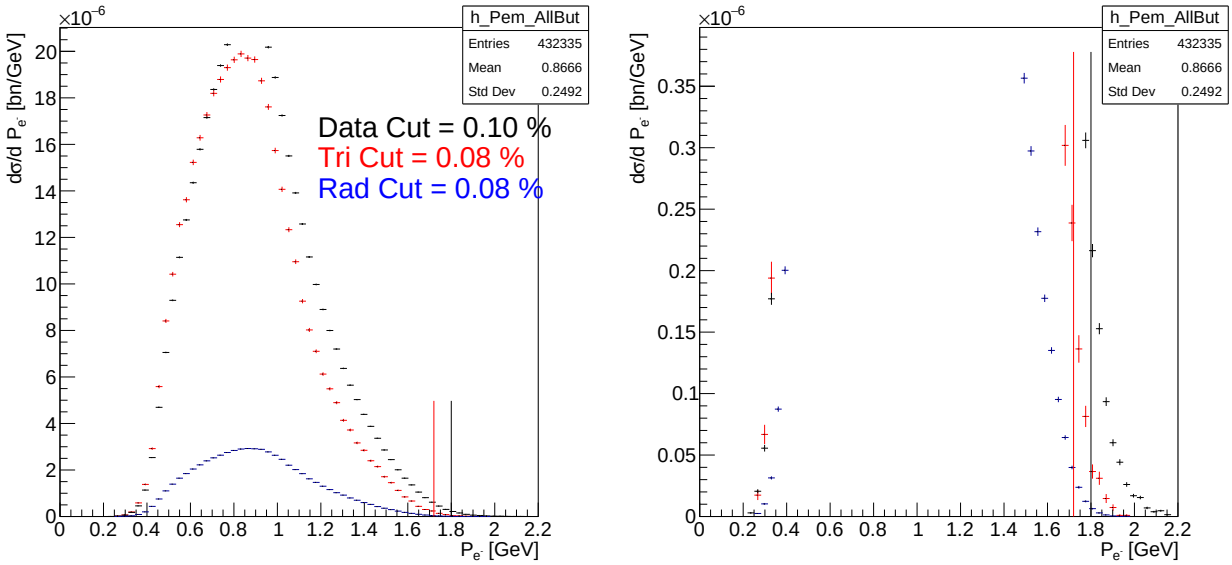


Figure 6: Electron momentum distributions obtained representing Data (Black points), Trident (Red points), and Raditative tridents (blue points). The vertical red line represents the cut. The right plot is the same distribution but zoomed that the tail of FEEs can be visible.

shown electron momentum distributions representing Data (Black points), Trident (Red points), and Raditative tridents (blue points). All these distributions passed all event selection cuts except the electron maximum momentum cut. Left and right figures are the same, the only difference is that in the right figure the vertical axis spans only small values region, where one can see the behavior of the high energy tail of the distribution. As one can see, when the rest of event selection cuts are applied, then there is no hint of the FEE that potentially could be misidentified as a trident electron. The cut is chosen such to keep 99.9% electrons from the given sample. The vertical red line represents the cut for MC samples, while the black line represents the cut value for data.

4.3 P_{Sum} Min cut

As it is explained in the introduction and in [3], A' and Radiative tridents have identical kinematics, while they are quite different from BH. We need to chose a phase space, where radiative tridents (signal) are maximized while Bgr (BH and converted WABs) are minimal. Unlike BH, radiative tridents are peaked at higher P_{Sum}, so we should chose a high P_{Sum} region. The Figure of Merit (FOM) is chosen as $\frac{S}{\sqrt{Tot}}$, where S is the number of radiative trident events, while "Tot" is the total number of observed events. In the data we of course don't have a way to know whether the given event is radtiative trident, BH or converted WAB event, and that is why we have to relay on MC in order to chose the P_{Sum} minimum momentum cut that maximize the FOM. In the left plot of Fig.7 shown difference MC

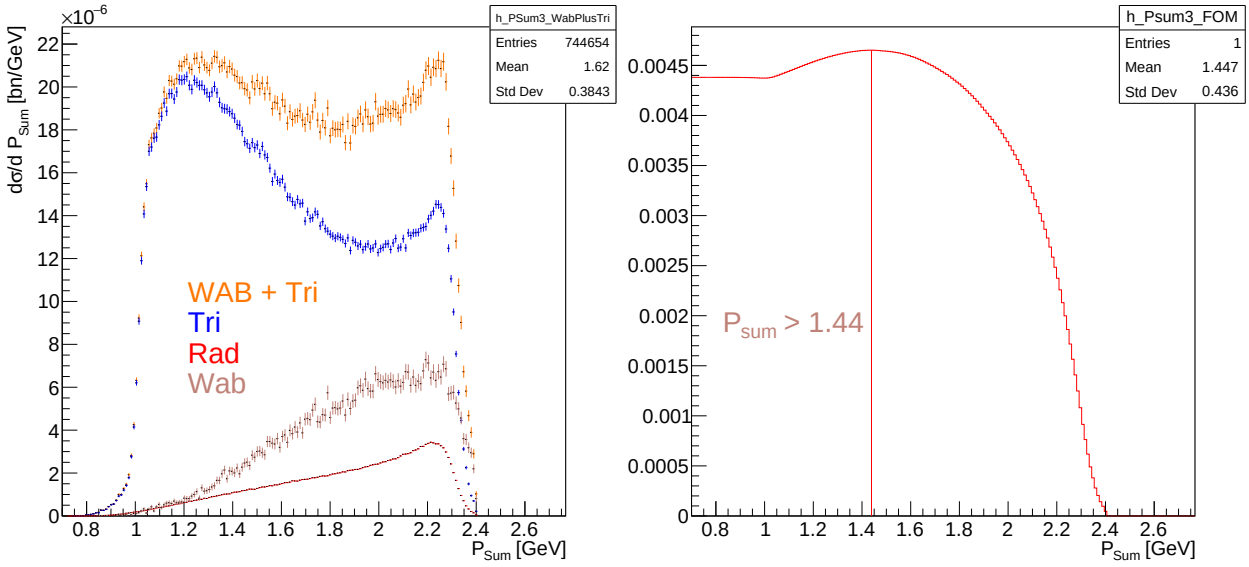


Figure 7: Left plot represents differential cross sections $\frac{d\sigma}{dP_{Sum}}$ as a function of P_{Sum} for different MC components, while the right plot represents the FOM as a function of the P_{Sum} Min cut. The vertical line on the right plot shows the cut value, where the FOM is maximum. Error bars in the left plot are square root of the square sums of statistical uncertainties and the generator cross section uncertainty.

components: Wab (brown), Rad (red), Tri (blue) and Wab + Tri (orange), when the rest of event selection cuts are applied. The right plot shows the FOM as a function of the P_{Sum} cut. The P_{Sum} value that maximizes the FOM is chosen for the P_{Sum} Min cut.

4.4 P_{Sum} Max cut

The purpose of putting a maximum limit on P_{Sum}, is to help to cut accidentals more, that could potentially pass the rest of event selection cuts. In Fig.8 shown P_{Sum} distributions for data and different MC samples. One can notice that, "Tri" and "Rad" are essentialy don't have tails, and more thean 99.9% are below 2.4 GeV, while Data and WABs have significant tail extending up to 2.7 GeV. This events above 2.4 GeV are from WABs. Please look the appendix B, for some more details on this (We still need to understand why cWABs have a longer tail, but other samples don't). We can see that

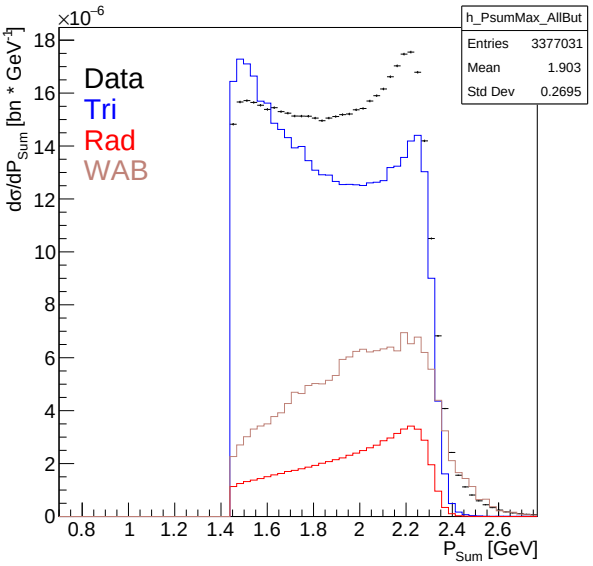


Figure 8: Observed cross section $\frac{\mathrm{d}\sigma}{\mathrm{dP_{Sum}}}$ as a function of $\mathrm{P_{Sum}}$, for Data and different MC components.

when the rest of event selection cuts are applied, "Tri" and "Rad" don't have any significant tail, so it was chosen to cut them at 2.4 GeV.

4.5 Two dimensional cuts

Some of event selections cuts described below are two dimensional cuts, i.e. the cut value depends on the value of another variable. In this analysis, these cuts are track-cluster matching cuts (time and coordinate matching), and for both time and coordinate matching cut limits are defined as 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 4.1). 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

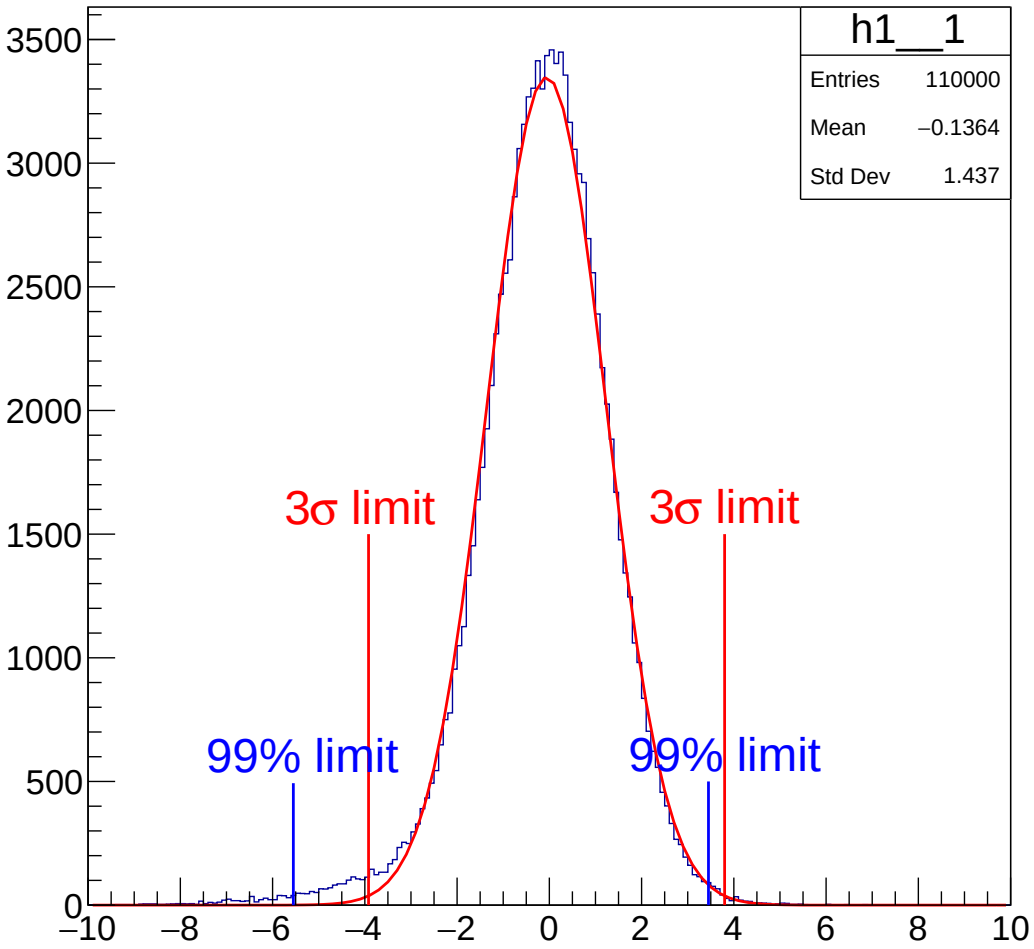


Figure 9: Illustration of 3σ cut limits vs 99% cut limits on a toy distribution.

will not keep 99.7% but rather might cut more events. As an example in Fig.9 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 of 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 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.

The number of events in the one dimensional histogram is not high, let say below 500. In this the number of events in each side of the tail will be 2 or less, and moreover, when the number of events in the histogram is less than 200 then 99% limit would correspond to less than 1 event to be in the tail. If we keep the 99% limits for all momentum bins, then for bins with small statistics (order of 500 events and less then), then the cut limits will have large fluctuations from bin to bin (momentum bin). For this reason the cut limits were tightened for momentum bins with less than 500 events. The purpose of this modifications is to make event selection cuts smoother as a function of momentum. Momentum bins with small number of events have either very low momentum (less than 400 MeV), or very high momentum (above 1.8 GeV), and they comprise less than 0.1% of events that pass all event selection cuts. 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%
$60 < N < 100$	95%
$30 < N < 60$	90%
$10 < N < 30$	75%
$0 < N < 10$	60%

Table 1: Cut limits for different statistic scenarios.

4.6 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.

4.6.1 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.10 shown Cluster-Track time difference as a function of particle momentum for electrons. Left hand plots represents electrons in the bottom half, the right hand plots represents electrons in the top half of the detector. Plots in the top row represents the data, and plots in the bottom represent MC (more specifically tri-trig sample). These distributions are constructed, when the rest of event selection cuts are applied, except the cluster-track time difference. On same figures, one can see regions marked by red "Boxes", which

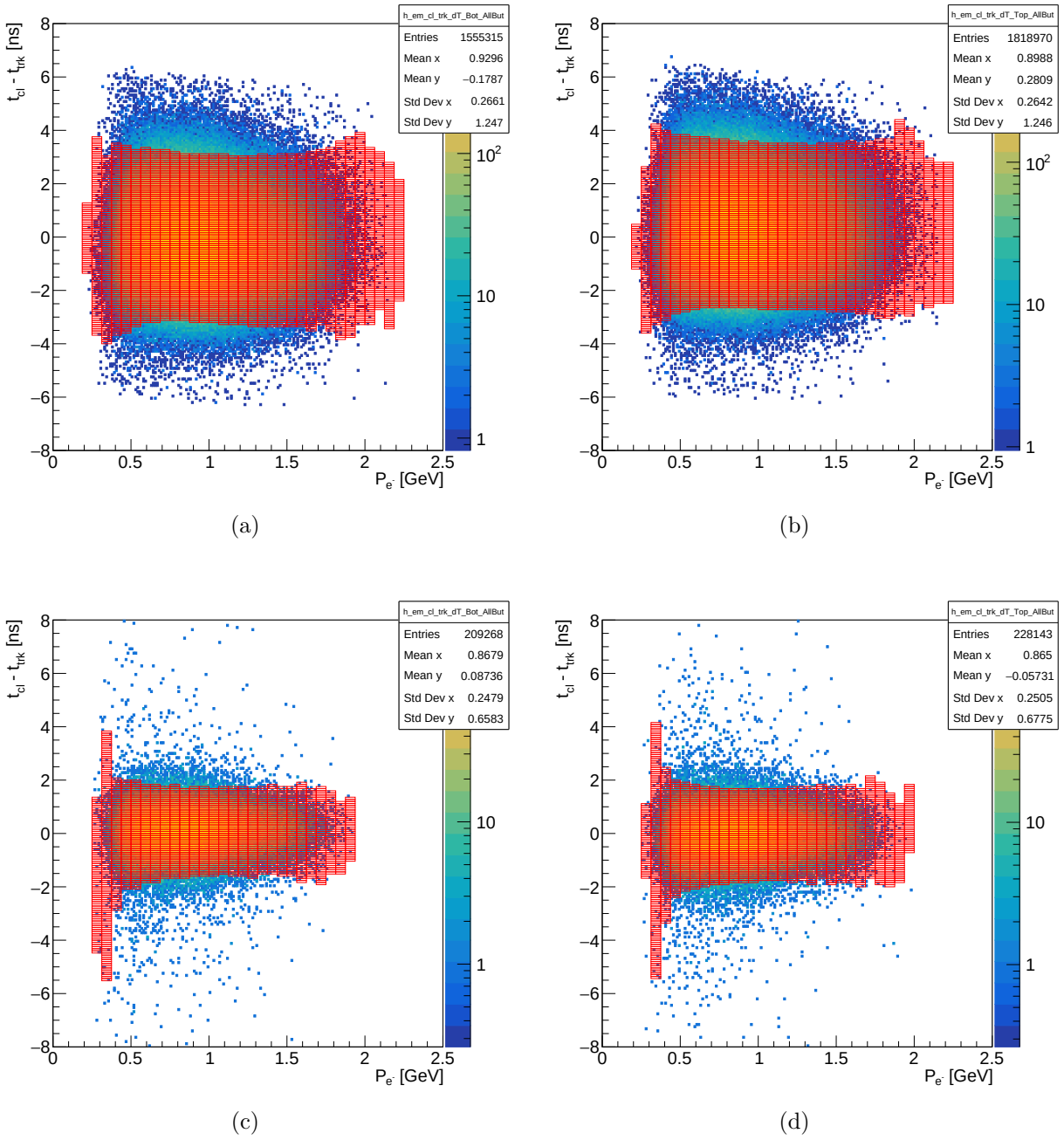


Figure 10: Cluster-Track time difference as a function of particle momentum. Left hand plots represent electrons in the bottom half, and right hand plot represents electrons in the top half of the detector. Plots in the top row represents the data, and plots in the bottom represent MC (more specifically tri-trig sample). The area marked by red "Boxes" represents the acceptance region.

represent 99% limits described in section 4.5. One can notice the MC time resolution is significantly narrow. Because of this, instead of a fix cut, for all data samples (Data, MC), we chose different cuts for Data and MC respectively, in other words these 99% cut limits for data is obtained using data events, while for MC 99% limits were obtained from MC sample. No difference is observed for positrons, and

we didn't put these plots in this document (although can be found in [5]).

4.6.2 Coordinate matching

Coordinate matching cuts are a little bit more complicated than time matching cuts. Initial studies showed that in addition to top/bottom difference, there is also significant difference depending on the track's charge, and whether the track has a hit in L6 or not. We will use only the horizontal "X" coordinate matching, not requiring anything on "Y". We haven't required "Y" coordinate matching, because there is non-negligible edge effects, which requires significant amount of efforts to properly take into account, while the cut itself, will not improve the track-cluster matching noticeably. In Fig.11

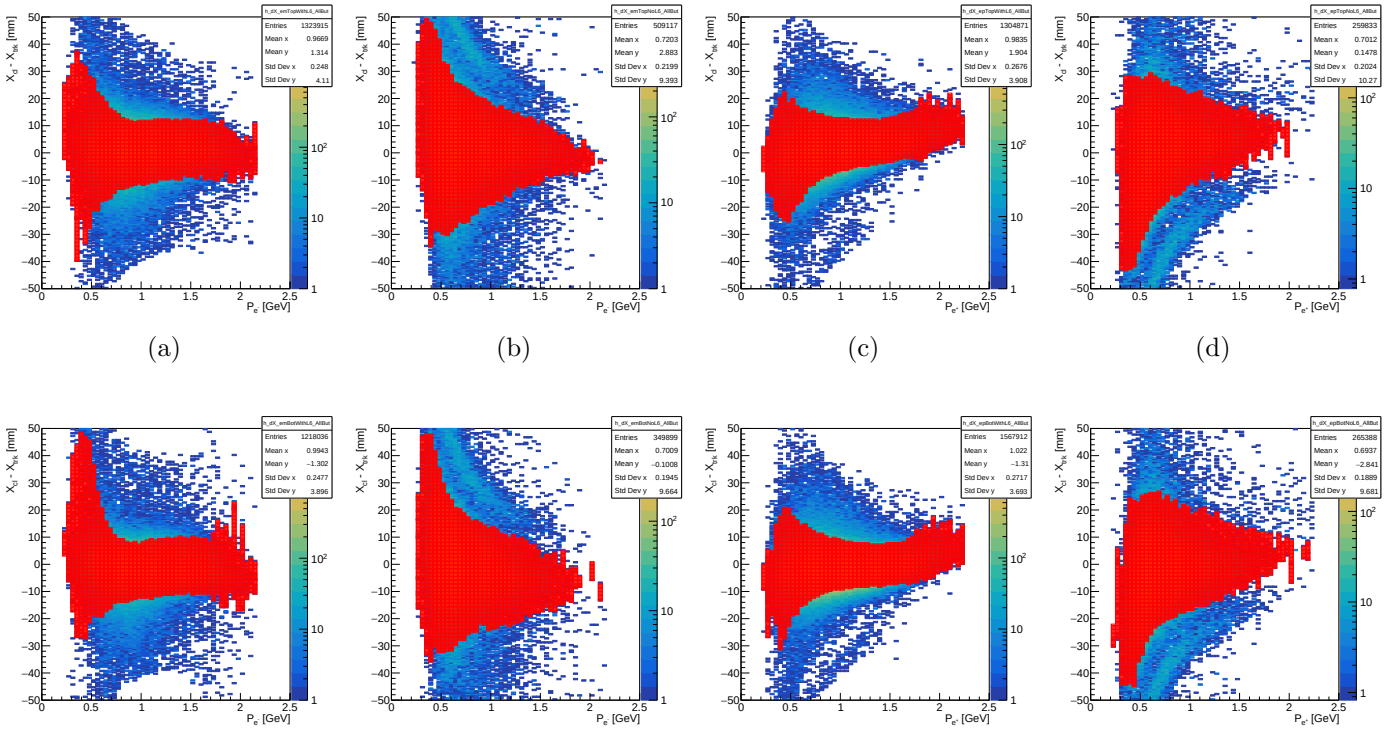


Figure 11: Cluster-track "X" coordinate difference as a function of momentum. This data represents the run 8099. Figures in the top row represent particles that are in the top half of the detector, and consequently figures in the bottom row represents bottom half of the detector. Figures a,b,e and f represent negative particles, and figures c, d, g and h represent positive particles. Figures a, c, e and g are tracks that have a 3D hit in L6, while b, d, f and h corresponds to tracks that don't have 3D hit in L6. Note: Each histogram is filled when the rest of event selection cuts are satisfied.

shown cluster-track "X" coordinate difference as a function of the particle momentum. These matching plots are categorized depending on the charge of the track, presence of a 3D hit in L6 and whether the track is in the bottom or top half of the detector. Caption of the figure describes the category of each plot. As one can easily see, distribution shapes significantly differ depending on the presence of a 3D hit in L6. They differ (not so strongly) also for negative and positive particles, and for bottom and top tracks as well. One also can notice that especially for "No L6" tracks, there is an outlier band. This fact is counterintuitive, since for a given momentum bin, one would expect a core (close to Gaussian),

with some continuously decreasing tails for the "ΔX distribution. The number of these events is of the order of 1% for the given category. We have performed initial study, to find out the cause of these events. No definite answer was obtained, however we observed that these events are not located in any specific part of the calorimeter, but rather they are distributed over all the calorimeter face. As an example in Fig.12 shown cluster "Y sv X" coordinate distribution of negatively charged outliers in the bottom half. Given the fact that this events comprise only order of 1% for "No L6" category (and sub percent level of overall data), it was decided to not further investigate, but rather just cut these events out. As in the case or cluster-track time difference cut, here also in Fig.11 by red "Boxes" shown the

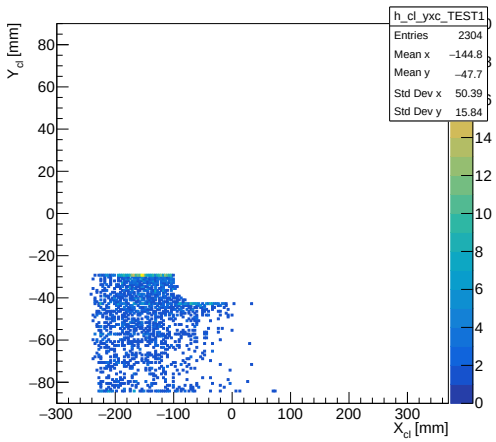
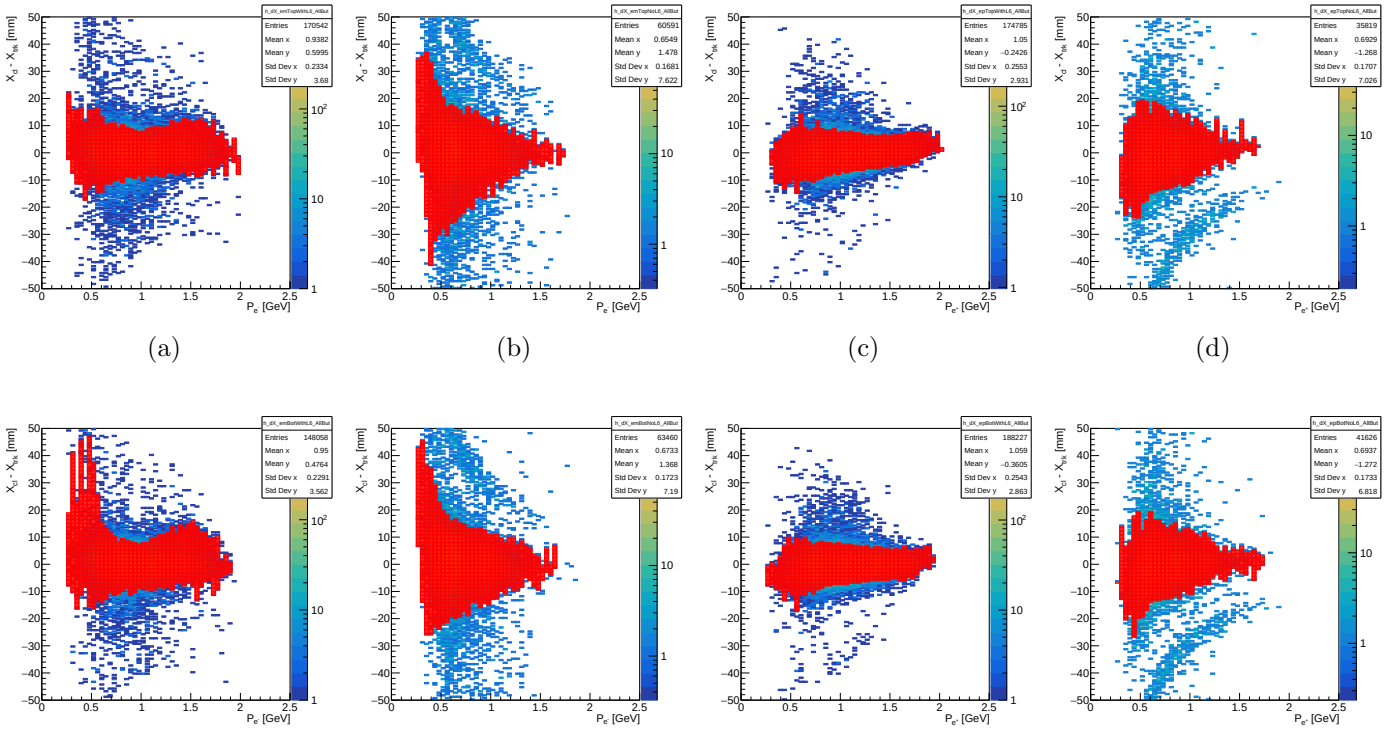


Figure 12: cluster "Y sv X" coordinate distribution of negatively charged outliers in the bottom half.

99% cut region after cutting outliers (need to add functions on the same plot, that shows how outliers are cut, but can be done during, or after the review).



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 13: Cluster-track "X" coordinate difference as a function of momentum. This data represents the tritrig MC sample. Figures in the top row represent particles that are in the top half of the detector, and consequently figures in the bottom row represents bottom half of the detector. Figures a,b,e and f represent negative particles, and figures c, d, g and h represent positive particles. Figures a, c, e and g are tracks that have a 3D hit in L6, while b, d, f and h corresponds to tracks that don't have 3D hit in L6. Note: Each histogram is filled when the rest of event selection cuts are satisfied.

4.7 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, 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).

Using Møller events (Reference to Møller section, but Møller selection is not complete yet.) we have not observed a local peak as a function of $\chi^2_{\text{Sum}}/\text{NDF}_{\text{Sum}}$, which suggests that we should not put any cut on this variable, and hence we have not applied any cut on it.

4.7.1 Selection of Møller events

This section is not yet finished

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.

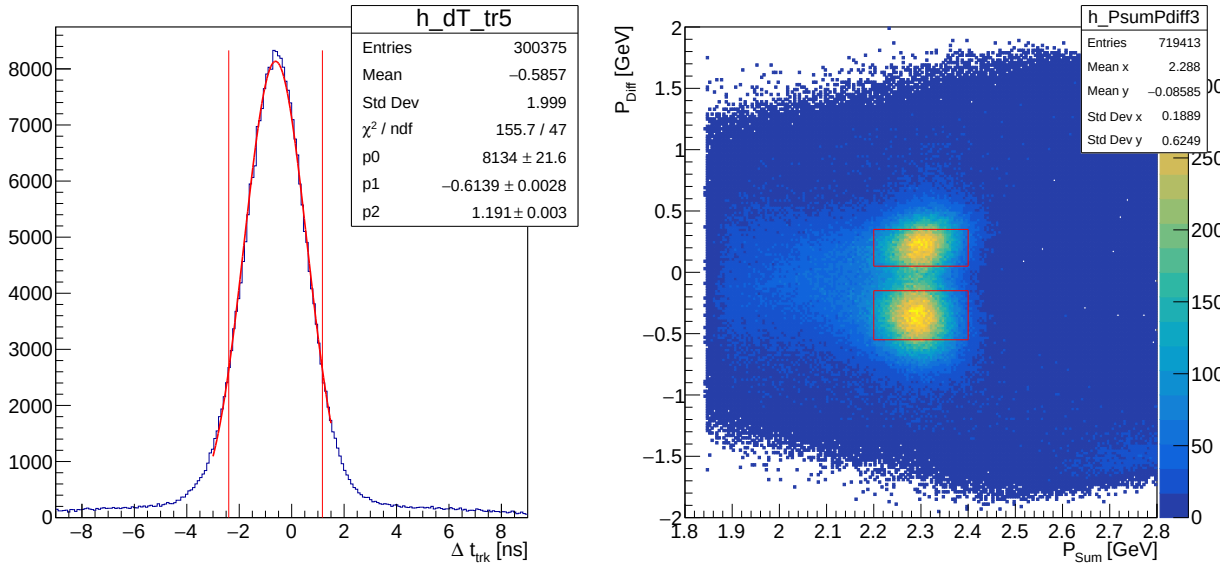
- Δt_{tr} time difference between top and bottom tracks. Cutting Δt_{tr} around 0, will suppress accidentals coming from different beam bunches.

- 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 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 cut is also based on the specific acceptance of Møller events. Only certain momentum configuration could potentially be inside the detector acceptance. This also will help to suppress accidentals.

- 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.

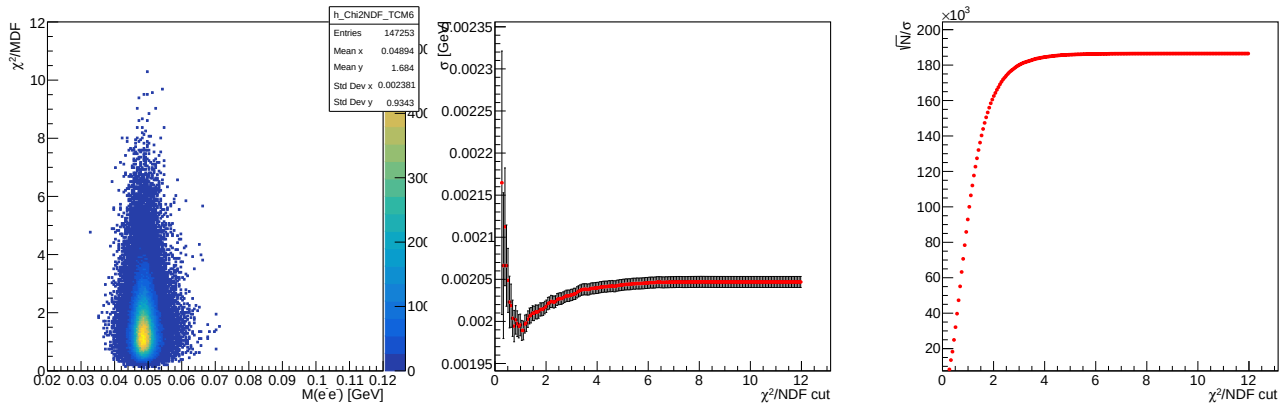
Similarly to the rest of event selection cuts for the trident final state, here for Møller event selection also, in order to understand where to put cuts on a given variable, cuts are applied to the rest of variables mentioned above, then the distribution of the given variable is studied. In Fig.14 shown Time



(a) (b)

Figure 14: 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.

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** .



(a) (b) (c)

4.8 Summary of Event selection cuts

This section summarizes the event selection cuts. The table 2 describes how much of the data is survived when a given cut is applied at the end, i.e. when the rest of cuts are applied before applying the given cut. As in the data we don't know the underlying process of the given event, we should focus on

Cut variable \ Data set	Data	Tri-beam	Rad-beam	Wab-beam	Tri + Wab
PsumMax	0.986362	0.998802	0.998741	0.955327	0.986328
PsumMin	0.687376	0.601244	0.865508	0.913744	0.664391
clDt	0.97363	0.995125	0.997946	0.9949	0.995062
Pem	0.997601	0.999063	0.999102	0.99763	0.998665
em_cl_trk_dT	0.989403	0.988463	0.995261	0.98833	0.988456
ep_cl_trk_dT	0.989087	0.989022	0.99504	0.964045	0.987673
dX_em	0.985935	0.982492	0.984343	0.981353	0.982432
dX_ep	0.984304	0.982944	0.983771	0.97479	0.982511

Table 2: The relative amount of events surviving each cut. The first column is the variable name on which the cut is applied, and the first row describes the data sample.

comparing the cut effects on "Data" and "Tri + WAB" sample which is our best known MC sample for the description of the data. In general all cuts throw consistent amount of data from the "Data" and "Tri + WAB" sample. The largest relative cut difference in between "Data" and "Tri + WAB" are from the cluster time difference cut "clDt" and the "PsumMin" cut. The discrepancy between this cuts is of the order of 2%. The discrepancy between the "Data" and "Tri + WAB" samples is in a sub percent level. The reason of difference for for "clDt" cut is that in MC cluster time resolution is significantly better than in data, and also non perfect time calibration of the data produces longer tails

in the data for the "clDt" distribution which is not present in MC (I will put a plot similar to the 4c for data and MC but with n-1 cuts, demonstrating the statement above.). There is also a significant difference in the data and MC, that could cause difference in the "PsumMin" cut. In particular the tritrig sample is generated with a PSum > 1GeV condition (Refer to Tongtong's MC samples section.) while in data there is no sharp cut on "PsumMin".

Momentum sum distribution for the final selected trident samples is shown in Fig.15. Data is

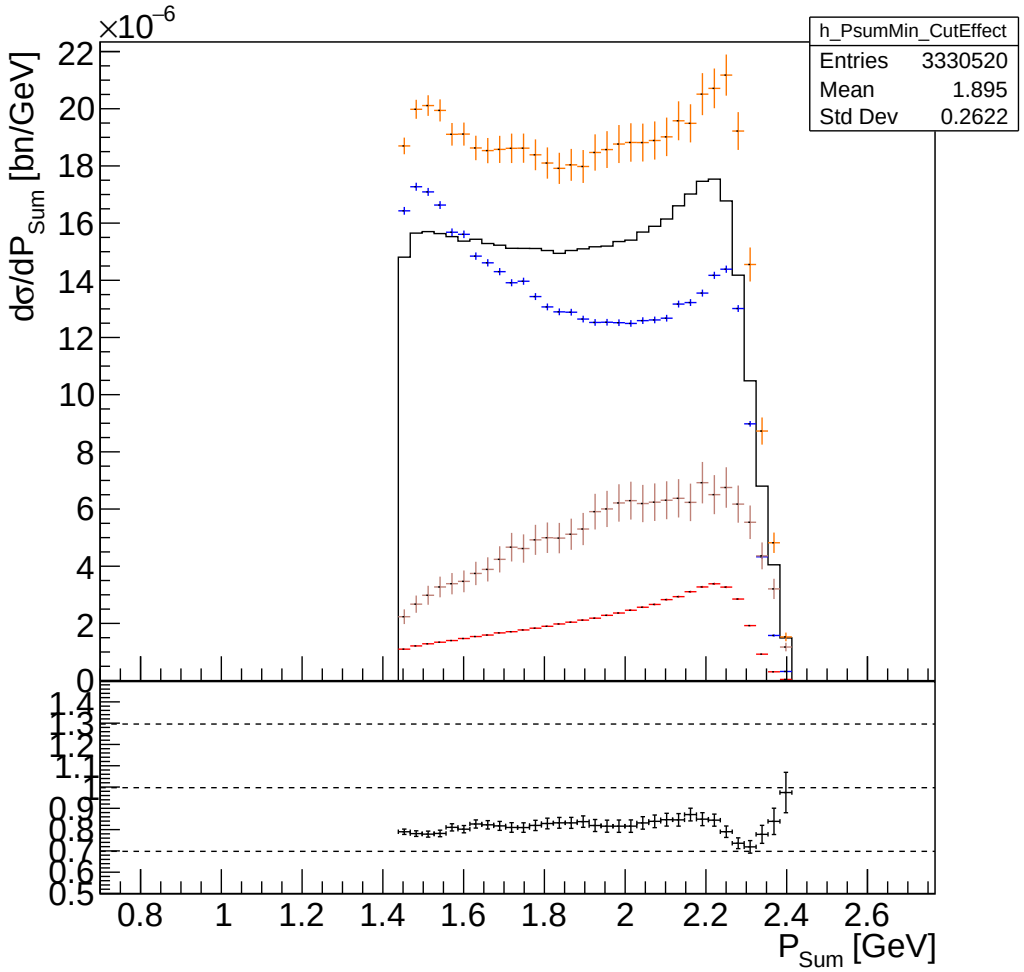


Figure 15: Differential cross section as a function of momentum sum. Data is presented by black curve, Blue markers represents "Tridents", Red markers represent Radiative tridents, Brown represents converted WAB and the orange color represents the sum of tridents and converted WABs (Tot). The ratio plot at the bottom is the ratio of data over the Tri + WAB.

presented by black curve, Blue markers represents "Tridents", Red markers represent Radiative tridents, Brown represents converted WAB and the orange color represents the sum of tridents and converted WABs (Tot). The ratio plot at the bottom is the ratio of data over the Tri + WAB. In general the data is about 15% - 20% below the MC prediction. One of the reason of this much discrepancy between the data and MC that we haven't yet took into account is the difference in the tracking efficiency between the data and MC. (Will be good if Matt G. can write few sentences here describing reasons for Data and MC eff difference, and why it is not so trivial to properly account it.)

With this event selection cuts, the 10% of the full 2016 data was analyzed, and the final invariant

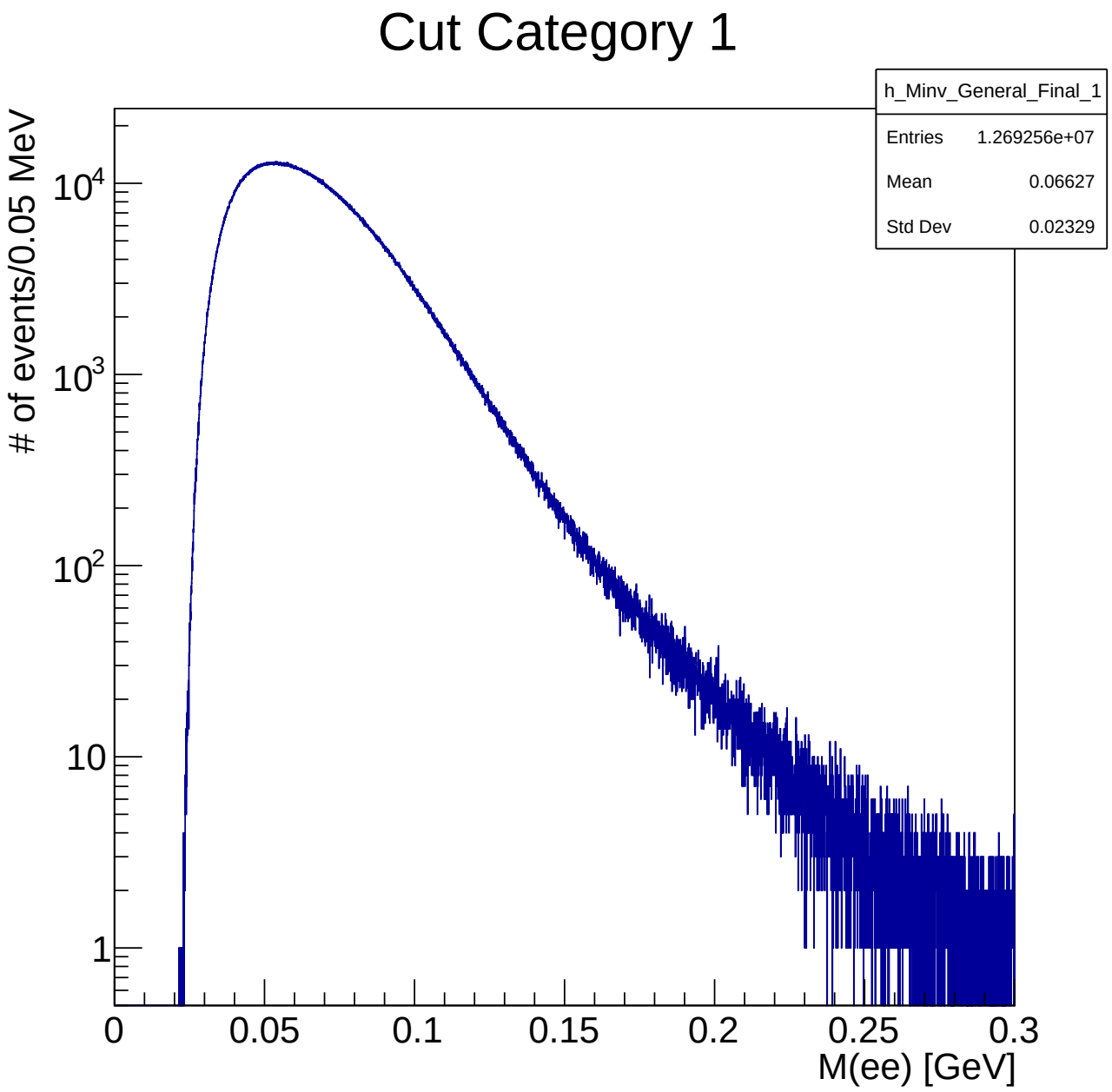


Figure 16: The Final invariant mass distribution representing the full 10% sample of the 2016 run.

mass distribution is shown in Fig.16.

5 Radiative fraction

6 Parametrization of Mass resolution

7 Run by run stability

8 Bump hunt analysis

9 Study of systematics

Here goes studies on systematics.

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

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

A.1 WAB Suppression cuts

In the section 2 it is described that, the two step process cWAB $eA \to eA\gamma(\to e^-e^+)$ can mimic the trident final sate. As the photon conversion can occur in the target, it also can occur in any layer of SVT too. While the photon conversion in layers 3, 4, 5 and 6 will not yield to reconstructed trucks (since tracking requires at least 5 3D hits), conversion in the first and second layer (if conversion occur in the axial sensor of layer 2) could cause both or either e^- or e^+ to be reconstructed. One of handles to suppress cWABs, is to require a presence of a hit in L1 from e^+ tracks. This should suppress cWABs significantly (photon conversions in L2 and in L1 stereo sensor), while should not significantly impact tridents that produced in the target. In this particular analysis, however, we will not require the presence of a hit in L1, and the reason is that at this moment we have significant discrepancy in the Data and MC L1 hit efficiency, which in turn translates to more discrepancy between data and MC rates.

A.1.1 d_0 cut

Another strong cut that allows to suppress cWABs is the cut on the distance of closest approach (d_0). The reason is that, if the photon conversion occurs not in the target, but in any layer of SVT, then

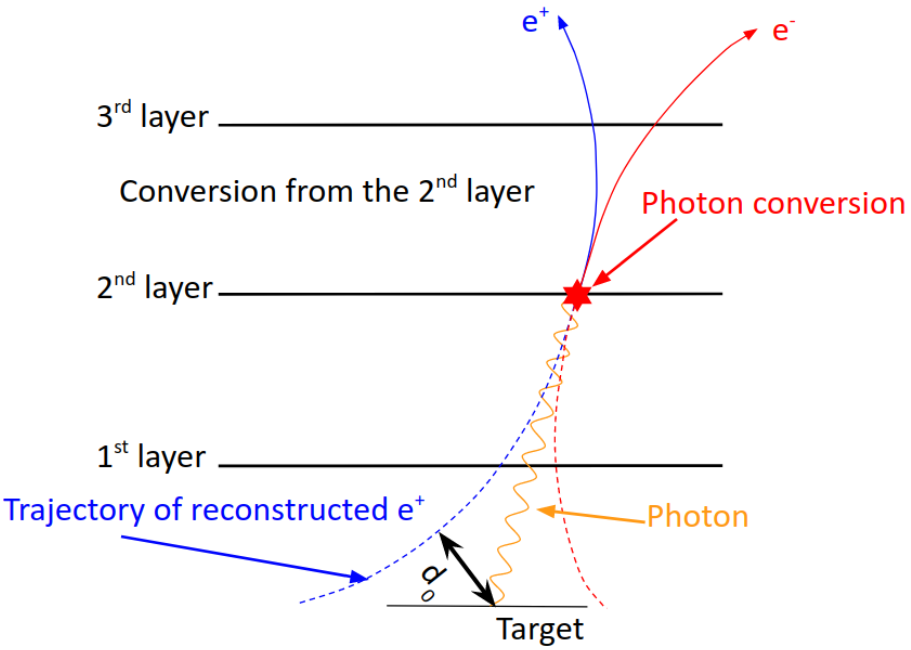
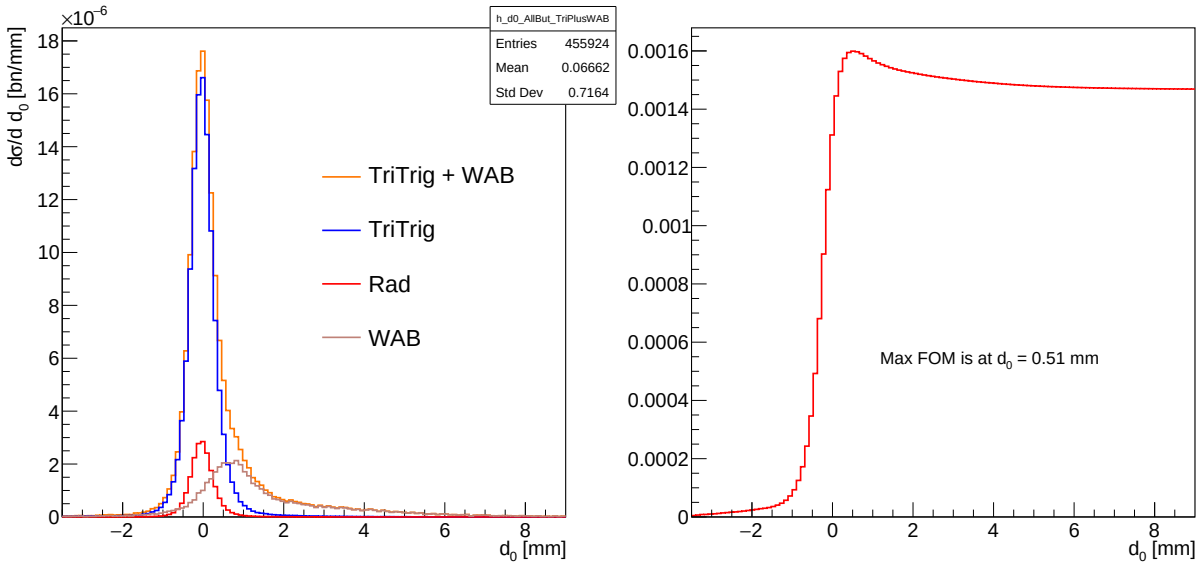


Figure 17: A sketch diagram representing a photon conversion in the 2nd layer of SVT. It also demonstrates that the reconstructed trajectory will be displaced from the photon production point.

the reconstructed trajectory will not pass through the beam-target interaction point, but rather will be systematically displaced from the interaction point in the target. This signature is one of important differences between the $eA \to Ae^-e^+e^-$ process and the two step process 1st: $eA \to Ae\gamma$ then 2nd $\gamma \to e^-e^+$. The cut value is determined using both, the data and MC simulations. The objective is to chose a cut value, that will maximize the reach of the experiment. Similarly to several of other cuts, the FOM is chosen as

$$\text{FOM} = \frac{\text{S}}{\text{Tot}} \tag{9}$$

where "S" is the number of Radiative events, and "Tot" is the total number of events, which is the sum of Rad. tridents, BH (including interference term between them) and cWABs. In data we don't know which of the abovementioned processes is responsible for the given event. Instead we will use the MC simulations for all abovementioned components. In Fig.18a shown differential cross section



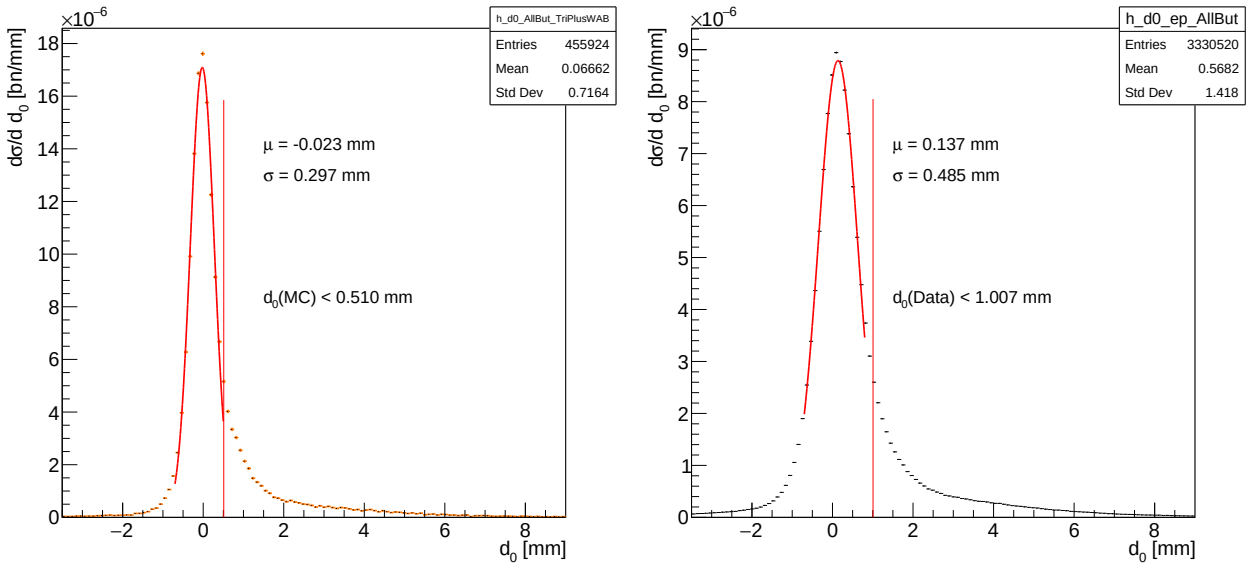
(a) (b)

Figure 18: Left: Observed cross differential cross section as a function of d_0 for different MC components (described in the text and in the figure). Right FOM as a function of d_0 cut value.

$\frac{d\sigma}{dd_0}$ for all MC components as a function of d_0. As one can note from the Fig.18a, d_0 is peaked at 0 for Rad and Tridents, but for cWABs the peak is shifted, and has a quite long tail on the positive d_0 side. The Fig.18b shows the FOM defined by eq.9 as a function of d_0 cut value. This value obtained by maximizing the FOM using only the MC will not represent the FOM for the data sample, because the d_0 resolution in Data and MC are significantly different. As one can see from the Fig.19, the d_0 resolution in data is about 1.6 times wider than the one for MC. Instead for the data sample, the d_0 cut is chosen in a way, that the distance between the cut value and the main peak is the same in the units of the d_0 resolution (σ_{d_0}). So in data the d_0 cut value is defined as:

$$d_0{}^{\text{cut}}(\text{Data}) = \mu_{d_0}(\text{Data}) + \sigma_{d_0}(\text{Data})\frac{d_0{}^{\text{cut}}(\text{MC}) - \mu_{d_0}(\text{MC})}{\sigma_{d_0}(\text{MC})} \tag{10}$$

Mean, σ and cut value on d_0 for data and MC are shown in the table.3.



(a) (b)

Figure 19: differential cross section as a function of d_0. Left figure represents MC data (Tritrig + cWAB), while right figure represents the data.

	MC	Data
mean of d_0 [mm]	-0.021	0.141
σ of d_0 [mm]	0.296	0.484
d_0 cut value [mm]	0.51	1.011

Table 3: Mean, σ and cut value on d_0 for data and MC.

B Data and MC comparison at high d_0

Although in this analysis we don't cut on d_0, however during the initial event selection studies, it was observed that, when we cut on d_0 (see appendix A.1.1 for details about the d_0 cut), the tail in P_{Sum} distribution vanishes, while when there is no cut applied on d_0, then "Data" and "WAB" tails are quite close to each other (see fig.8). To check, if at high P_{Sum} the data represents the WABs, rather than something else, we have applied a reverse cut on d_0, i.e. selected events with $d_0 > 2.25$mm, and applied the rest of event selection cuts (exept cuts on P_{Sum}). As you can see from the Fig.18a and Fig.19b, this high d_0, could have data and WABs (contributions from "Tri" and "Rad" are very small). In Fig.20

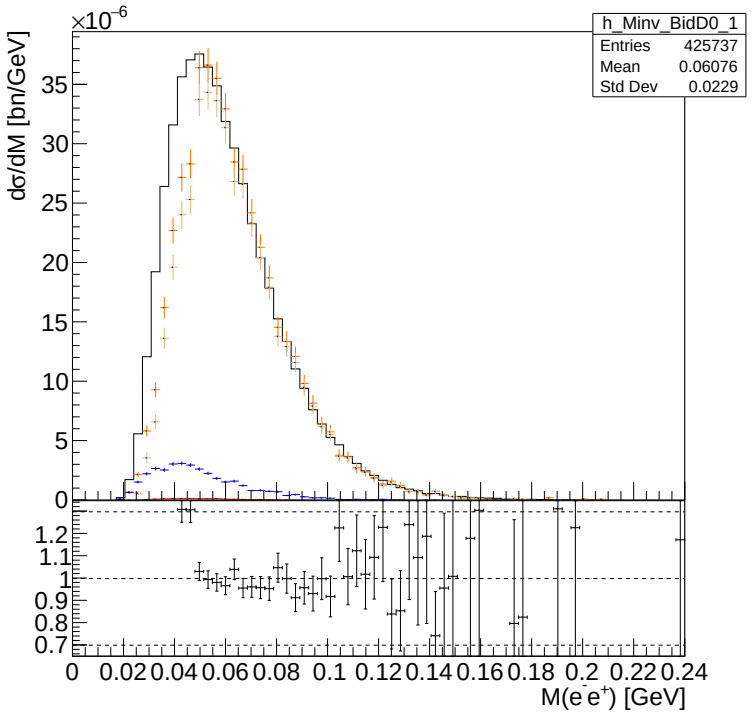
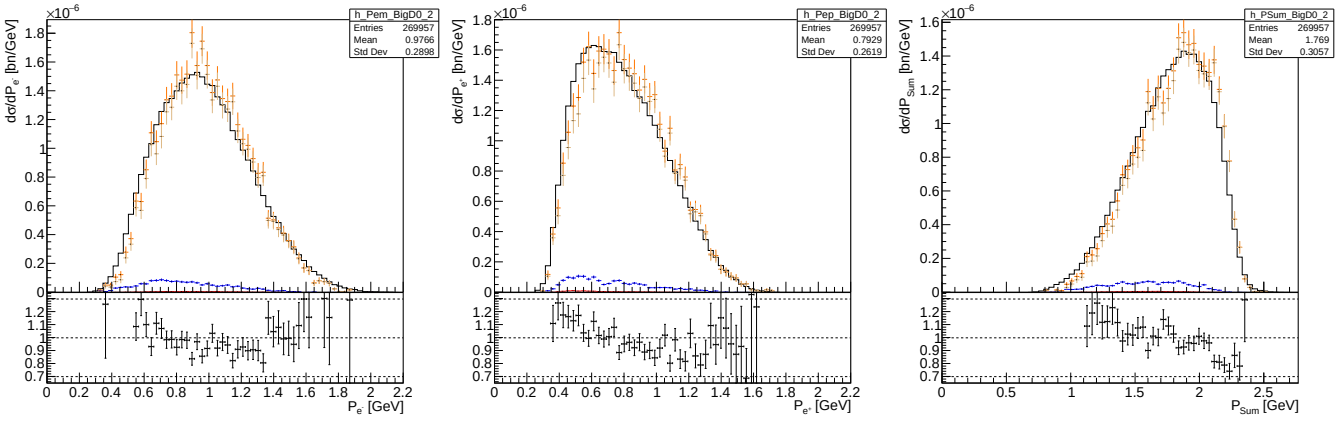


Figure 20: The differential cross sections as a function of the lepton pair masss, when a reverse d_0 cut is applied. Black is tata, Brown is WAB, and by Orange color represented the 'Tri" + "Rad" component. Radiatives are represented by the Red color, however it has negligibly small contribution (almost unnoticeable).

shown data and MC comparison for such high d_0 events. Above M > 50 MeV Data and MC agree to each other reasonably well. The discrepancy at small mass cound be because of the small differences in the real and MC acceptances. In Fig.21 shown data and MC comparisons of electron momentum (left), positron momentum (middle) and the momentum summ (right). For high masses M > 50 MeV, WAB mc describes data reasonably weel. It is not perfect, but it is better than the Data and MC comparison for the full trident sample. This is not a poof that in all kinematic region (even smaller d_0 values) WAB will describe the data, however hit hints that the high end tail in the P_{Sum} distribution is from the WAB sample.



(a) (b) (c)

Figure 21: Comparison of Data and MC, for Electron Momentum (Left), Positron momentum (middle), and the momentum sum (right). These are events with M $>$ 50 MeV.

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

[5] Additional materials/plots for event selection analysis. https://github.com/rafopar/BumpHunt_2016/blob/master/pass4/Latex/EvSelectionPlots.pdf