Ingredients for Analysis

Matt Graham
HPS Analysis Workshop
October 18, 2022

Introduction

- Basic recipe for bump hunt or vertexing searches
  - Selection: we have base selections from 2016; likely could be improved for some gain
  - Radiative fraction for A' production (m,ε): this comes from MC…is it reliable? Cross-checks…
  - Mass resolution: does MC match data (no, it does not)
  - Acceptance x efficiency vs Z (displaced vertex only)
  - Signal extraction/Limit setting: fit in mass for BH; cut-and-count (basically) for vertexing (this should be revisited)
- Corrections needed for MC usage
  - Mass resolution – Rafo has a nice recipe for smearing momentum resolution and so far it's good enough. Will it continue to be?
  - Hit and Tracking efficiency
    - Tracking was generally "ok" but hit efficiency in layer 1 was bad
      - Implemented slope-based L1 hit killing for 2016
    - This gets more complicated for 2019/21: Kalman, more layer combos
  - Beware of cutting too hard on variables where MC and data don't agree well…the MC does a pretty good job replicating most observables but it's not perfect.
  - Hopefully, we can improve the data calibration (energies, alignment, timing) and apply smallish smearing (e.g.) at appropriate places in the MC enough so that any post-reconstruction data and MC match up even better this cycle…discussion in MC talk

Backgrounds and how to reduce them

random garbage
→
Track Chi Squared
Vertex Chi Squared
Cluster Match
Very high/low momentum

tridents
cWABs
accidentals
→
High pSum cut (tridents)***
Cluster & track time (accidentals)
Layer requirements or splitting (cWABs)
Positron d0 (cWABs)

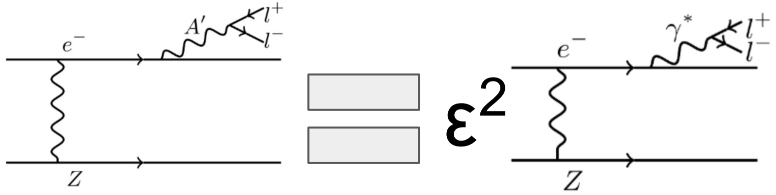
*** not for SIMPs!

hit miss-ID
hard scatter
double hard scatter
downstream trident production (TBD)
→
Hit isolation in first layer(s)
Track impact parameter
V0 projection to beamspot
Unbiased track projection at first layers (TBD)
Fit tracks with BS constraint (TBD, unformed)

Handles for separating signal from background

- The most important handles we have are pretty much in descriptions…bump hunt, displaced vertex…looking for a peak in mass and/or displaced vertex is very powerful. Those, along with high pSum cut (for nominal A’ searches) are the main drivers.
- We may be leaving out a lot of discrimination power, particularly for displaced vertex
  - many of the typical variables we look at have some correlations, some of them funny.
  - An MVA is the way to go but do we have good training samples?
- Always keep in mind systematic errors…

$$\left.\frac{d\sigma_{A'}}{dm}\right|_{m=m_{A'}} = \frac{3\pi m_{A'}\epsilon^2}{2N_{\text{eff}}\alpha}\left.\frac{d\sigma_{\gamma^*}}{dm}\right|_{m=m_{A'}}$$

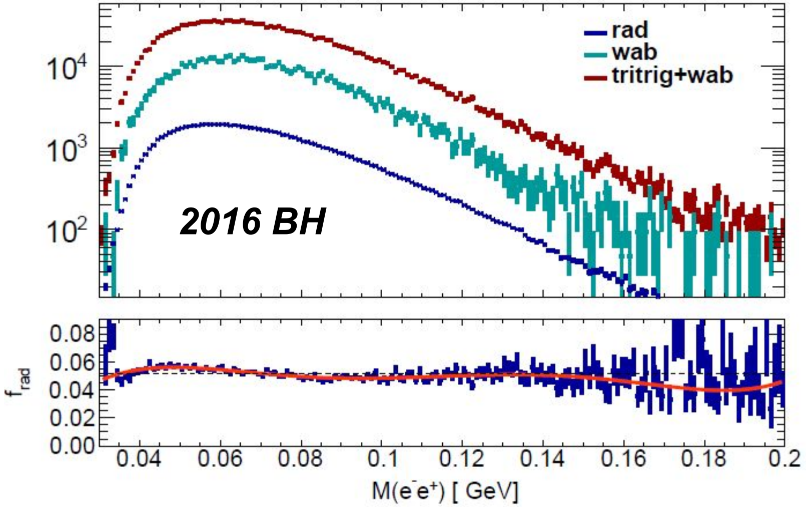


We substitute radiative rate by fraction of radiative events (f_{rad}, from MC) and the total rate observed (N_{bkg}, from data)

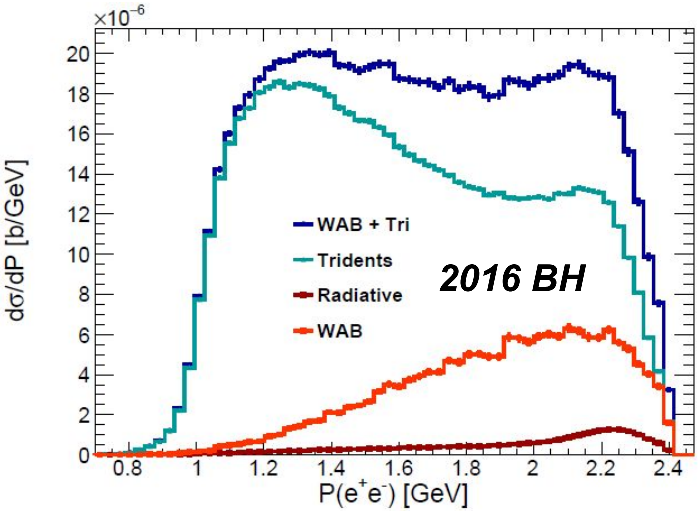
$$f_{\text{rad}} = \frac{\frac{dN_{\gamma^*}}{dm}}{\frac{dN_{\text{bkg}}}{dm}} = \frac{\frac{dN_{\gamma^*}}{dm}}{\frac{dN_{\text{tri}}}{dm} + \frac{dN_{\text{wab}}}{dm}}$$

Why? Could just get radiative rate from MC but that forces you to get absolute acceptance and efficiency correct (plus believe the MC generator). Using the fraction, many of these potential systematic effects cancel.

$$\epsilon^2 = \frac{2\alpha N_{\text{sig}}^{\text{up}}}{3\pi m_{A'} f_{\text{rad}} \cdot \frac{dN_{\text{bkg}}}{dm}}$$



Another plot on sample composition

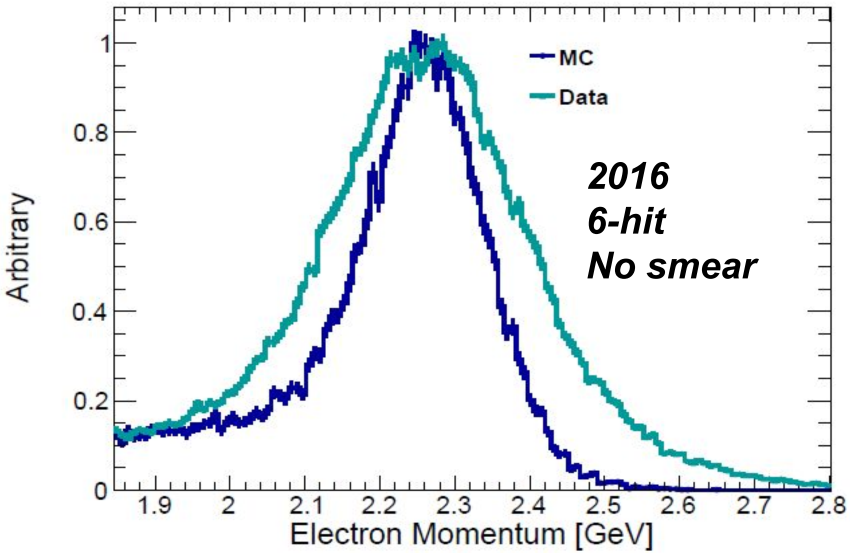


This shows how the different categories of events look like before the pSum cut … note how the cWAB shape looks similar to the radiative (i.e. signal)

Momentum Resolution

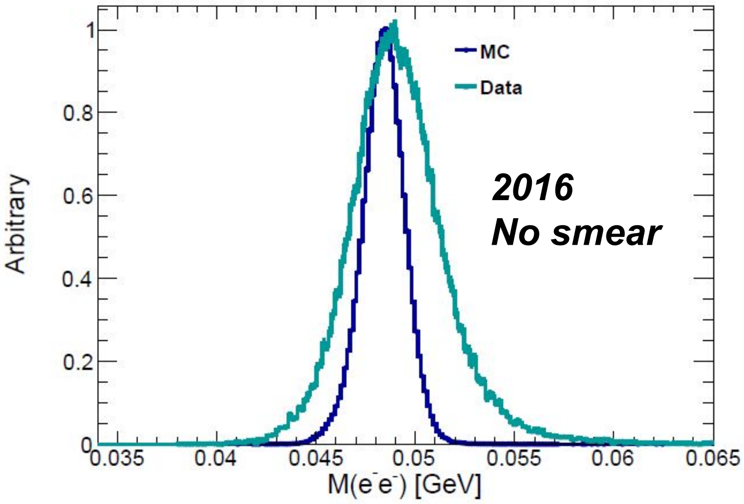
For 2016 our momentum resolution in data was worse than MC (and likely that will be the case for 2019/2021). We (Rafo) compared FEE momenta for MC/data separating top/bottom and 5-/6-hit tracks. We then smeared all MC tracks after reconstruction by the fractional differences in resolution.

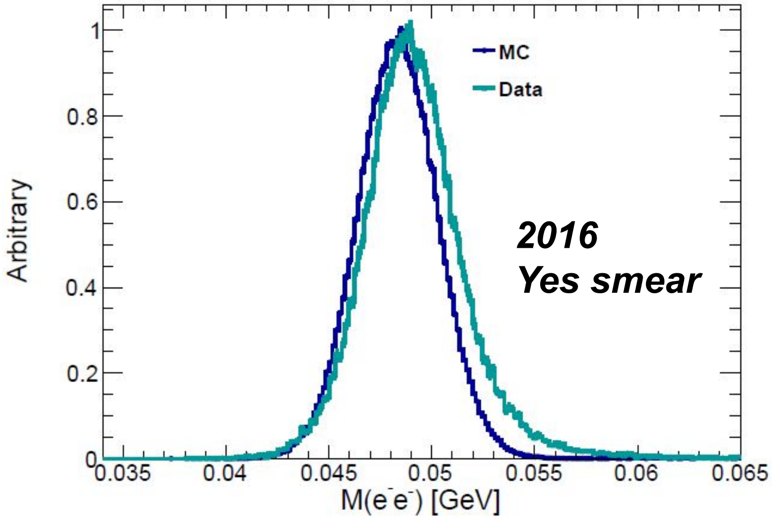
Variable \ Data Sample	Data	MC
μ_{Top}5hits [GeV]	2.262 ± 0.0022	2.246 ± 0.0022
μ_{Top}6hits [GeV]	2.285 ± 0.00071	2.255 ± 0.00081
μ_{Bot}5hits [GeV]	2.251 ± 0.0020	2.245 ± 0.0022
μ_{Bot}6hits [GeV]	2.254 ± 0.00072	2.260 ± 0.00069
σ_{Top}5hits [GeV]	0.182 ± 0.0033	0.099 ± 0.0016
σ_{Top}6hits [GeV]	0.130 ± 0.00089	0.083 ± 0.00065
σ_{Bot}5hits [GeV]	0.170 ± 0.0027	0.099 ± 0.0017
σ_{Bot}6hits [GeV]	0.131 ± 0.00079	0.082 ± 0.00057
$\Sigma^{\mathrm{smear}}_{\mathrm{Top}}$5hits [%]	N/A	6.733 ± 0.1632
$\Sigma^{\mathrm{smear}}_{\mathrm{Top}}$6hits [%]	N/A	4.358 ± 0.0485
$\Sigma^{\mathrm{smear}}_{\mathrm{Bot}}$5hits [%]	N/A	6.156 ± 0.1415
$\Sigma^{\mathrm{smear}}_{\mathrm{Bot}}$6hits [%]	N/A	4.556 ± 0.0431



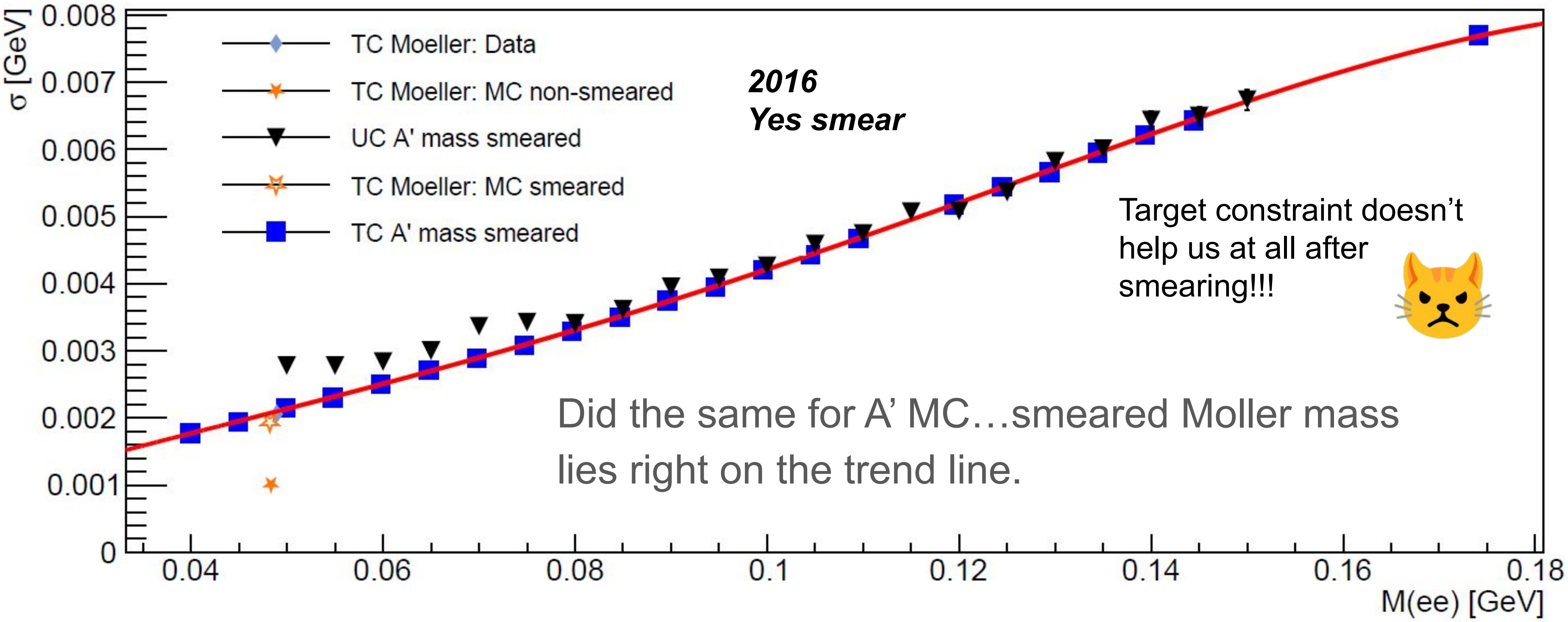
Moller Mass Resolution

…looking at the moller (e^-e^-) mass shows similar worse resolution in data compared to MC out of the box. Just by smearing the MC track momenta, we see much better agreement in the mass resolutions.





Mass resolution: smeared A’ and Mollers 2016



Target constraint doesn’t help us at all after smearing!!!

Did the same for A’ MC…smeared Moller mass lies right on the trend line.

Mass resolution 2019/2021

- Moller rate (in acceptance) decreases with increasing beam energy
  - Opening angle decreases due to increased boost
- For 2019 (4.5 GeV) we shouldn't have any to speak of and even for 2021 (3.7 GeV) there is only a sliver of phase space that we'd see.
- For 2021, we took a dedicated Moller run (2.2 GeV) so that will help a lot
- A possibility for momentum resolution: 3-prong tridents, tying the total momentum to pBeam
  - Fit pX,pY,pZ for each track
  - Constrain 3-track momentum sum to pBeam (vector)
  - Inputs are the track parameters + uncertainties
  - We haven't tried this at all…would give us e^- and e^+ resolutions vs momentum if it works
  - Could use Norman's 3-cluster exclusive sample, looks clean

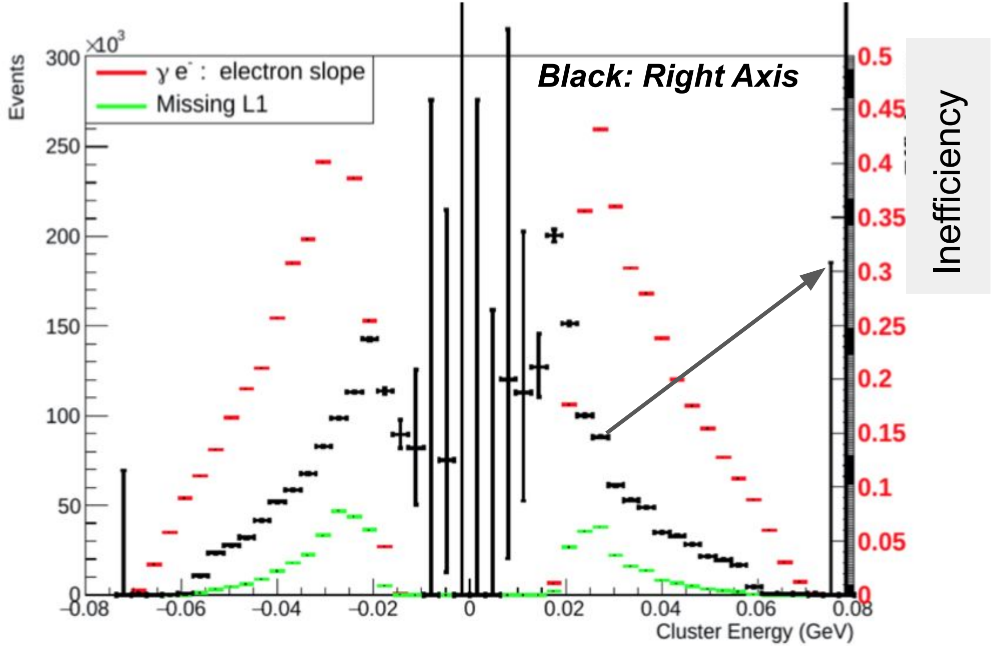
L1 efficiency and hit killing: 2016

For 2016, we were primarily concerned about MC not reproducting the hit efficiency in module (stereo/axial pair) 1. We took this into account by taking well reconstructed WABs (γe⁻) and taking the ratio of electrons without an L1 hit to the total reconned versus tanlamba.

We then used this inefficiency to kill L1 hits on tracks based on their tanlambda.

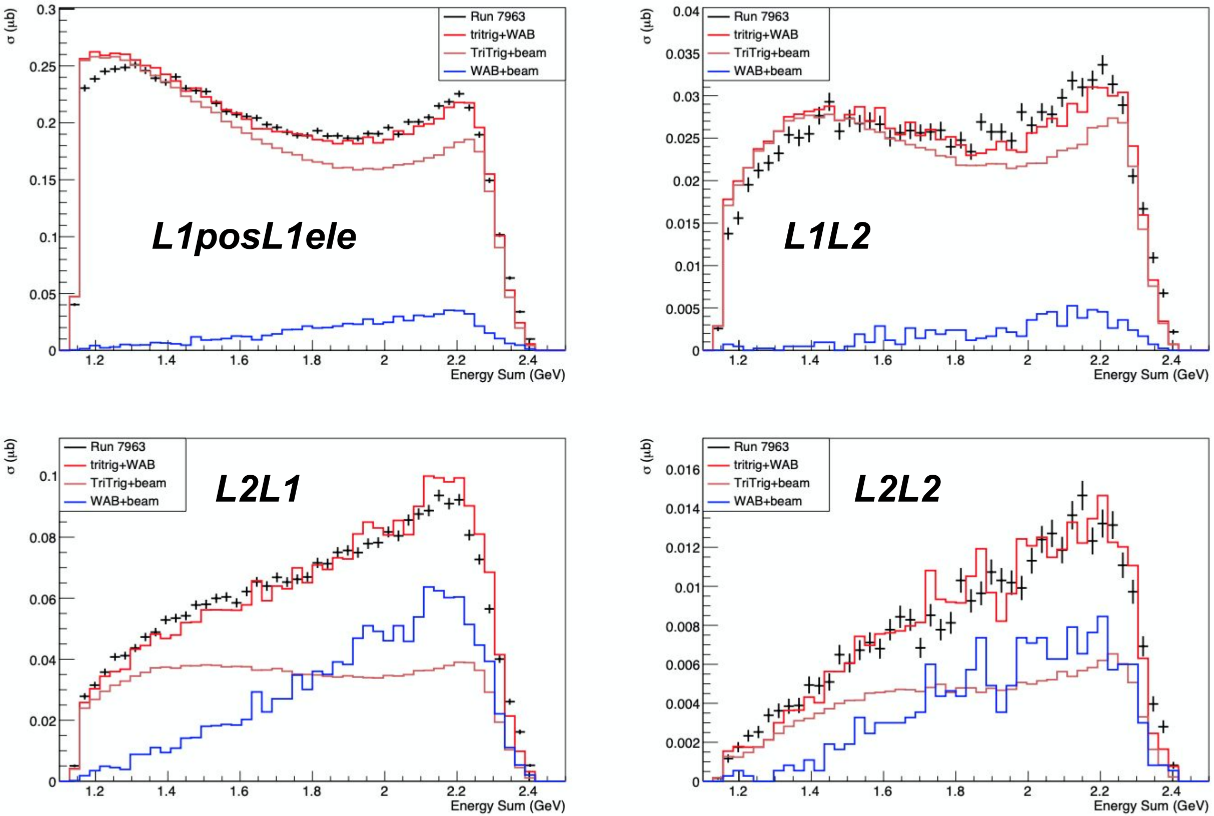
For vertexing, we actually killed hits and then re-reconstructed.

We should do something smarter moving forward!



Above is data…the MC has ~0 inefficiency

First hit layer combinations MC & Data: 2016

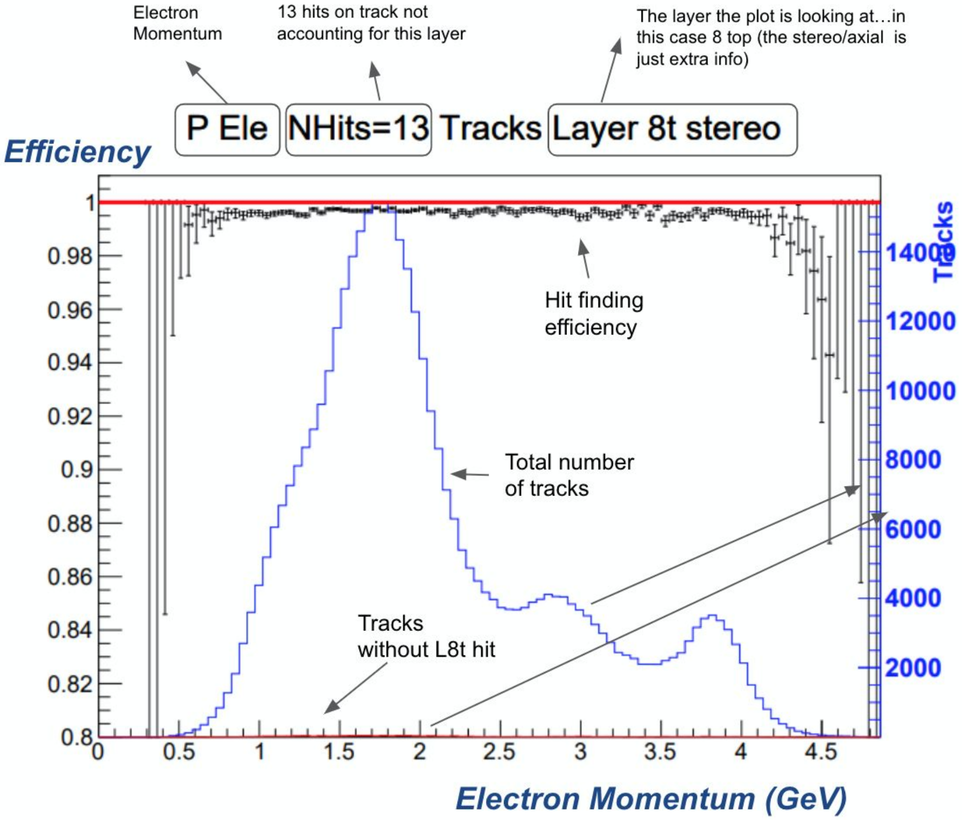


This is after slope-based L1 hit killing is applied and also with MC scaled by 0.88.

Note the different contributions of cWABs for the different categories.

Tridents/WABs Note

Hit-on-track efficiency: 2019/21



Code exists to compute the hit-on-track efficiency using the kalman tracking. What we should do is apply the MC/data inefficiency during MC readout.

…have to admit I tried this on 2016 data years ago and was not happy with results…we should try again.

[SVT Efficiency Talk at May 2022 Workshop]

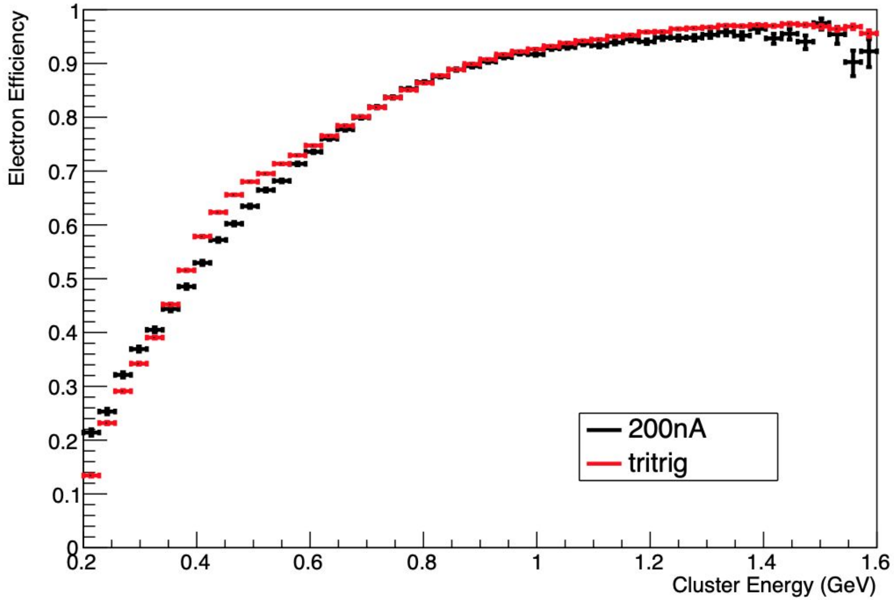
Track Finding Efficiency: 2016

We have never used the track finding efficiency for any corrections, but it is a very important cross-check.

Here I show electron efficiency vs cluster energy for events that have positron-side and electron-side clusters (top/bottom) and a positron track.

There are other ways to do this:

- WABs with 2 clusters adding up to beam energy
- 3-prong tridents with 3 clusters adding up to beam energy



[Tridents/WABs Note]

Improvements to analysis to consider

- Separate data set into hit content, particularly L0/L1/L2
  - Different mass and vertex resolutions, WAB contamination
  - For vertexing this is imperative…for bump hunt?
- Use ECal to improve energy (→ momentum → mass) resolution
  - This definitely can help for fiducial clusters…not sure about tracks near edge of ECal?
- ML fit for vertexing (1d in z-vertex or 2d including mass or ???)
- Use 3-prong tridents to their full extent?
  - Currently we don't even reconstruct these as a composite state
  - same-half/opposite-half electron events have different WAB contamination
    - Which could help us nail down the WAB rate…we've studied this before but didn't use it
- Some other clever ideas I'm sure other people have.