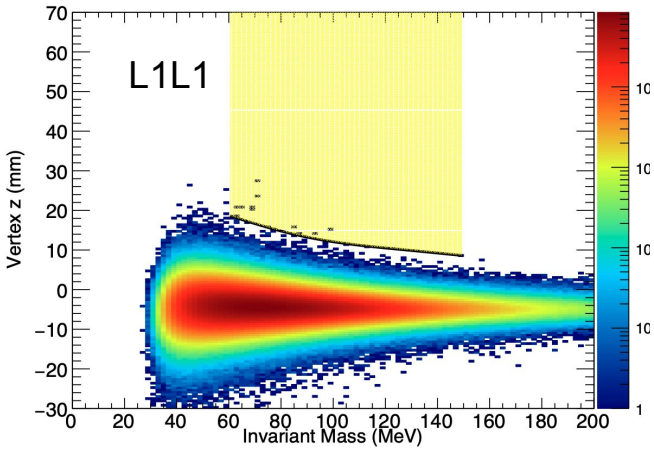
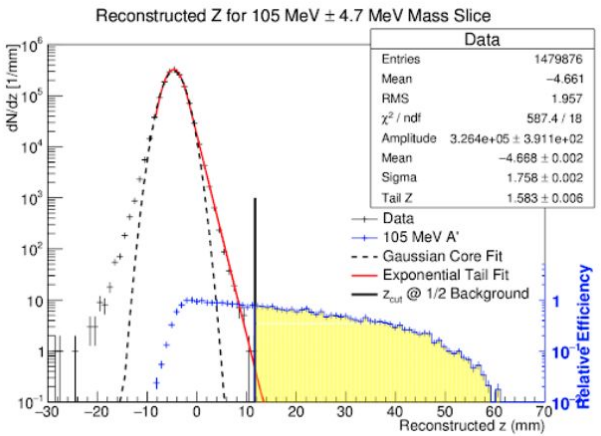
New(?) Ideas in Vertexing Analysis

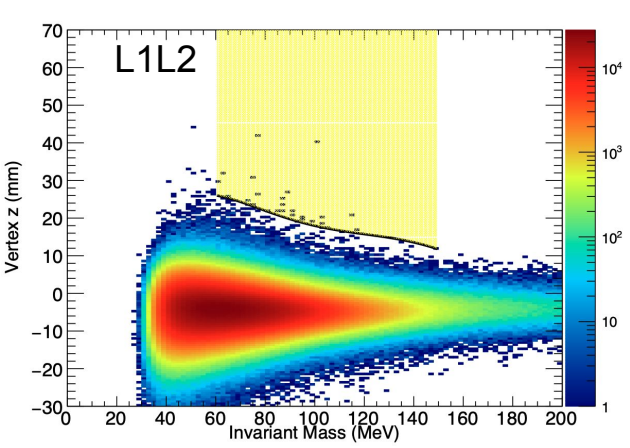
Matt Solt
May 3, 2022

Introduction

- “New(?) Ideas in Vertexing Analysis” -> “Old ideas that we were too busy to implement that we should probably be reminded of”
- Improvements on event selection, background reduction, systematics, limit setting, signal searches, etc.
- Some improvements are for 2019/2021, but some for SIMP searches in 2016







Selection

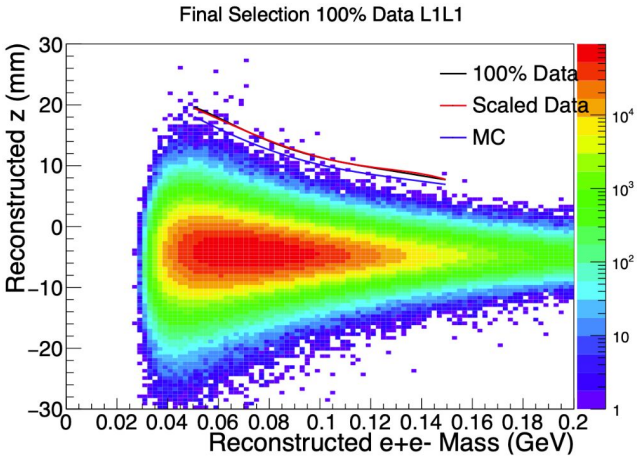
We have 2 main backgrounds:

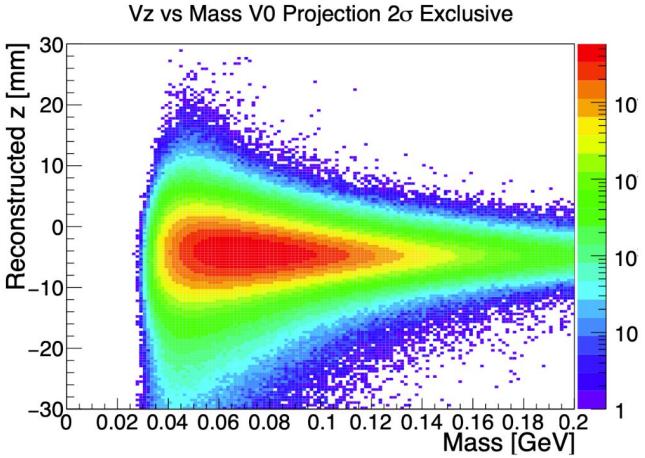
- Coulomb Scattering
- Mis-tracking

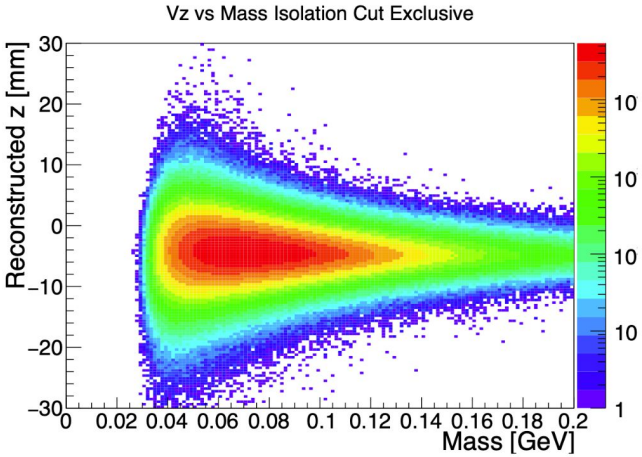
3 main handles for rejection:

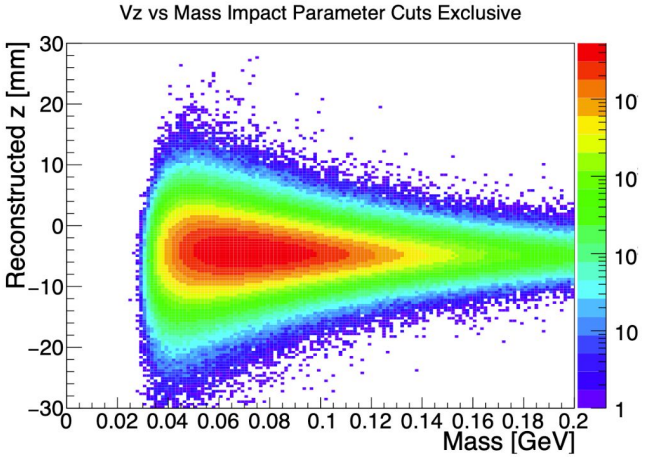
- V0 projection
- Impact parameters
- Isolation Cut

Cut Description	Requirement
Layer 1 Requirement	e^+ and e^- have L1 hit
Layer 2 Requirement	e^+ and e^- have L2 hit
V0 projection to target	Fitted 2σ cut
Isolation Cut	Eq. 20
Impact Parameters	Refer to text
Track Shared Hits	$= 0$
Vertex Multiplicity	$= 0$





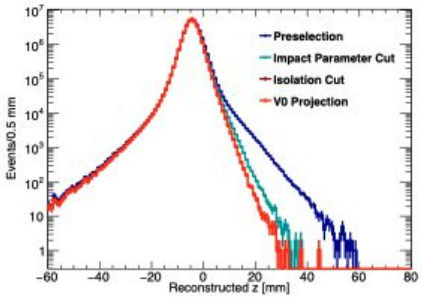




Improved Selection (I)

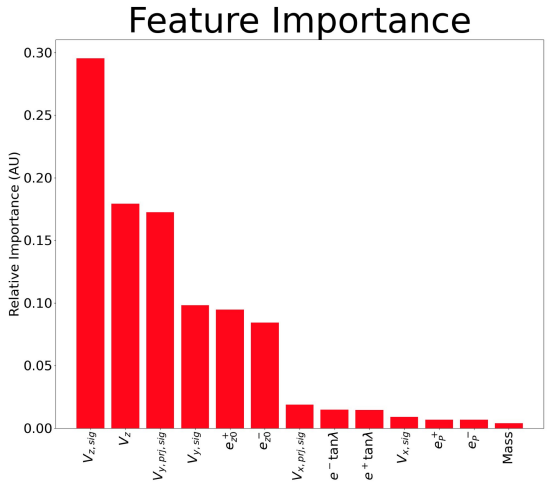
- Procedure of the tight selection needs improvement
  - **V0 projection** - run-dependence should be done at reco level
  - **Impact parameters** - parameterized as a function of mass and z (can get very messy)
  - **Isolation Cut** - effective but simplistic. Can we do this with improved tracking algorithms?
  - Track/vertex multiplicity needs a more careful look
  - Ideally, these selections would correctly incorporate event-by-event errors correctly
  - These selections are also not currently optimized
  - **Biasing** of scattering (and WABs) may be necessary to avoid large scale MC generation
  - Also, better sharing of cuts with resonance search for simplicity (e.g. timing cuts, cluster/track match, etc.)

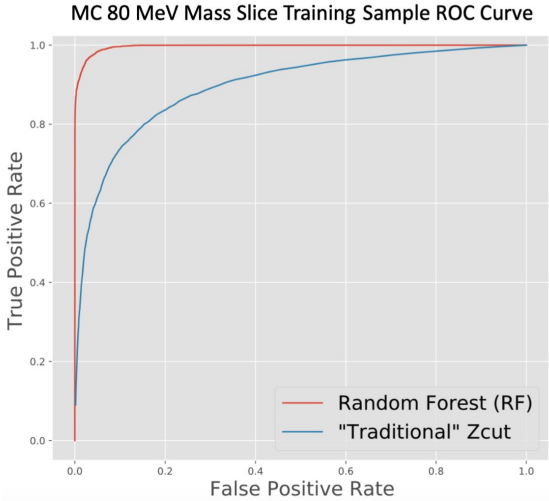
Cut Description	Requirement
Layer 1 Requirement	e^+ and e^- have L1 hit
Layer 2 Requirement	e^+ and e^- have L2 hit
V0 projection to target	Fitted 2σ cut
Isolation Cut	Eq. 20
Impact Parameters	Refer to text
Track Shared Hits	$= 0$
Vertex Multiplicity	$= 0$



Improved Selection (II)

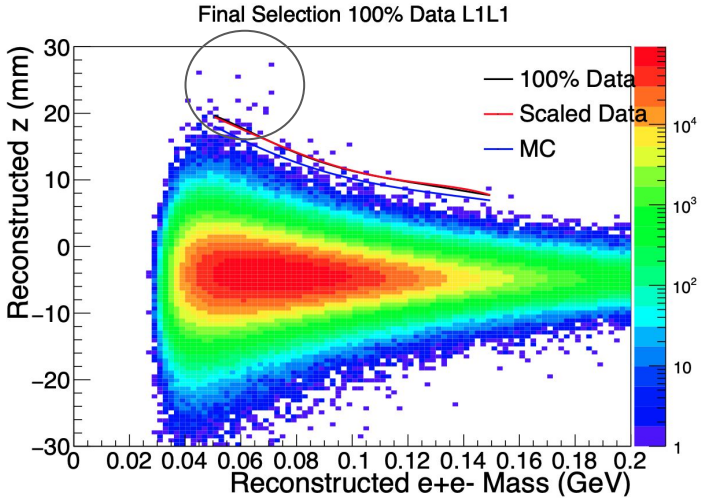
- A simple ML algorithm can potentially make the selection easier and more efficient
  - Some success with a random forest, a BDT may also work. Deep NN is unnecessary.
- Pros:
  - More efficient background rejection
  - Easier optimization of selection
  - (Probably) better result if done correctly
- Cons:
  - Train on simulation, so it has to match the data
  - Analysis procedure downstream must be changed
  - Systematic uncertainty estimates are more difficult

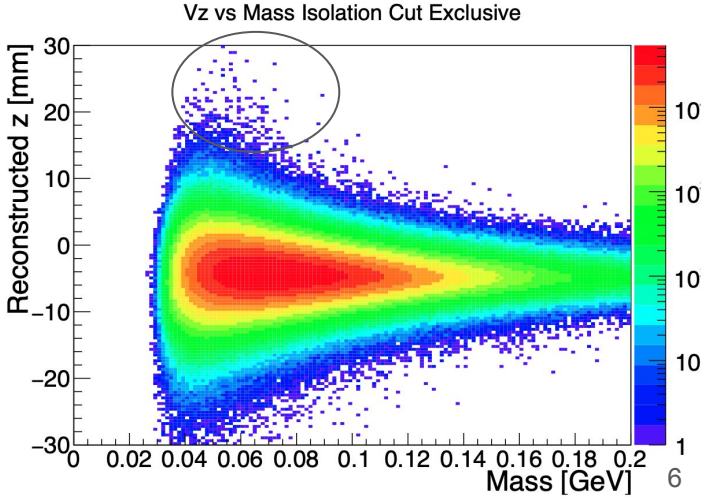




Improving Mis-tracking

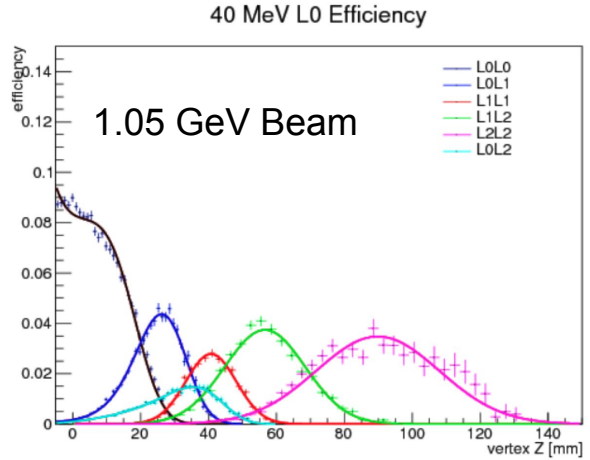
- The isolation cut is effective, but suboptimal
- I have shown some evidence that low mass background could be mis-tracking (with post-reco hit inefficiencies)
- Reducing mis-tracking background
  - Hit efficiency effects? Can be mitigated by improved pulse fitting (understanding where fits fail)
  - Tracking algorithm issues? Kalman filter approach

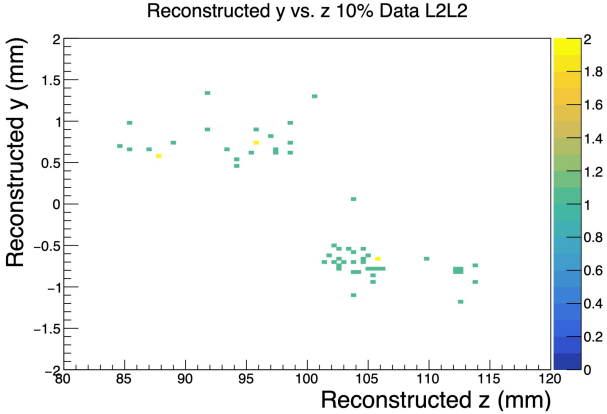




More “Categories” with Upgraded Detector

- With an upgraded detector: L1L1 + L1L2 -> LiLj i,j ∈ [0,2]
  - The Kalman filter no longer requires 3D hits. More categories?
  - Tune cuts in each individual category? Combine in the last step? Etc?
- We don’t understand our WAB rate in the inactive Si
- We now also have to worry about trident production in the Si
  - This has been shown in both L1L2 and L2L2, but not in our signal region yet





Improving Systematics

- e+e- Composition systematic is dominated by WAB uncertainties
- Analysis cuts systematics are very conservative estimates and need to be improved
- Dominated by target position uncertainty (+/- 0.5 mm)

Systematic Description	L1L1 Value	L1L2 Value
e^+e^- Composition	~7%	
Mass Resolution	~3%	
Analysis Cuts	~8%	~13%
A' Efficiency	~5%	
Total in Quad	12%	16%
Target position	~5-10% (m/ϵ dep)	

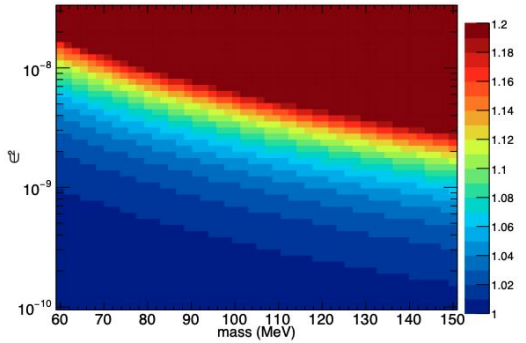
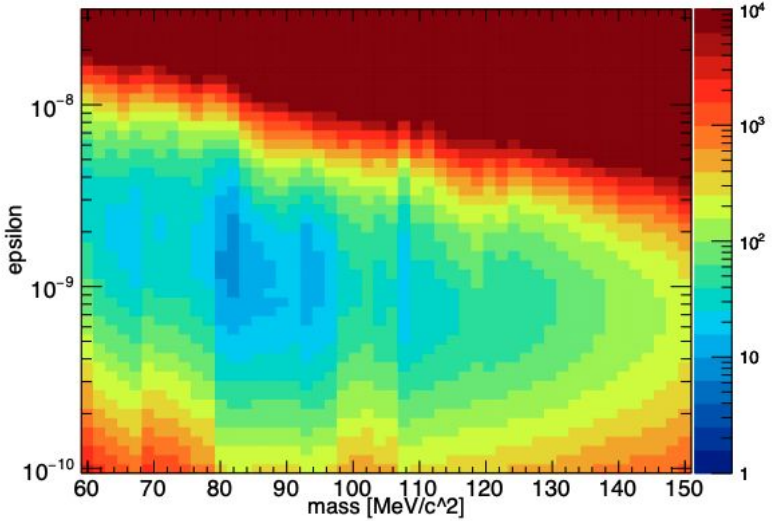


FIG. 32: The ratio of the limit for the L1L1 category from the target 0.5 mm upstream of the nominal position to the target at the nominal position using the Optimum Interval Method.

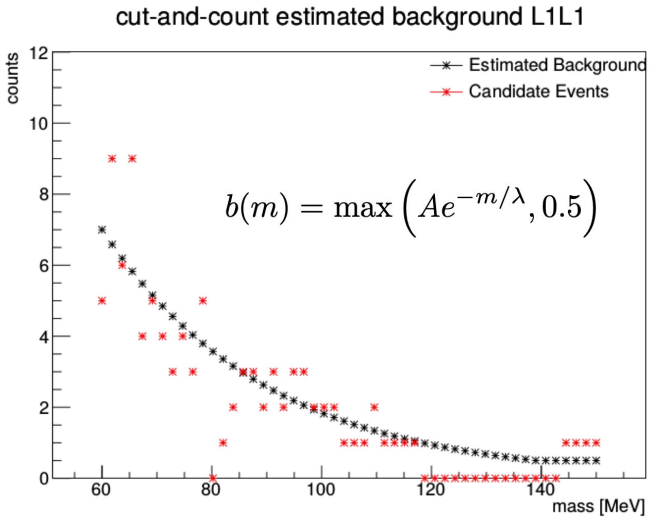
Limit Setting Procedure

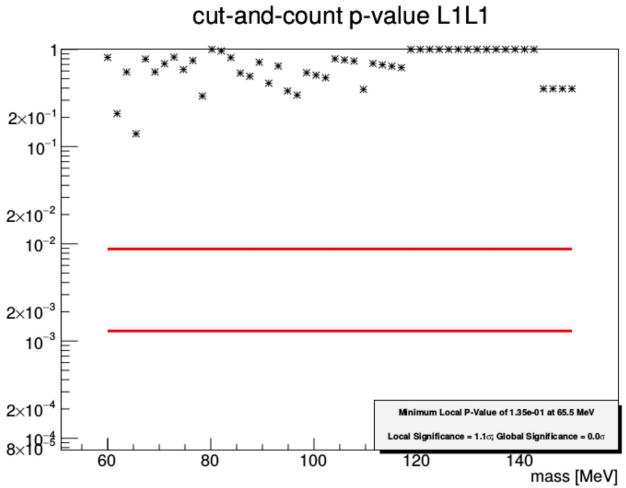
- We use the optimum interval method (OIM), but it may be suboptimal
  - Generally a conservative limit setting method
  - Penalizes for the number of intervals it searches, thus not good for large background mass slices
- The way we combine L1L1 + L1L2 may also be suboptimal



Signal Searches

- Eventually, we need a procedure for actually searching for a signal (not just exclusion)
- A basic cut-and-count analysis in 2016 L1L1
- A simple 0.5 background estimate gives a ~4σ global excess…
- A more reasonable unbiased background fit gives a 0.0σ global excess
- Plenty of room for improvement





Summary/Conclusion

- Improved event selection
  - There are simpler ways to incorporate these selections, particularly the “tight” selections
  - ML approach?
- Reduce mis-tracking
  - Reduced hit inefficiency, Kalman filter, improved hit fitting, etc.
- More complicated A’ tracking “categories”
- Improving systematics
  - Target positions, analysis cuts, radiative fraction, etc.
- Limit setting procedure and signal searches
- Opportunities in both SIMP searches (2016) and data with the upgraded detector to incorporate these improvements