Extending SIMP Search to L1L2

Tom Eichlersmith

he/him/his

University of Minnesota

eichl008@umn.edu

July 16, 2024

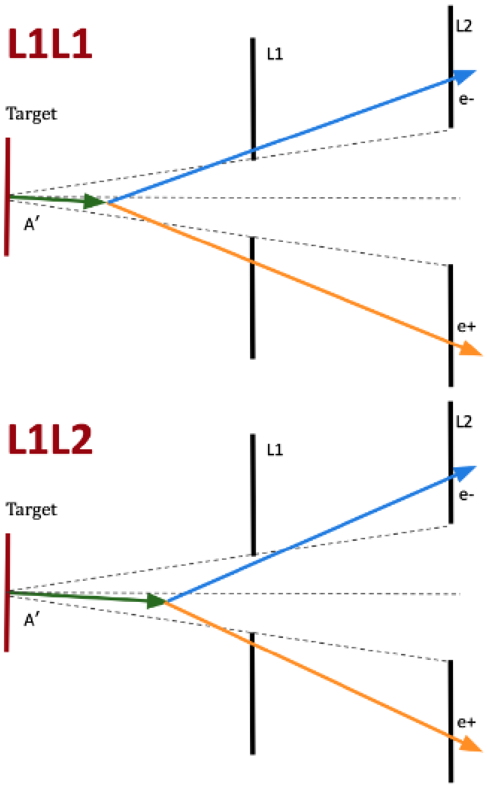


Fig 30 from PRD

Extend

- Investigate L1L2 “reconstruction category”
- Combine reach from both reconstruction categories
- **Potential** KF Tracking opens door to re-defining reconstruction categories.
  - e.g. if one track has 1/2 sensors in L1 and 2/2 sensors in L2, is the vertex “closer” to L1L2 or L1L1?

Expectations

- Reach lower ϵ^2 via larger displacements
- Cut tightening due to loss in reconstruction precision

Strategy

Samples

Same data and simulation samples as Alic's SIMP (L1L1) search.
(although partitioning of these samples is different)

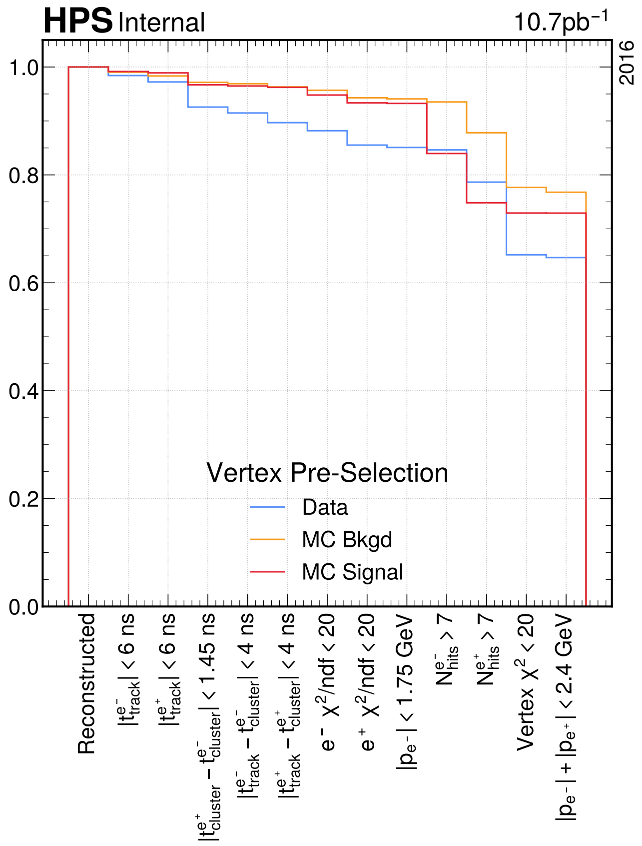
Selections

Rely on Alic's thorough study and validation, copy pre-selection and start with same final selection variables.

Search and Exclusion

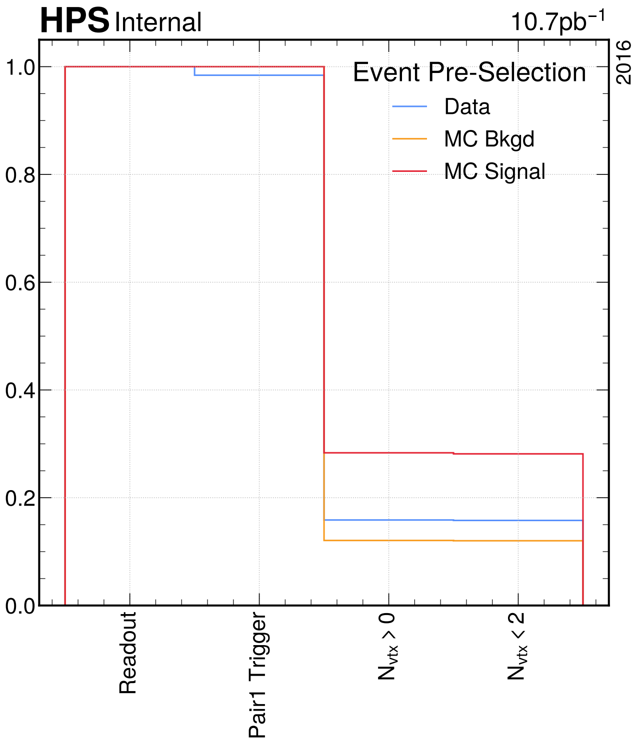
- Search for excess in $|m_{\text{reco}} - m_{\text{true}}|/\sigma_m$ vs $\min(|y_{0,e^-}|, |y_{0,e^+}|)$ space
- Exclude by using OIM on the z distribution after final selections

Pre-Selection on Vertices



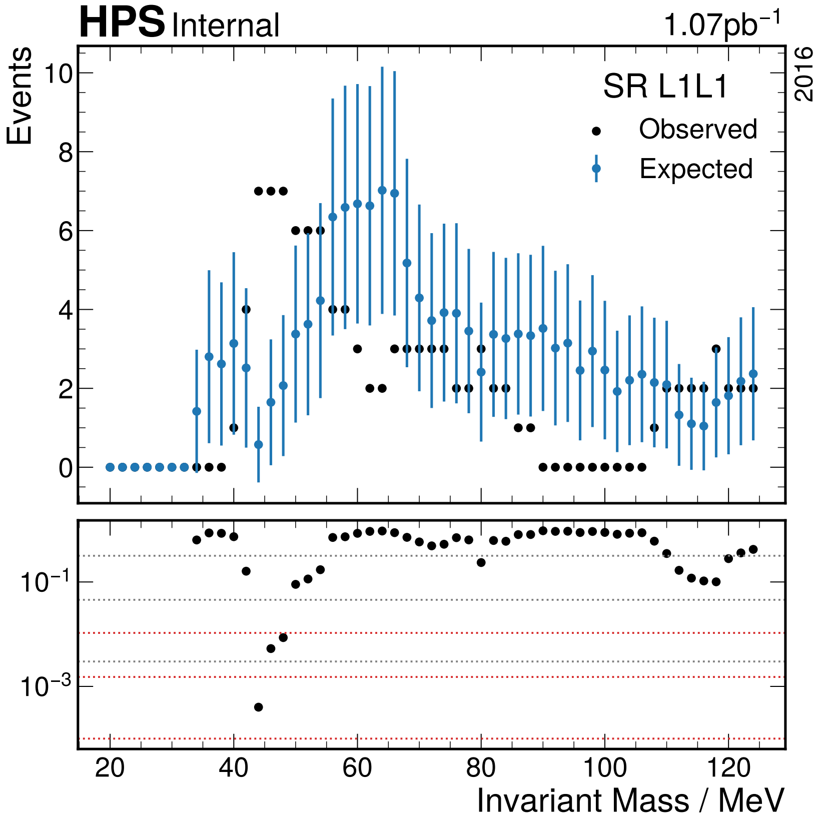
- Same Pre-Selection on vertices as developed and validated by Alic
- Seeing same efficiencies as documented within Alic's SIMP (L1L1) note

Pre-Selection on Events



- Similar to first stage of Alic's event selection, although dropping reconstruction category requirement
- Largest effect is requiring at least one pre-selected vertex

L1L1 Search on 10% Data

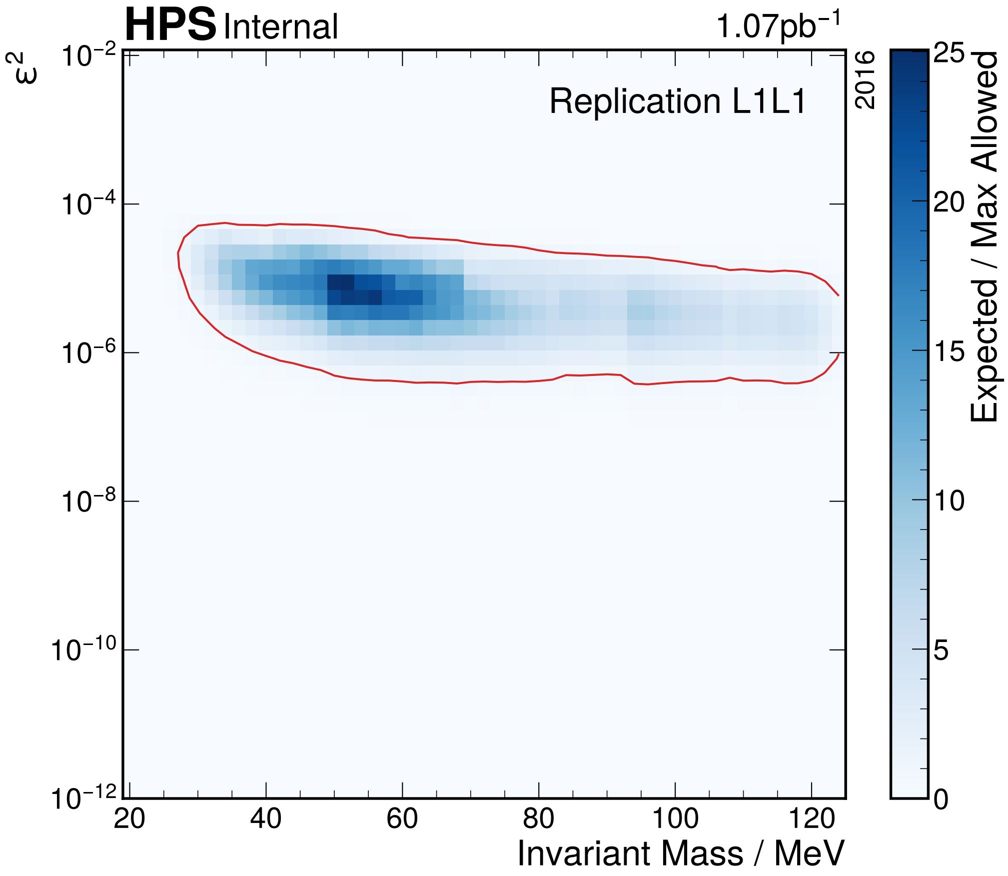


- Copying Alic's cuts and search strategy
- Seeing some funky-ness $\sim$ 44 MeV but only barely passed 2σ globally in p-value (bottom panel, local σ in gray, global in red)
- Choose to move towards exclusion instead

Difference with Alic

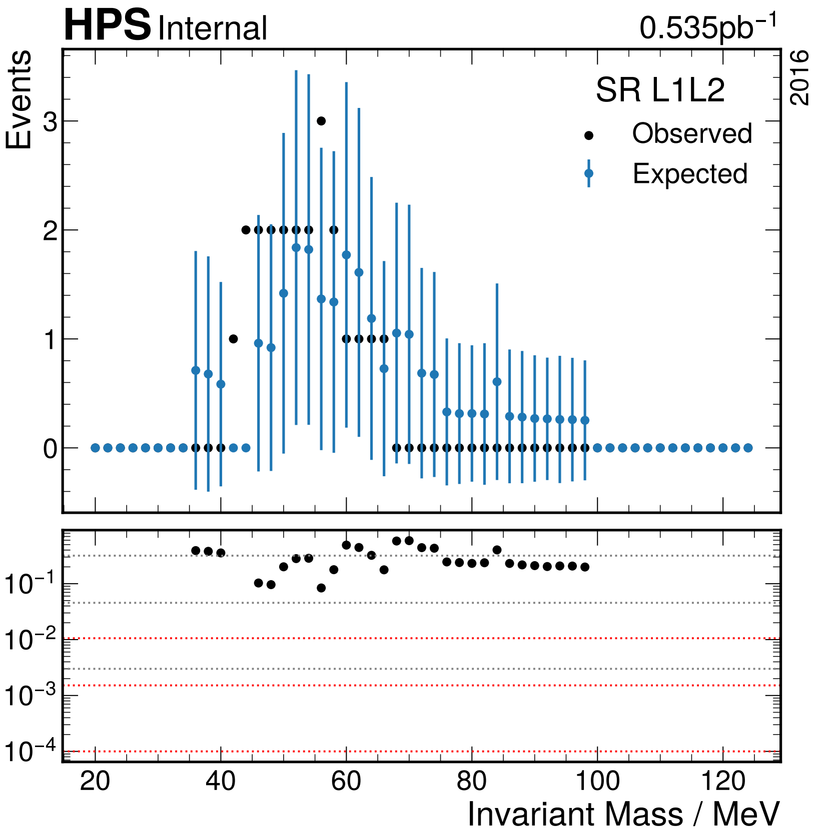
Learned this morning that the 10% data sample I'm using is slightly larger than the "10%" data sample Alic has been using.
Additionally, this plot is *not* using the same edges for the ABCD method as Alic so it is not an identical replication.

L1L1 Exclusion on 10% Data



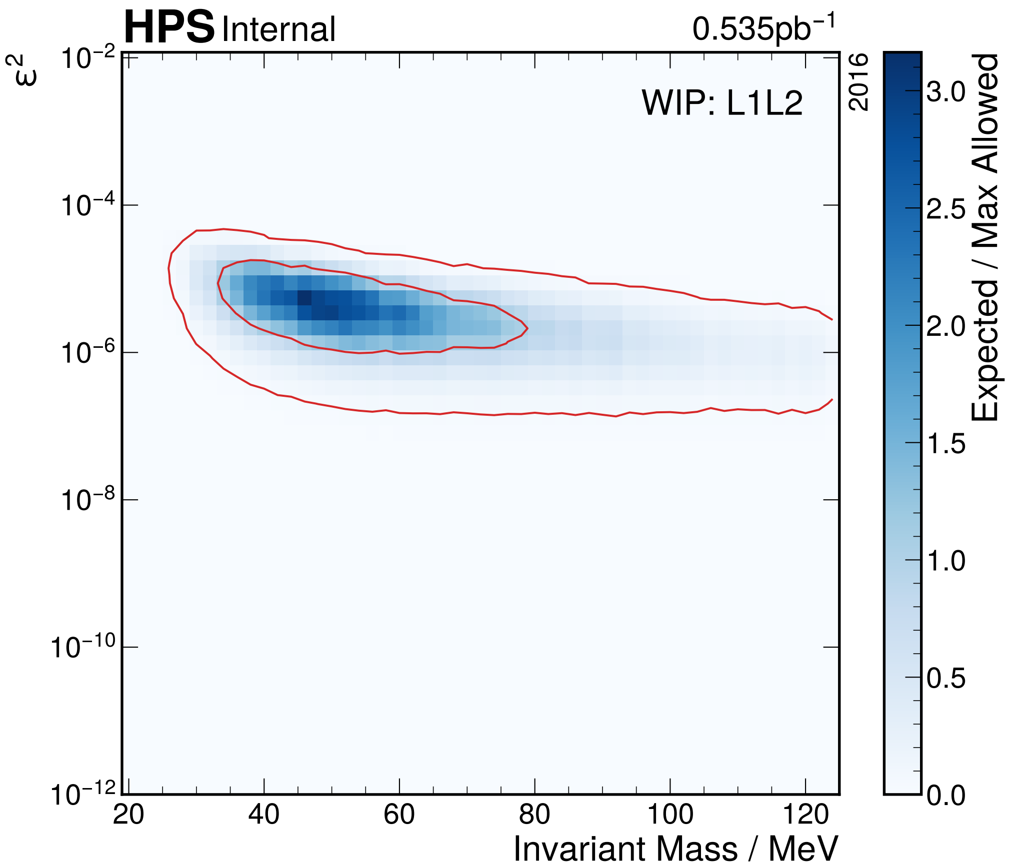
- Copying Alic's cuts and search strategy
- Observing qualitatively similar shape as Alic's code
- Calling this good

L1L2 Search on Run 7800



- Tightened min-y_0 cut by 0.5 mm
  - Observing that this is almost certainly too tight
  - Still helpful for a first pass and testing of code
- Not seeing an excess in Run 7800 $\rightarrow$ estimate exclusion due to Run 7800

L1L2 Exclusion on Run 7800



- Can see *slight* lowering (about one bin in ϵ^2)
- Both contour at 1 (excluded by Run 7800 alone) and 0.05 (approximate exclusion of 100% data set assuming the maximum allowed does not change)

Future Work

Test

- Compare my search/exclusion results to Alic's in L1L1 to validate code

Refine

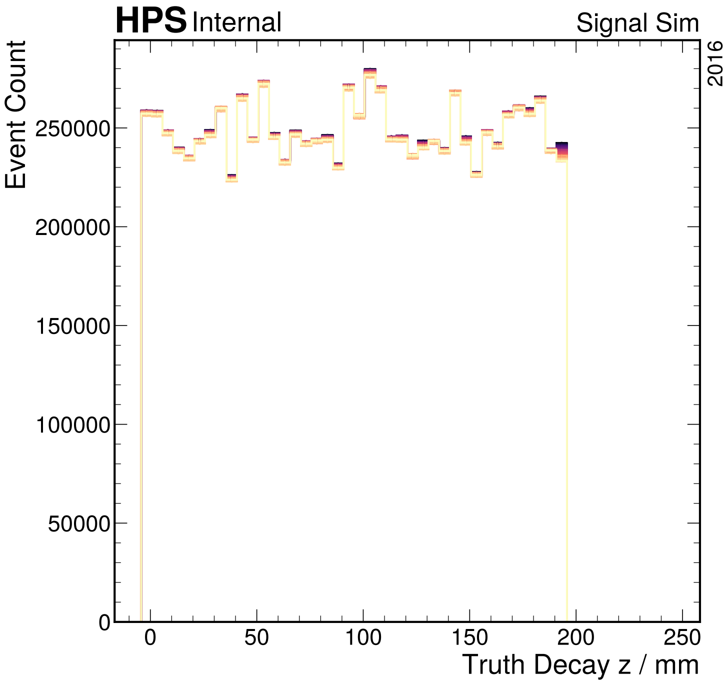
- Increase number of ϵ^2 test points in exclusion

Optimize

- Utilize Punzi Figure of Merit within each z bin to optimize cuts for each test mass
- Optimization done with simulation samples (background and signal)

Questions

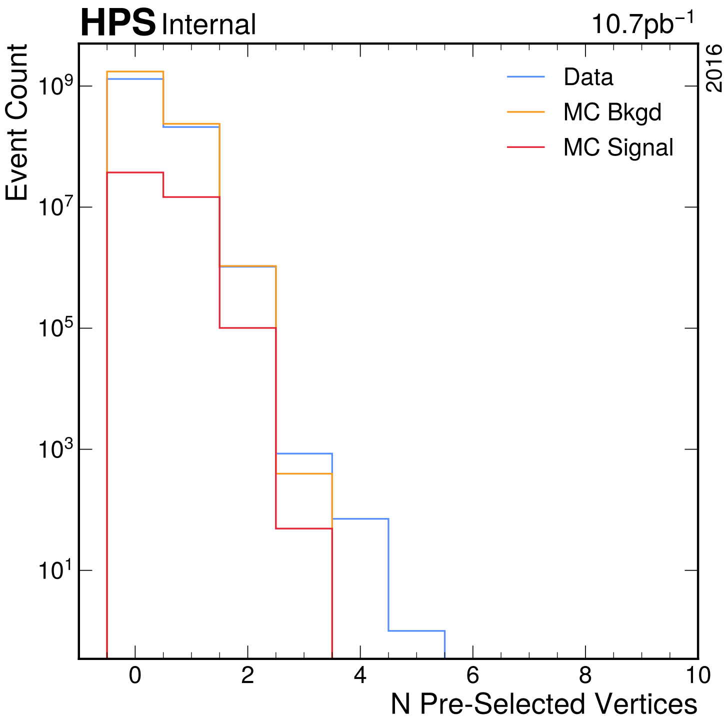
Displacement of SIMP Signal Events



One of the first things I did to make sure samples didn't need to be created.

- Different colors correspond to different mass points
- Truth-level decay vertices sampled out until ~ 200 mm
- Close to the same z position as L1, so I think these samples can be faithfully used to study the L1L2 selection
- Idential distribution across colors makes me think that the random seed determining the decay length was not changed, but I think that is okay since we normalize by this distribution during exclusion estimates anyways

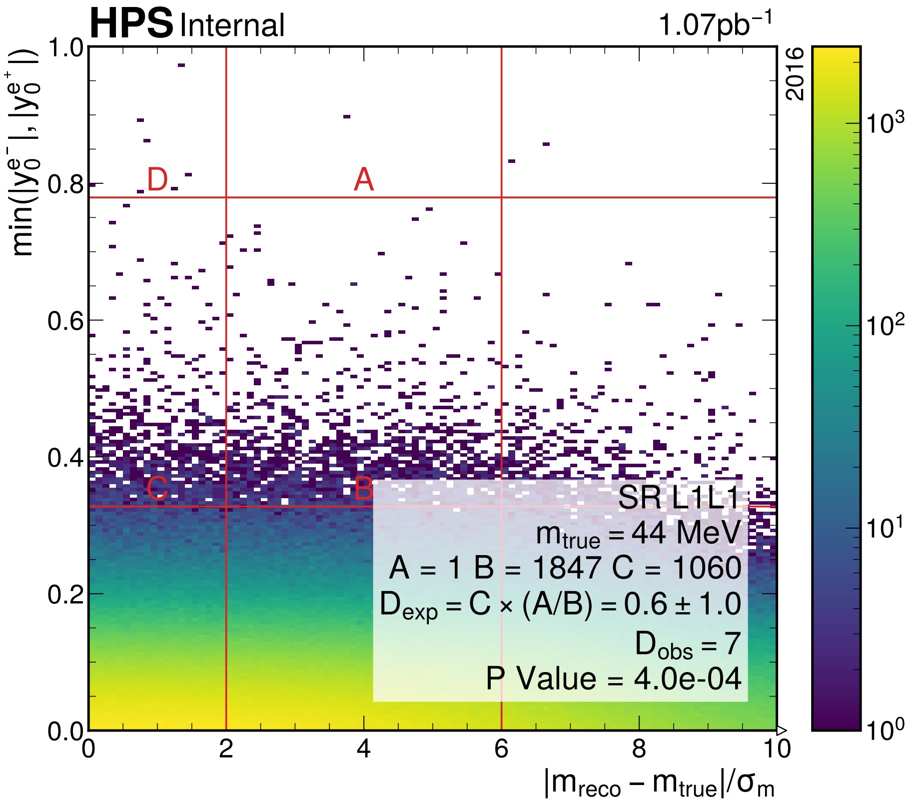
Number of Pre-Selected Vertices per Event



Event Pre-Selection basically amounts to choosing the events falling into the $N = 1$ bin.
Data is the only sample which has the additional requirement of the Pair1Trigger which has a small effect.

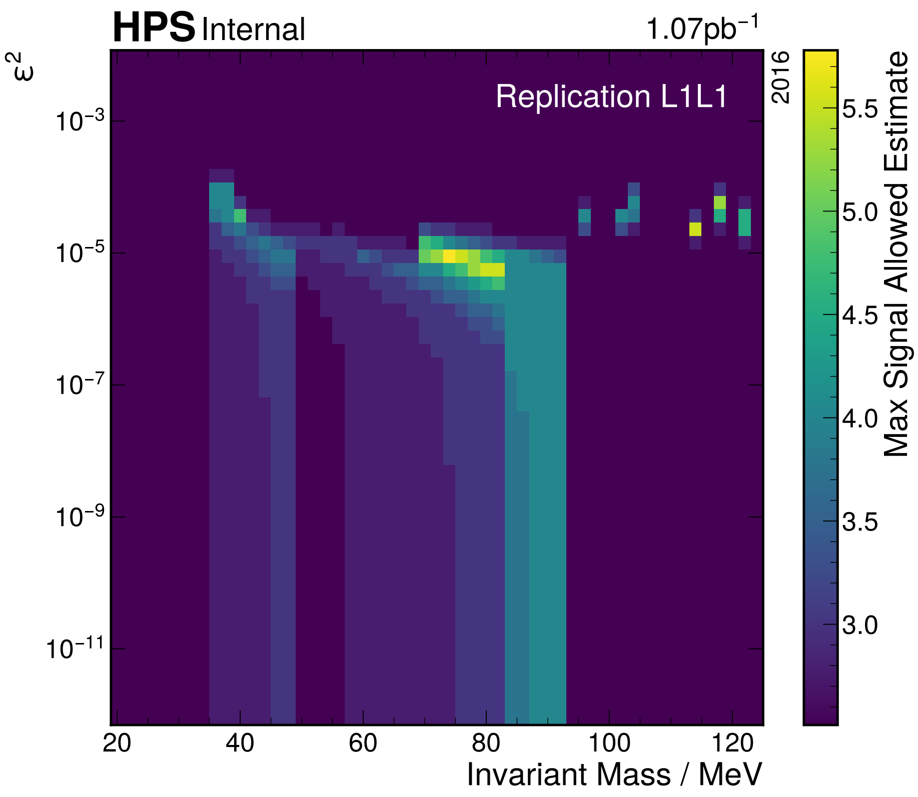
Search Procedure

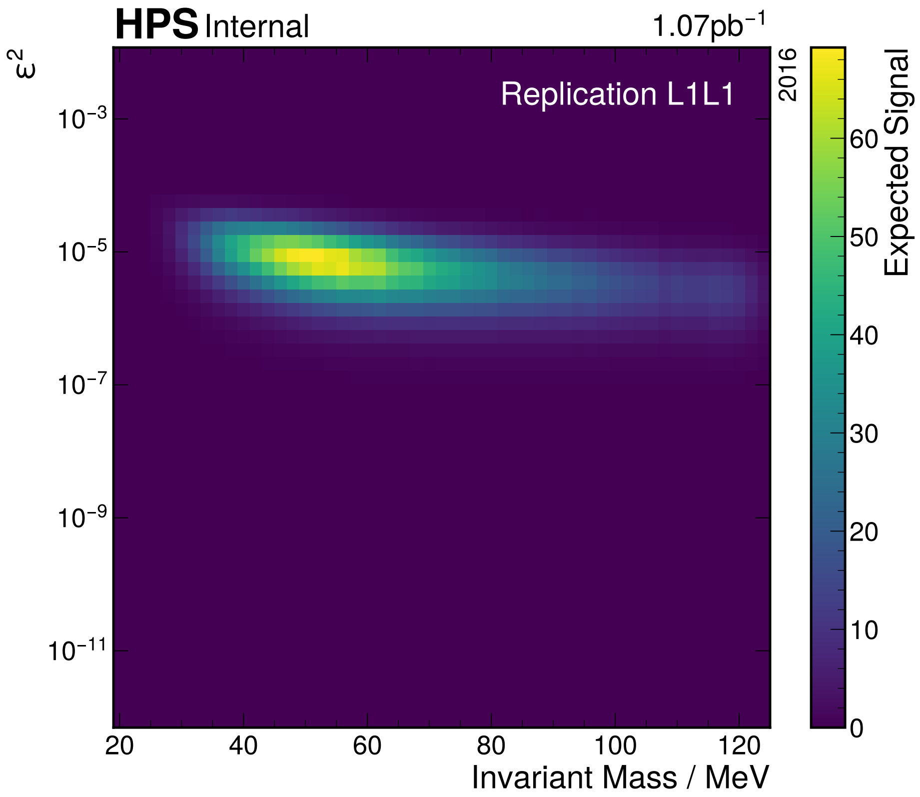
and 44MeV L1L1 Search on 10% Data



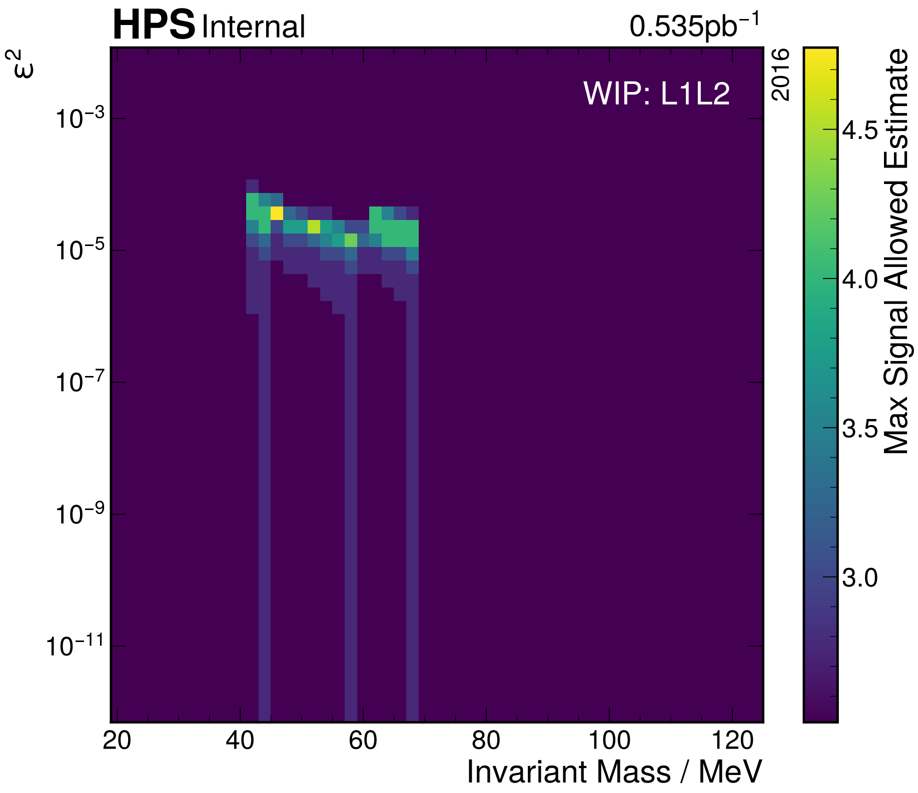
1. Fill histogram with data
2. Set mass pull edges at 2 and 6
3. Set upper min-y_0 edge at optimized cut value
4. Lower other min-y_0 edge from the cut value until there are at least 1k events in region C
5. Calculate expected number of events in D and compare to observed number of events
6. Estimate p-value by throwing toy experiments in A (Poission), B and C (Normal) and re-calculating D from these toys

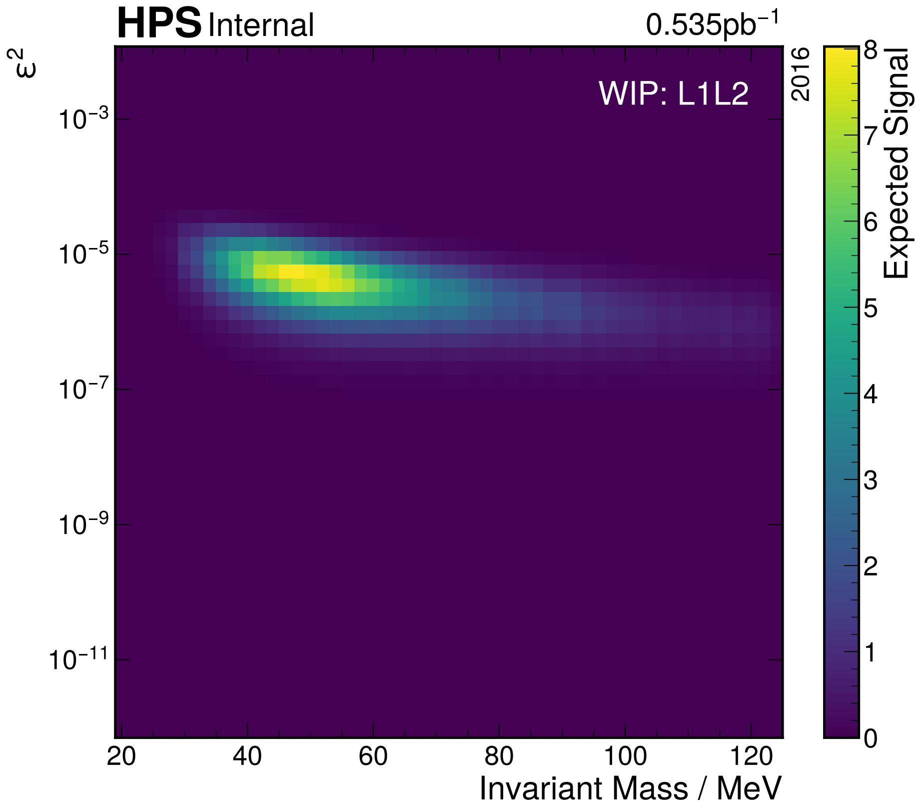
Max Allowed and Expected Signal for L1L1 10%





Max Allowed and Expected Signal for L1L2 Run 7800





Punzi Figure of Merit

Basis

Proposed in ▸ PHYSTAT2003 by Giovanni Punzi where a FoM is designed to be maximized while improving *both* search and exclusion potential.

$$f_{\text{punzi}} = \frac{E}{\frac{a}{2} + \sqrt{B}}$$

where E is the signal efficiency, B is the background yield, and a is the desired confidence level of search or exclusion (in number of σ, currently using 3).

Two main benefits (from my perspective)

- Does not diverge as $B \to 0$
- Does not require knowledge of absolute rate of signal