Low-Pt Tracking for ATLAS Analyses

Patrick McCormack

~CTD draft slides (meant to be 15 min talk)~

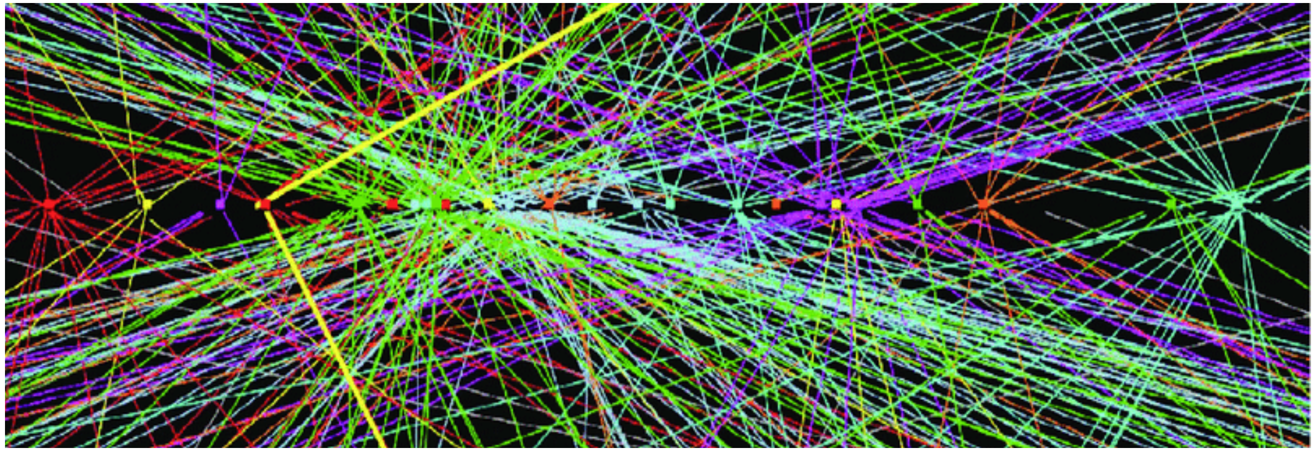
March 27, 2019

A quick definition

- The current ATLAS tracking framework attempts to reconstruct particles with Pt down to 500 MeV, so for the rest of this talk, “low-Pt tracks” will refer to those with Pt < 500 MeV

Why not low-Pt tracking?

- Physically, the particles from pileup interactions and from the underlying event (UE) of high-q^2 pp interactions at the LHC are in fact MOST LIKELY to have Pt < 500 MeV
  - For most analyses, these particles are not interesting- we typically care about objects with Pt >~ O(10 GeV)
  - Reconstructing low-Pt tracks would make the below event display even busier
- There is also a steep reconstruction time and storage cost when attempting to go to low pt



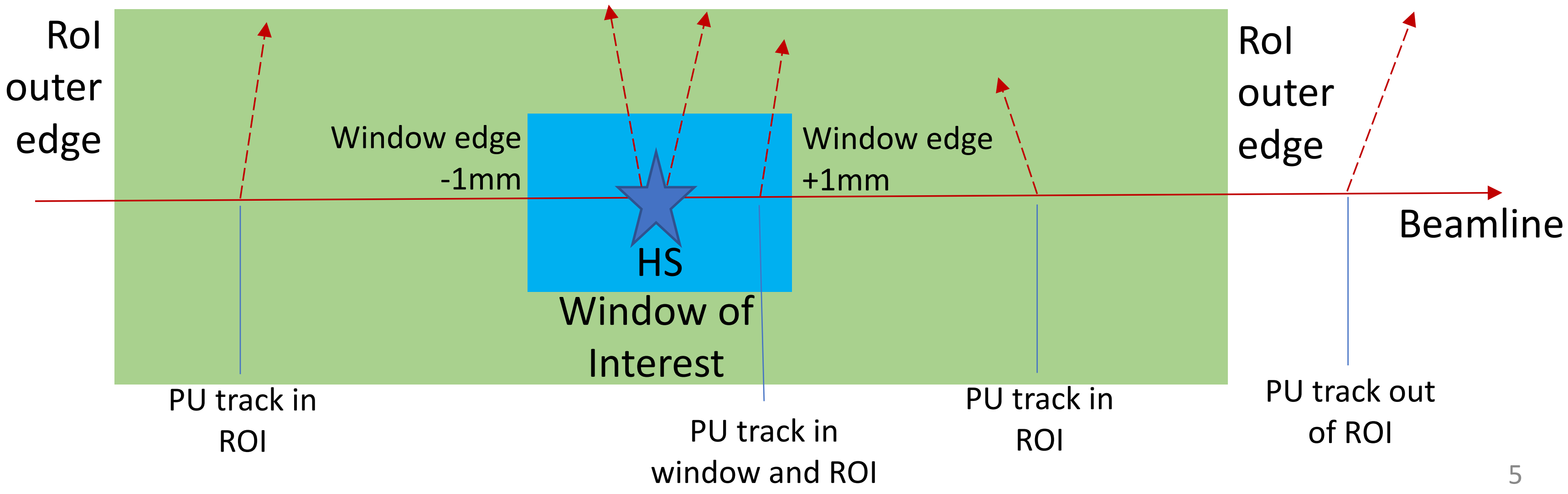
So why low-Pt tracking?

- That being said, there ARE ATLAS analyses where low-Pt track information can be useful!
  - Searches for photon-induced physics can use better reconstruction of an UE to help distinguish between photon- and parton-induced interactions
  - Charm tagging can be improved, as D meson decays often result in low-Pt tracks
  - Some SUSY models predict low-Pt tracks (e.g. small chargino-neutralino LSP mass splitting)

Implementing low-Pt tracking

- Current setup uses two passes
  1. Standard tracking, currently down to 500 MeV
  2. Using **leftover hits**, reconstruct tracks between 100 and 500 MeV
- To reduce reconstruction time and combinatorics, in the second pass, only consider track seeds pointing into a region of interest (RoI) near the hard scatter (HS). RoI position can be chosen from reconstructed objects (data and MC) or from a truth-level position in MC

Tracks are reconstructed for seeds that point into the RoI (green region). However, the tracks that are most relevant for analysis are those near the hard scatter (HS), in what I will call the “window of interest” (blue region). Most of the following plots are only of tracks in the blue region.



Performance metrics- 3 key features

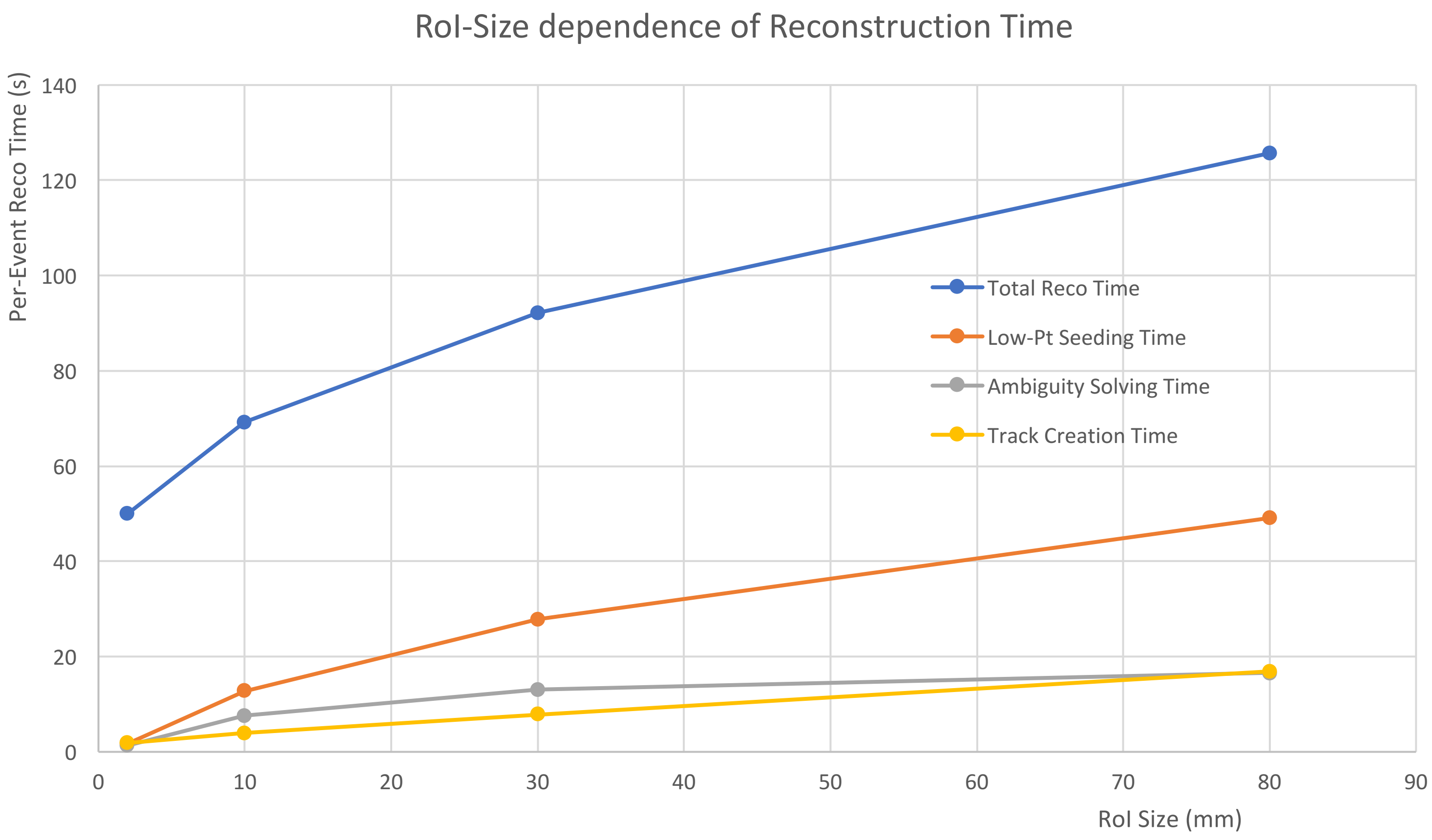
1. Reconstruction time. How much longer does the reconstruction take when low-Pt tracking is included?
2. Efficiency of reconstructing charged particles (truth-level objects)

$$eff = \frac{\#charged\ particles\ with\ at\ least\ one\ reconstructed\ track}{\#charged\ particles}$$

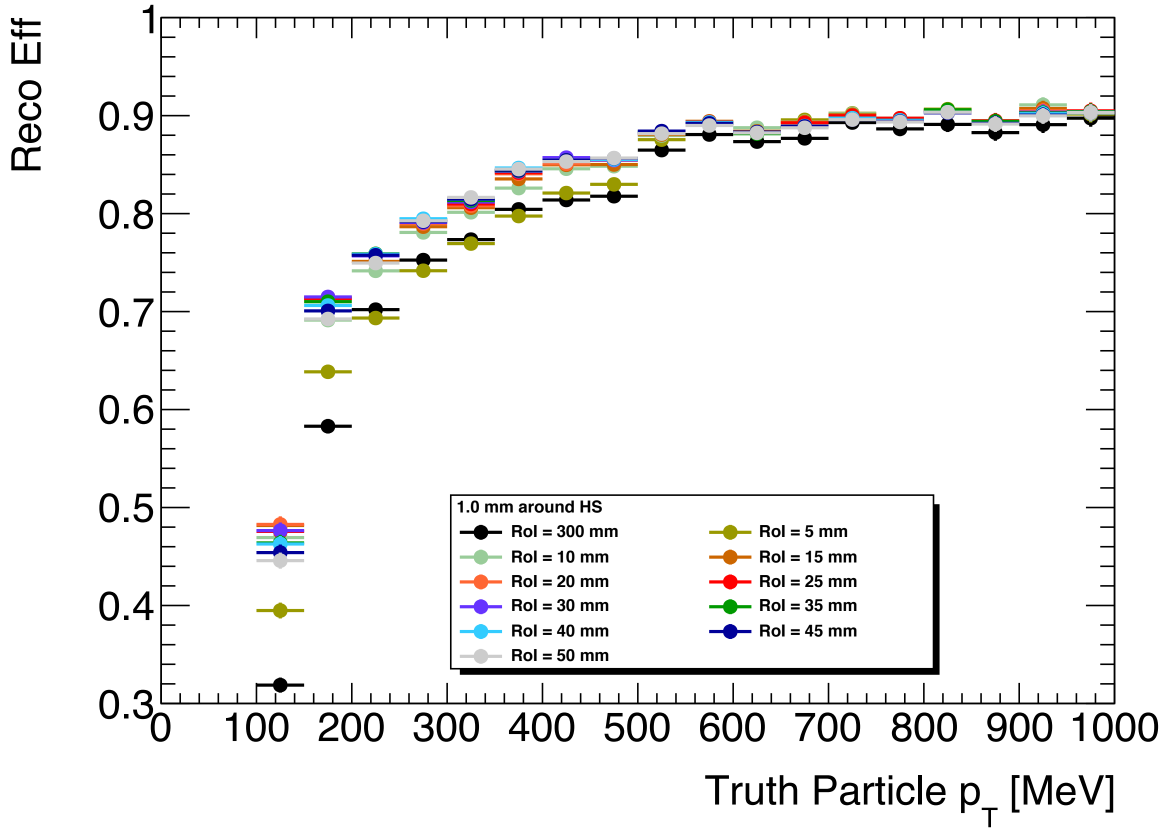
3. The fake rate, or alternatively the raw number of fake tracks in the window of interest. A fake track is a track constructed from detector hits created by multiple different charged particles- i.e. it does not correspond to any truth level object’s trajectory

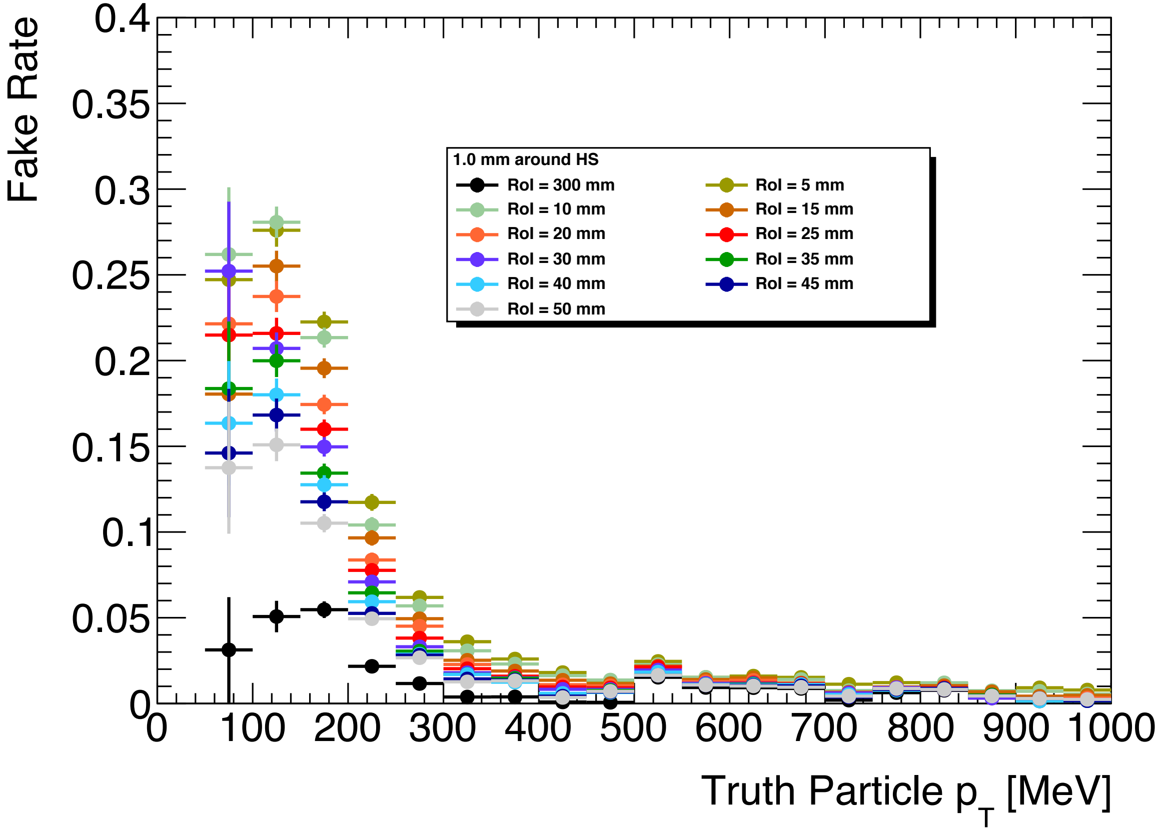
Reconstruction time

- Right: reconstruction time for various components of tracking as a function of RoI size.
  - The yellow, orange, and grey correspond to the **low-Pt tracking pass times only**
  - For reference, the generic tracking times average
    - 14s for seeding
    - 2.8s for ambi solving
    - 6.0s for track creation
- A single pass setup that reconstructs tracks down to 100 MeV with no RoI:
  - Tentatively 5x slower than 30mm RoI time
- The times are highly pileup-dependent. What is plotted on the right is averaged over LHC Run 2 pileup conditions



Efficiency and Fake Rate





Plots made using an ATLAS WW sample with full PU truth info

RoI for track seeds given by legend, but only tracks/truth particles within 1.0mm of di-leptons plotted

Considering Performance

- A small RoI (< 10 mm) **is fast**, but has lower efficiency and **high fakes**
- A large RoI (>100 mm) has **low fakes**, but has generally lower efficiency and is **time intensive**
- A medium RoI size (~30 mm) has **close to maximal efficiency**, at the cost of medium fakes and a moderate reconstruction time impact

- We adopt 30mm as a tentative default, though this is an adjustable parameter when the algorithm is called

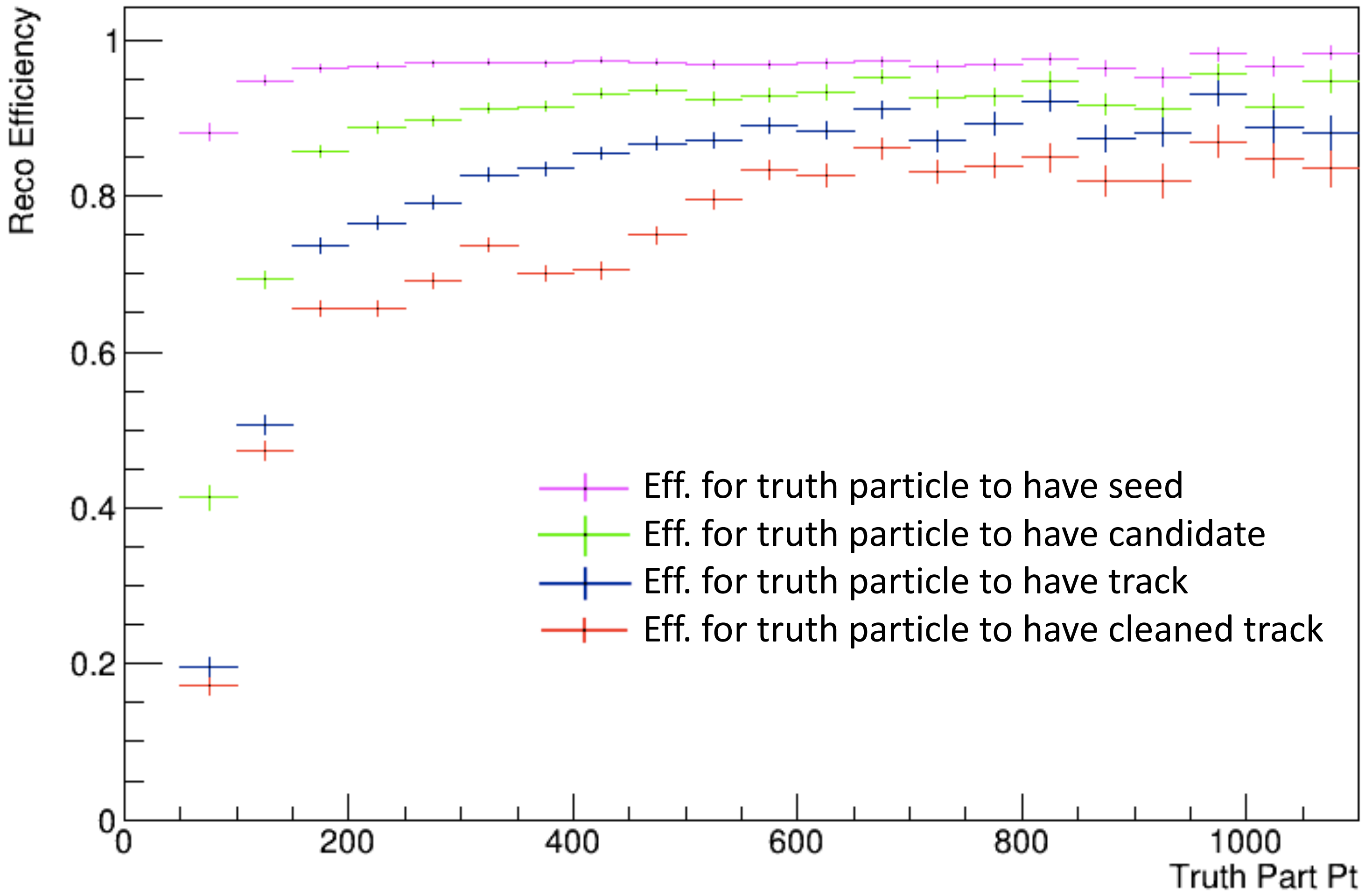
A deeper look: seeds and candidates

- We essentially have a 4 stage process
  - Create seeds for low-pt tracks that fall within RoI
  - Create candidates from the seeds
  - Create tracks from the candidates
  - Apply cleaning cuts to get a final set of tracks (see backup)

  - max number of SCT Holes 1
  - min number of Pixel Hits 3 for $200 < p_T/\text{GeV} < 350$
  - min number of Pixel Hits 4 for $350 < p_T/\text{GeV} < 500$
  - If IBL hit is expected: at least 1 IBL hit; if no IBL hit is expected: a Layer-0 hit if expected
- Questions we want to answer
  - Where is our main loss of efficiency?
  - How many seeds/candidates/tracks do we get per truth particle?
  - Where do fakes come from?
- Most of the following studies were performed in a Zmumu sample with a 35mm beamspot and a filter restricting the number of truth-level charged particles with Pt > 500 MeV from the hard-scatter to be less than 11

Seeds and candidates- efficiencies

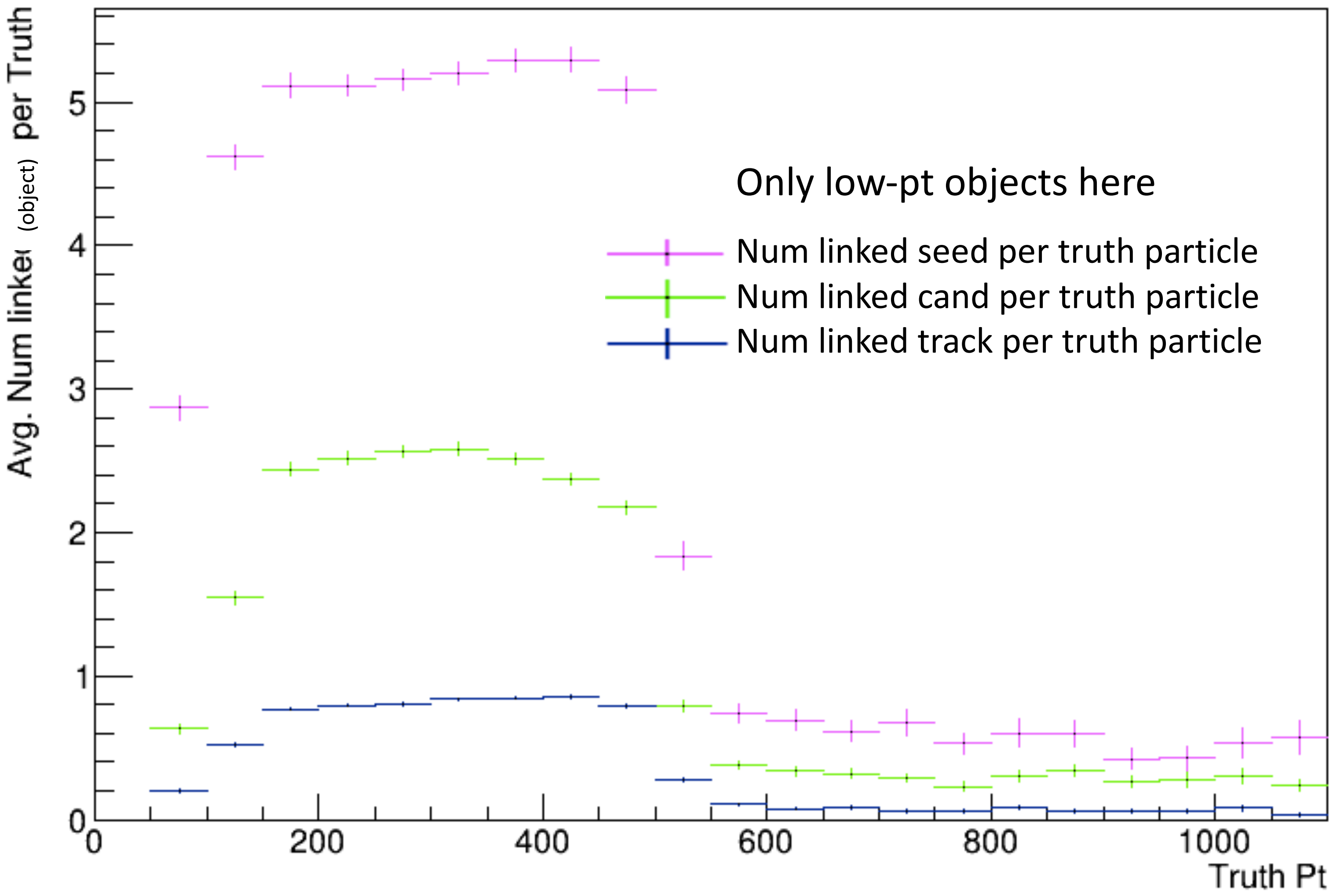
- A high fraction of truth particles have at least one linked seed
  - Linking here is done in the traditional way- a weighted average of the truth particles that contribute to hits in the reconstructed object
- For most bins, there is a consistent drop in eff. from seed to cand to track
  - Larger drops at lower pt



"Efficiency" is defined as the fraction of truth particles within 1mm of the event's hard scatter which have **at least one** linked object of the type in question (seed/cand/track)

Seeds and candidates- per-particle production

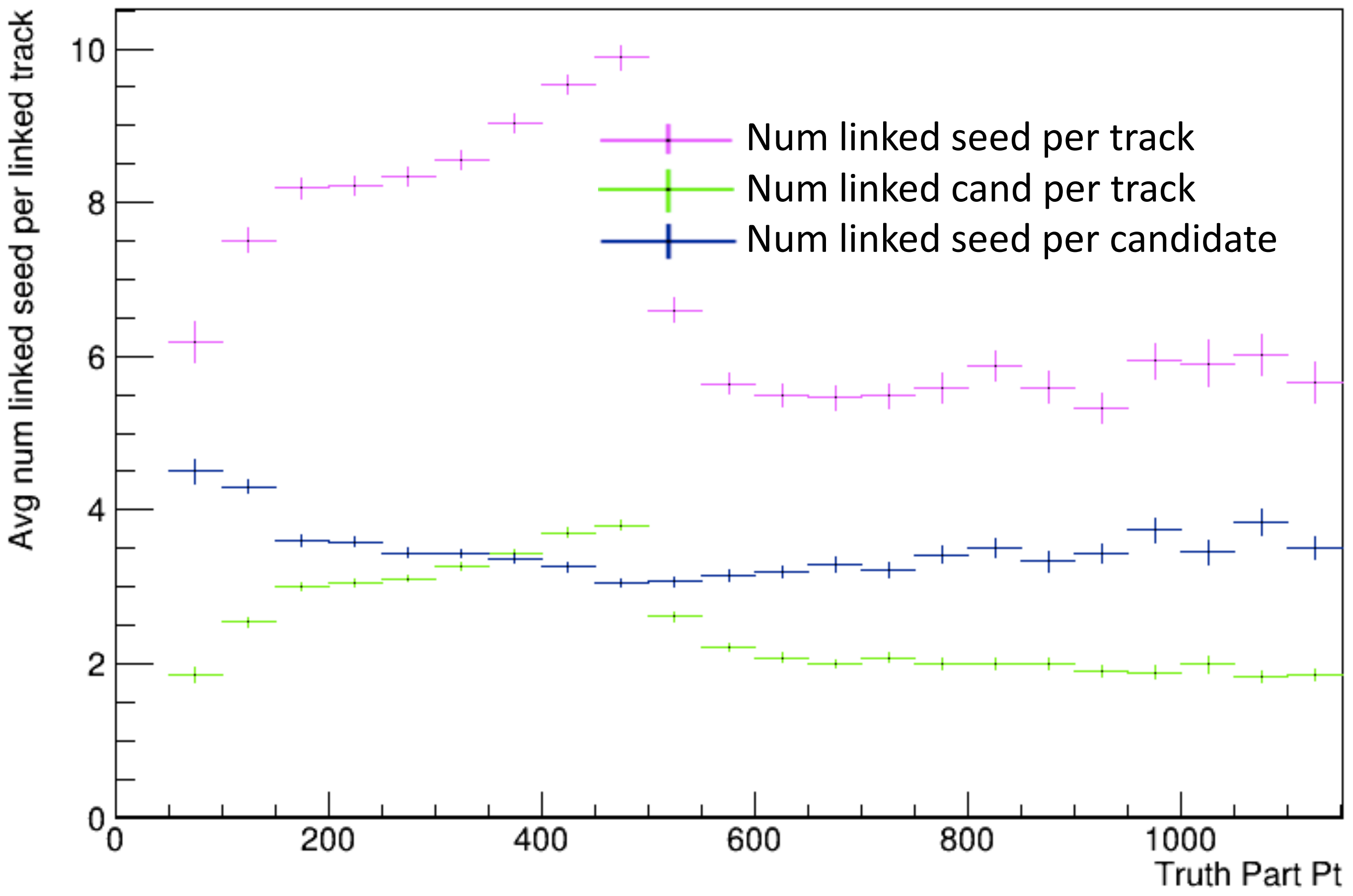
- Looking at low-pt objects alone, we see that the number of seeds/cands/tracks per truth particle is fairly constant between 150 and 500 MeV, with a steep dropoff below 150 MeV
  - One would really expect the seeds to “blow up” in the low-pt region due to combinatorics, but there are internal cleaning cuts which prevent this



Seeds and candidates- along the chain of reconstruction

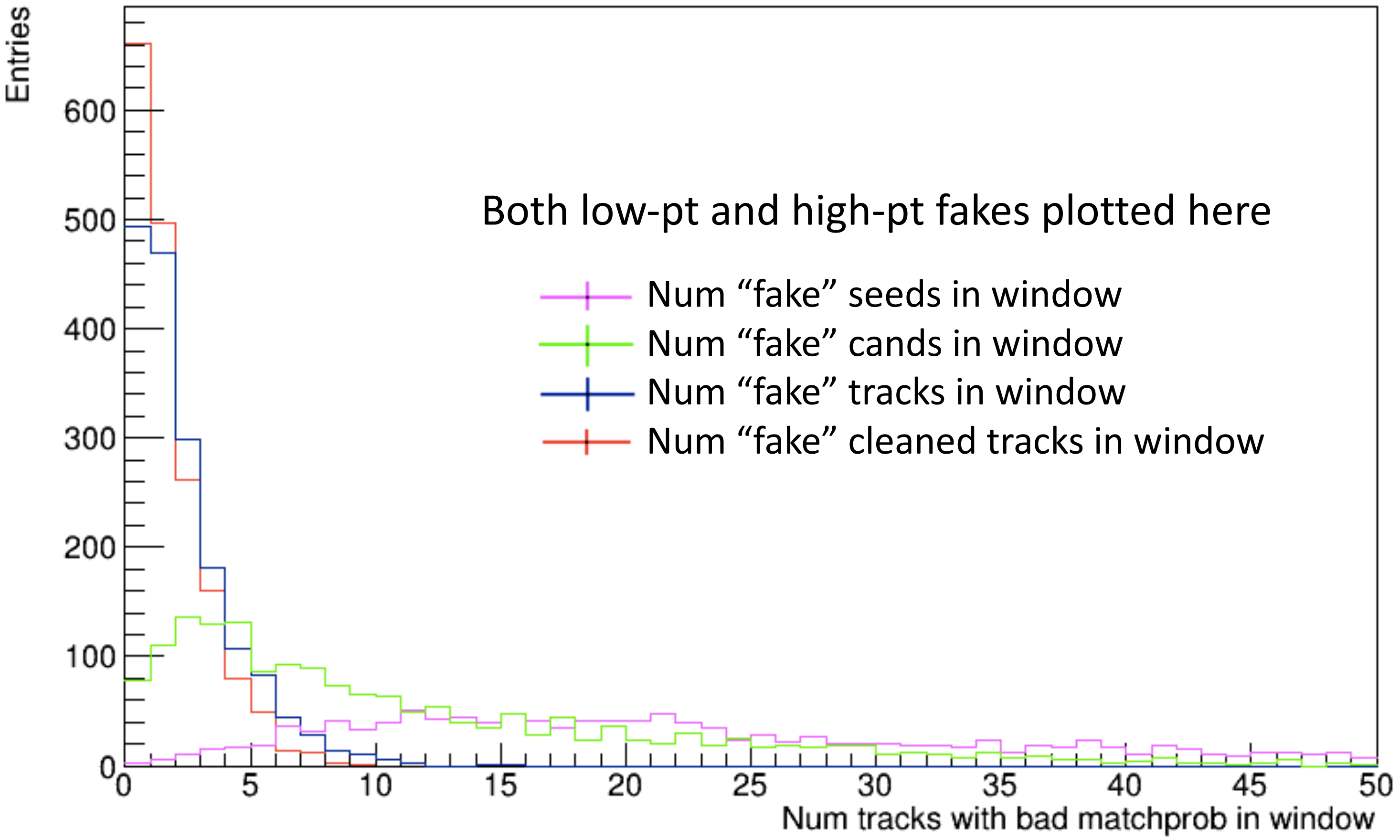
- Above we looked at the seeds/cands/tracks per TRUTH
- Now look at seeds/track, cands/track, and seeds/cand
  - Traces the first 3 steps in track reconstruction (plotted on right for both low-pt and high-pt objects)
  - Seeds/track and cands/track look as expected from previous slide
  - The number of seeds per candidate is relatively flat, but rises at lower pt- fewer seeds are progressing to the next step in that regime

Both low-pt and high-pt objects plotted here



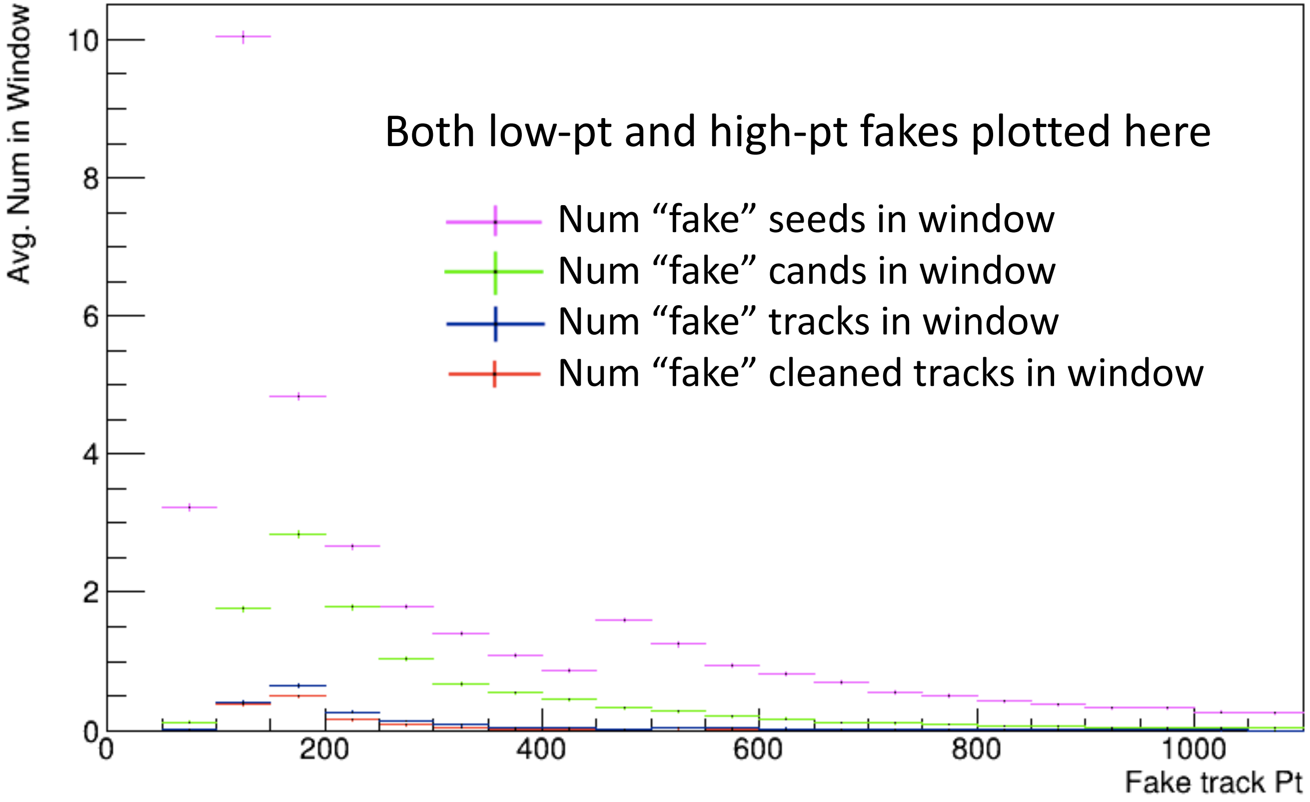
Seeds and candidates-per-event fakes

- The number of fake objects (within 1mm in z of HS and within 1mm of d0) is very important to yy->WW analysis
- By the cleaned step of track production, 0 fake tracks is the most likely outcome



Seeds and candidates- fakes vs Pt

- The problematic bin for cands, tracks, and cleaned tracks is 150-200 MeV
- Interestingly, the most problematic bin for fakes is 100-150 MeV (by a factor of about 2 over the next bin)
  - It seems that there is good reason that seeds should be less likely to be accepted below 200 MeV, and especially below 150 MeV



Seeds and candidates: quick conclusions

- Ultimately, it looks like much of the drop in efficiency below 150 MeV can be contributed to not only a **lack of seeds** in that regime, but also a **general low-quality of seeds**
  - For truth particles with Pt < 150 MeV, there are fewer seeds produced, despite it being combinatorially favorable to have many such seeds. Built in restrictions on seed production seem to cause this
  - However, the lack of very-low pT seeds is well justified- even in the 150-200 MeV track bin, we start to see a steep increase in the rate of fakes. Half of events are expected to have a fake track in the window with Pt between 150-200 MeV even after cleaning cuts!

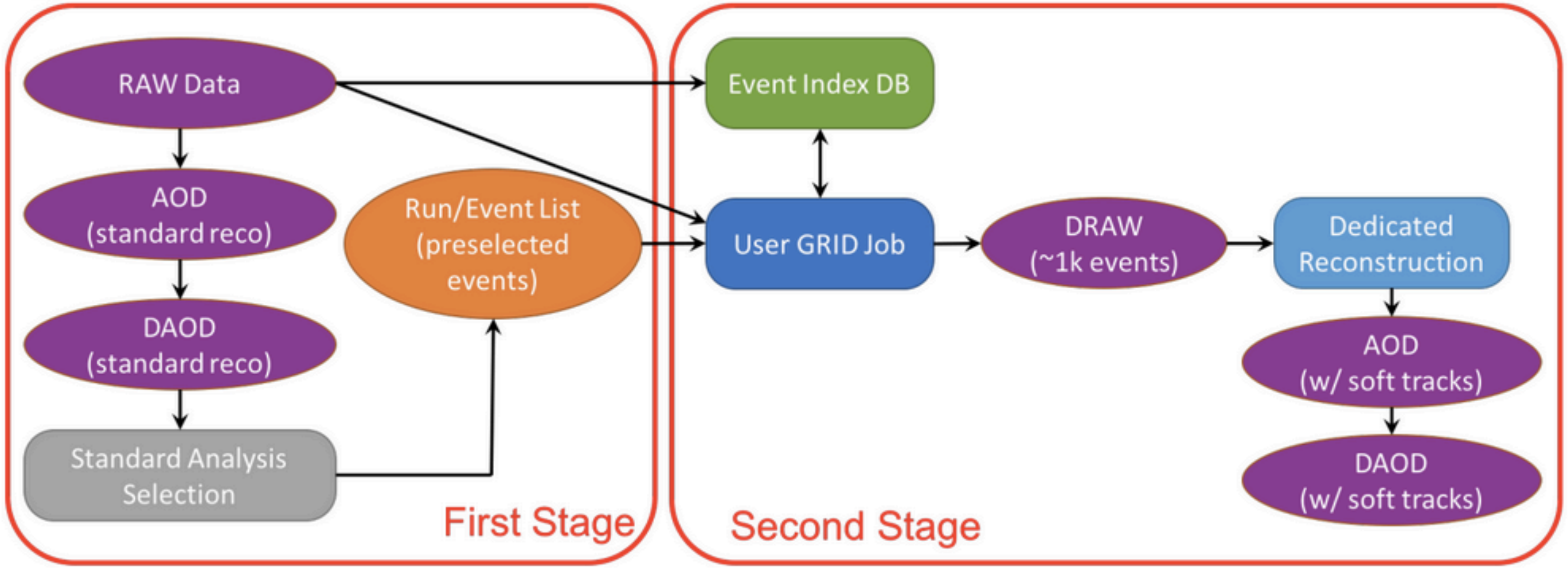
Conclusions

- Low-Pt tracking can be a useful tool for many analyses
  - We have created a framework in the ATLAS software that can be applied on events of interest to specific analyses, making it very portable
- Extending tracking down to ~200 MeV can actually be done in generic run 2 pileup conditions with relatively high efficiency and low fake rate
  - In the very-low Pt regime, fake rate tends to become significant. Right now, built in cuts in our seeding algorithm restrict the number of seeds produced there; without these cuts, the fake rate would almost certainly worsen
- More studies will be needed if we want to aggressively reduce the fake rate in the very low-Pt regime

Backup

A word on selecting events

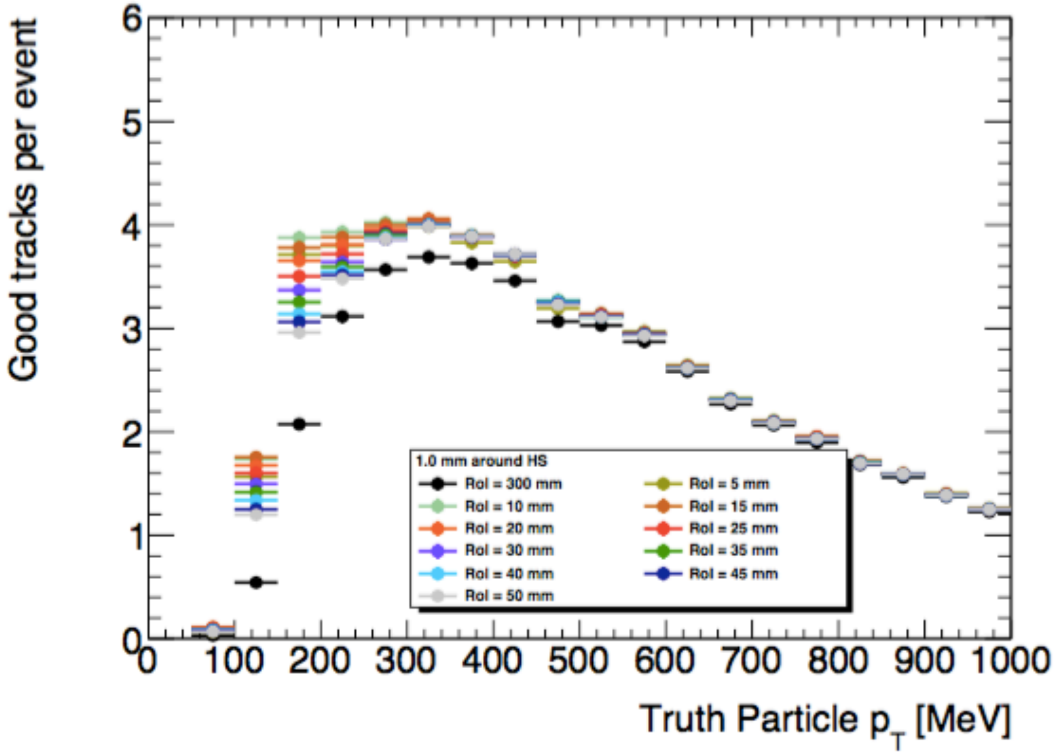
- It is clear that investigating a low-pt tracking setup is worthwhile
- But with that said, we DO NOT suggest a full reprocessing of data in order to run low-pt tracking
  - Analyses interesting in low-Pt tracking can normally make a list of interesting events (before applying low-Pt tracking)
  - Small dedicated DRAW dataset is produced using these events, and low-pt tracking is run on these events

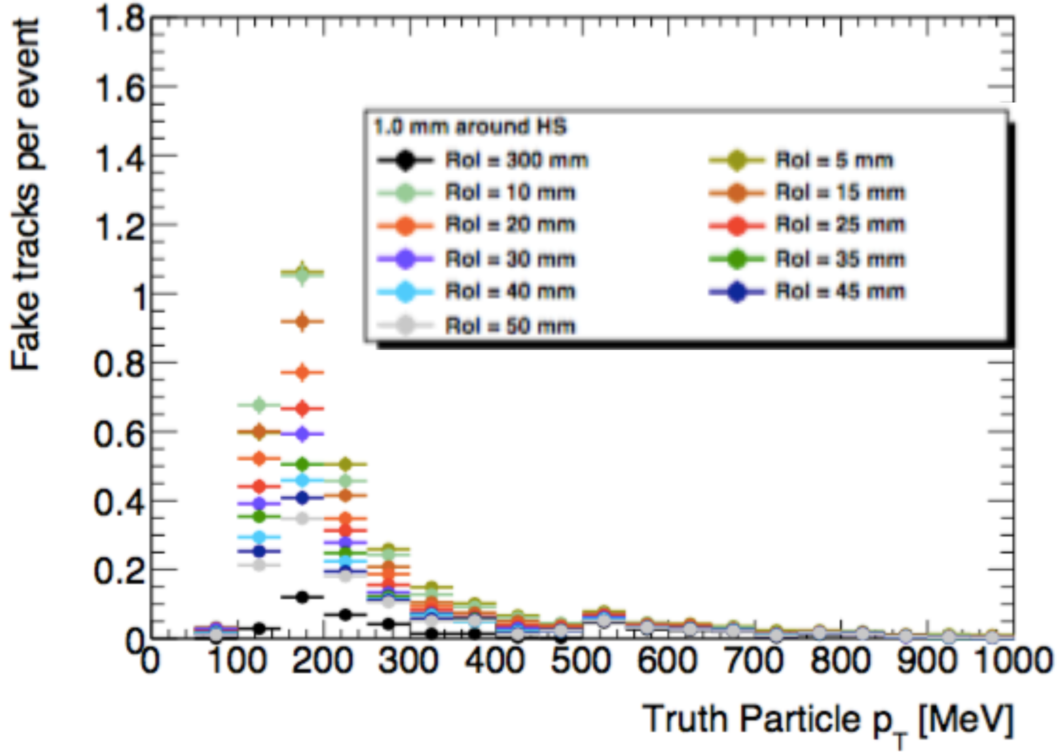


This overall process can in principle be applied to any analysis with unique reconstruction needs

Initial study: RoI size and cleaning cuts

Raw values of added tracks very important for us

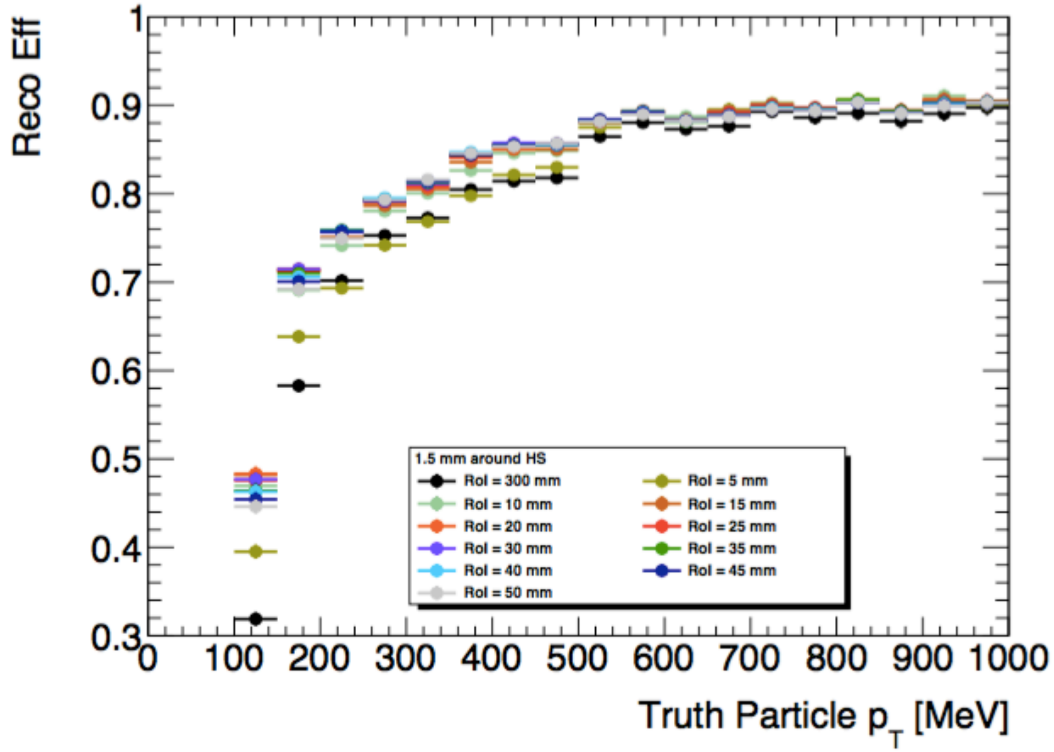


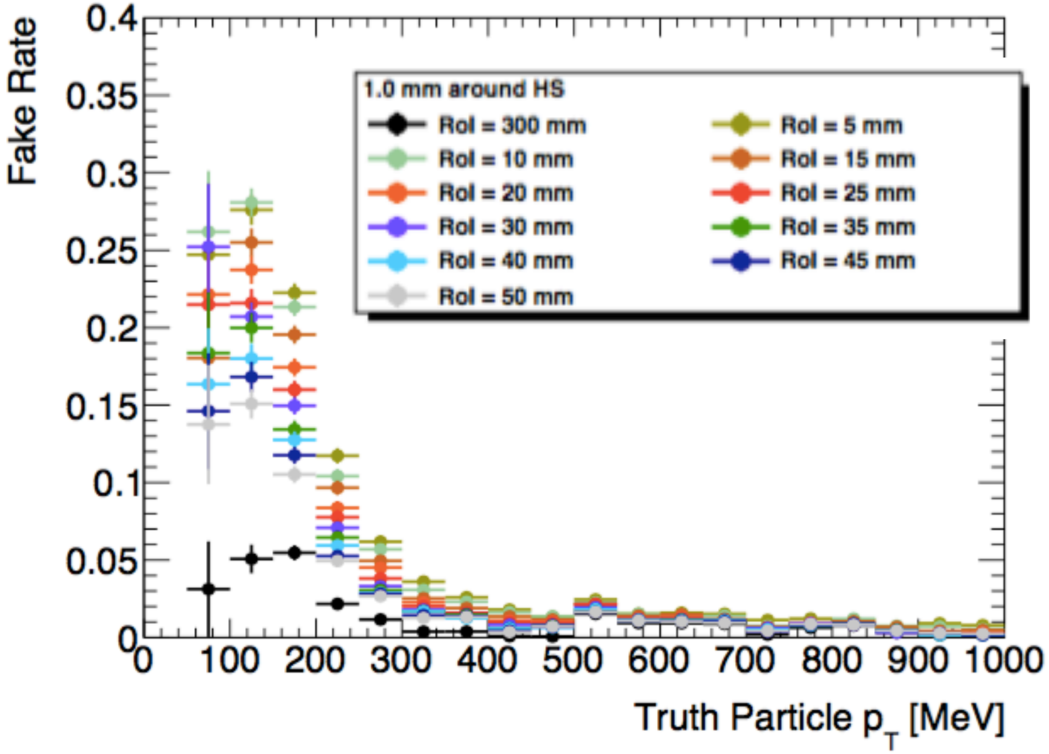


Plots made using an Inclusive WW sample with full PU truth info

RoI for track seeds given by legend, but only tracks/truth particles within 1.0mm of di-leptons plotted

Low-pt tracking results



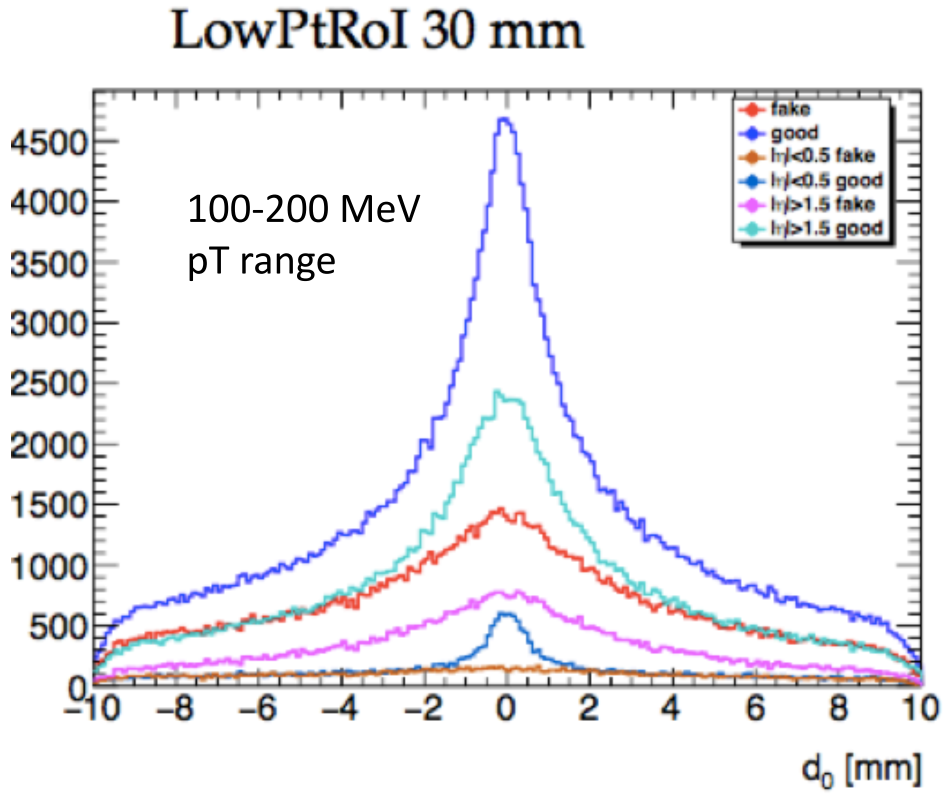


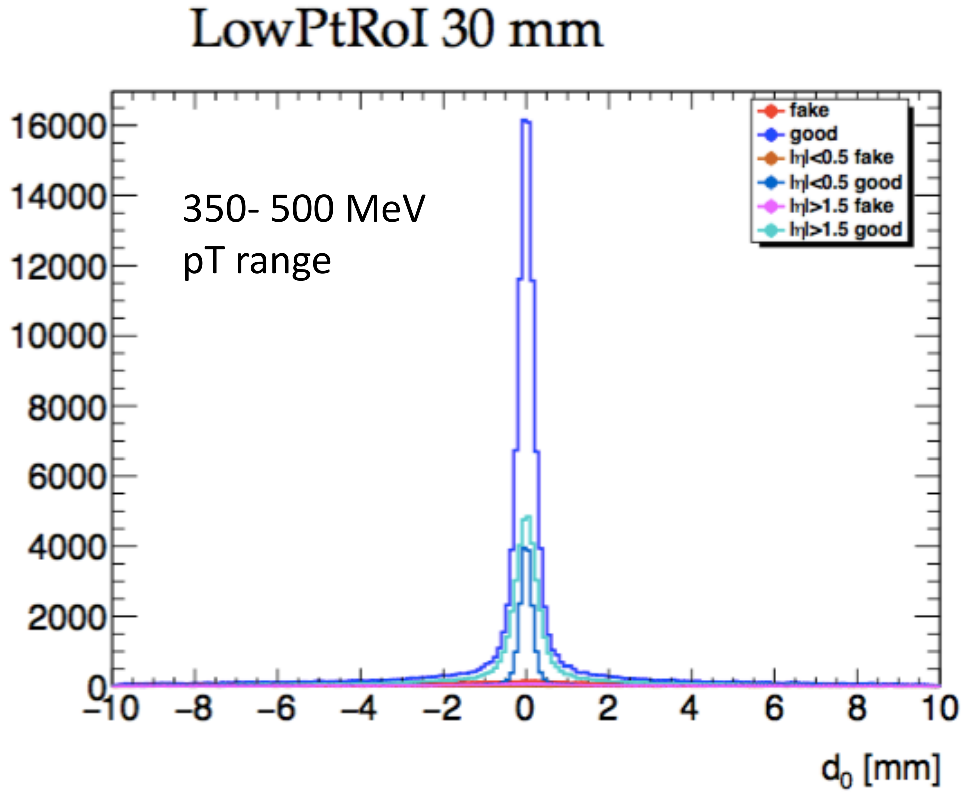
Plots made using an Inclusive WW sample with full PU truth info

RoI for track seeds given by legend, but only tracks/truth particles within 1.0mm of di-leptons plotted

Offline track selections

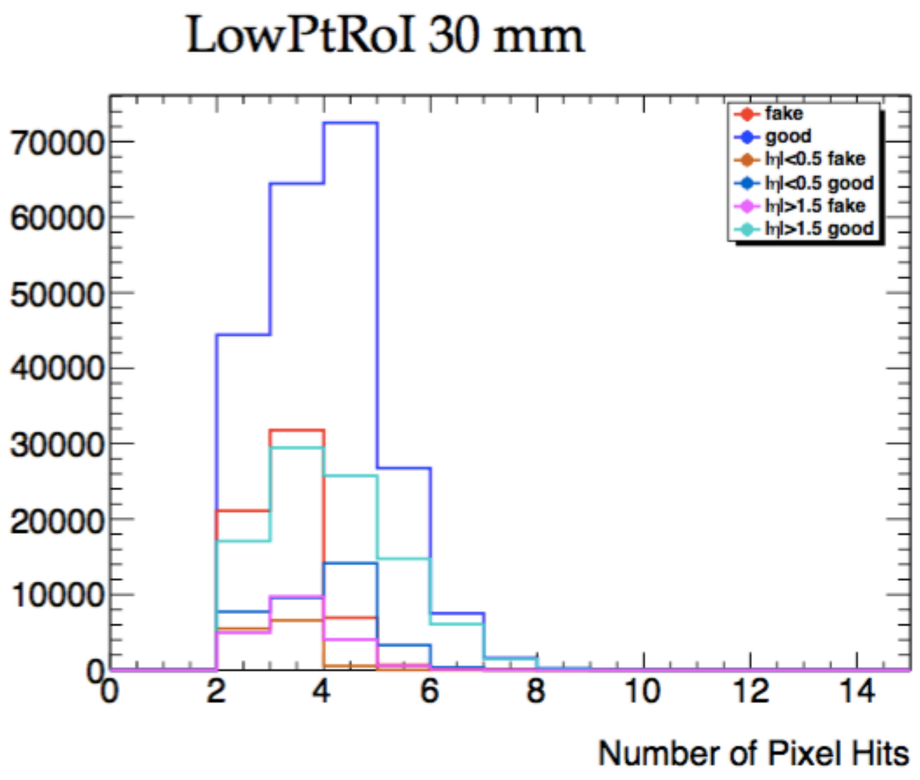
- Different regions of track phase space have different overall quality
  - E.g. lower pt, higher eta, and greater d0 all associated with great fake rate

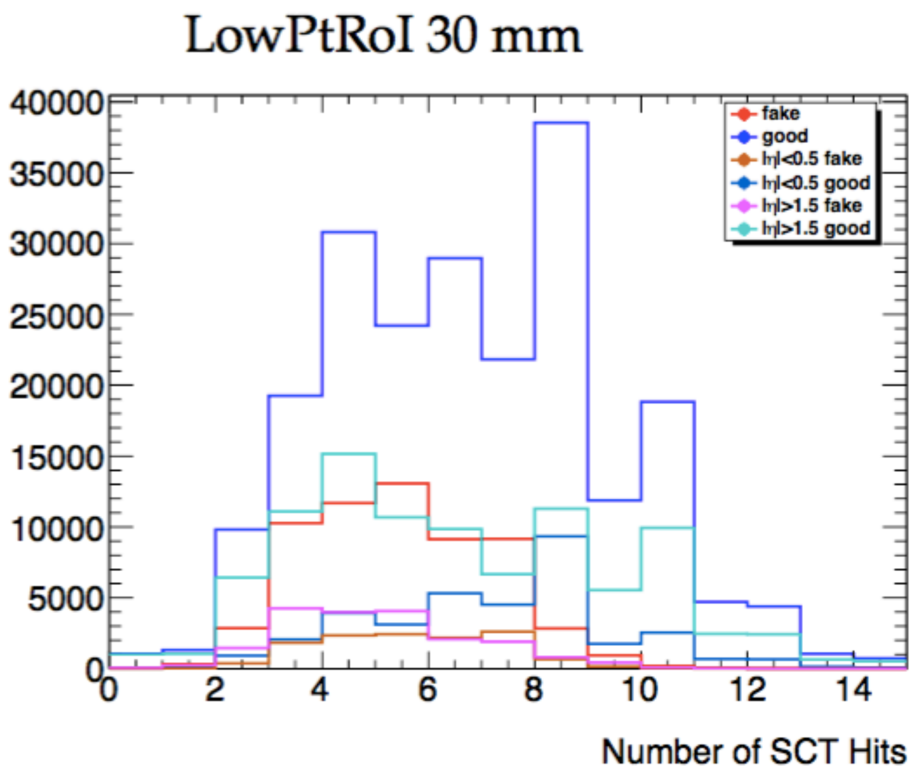




Offline track selections

- Numbers of pixel and SCT hits also give valuable information

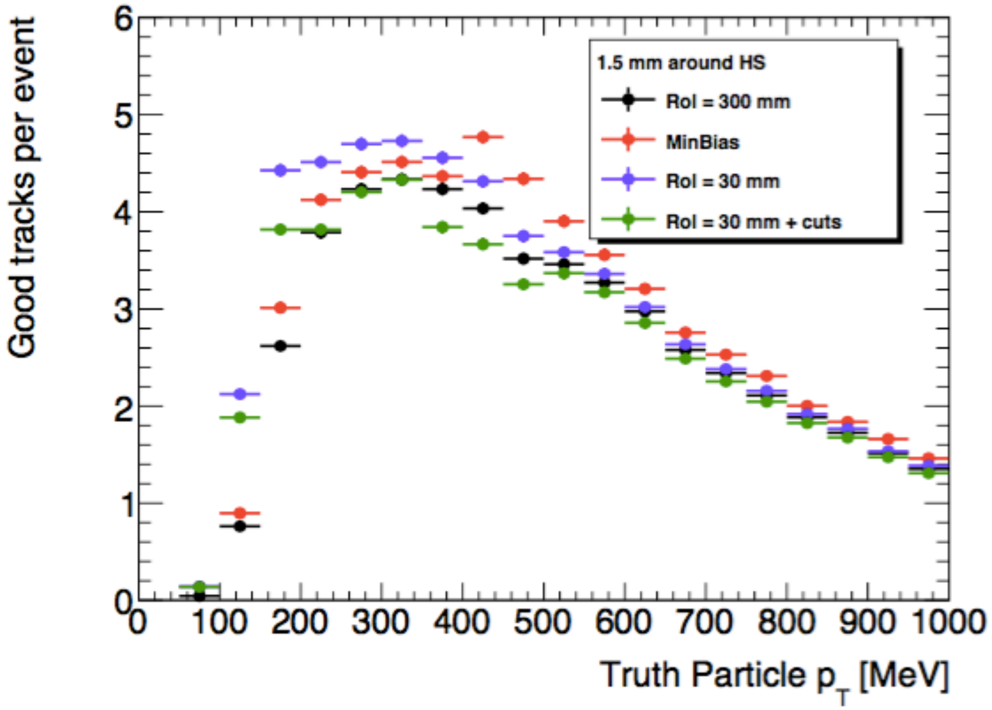


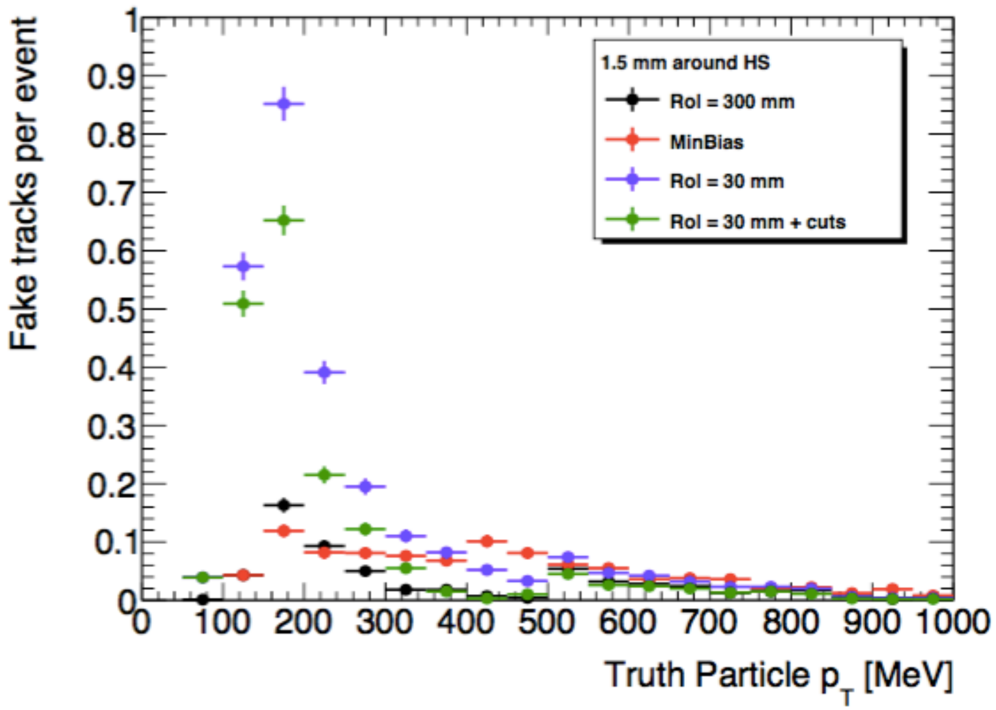


Plots of the number of Pixel and SCT hits for 200-350 MeV pT slice.
Fake tracks tend to have fewer hits.

Offline track selections

- With a sample set of cuts like (inspired by MinBias set of cuts):
  - max number of SCT Holes 1
  - min number of Pixel Hits 3 for $200 < p_T/\mathrm{GeV} < 350$
  - min number of Pixel Hits 4 for $350 < p_T/\mathrm{GeV} < 500$
  - If IBL hit is expected: at least 1 IBL hit; if no IBL hit is expected: a Layer-0 hit if expected



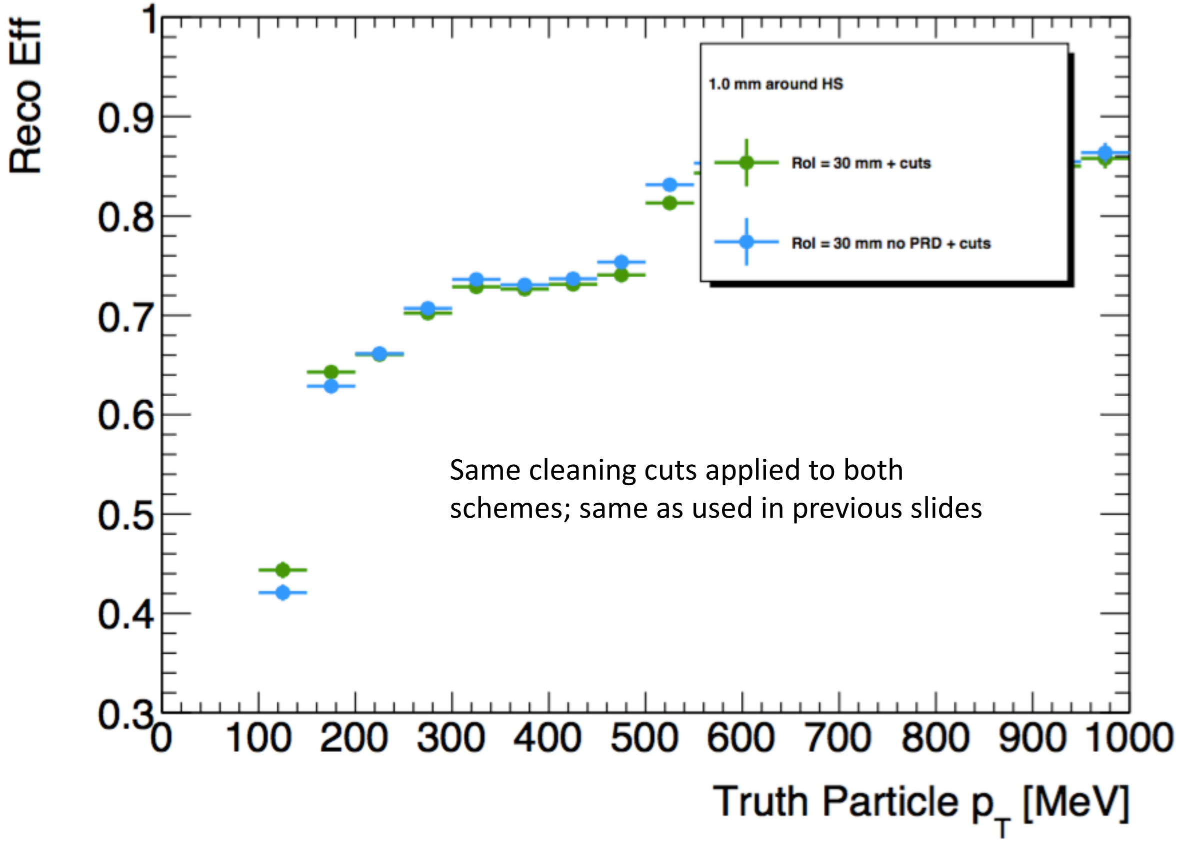


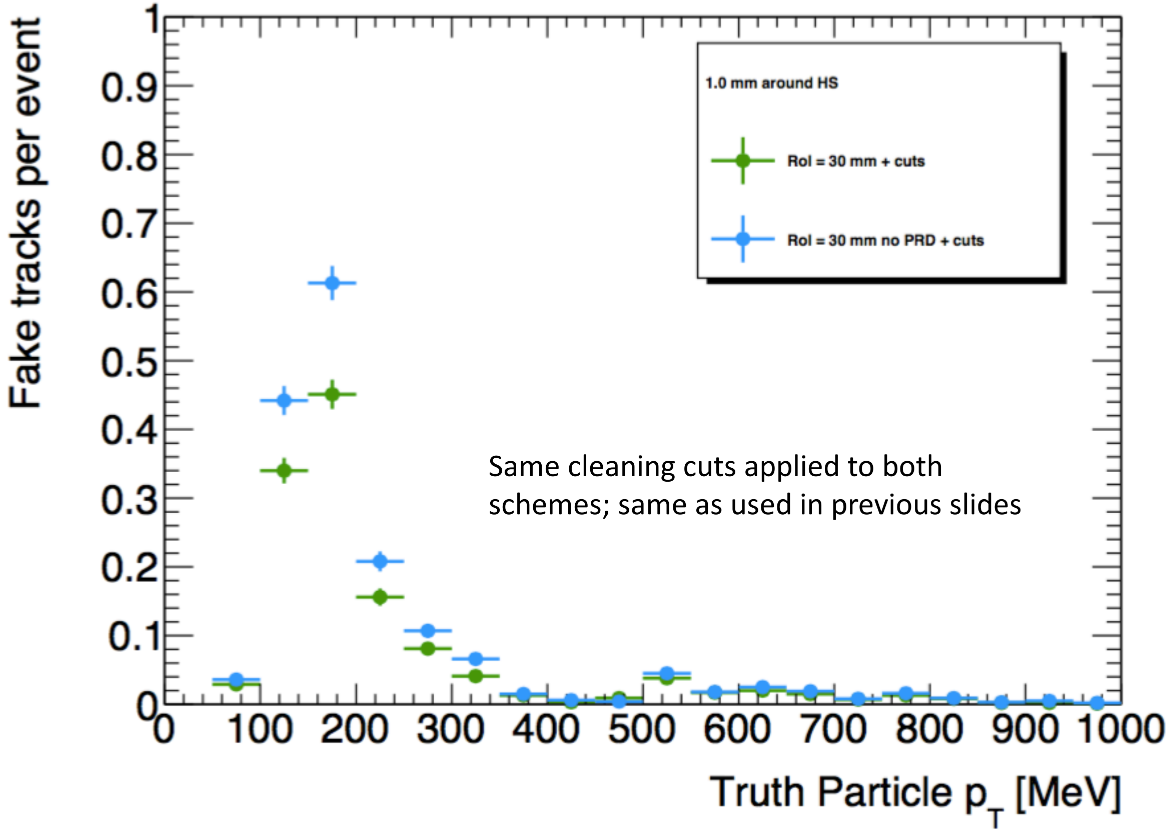
Seeds and candidates- no hit-removal study

- Another question worth asking is should we retain all inner detector hits for the low-pt tracking step?
  - Perhaps removal of hits used in generic tracks leads to lower efficiency
  - Also, a lower fake rate had been observed in low-pt tracking studies in low-pileup minbias samples, which used a single pass to recontruct both low- and high-pt tracks
  - More importantly: many low-pt tracks have at least a couple of pixel+SCT holes, which restricts how effective cleaning cuts can be (see backup). Will low-pt tracks that have all ID hits available to them have fewer holes?

Seeds and candidates- no hit-removal study

- See below the efficiency and expected per-event fake contribution
  - Green shows our main setup, blue shows the setup WITHOUT hit removal
  - WITHOUT hit removal, little effect on efficiency with some **increase** in fakes





Seeds and candidates- no hit-removal study

- However, a main point here was to check the effect on track quality
  - The scheme WITHOUT hit removal, blue and teal (for fakes), seems to have generally similar quality to tracks from the default scheme with hit removal