Reconstruction Validation for HD Production

Ryan Cross

Upcoming Production

There is an upcoming production for the DUNE Far Detector, specifically the horizontal drift detector.

There is much more information in Dom's initial slides, linked below, but the short version is that we've not had a production since 2018, and lots has changed since then, so it would be useful to have new production files as they are used so heavily as reference points.

Also as part of this updating process, large parts of the code base have been tidied up, de-cluttered, and made more modern by including advances by PD-SP, rather than relying on TDR-era bits (or 35-ton era).

Since this is such a large update, the validation of some reference files is very useful, before a larger, full stats production begins.

There is numerous samples of around 15,000 events (forward going mu, APA crossing mu, `nu`, `nue`, `nutau`) that are fully reconstructed, and can be used to help verify the production. Check Dom's slides if you'd like more info on that!

Dom's slides are here.

Pandora CI Configuration

As part of verifying the production, I can re-use some of the machinery that was put in place for a high-stats Pandora CI configuration.

Earlier in the year, with lots of help from Vito, I added some reconstruction level plots that let us keep an eye on reconstruction metrics, as a configuration in the existing CI infrastructure.

Using the Pandora analyser module (from Maria Brigida Brunetti @ Warwick), we can plot out reconstruction metrics, and keep an eye on how changes to Pandora and similar alter the reconstruction performance as dunetpc updates.
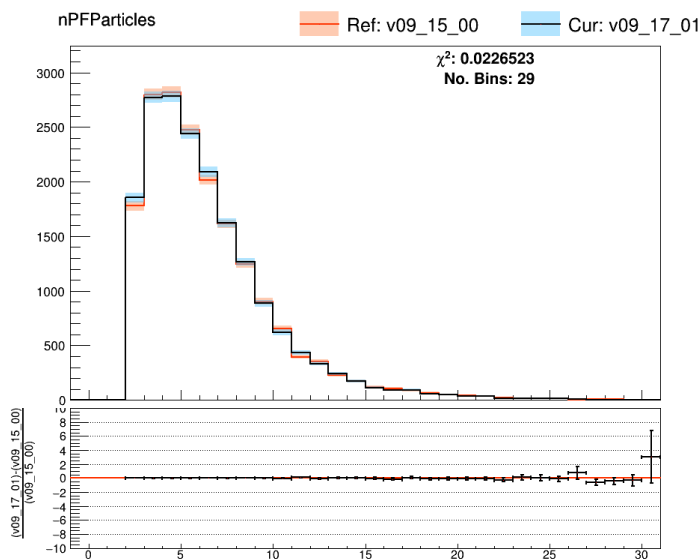
Older slides on this setup here.

Process

The process here is pretty simple, but I've included it for completeness.

First, run the analyser module over the already made files, and store the result.

Secondly, run the full chain of FHiCL files, from `G4` → `detsim` → `reco` → `ana`, on an older non-refactored version of dunetpc, but using the same events.

We can then compare the plots using the CI scripts, to get a comparison where the events themselves match, but the sim/reco module they use is across 2 versions of dunetpc.

I'm going to talk through some of the more interesting plots, though its not fully understood where some of the features come from yet, so further work will be needed to understand some of the plots.

Rolling up Showers

The first issue encountered is that showers aren't being rolled up, whereas they are for the current standard workflow files.

That makes some of the plots look very different, which is unfortunate, but does at least show one issue that needs updating. Luckily, that is just a FHiCL parameter being flipped.
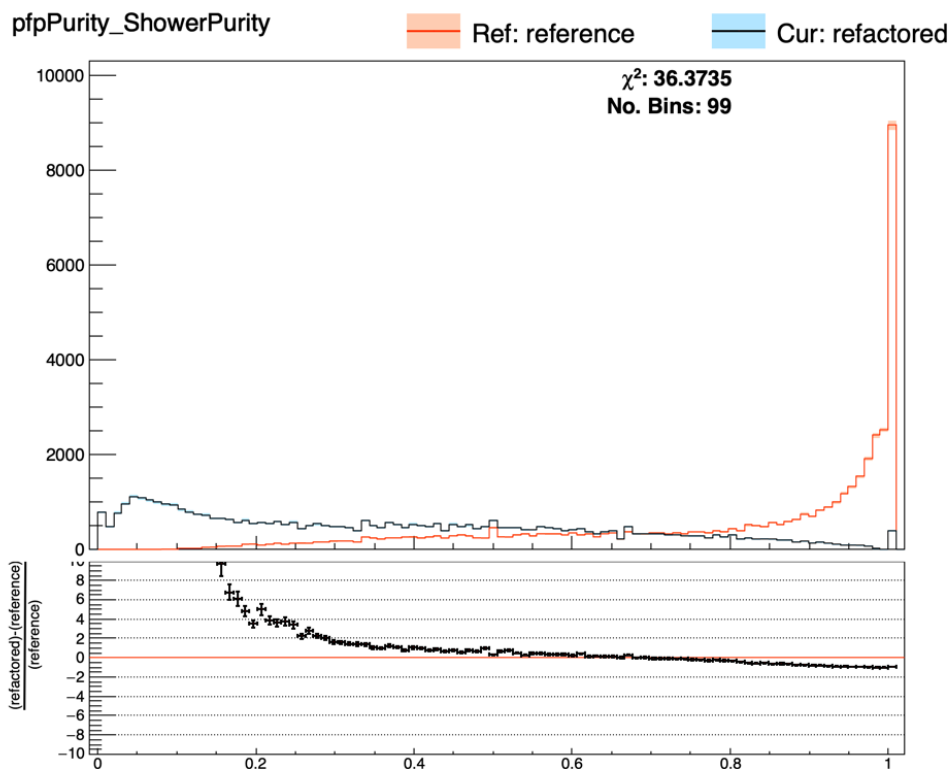
Number of `PFParticles`

The broad number of `PFParticles` has changed somewhat, though I've not been able to pin-point exactly what has caused this change yet.

To be clear though, some changes should be expected, a lot of things have changed so Pandora's results will have changed as well, we just need to verify that this is an understandable change, and not a regression.
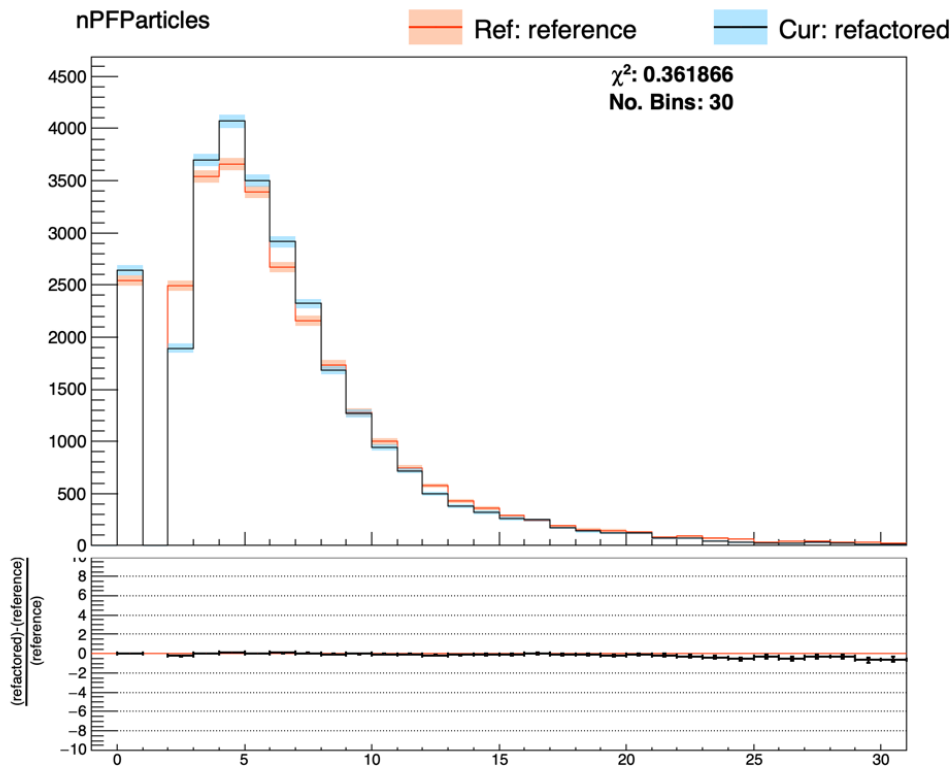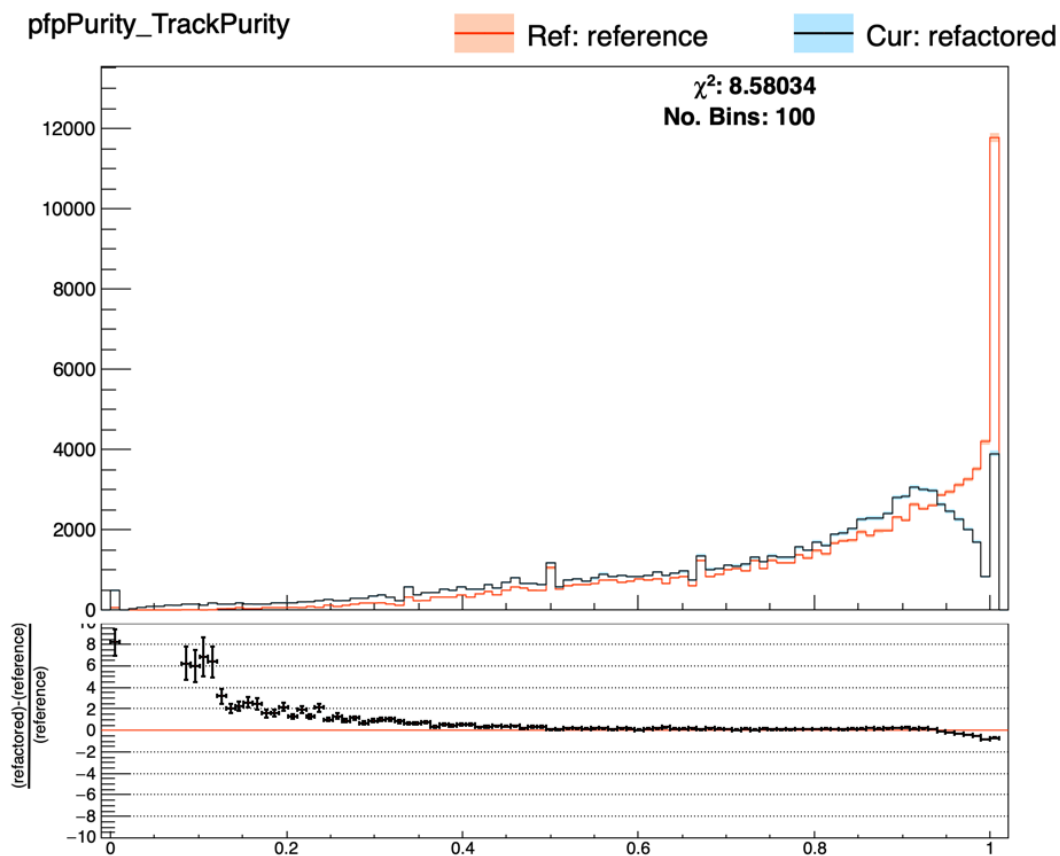
Tracks

A more confusing issue is that the tracks have seemingly also been reconstructed differently. This could be a knock-on effect of the shower rolling up causing issues with the generation of metrics (though I'm not sure how they could cause issues for tracks).
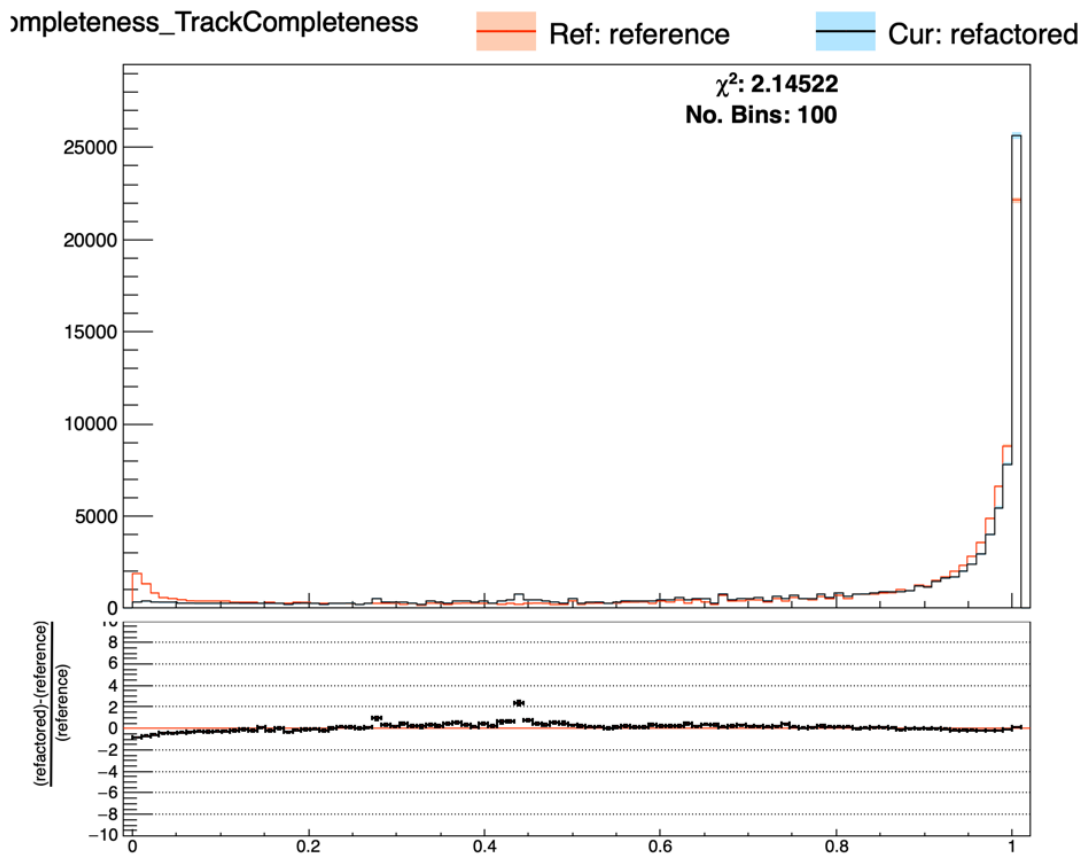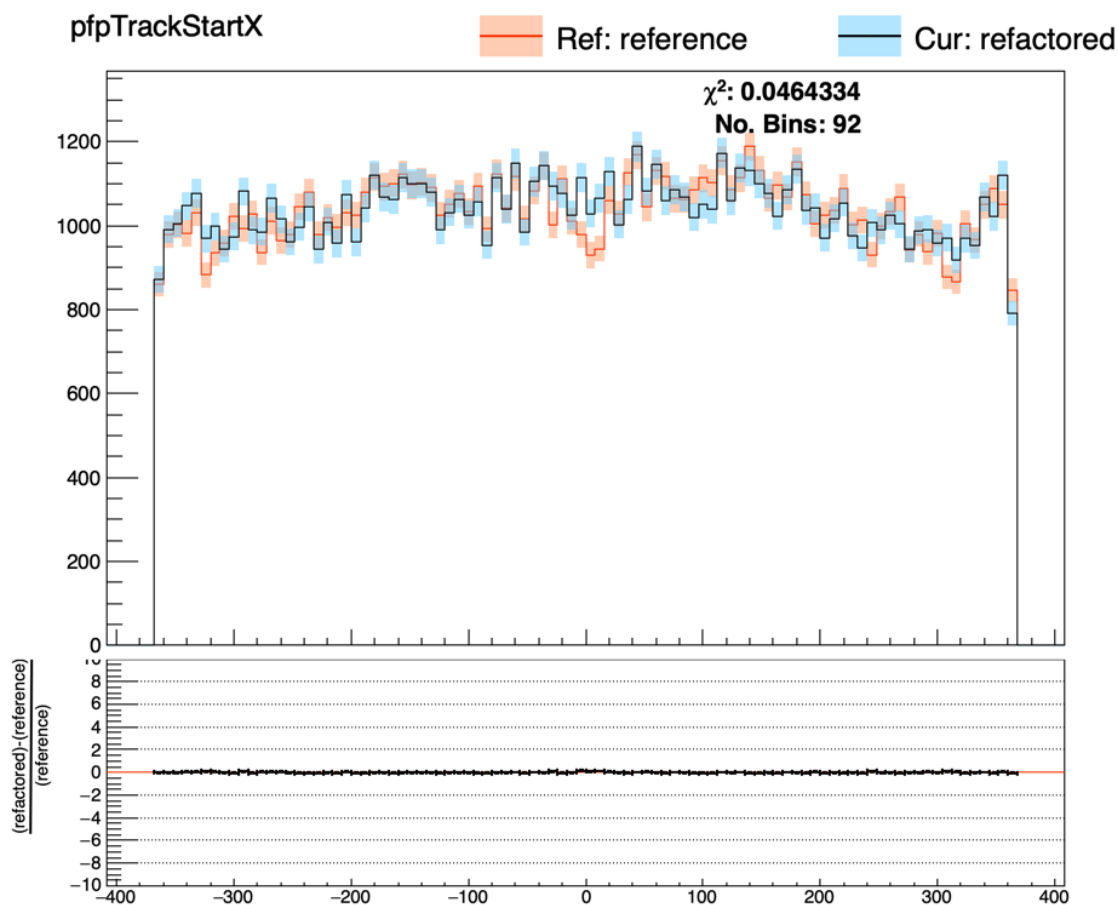
Tracks

A more confusing issue is that the tracks have seemingly also been reconstructed differently. This could be a knock-on effect of the shower rolling up causing issues with the generation of metrics (though I'm not sure how they could cause issues for tracks).
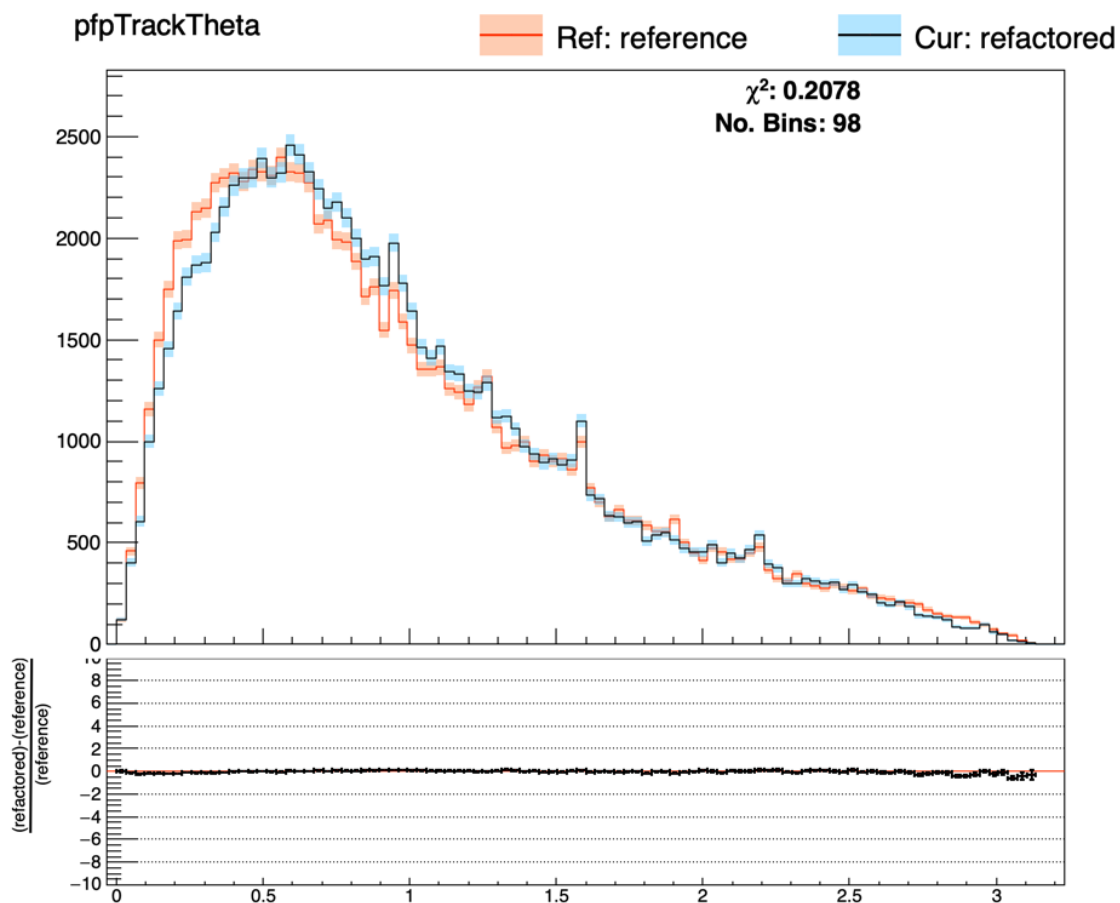
Track Properties

On a better note, looking at the broad properties of the tracks, such as their direction, starting positions, angular distributions, things are in mostly good agreement.
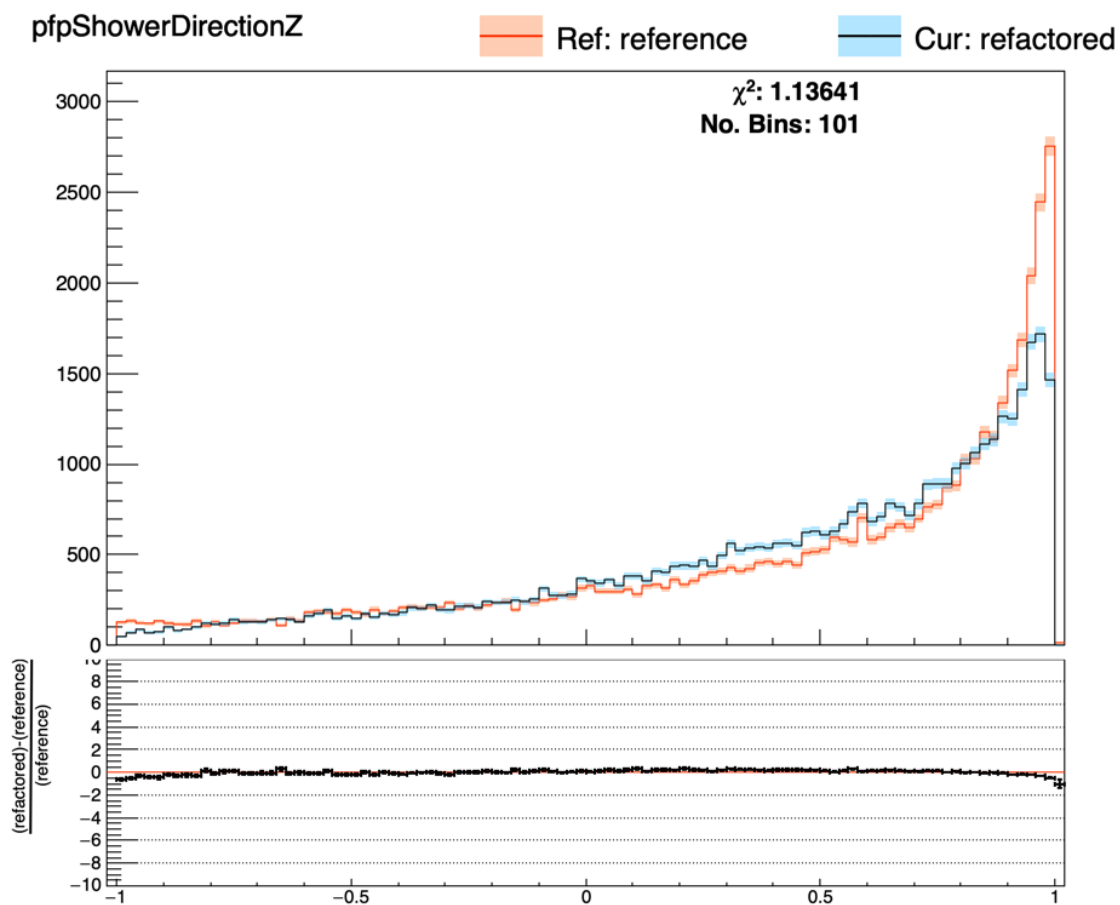
Track Properties

Though there is some slight change in the track angle plot that will need to be checked.

Shower Properties

For showers, its a bit of the opposite story, where any plot that requires a rolled up shower is currently broken, but the plots that don't require that (or are less sensitive to it) look normal.
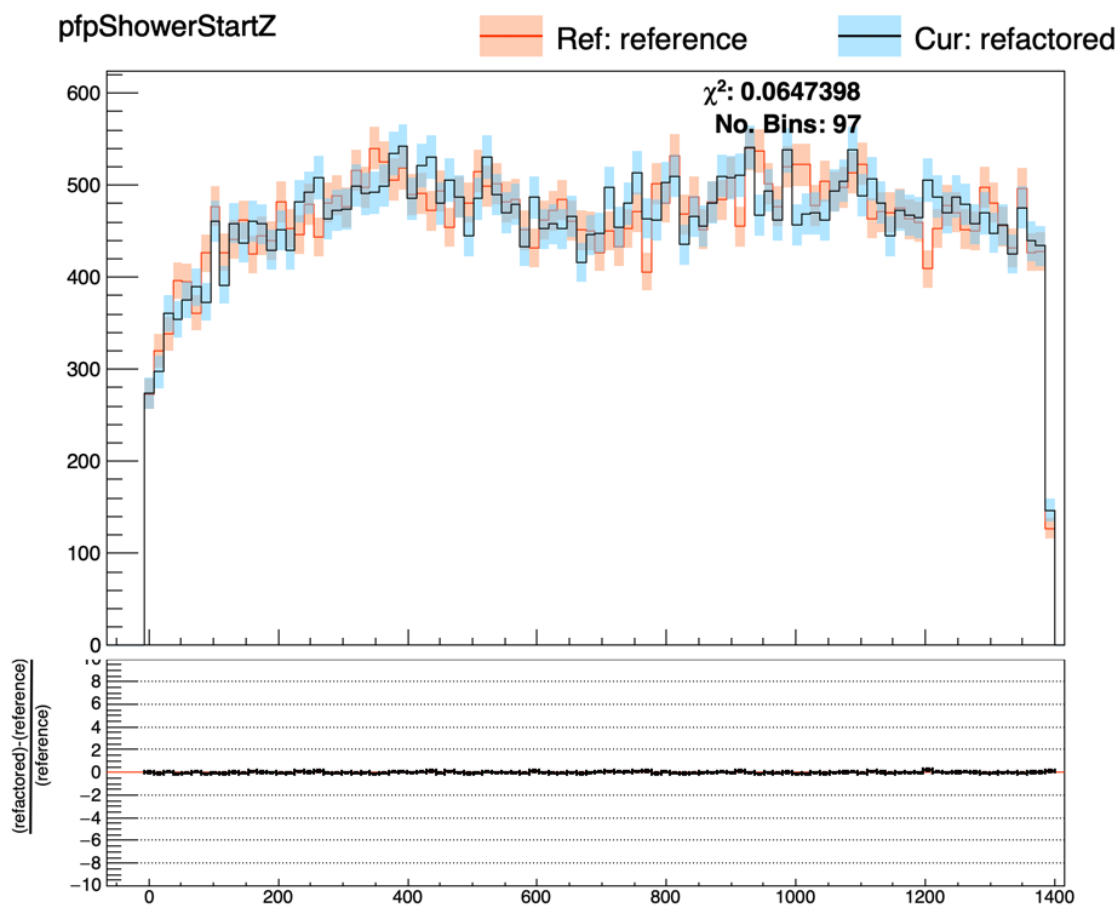
Shower Properties

For showers, its a bit of the opposite story, where any plot that requires a rolled up shower is currently broken, but the plots that don't require that (or are less sensitive to it) look normal.

Points of Interest

There is obviously still some differences that we don't yet understand, so it would be good to get a handle on them.

- The showers broadly look good, so hopefully just re-enabling rolling up the showers will fix issues there.
- The issues with tracks will require a bit more investigation to work out what is happening.
  - Firstly, probably makes sense to do some easy digging and flick through a few event displays to check things by eye.
  - Then try and dig into the metrics and find a case which does well before and badly after, to see if there is any discernible difference, or just an issue with the metrics.
- Once those two issues are sorted, will allow some other produced plots to be more useful (i.e. comparison of vertex position, which is currently harder to read when there is so many showers etc).
  - Stuff like the vertex placement is especially interesting as it relies on a BDT, so it will be good to see how that compares when key parts of the sim/reco chain are changed.

Reconstruction Validation for HD Production

Ryan Cross

Backup Slides