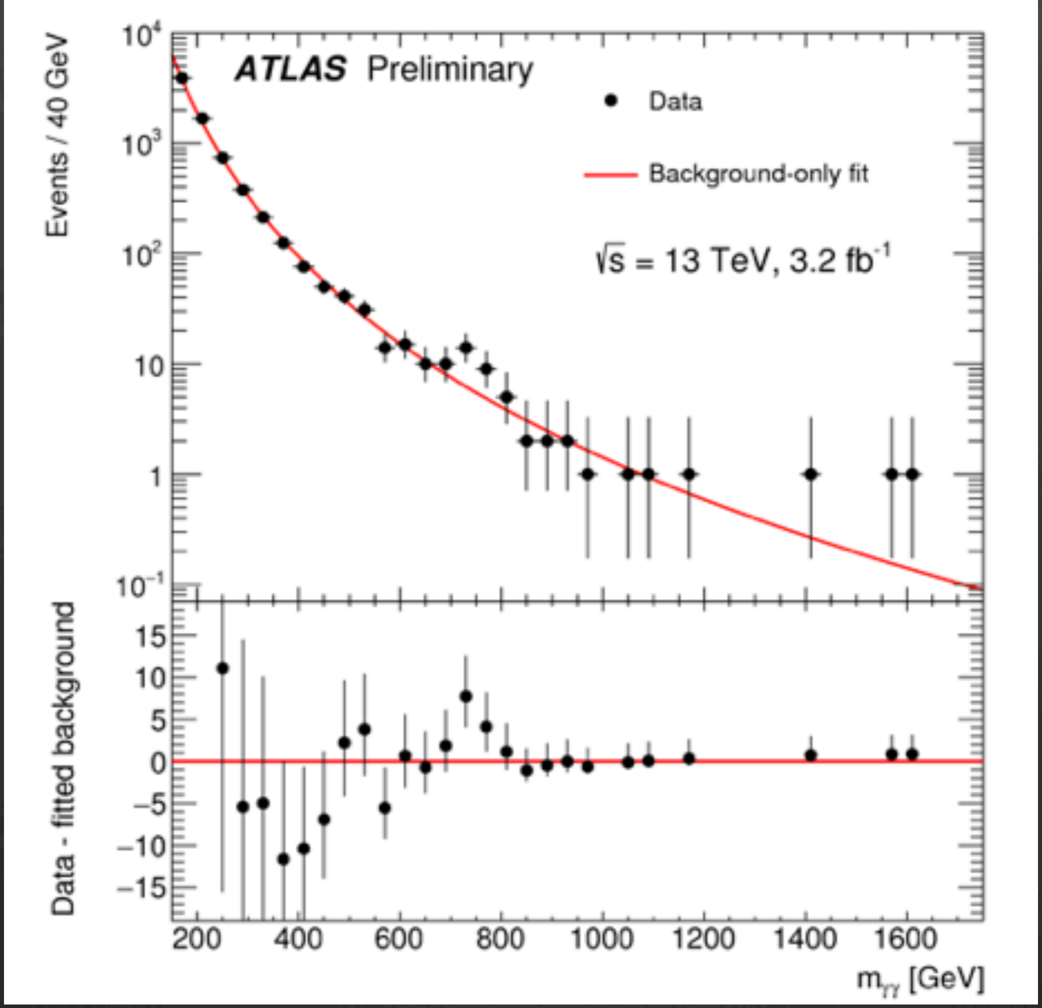
Isolation Studies for Diphoton Background

Alan Ottenstein

- The overarching goal of my project is to help in the search for undiscovered particles (maybe one with a mass of around 750 GeV..), particularly ones that decay into two photons.



http://www.physicsmatt.com/blog/2015/12/16/diphotons-at-750-gev

- At every collision, ATLAS (let's not worry about any other detectors) detects TONS of particles, and they try their best to identify what everything is.
- If you want to look for a particle which decays into two photons, you only want to look at the photons the detector finds.
- To "only look at photons," you have to get rid of any data that doesn't "look" like a photon.
- This is typically done using cuts on a variety of variables measured by the detector.
- This isn't perfect, and often you'll have hadronic jets "faking" as photons, i.e. they pass all the cuts, but still are not photons. We'll talk more about this later.

Background!

- How can we identify a "new" particle?
- We try to understand what our data should look like according to the standard model as best we can, so any deviations from this are obvious!
- In proton-proton collisions, photons with a wide range of energies are produced, and histograms of these photons with respect to energy, transverse momentum, etc. usually have fairly smooth lines.
- These photons are what we want our background to consist of.

Resonances!

- But we're looking for resonances, which are bumps that deviate from our smooth photon plots, and come from a particle being created and decaying into a pair of photons.
- So as I said before, the better we understand the background, the more easily we can identify resonances, and thus new particles!

Sounds easy enough right?

- Well, we have a theory that’s supposed to describe how our particles interact, The Standard Model, so shouldn’t we be able to use that to determine how the background should look?
- Unfortunately this, like many other things in life, is easier said than done.
- Things get complicated fast, and we don’t completely trust our simulations, and don’t want to rely on them to completely describe the background.
- A lot of techniques using both simulations and data are used.

Brief Recap

- So, given all the data we get from ATLAS, we want to throw out everything that isn't a photon, and keep everything that is.
- Then we want to divide these photons into background, i.e. photons created directly from the proton-proton interaction, and resonances, which are photons that result from a massive particle being created, and then decaying into a photon pair.

Jets!

- However, the "photons" we detect aren't always photons… there are always hadronic jets which "look" too much like photons to the detector to remove them from our data, i.e. they pass all of the cuts we apply to restrict our data to photons.
- So, what we want to do is try to a) make better cuts to minimize the number of jets, and b) figure out just what the contribution of jets is to our background. This is what I am working on.

Isolation

- A good way to decide if you've detected a photon or a jet is to look at isolation, which is basically a measure of how many other particles are near the particle you're looking at.
- The two isolation variables I look at primarily are Topoetcone40, which is the amount of energy measured by the calorimeter in a region around the particle, excluding the energy of the particle itself, and Ptcone20, which is the amount of momentum associated with tracks in the inner detector in a region around the particle.
- These isolation variables tend to be high for jets.

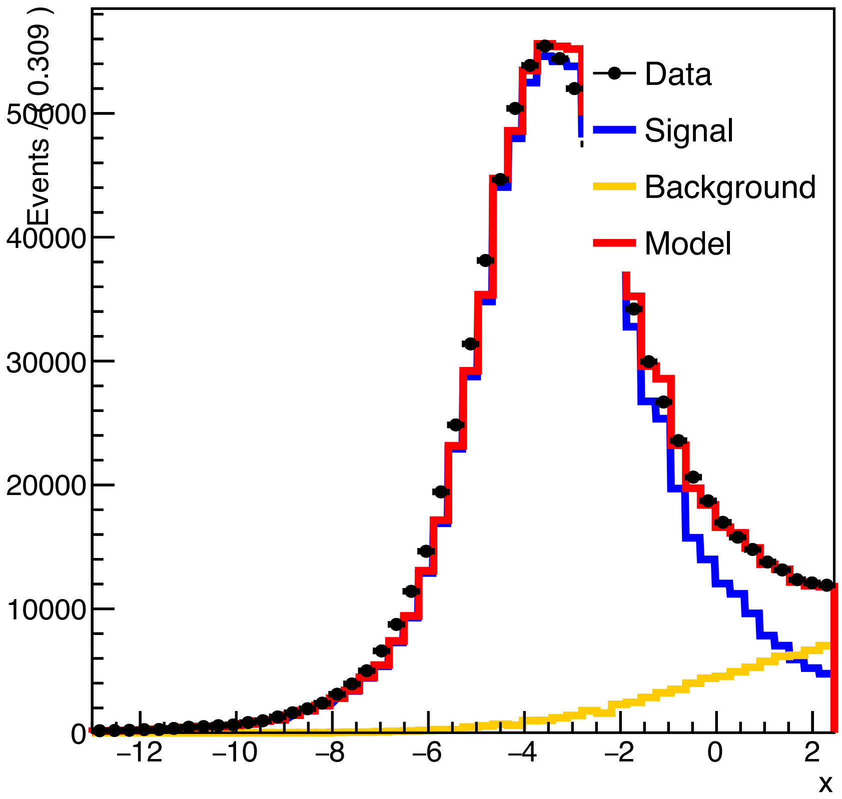
ME

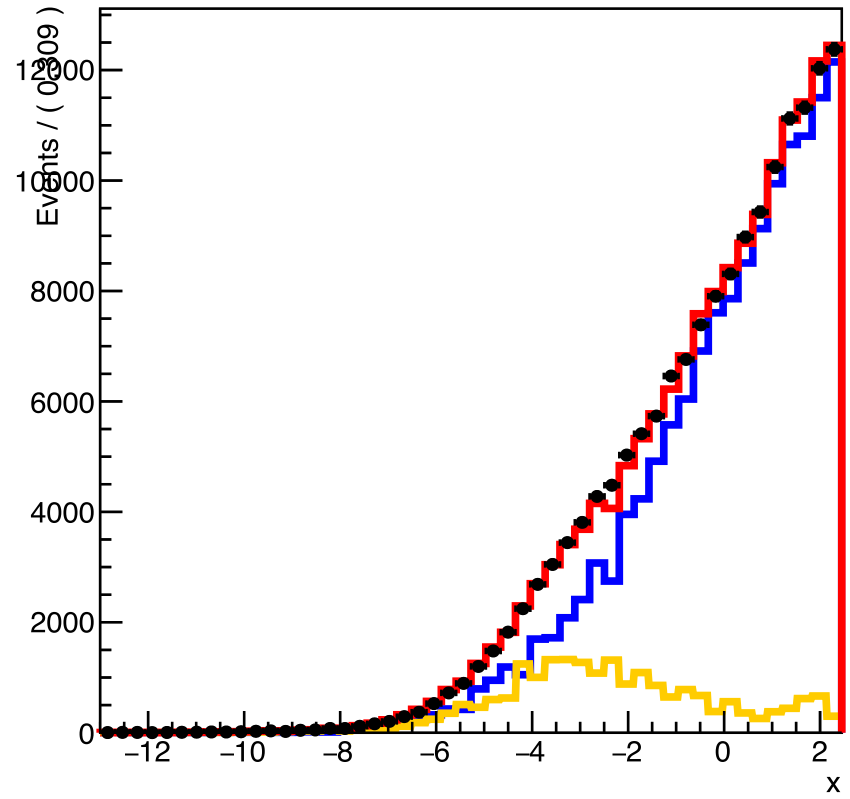
- I’m trying to figure out a) what is the “best” isolation variable to cut on, and b) what the cut should be.

Dynamic Template

- I plot histograms with respect to isolation variables of both data and MC simulations for tight (really "looks" like a photon, based primarily on shower shape, I think) and anti-tight regions (fails the tight cut in certain ways, we also require "LoosePrime4").
- Then I take these four histograms (always titled mcA, mcB, dataA, and dataB), put them into a root file, and run what's called a "dynamic template," created by Bruno Lenzi, on them, which runs a fit and gives a "purity" number which is the percent of photons in the tight data region.

Here's an example of what a template fit might look like.

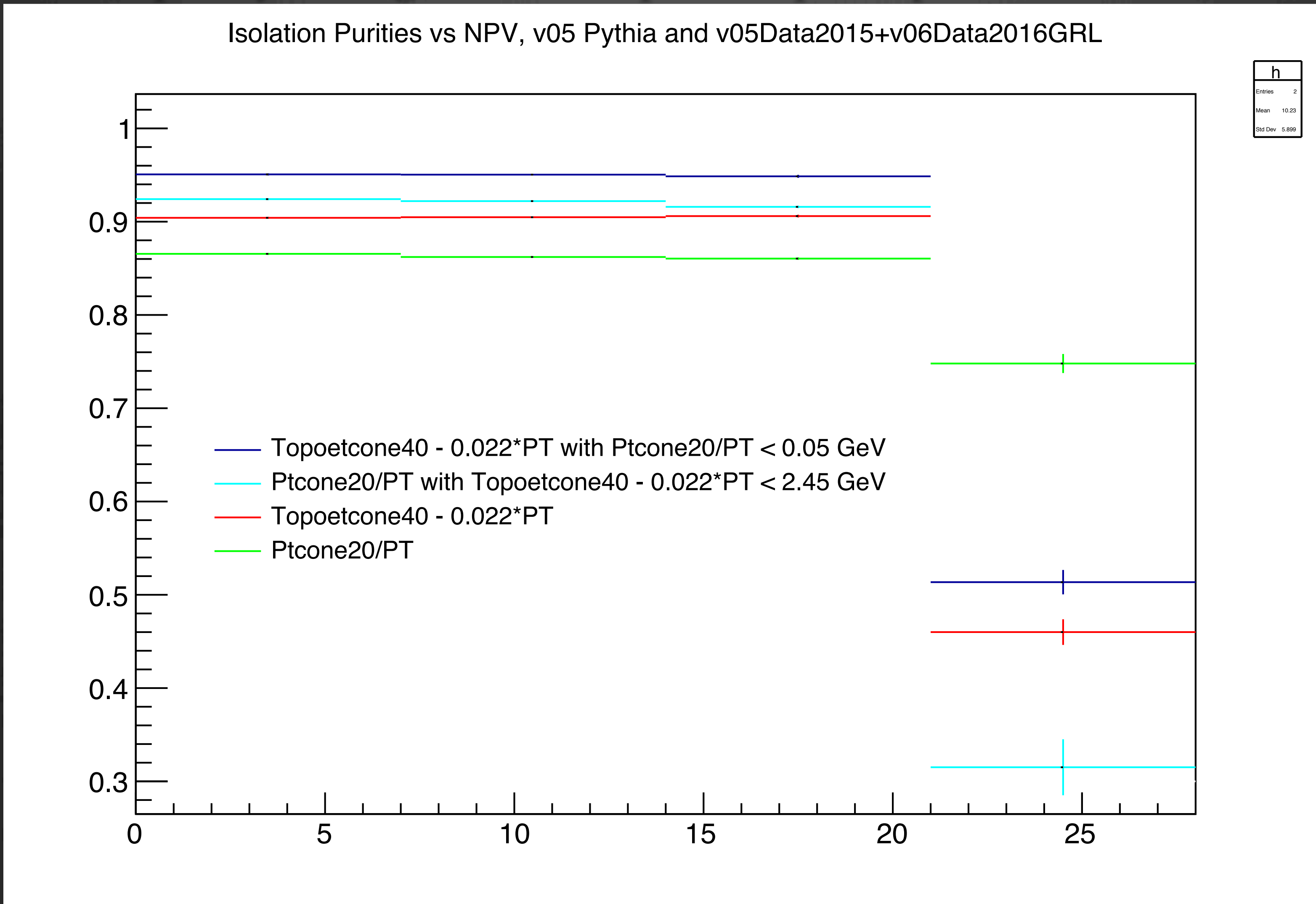




- The range for which I plot the isolation variable acts as a cut on the variable.
- So I can run these dynamic template fits and get purity numbers for a variety of isolation variables, and for a variety of histograms ranges (so a variety of different cuts on the isolation variable) for each variable.
- In addition to Topoetcone40 and Ptcone20 that I mentioned before, there are also Topoetcone20 and Ptcone40.
- The idea is to do this a lot, and see with which variable, and with which cut we get the best purity, i.e. which cut is the best at getting rid of jets and keeping photons.
- Sometimes we also apply a cut on Ptcone20 when we plot with respect to Topoetcone40, and vice versa.

Pile Up

- Lately I've been adding cuts on "Pile up", and plotting how the purity changes with respect to pile up.

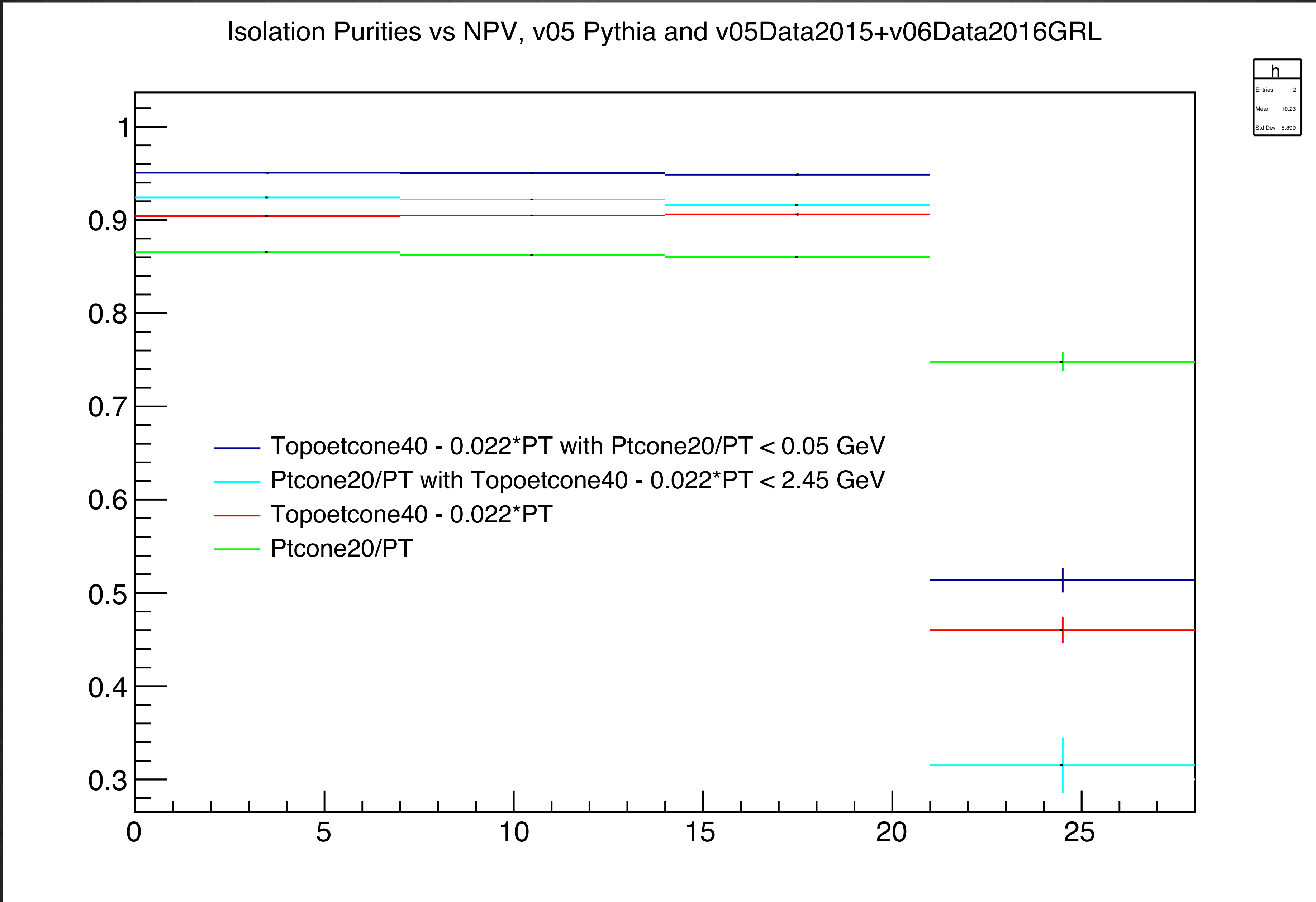


So What is Pile Up?

- My understanding is that pile up refers roughly to the number of events that happen at the same time.
- Protons go around in bunches, so when two bunches collide, the particles the detector detects will typically not all come from the same proton-proton collision.
- Two measures of pileup are “mu,” I don’t know what this means, and “npv,” the number of primary vertices.

- Data and Monte Carlo when I get them are in the form of ntuples.
- I use "Single Photon Ntuples," in which each "photon" is stored as an ntuple, which includes the energy of the photon, its transverse momentum, its Topoetcone40, and MANY other things.
- Among these stored values are mu and npv.
- The npv value for a photon is the number of primary vertices (basically the number of collisions) that happened at the same time as the collision that created the photon.

Here's that plot again



The Plot

- The x-axis is NPV, the y-axis is purity.
- You may notice that instead of plotting just Ptcone20 and Topoetcone40, I plotted Ptcone20/PT (Transverse momentum) and Topoetcone40 – 0.022*PT.
- That's because those are values that have been cut on in the past, and my advisor recommended using them.
- Additionally, the isolation cuts I did were using past cut values, < 2.45 GeV for the Calo isolation (topo…) and < 0.05 GeV for Ptcone20/PT.

Stability

- So what's the point in finding the purities for different NPV values?
- In addition to the goal of getting as high of a purity as possible, we want our isolation cut not to be affected by pileup.
- As can be seen in the plot, the purities are fairly stable, until the fourth region, npv > 21.
- This is a problem I'm working on fixing, but I'm not completely sure it's fixable.

More Problems

- The dark blue and light blue lines in the plot I showed you are applying the exact same cuts, but they have different purities.
- Not knowing how the template works, I didn't think this was that big of a deal at first, I don't know what impact the curve shapes have.
- But everyone seems pretty upset that those lines are different.

The Process

- The way things have been going is I do a lot of template fits, and send my fits and the purities to my advisor, Bruno (creator of the dynamic template), and other higher up people, and someone says "why is this like this?" about something, and we decide
- a) I made a mistake, and I go fix it and come back
- b) there's a problem with the template, and Bruno fixes it.
- c) there's a problem with the ntuples I'm using, and Magda (creator of the ntuples) fixes them.

The Plan

- I also plan to make similar plots as I did with pile up with different cuts on pseudorapidity (eta) and for both converted and unconverted photons.
- Converted photons are detected as an electron-positron pair, unconverted photons are detected as photons.
- Then I will start messing around with Topoetcone20 and Ptcone40 as well, and try a lot of different cut values.
- And we'll see how high I can get my purity values to be, and how stable with respect to changes in other variables they will be.