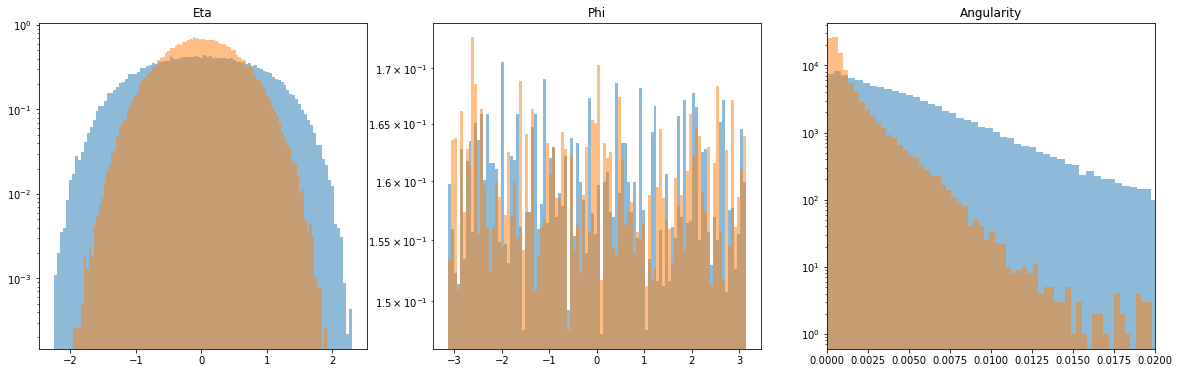
Data taken at the LHC in Geneva collects and calculates all sorts of metadata along with the actual observable that scientists are looking for. In the search for the Higgs Boson, this metadata is very useful in making the measurement and statistical sensitivity greater specifically for the Higgs. The dataset given to us for our Lab the next 5 weeks includes all of this metadata for which I will come up with a way to make certain selections that increase the sensitivity of our data. This metadata includes values of “sub-jetiness”, sub-jetiness ratios, jet size, calculated mass, transverse momentum, geometrical attributes, and more.

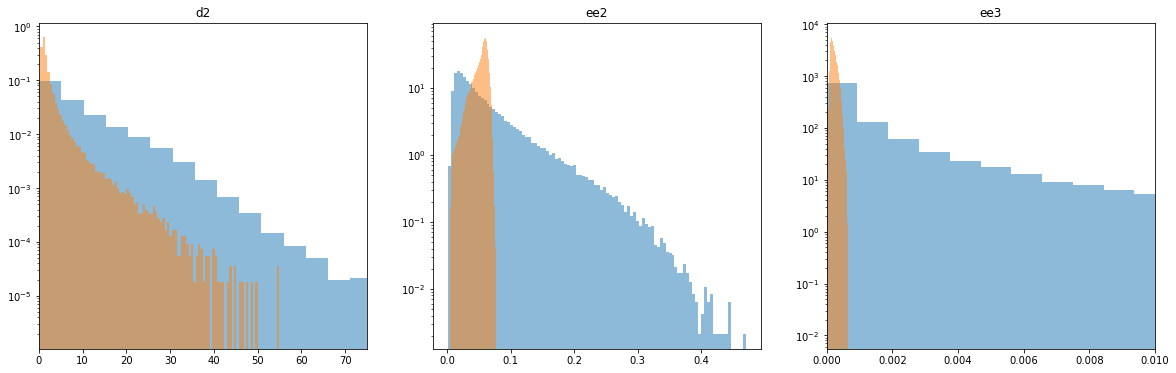
Here, I will include some useful plots that will help me make my decisions on what data to keep and remove. \*\*\* histograms shown are of background (blue) and simulated signal (orange) \*\*\*



The above plot shows me that there is discriminatory features in Eta and Angularity, but nothing discernable in Phi. This will allow for cuts to be made in Eta and Angularity. Eta appears to be Gaussian and Angularity looks to be Rayleigh or exponential (further research must be done). No easy cuts for Phi or Angularity.

Kept after cuts so far:

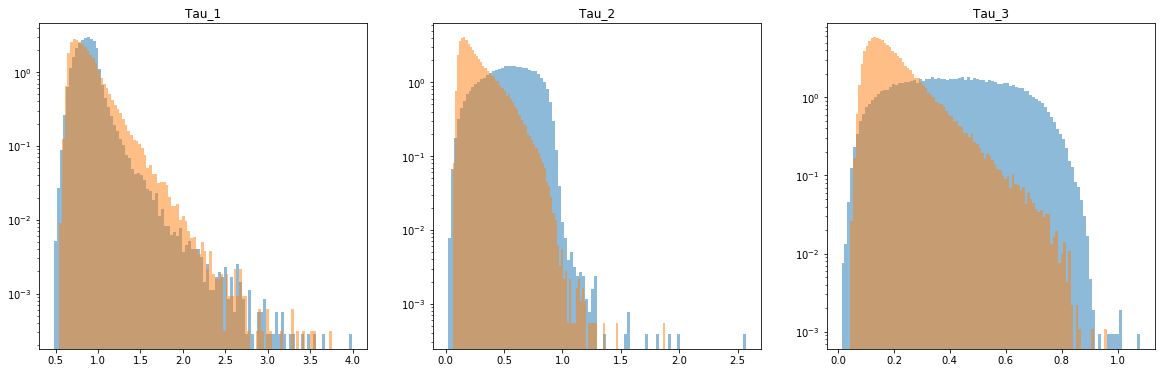
* Eta cuts from -2 < eta < 1.9



These plots describe a noticeable difference between the background and the signal. Here 'd2' shows a peak at a different value than the background, also that peak is much tighter and therefore more concentrated. 'ee2' has completely different distributions for the two plots. The distribution for the signal appears to be an "extreme value Type I" distribution and the background appears Poissonian or a gamma distribution (more research). 'ee3' has a similar situation to 'd2' in which there is a definite peak of greater concentration that will allow for easy cuts.

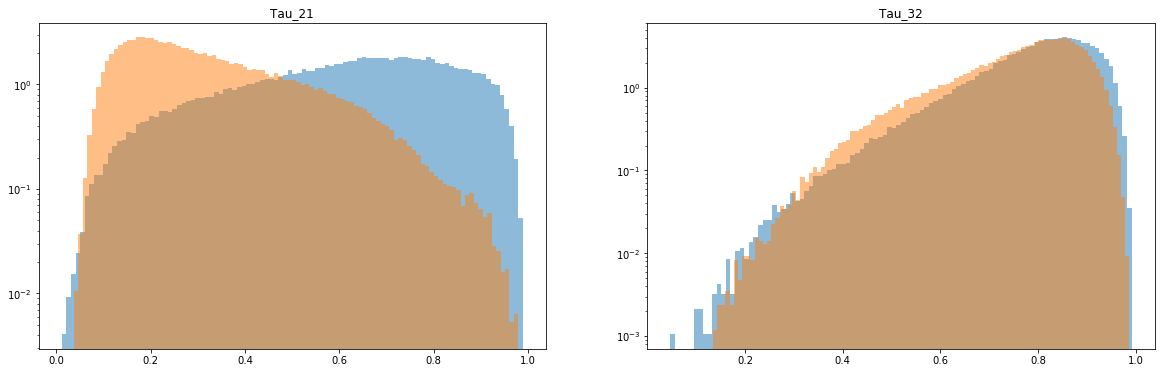
Kept after cuts so far:

* Eta: -2 < x < 1.9
* ‘d2’: x < 50
* ‘ee2’: x < ~.08
* ‘ee3’: x < ~.001



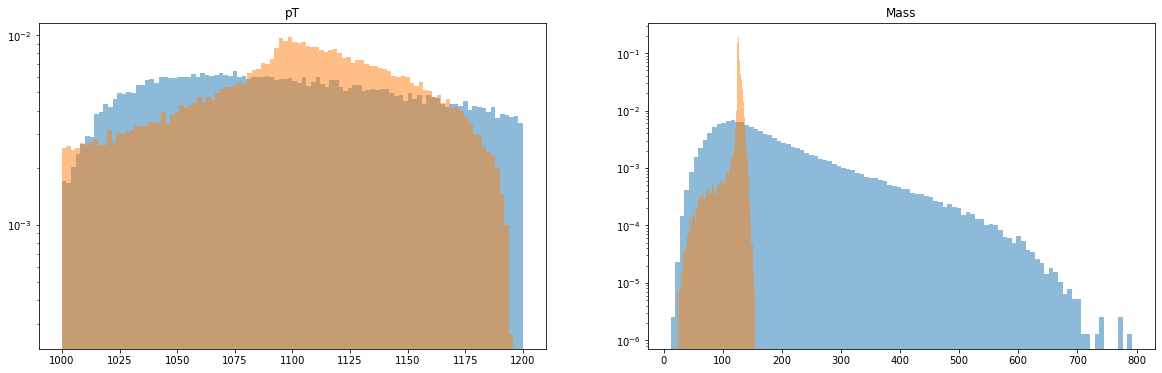
Tau\_1 appears to have relatively the same distribution for the background and signal, but Tau\_2 and Ta\_3 have very different distributions. The signal distributions seem to be of the same shape, which looks like a chi or chi^2 distribution (more research needed). Not sure what the background distributions are as they appear to plateau on the log and normal scales. Possible ways to filter out data would be to change the sensitivity for certain values of sub-jettiness.

NO NEW CUTS.



Tau\_32 looks extremely similar in both the background and signal datasets, probably not great for filtering. But, Tau\_21 has extreme differences in the peak value. This could be used to change sensitivity of the sets. The distributions look like "Extreme Value Type I" distributions. These distributions do not provide easy cutting parameters, but other manipulations may prove to come in handy for the next couple of labs.

NO NEW CUTS.

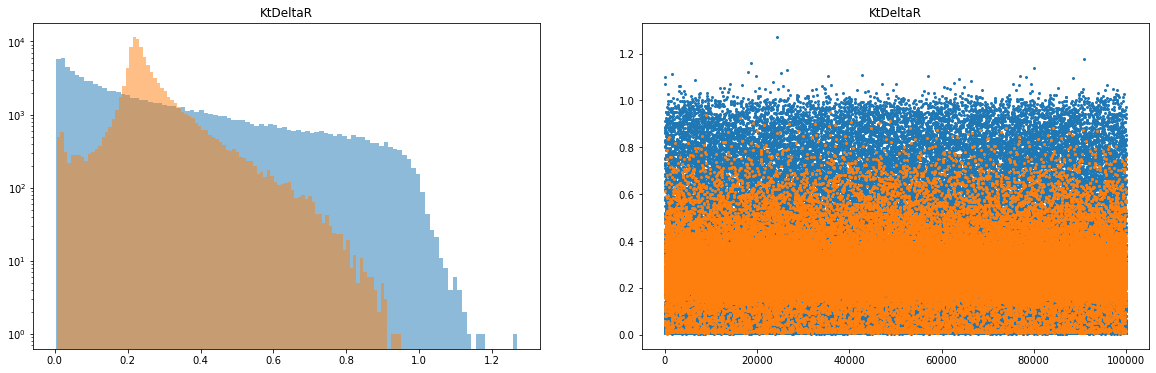


The Mass of the signal should follow a distribution composed of both a Landau and a Gaussian.

These are the two most interesting plots to me. The momentum measurement distribution looks like a combination of distributions overlayed each other. The curvature of the pT distribution is like something I have never seen before. Also, after some reading, I found that the mass of the signal should appear as the Landau and Gaussian Distributions combined. Mass cuts are going to be the most crucial part of finding the higgs.

Kept after cuts so far:

* Eta: -2 < x < 1.9
* ‘d2’: x < ~50
* ‘ee2’: x < ~.08
* ‘ee3’: x < ~.001
* Mass: x < ~160

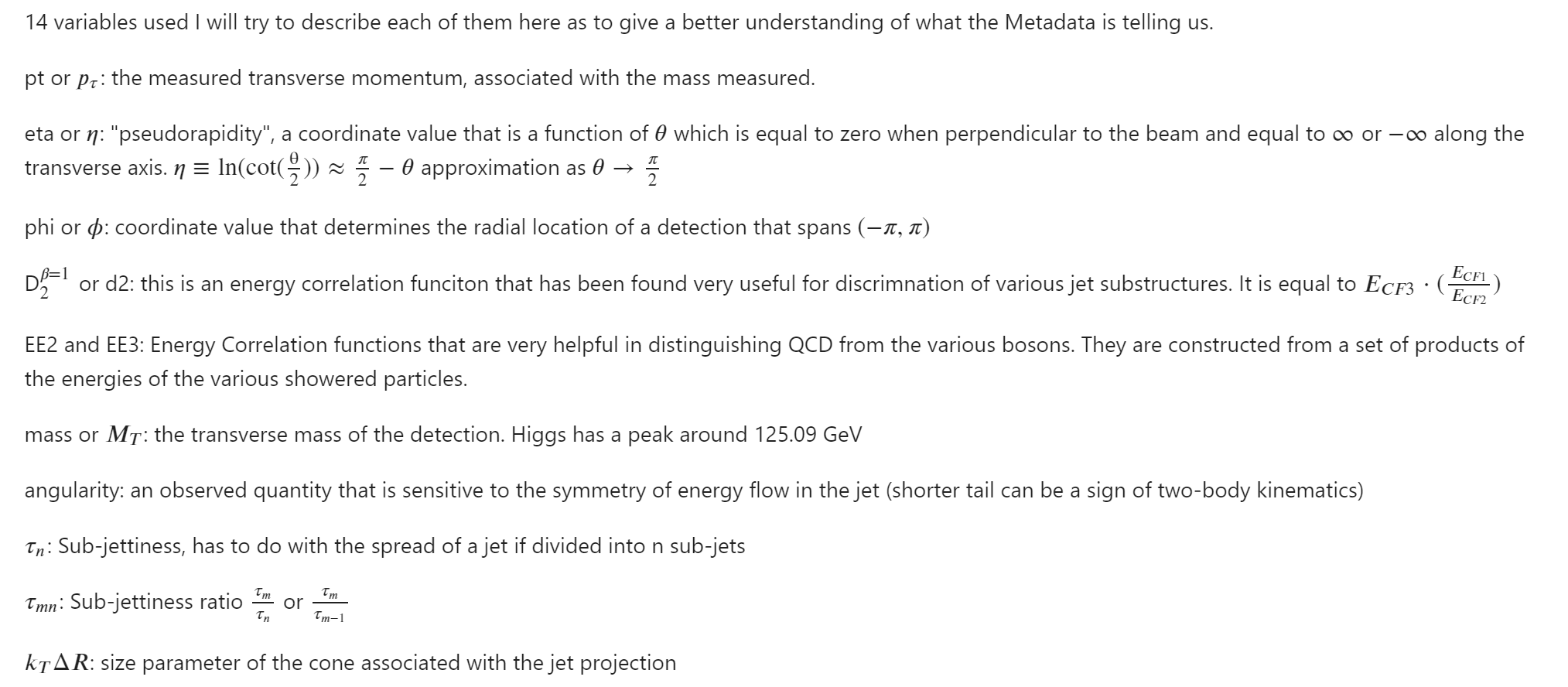


This peak is very curious. I can see that some cuts from this data can be taken out but I don't think much of a change will come from that. This also looks like it could be from multiple different distributions and further investigating must be done.

Kept after cuts so far:

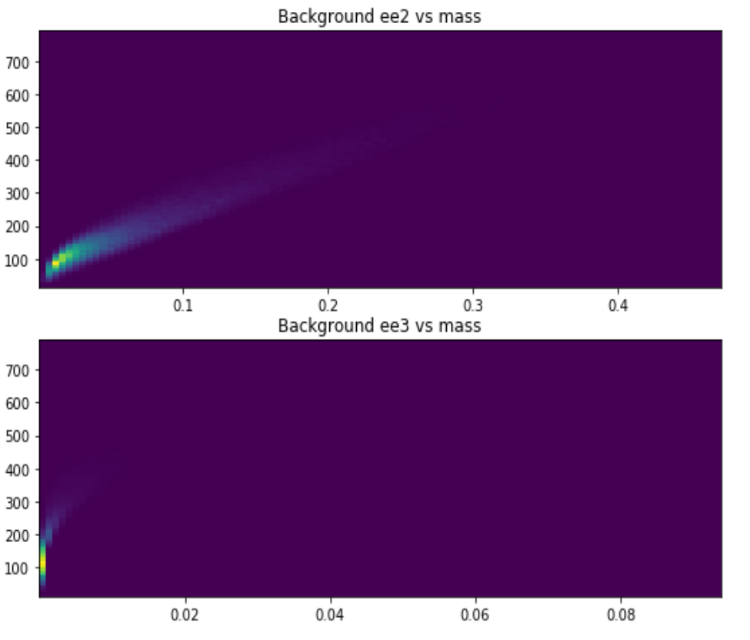
* Eta: -2 < x < 1.9
* ‘d2’: x < ~50
* ‘ee2’: x < ~.08
* ‘ee3’: x < ~.001
* Mass: x < ~160
* R: x < ~.9

Here I describe most of the 14 variable of metadata in attempts to understand their distributions and the physics behind what they mean.

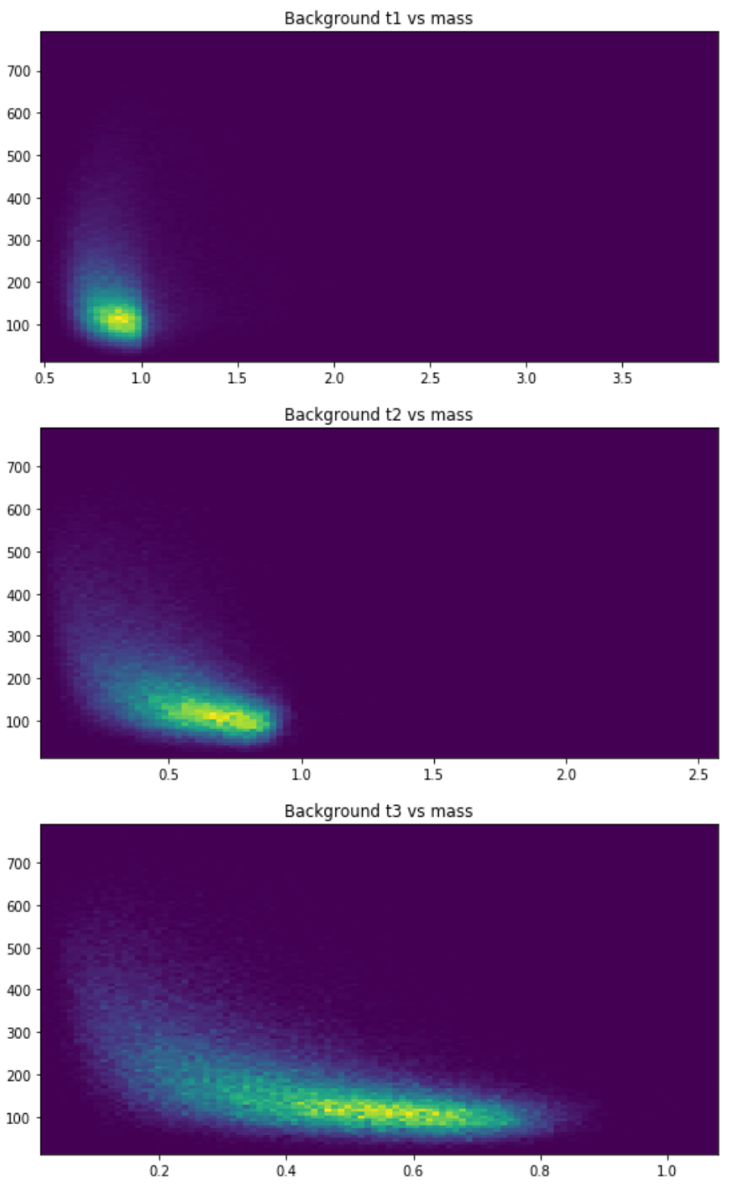


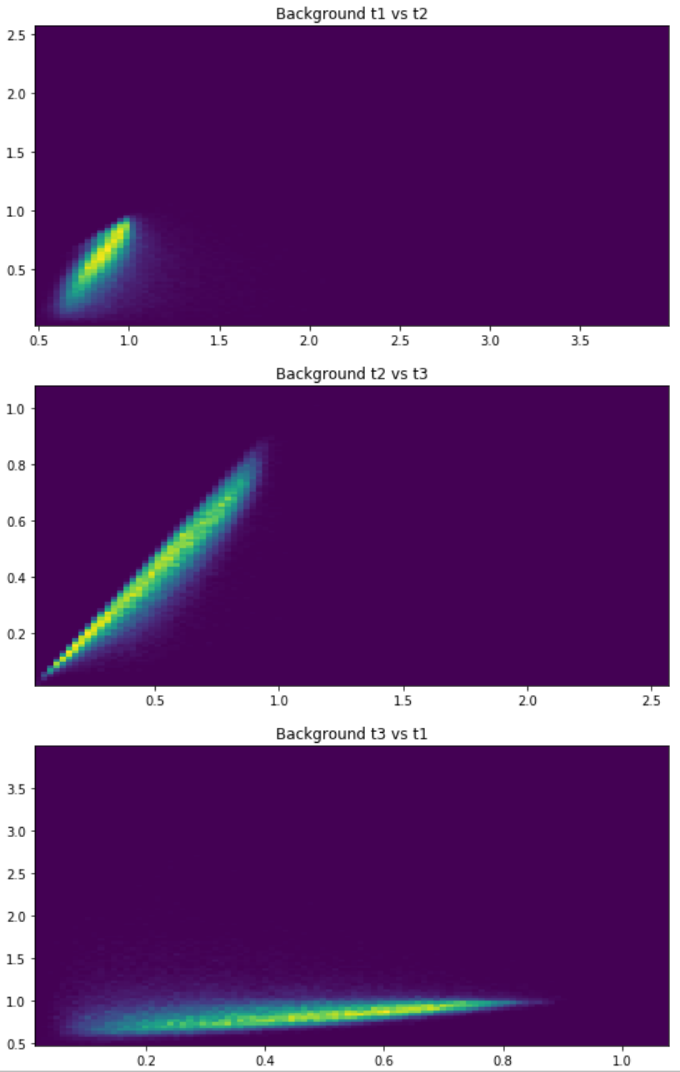
1. Above I described the shapes and peaks of both the background and signal data in attempts to find easy comparisons that we can filter out of our background. Most of the distributions were noticeably different, some were very similar, and ones like mass, ee2, and ee3 provided easily executed cuts from the background in order to reduce the background signal.

Most of the correlation checking I did was between the different parameters and the corresponding mass of the detection. These plots showed that there were some correlations between different values and the mass of the detection. This leads to the opportunity to find correlations between multiple metadata variable to further increase the sensitivity.



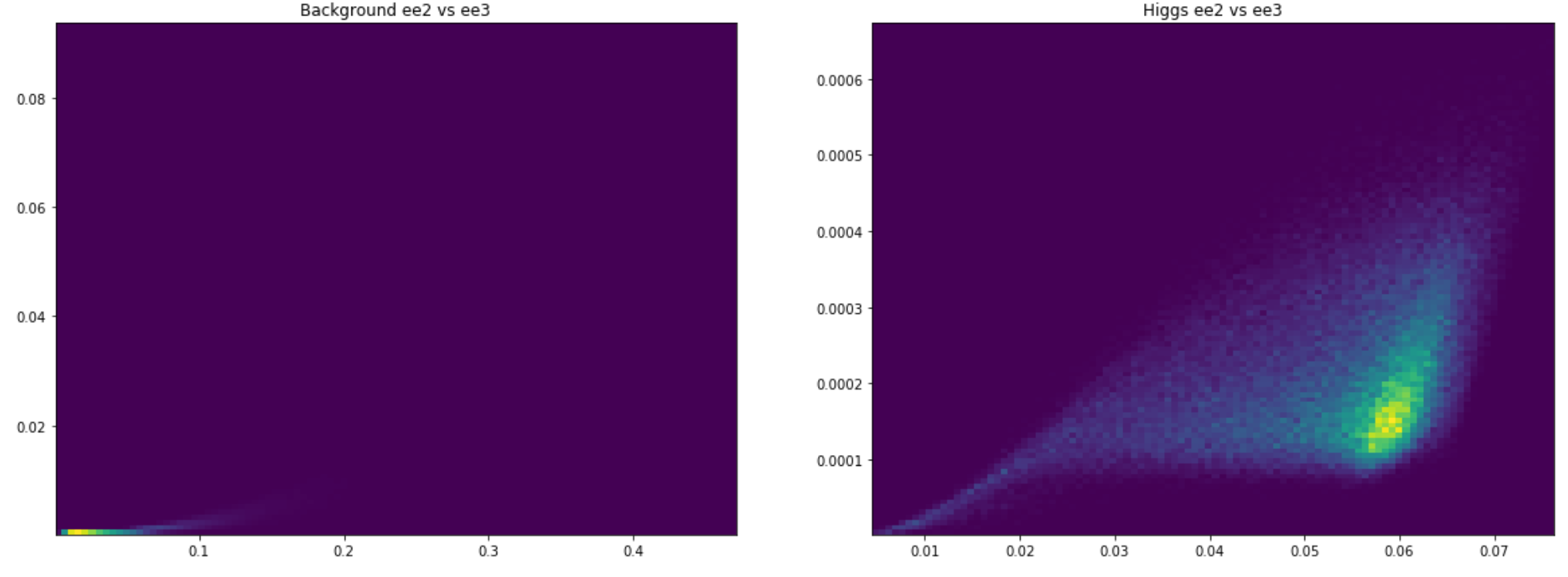
(Left) you can see that the mass increases with the value of ee2. It follows a similar trend with ee3 but at a much faster rate.

(Right) These plots show that for each level of sub-jetiness, there is a similar distribution shape when viewed against mass. With each value of sub-jetiness the distribution reaches to a smaller and smaller value but the shape is the same.

This shows that the sub-jetiness levels are correlated with one another. The most striking correlation is that between t2 and t3 which trends at or just below a 1:1 ratio.

A similar trend follow between t1 and t2 although ratio value larger than 1 are reached in this case.

The correlation between t1 and t3 is harder to distinguish although it does also appear to have a linear trend.



^^^ Interesting plot of how ee2 and ee3 are correlated. This asymmetric funneled cloud brings up a lot of questions.

The biggest change that I realized when looking at the correlations of the background compared to that of the actual signal is that the background has far wider spreads and appear much more like a cloud of data. The signal values have pretty strong convergence zones which are nice because that allows for more opportunity to eliminate data that doesn’t match from our background.