Weekly Report

Shudong WANG

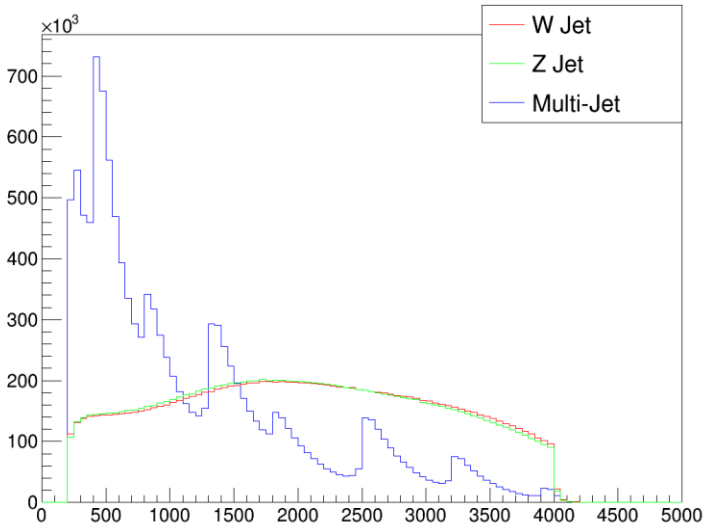
Brief Report

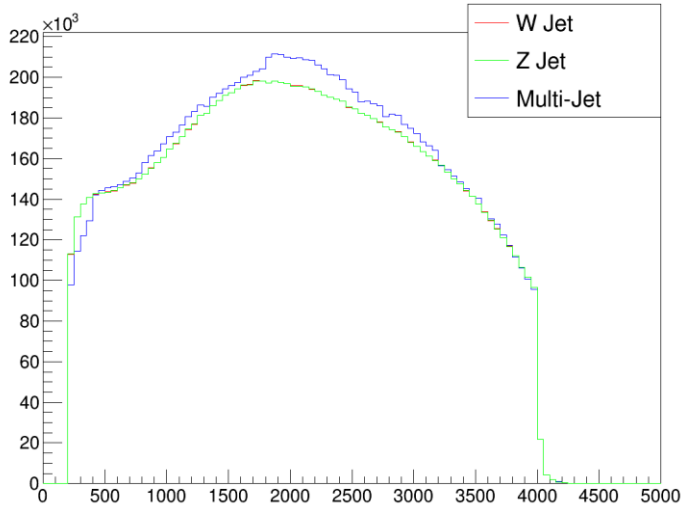
- Started to work on my QT, it’s about analyzing test parameters of ITk components, will be shown on the webpage I developed.
- Still generating CEPC ttbar samples, it has took way more time than I expected.
- Tried Z tagging and W vs Z classification using ParticleNet.
- Tried W tagging using ParticleTransformer.

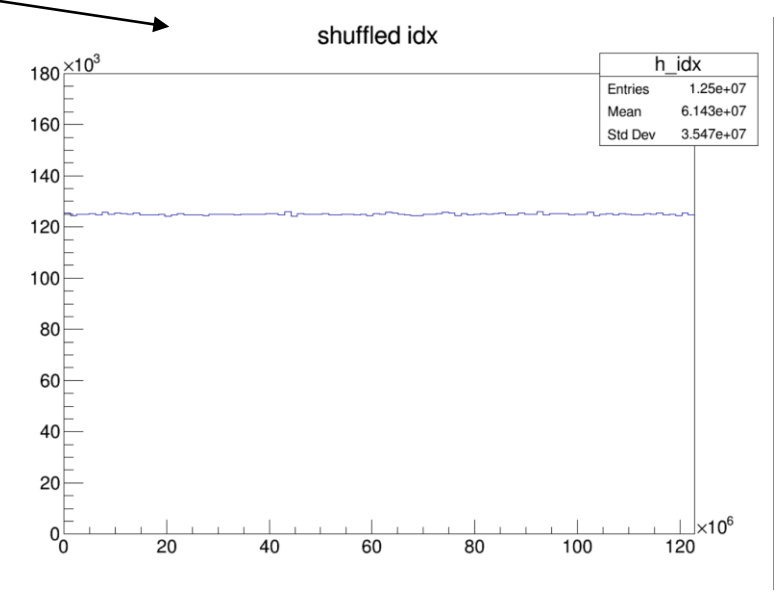
Constituent based W/Z boson tagging

- Problem found

It is a NOT Result of biased sampling ?





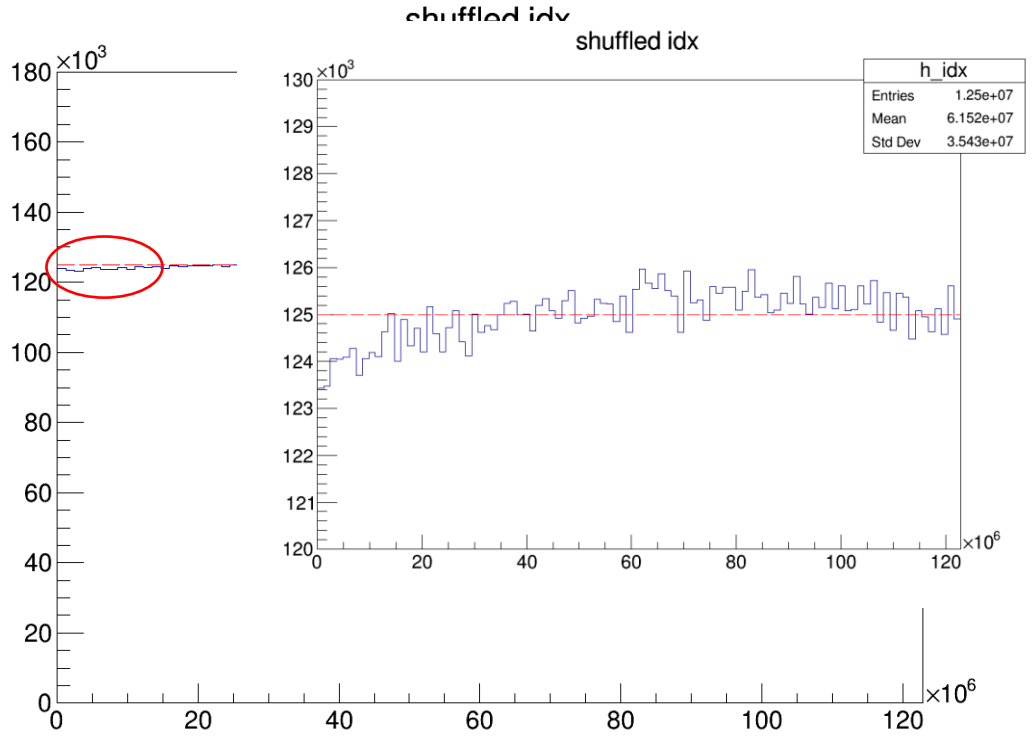


- Weights gained by using all sig(13M+) & bkg (1.2B+) events, then do the sampling to get 1:1 sig/bkg ratio (12.5M: 12.5M) and apply weights*sf as training weights.
- Won't affect the result much, but need to be improved.

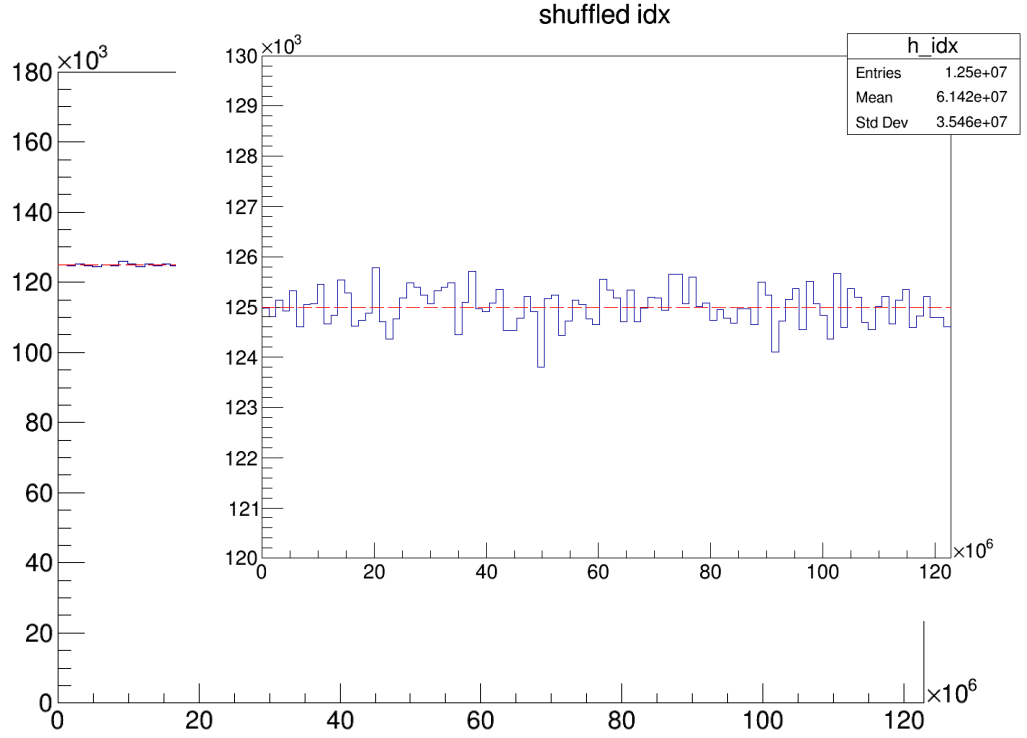
Constituent based W/Z boson tagging

- Use an old sample which is not uniformly sampled

It is a Result of biased sampling !



shuffled only once



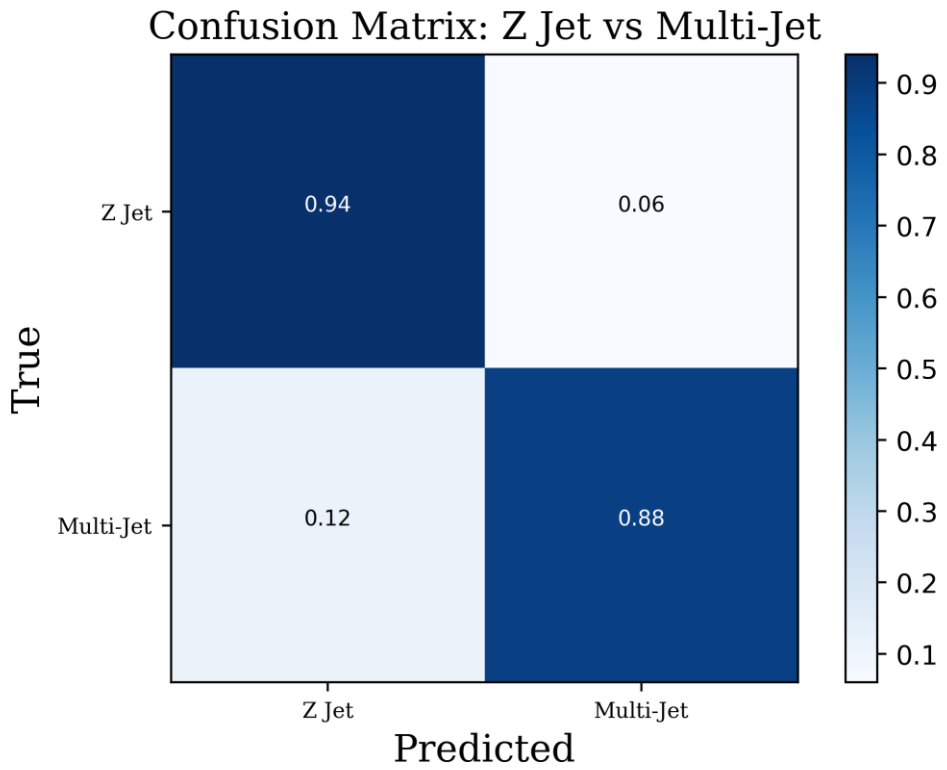
shuffled 4 times

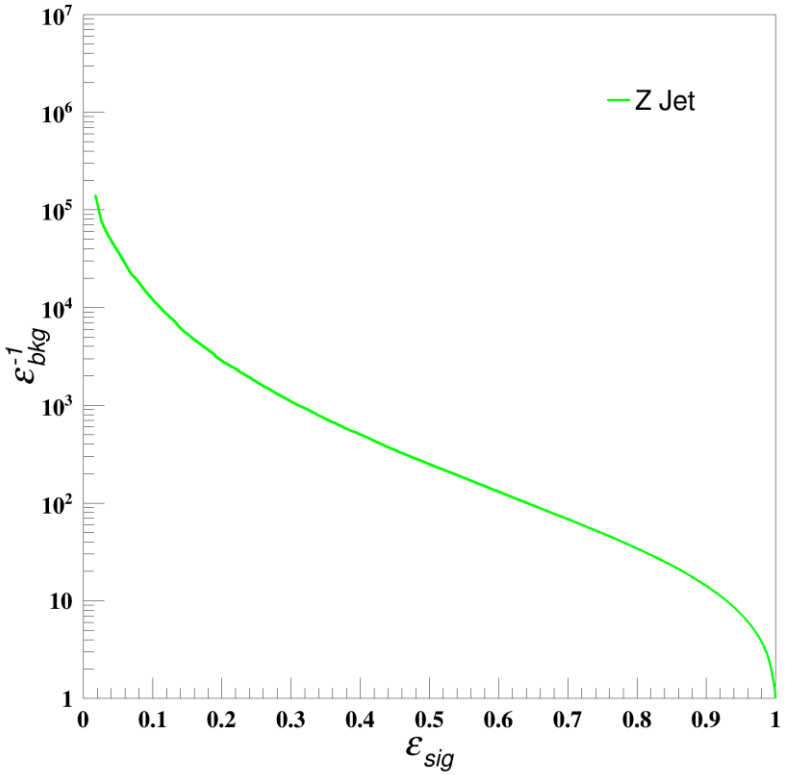
Constituent based W/Z boson tagging

- Overfitting problem I mentioned last week.
- Just a stupid mistake.
- To uniformly sample data, you need to randomly access entries.
- Random access of .root file is extremely slow, most of the time is spent on decompressing and loading baskets.
- To overcome this, with Shuiting's suggestion, I shuffled all the indexes first and take the firsts 12.5M shuffled indexes, then sort these 12.5M indexes and take the entries correspond to these indexes.
- I didn't shuffle these 12.5M entries and the order of indexes is related with the pt of QCD jet samples.
- So when I split the sampled dataset into train and validation sets, their pt ranges are different, no wonder I met overfitting problem!

Z tagging with ParticleNet

- Z tagging

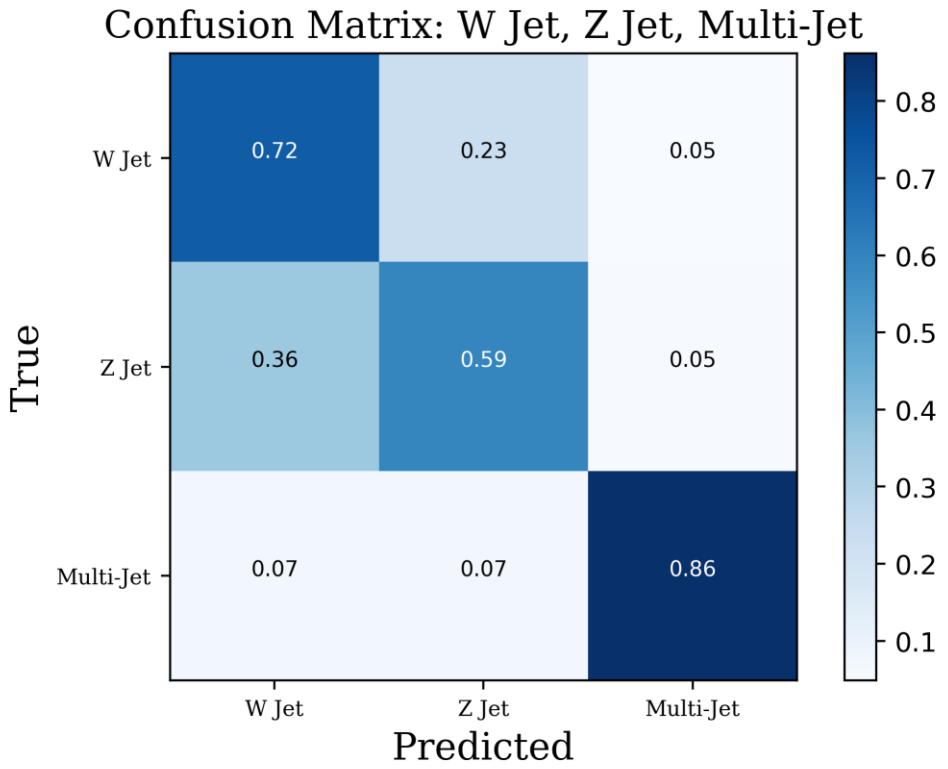


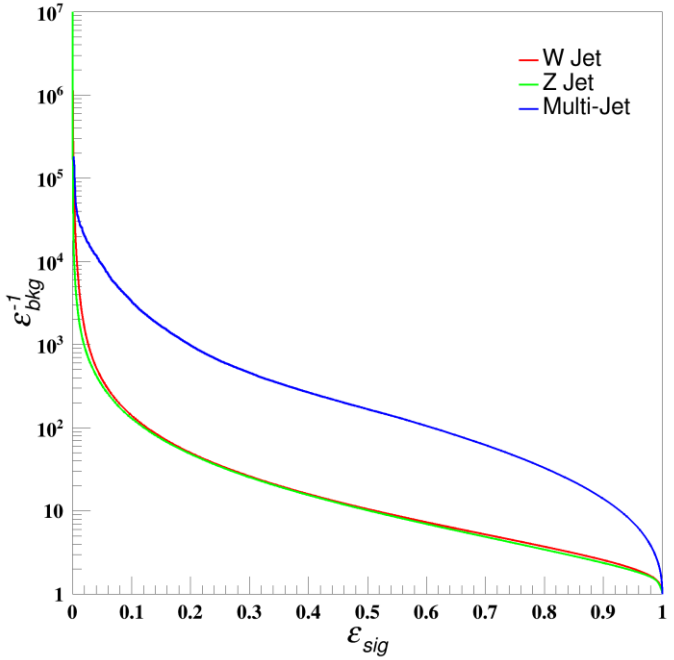


ΔR

W, Z and Multi-Jet 3-category classification

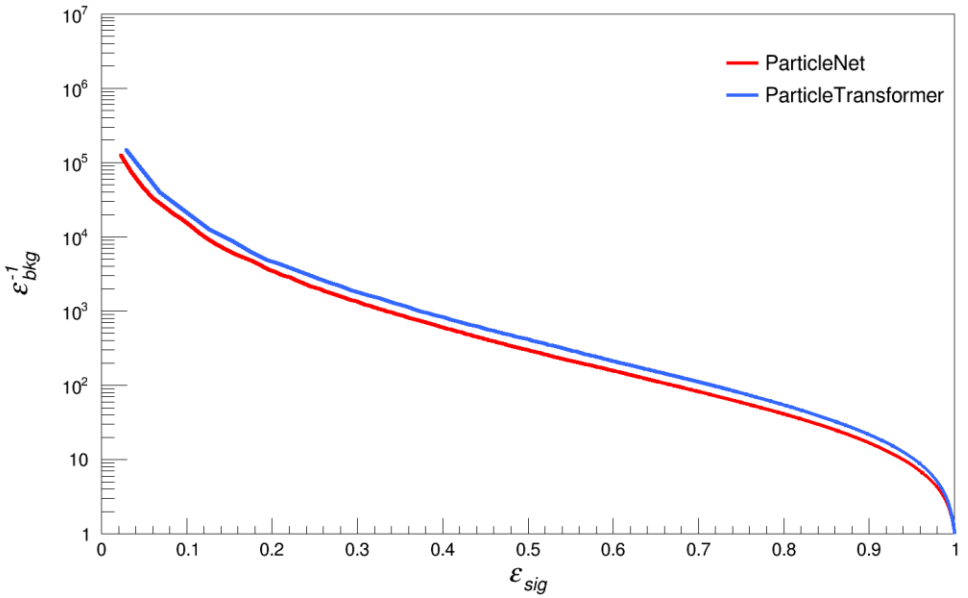
- 3-category





W tagging with ParticleTransformer

- W tagging with ParticleTransformer



ParT

AUC: 0.980, ACC: 0.928

$1 / \varepsilon_{bkg}$ @ $\varepsilon_{sig} = 50\%$: 415

$1 / \varepsilon_{bkg}$ @ $\varepsilon_{sig} = 80\%$: 54.6

ParNet

AUC: 0.975, ACC: 0.918

$1 / \varepsilon_{bkg}$ @ $\varepsilon_{sig} = 50\%$: 299

$1 / \varepsilon_{bkg}$ @ $\varepsilon_{sig} = 80\%$: 41