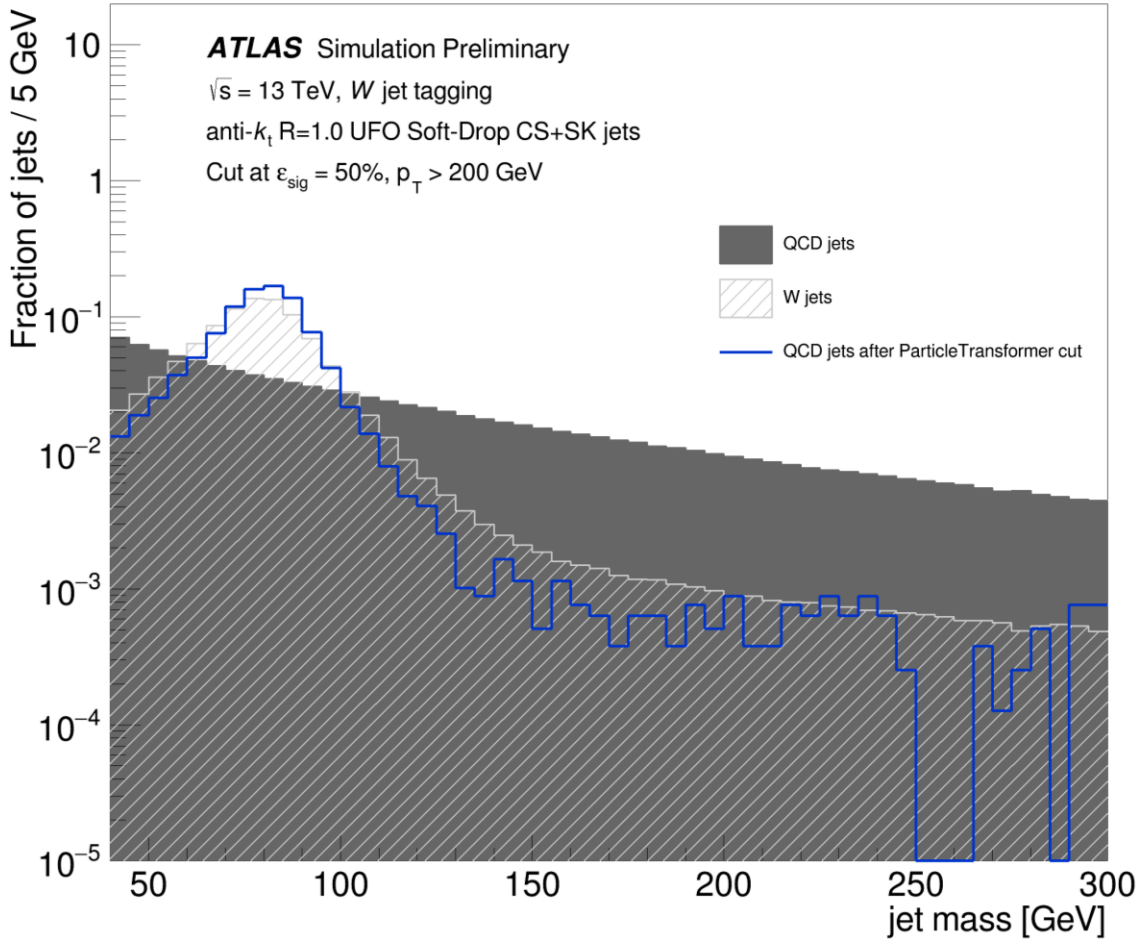
Weekly Report

Shudong WANG

Progresses

- **W tagging - Mass deccorelation**
- **Mass scuplting**



Progresses

- **W tagging - Mass deccorelation**
- **Distance correlation (DisCo) for decorrelation**

Distance Correlation

$$x_{jk} = |X_j - X_k|$$ **Distances of all examples in batch for classifier output**

$$y_{jk} = |Y_j - Y_k|$$ **... for variable to decorrelate**

$$\hat{x}_{jk} = x_{jk} - \overline{x}_{j\cdot} - \overline{x}_{\cdot k} + \overline{x}_{\cdot\cdot}$$

$$\hat{y}_{jk} = y_{jk} - \overline{y}_{j\cdot} - \overline{y}_{\cdot k} + \overline{y}_{\cdot\cdot}$$ **Center distributions**

$$\mathrm{dCov}^2 = \frac{1}{n}\sum_j \sum_k \hat{x}_{jk}\hat{y}_{jk}$$ **And calculate average product per batch**

Some nice properties:
- Zero iff X,Y are independent; positive otherwise!
- Computationally tractable!
- Doesn't require binning!

$$\mathrm{dCorr}^2(X,Y) = \frac{\mathrm{dCov}^2(X,Y)}{\mathrm{dCov}(X,X)\mathrm{dCov}(Y,Y)}$$

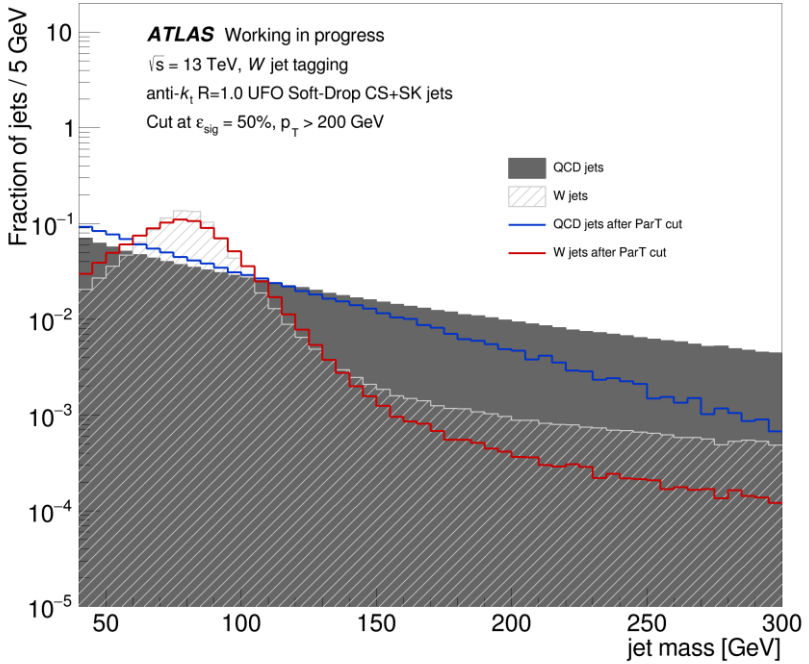
- How to use this for decorrelation?
- add dCorr² to the loss function

$$L = L_{classifier}(\vec{y}, \vec{y}_{true}) + \lambda\, \mathrm{dCorr}^2_{y_{true}=0}(\vec{m}, \vec{y})$$

- λ : a hyperparameter that controls the tradeoff between performance and decorrelation
- $\vec{y}$: output of the NN on a single minibatch
- $\vec{y}_{true}$: true labels
- $\vec{m}$: masses
- y_{true}= 0 : the distance correlation is only calculated for the subset of the minibatch that is background

Progresses

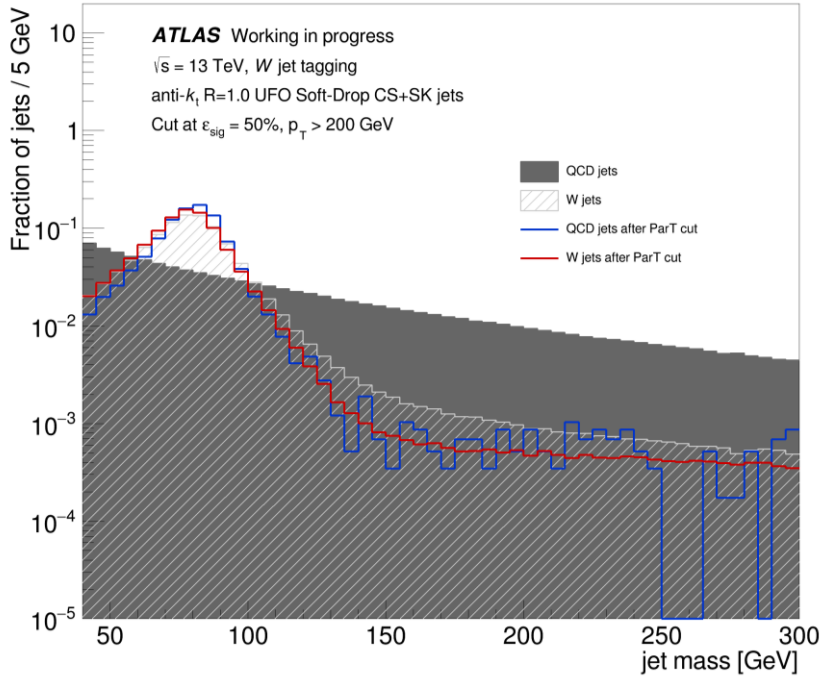
- **W tagging - Mass deccorelation**
- **Distance correlation (DisCo) for decorrelation**
- ParticleTransformer as baseline (trained for only 1 epoch to scan λ)
  - $\lambda = 300$



ε_{bkg}^{-1} =13.67

- seems working… but not very well
- needs to train for more epoch, very time consuming

- original



ε_{bkg}^{-1} =415

Progresses

- **VBF Validation**
- Samples for all ML, yyML and bbZZ_4l channels with 6 couplings generated.
- Validation plots have been drawn.

- **ITk**
- Get rid of maintenance & development of ITk Production Flow Page
- Module wire-bonding result histograms for barrel module PRR, updated