Significant improvement of Quark/Gluon separation with the ATLAS detector at the LHC

(LHC-ATLAS実験におけるクォーク・グルーオンの識別の向上)

LU CHENG YU (ルー)

Univ. of Tokyo, ICEPP

February 21, 2022

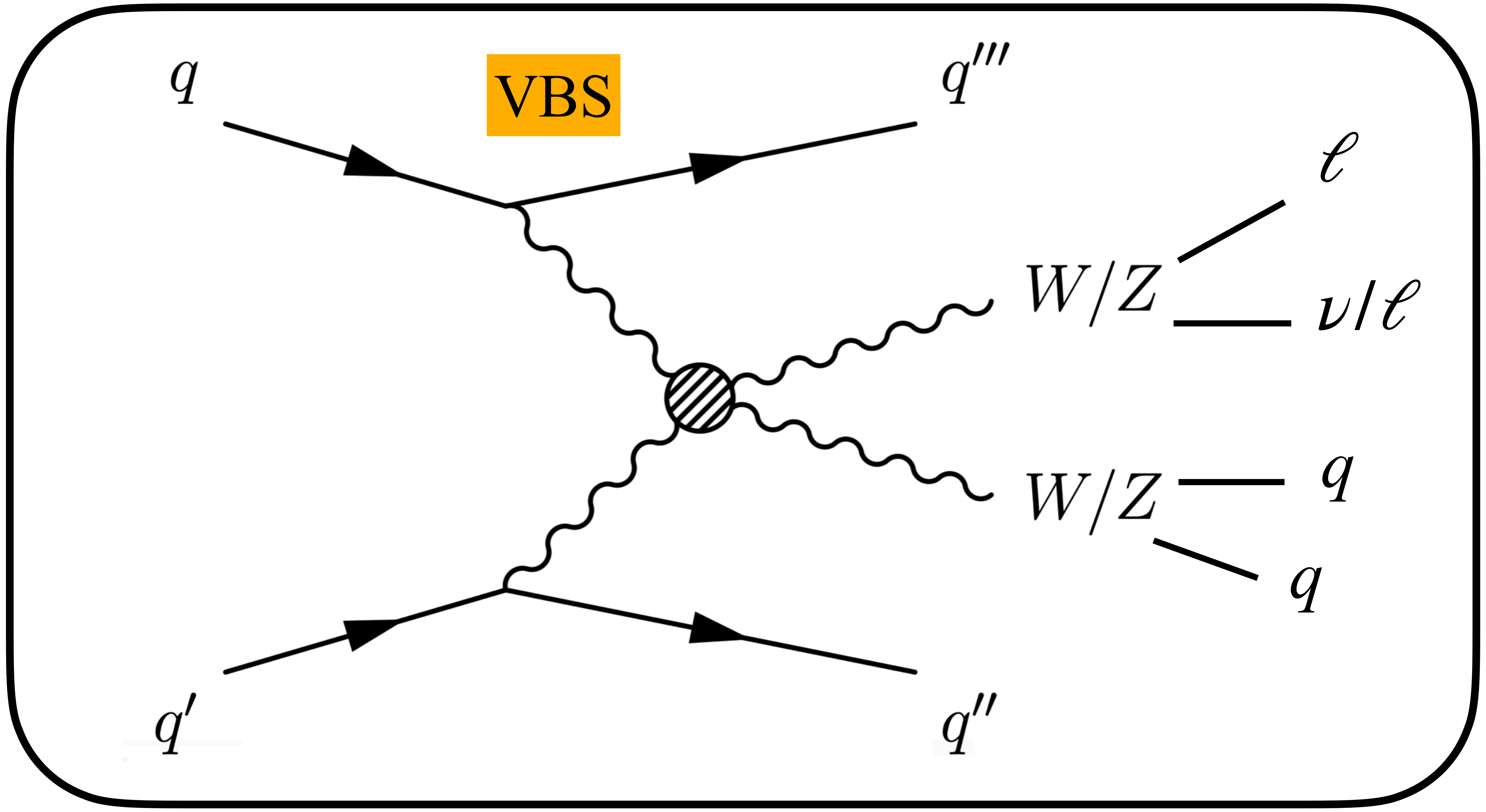
東京大学
素粒子物理国際研究センター
International Center for Elementary Particle Physics
The University of Tokyo

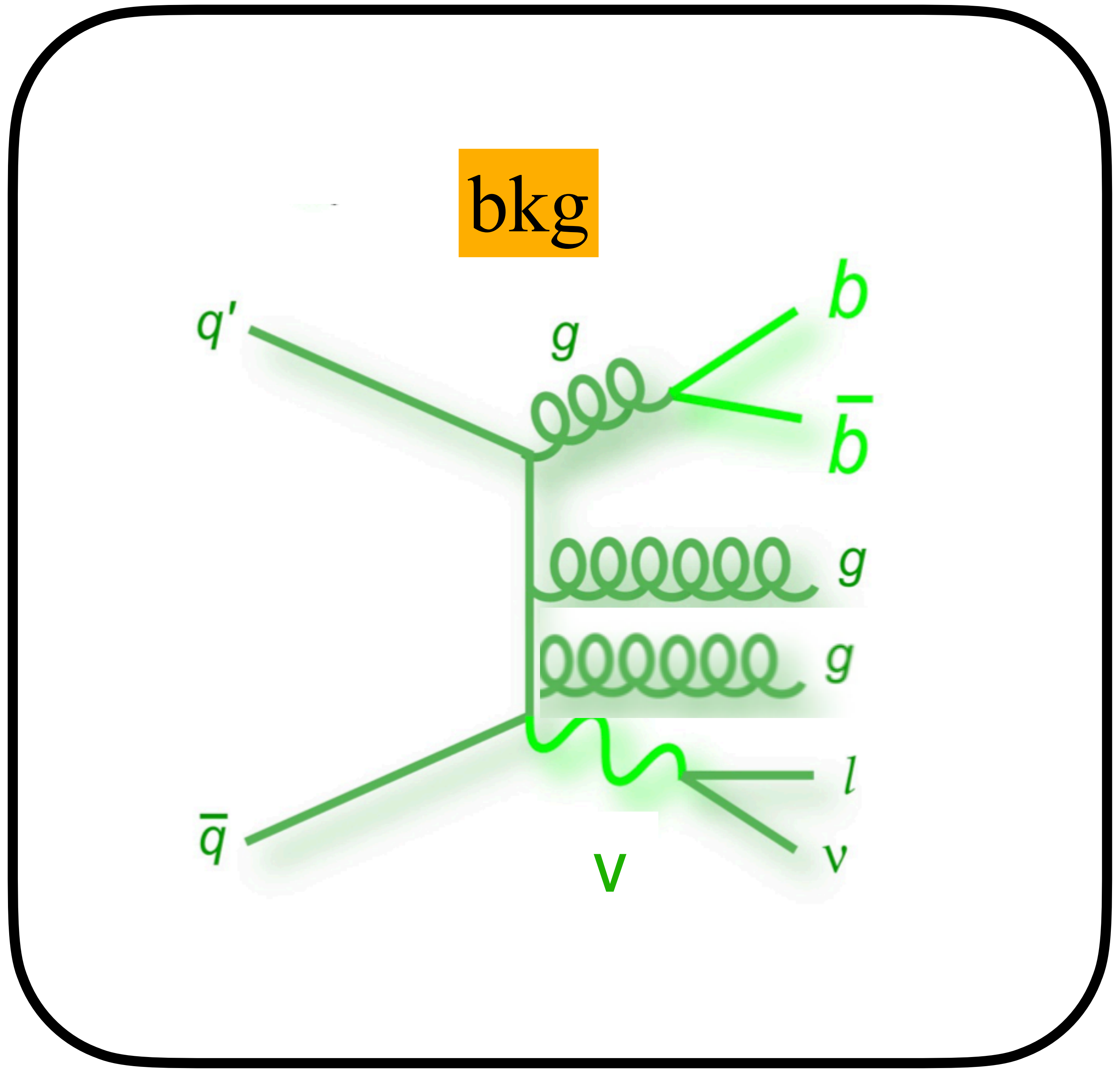
ATLAS EXPERIMENT

Outline

1. Motivation
2. q/g Tagging
    1. Models and inputs
    2. ROC Performance
3. Application to 1-lepton semileptonic VBS
4. Conclusion

Motivation





- Vector Boson Scattering (VBS) is a sensitive probe to examine the electroweak symmetry breaking (EWSM) in SM and BSM physics.
- The semileptonic VBS signal events are characterized by four quark jets, while background events have 2-3 gluon jets.
- It is essential to develop a tool to separate the quarks and gluons, known as Quark/Gluon tagging (q/g tagging).

Quark/Gluon

- Compared to quark, gluon has a larger color factor, producing more particles in the detector.
- However, actually it is very harsh to separate the quark and gluon because of their similarity, especially in the low p_T region.

ATL-PHYS-SLIDE-2017-781

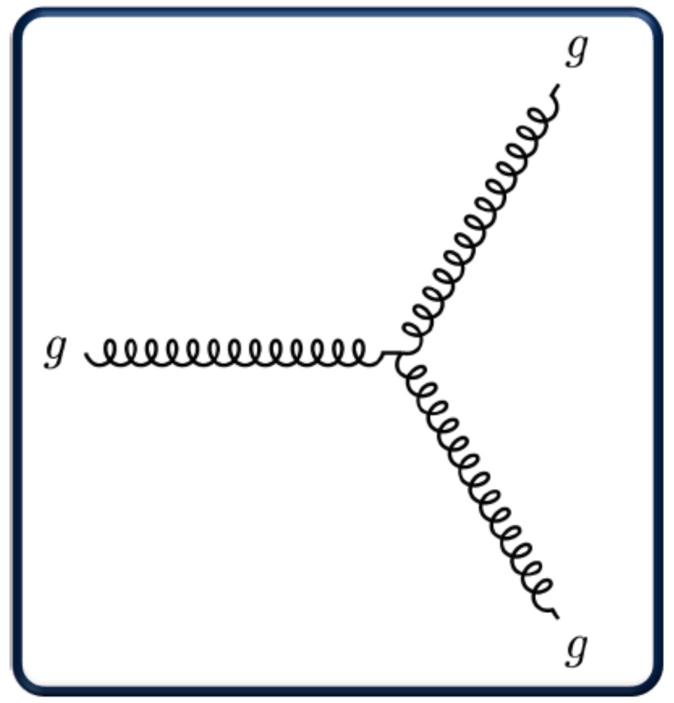
Gluon

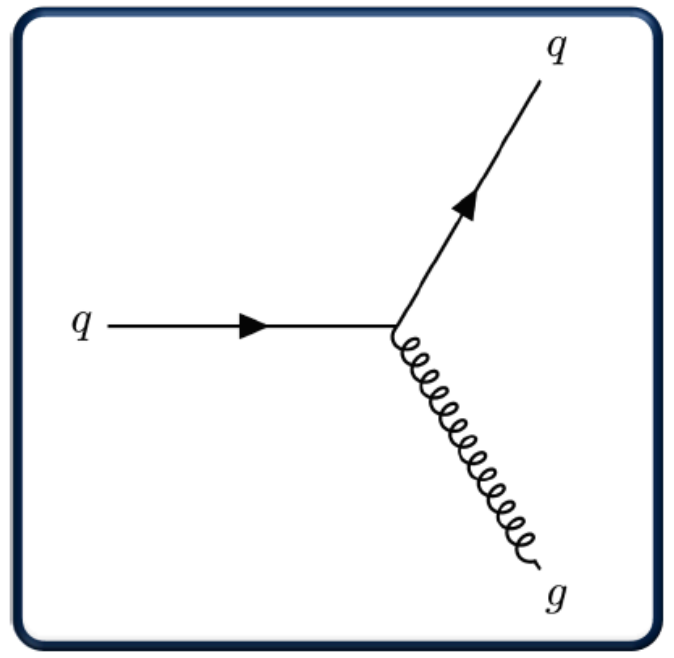
$$C_A \equiv N_C = 3$$

Quark

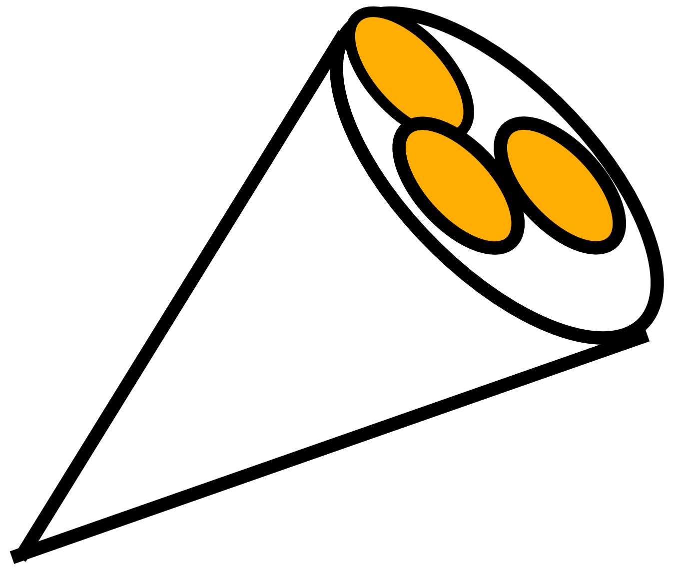
$$C_F \equiv \frac{N_C^2 - 1}{2N_C} = \frac{4}{3}$$

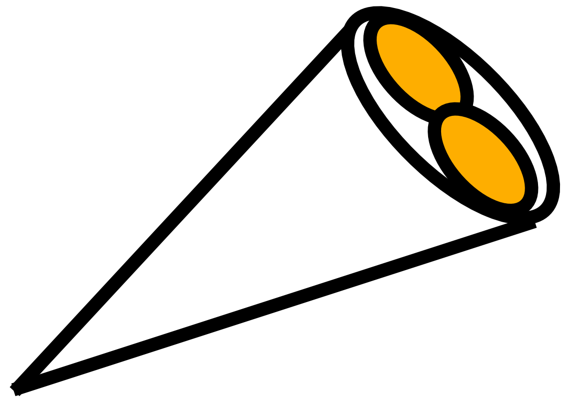
$$\frac{C_A}{C_F} = \frac{9}{4}$$



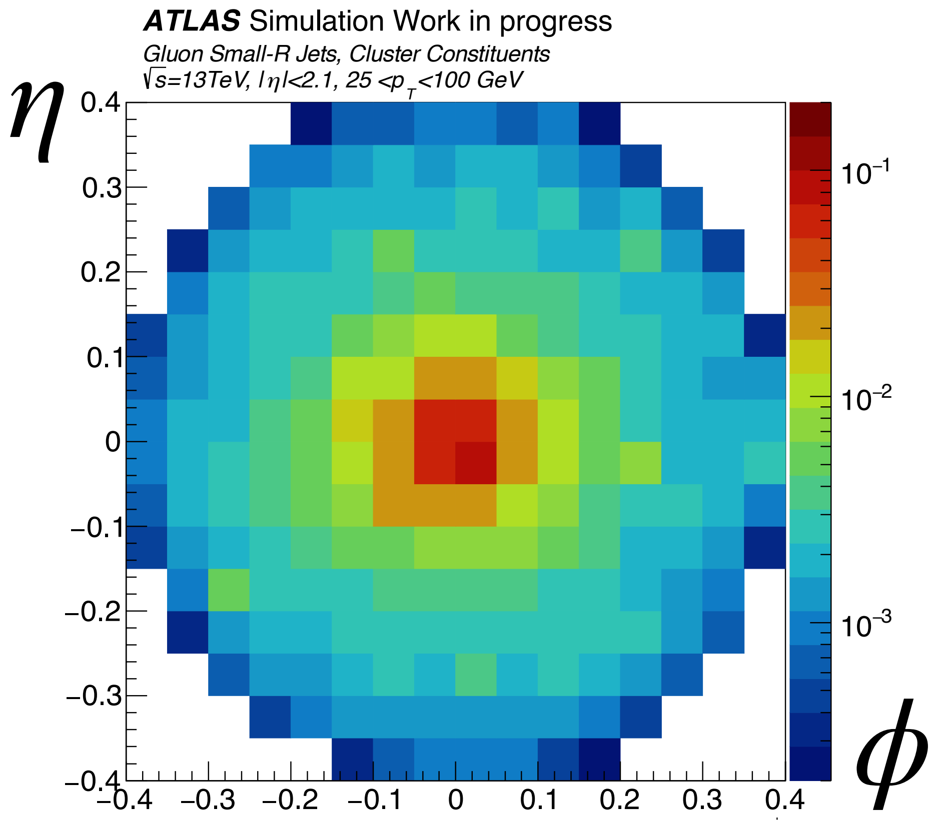


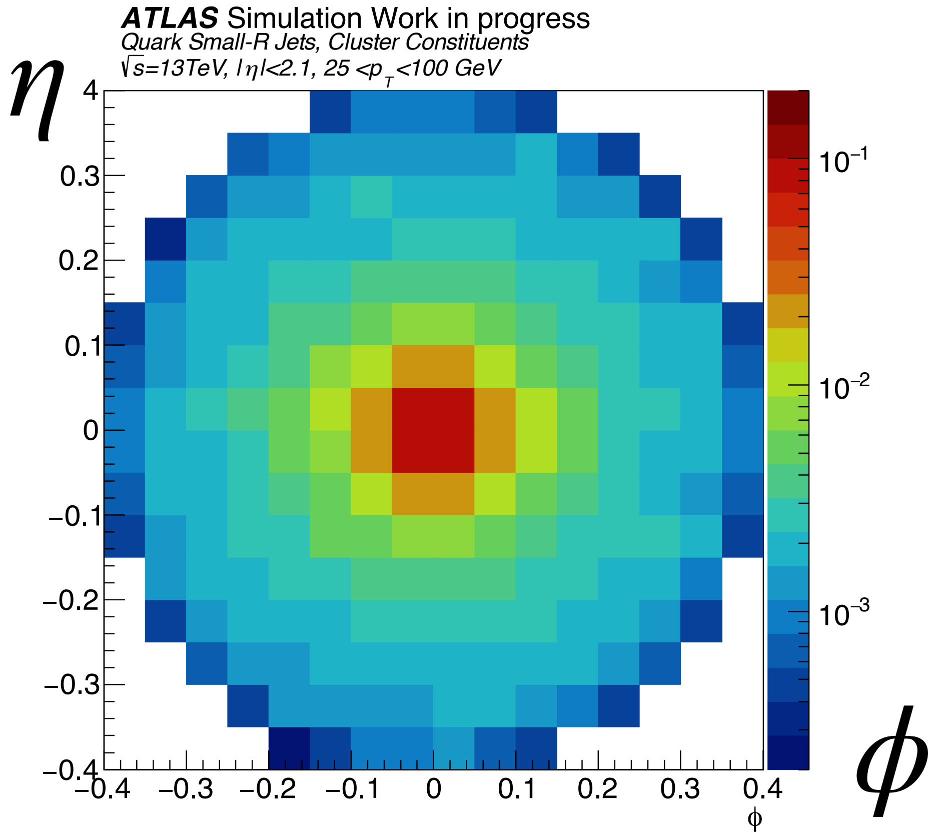
Jet



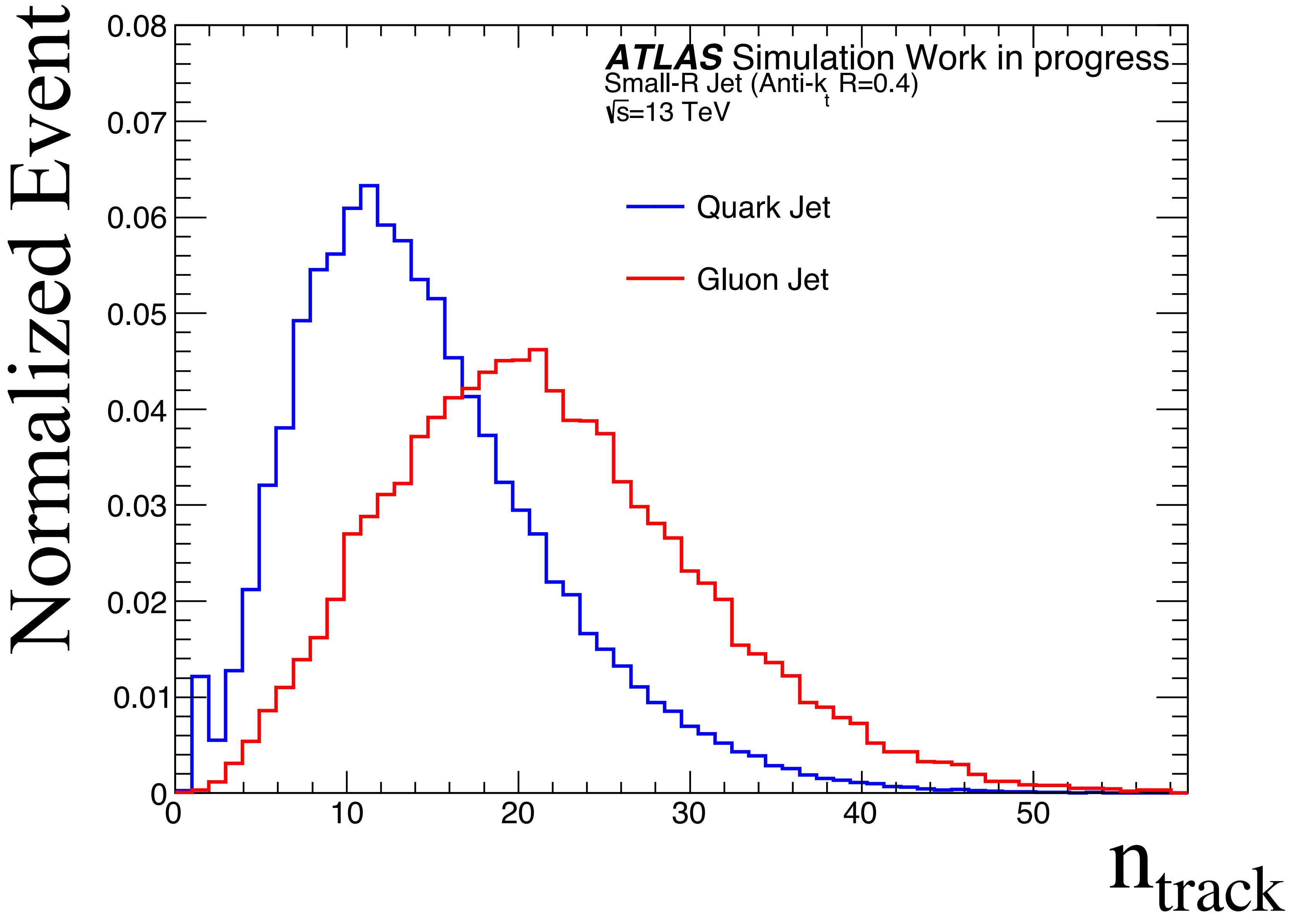


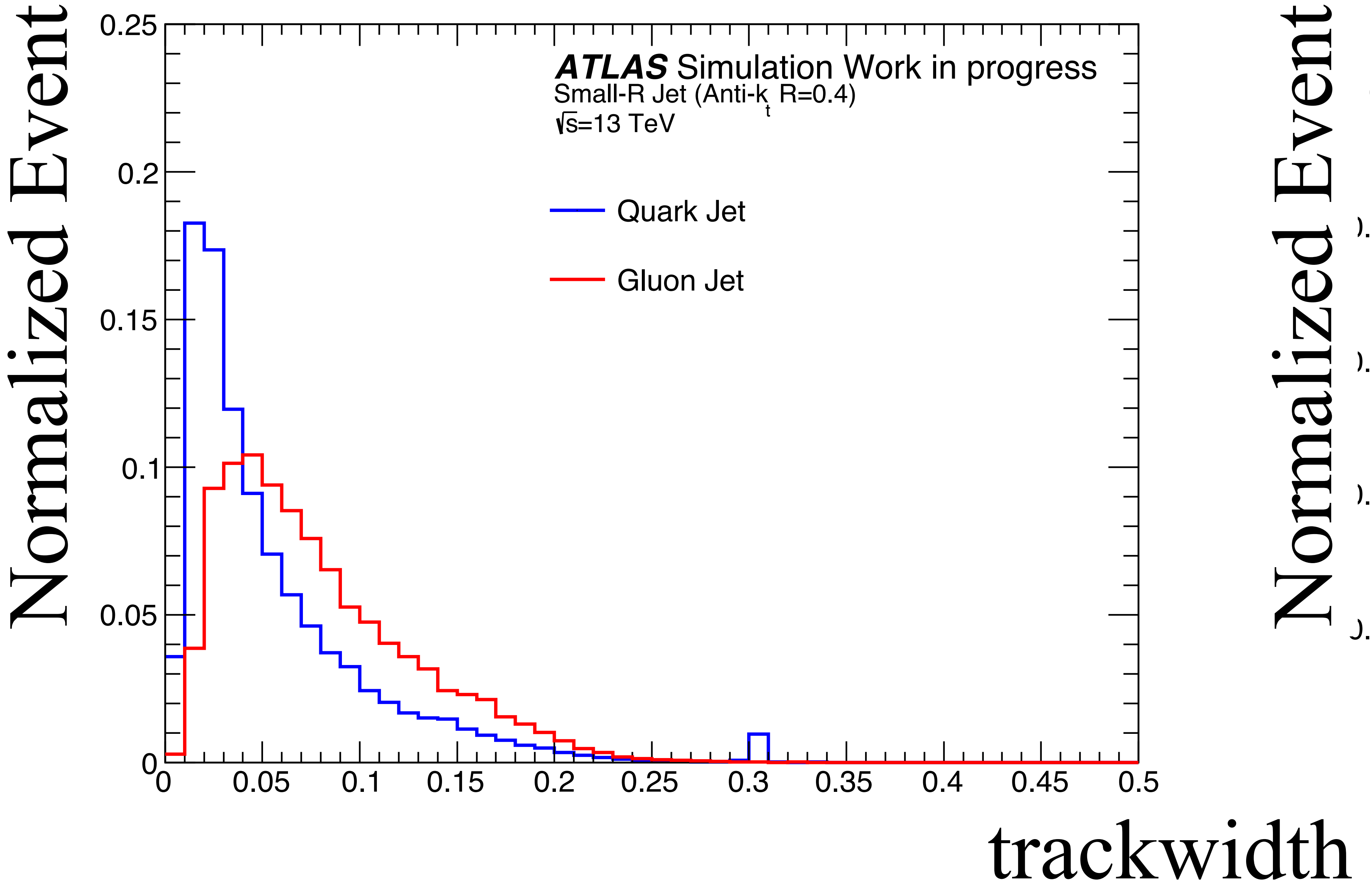
Average Intensity ($25 < p_T < 100$ GeV, $|\eta| < 2.1$)

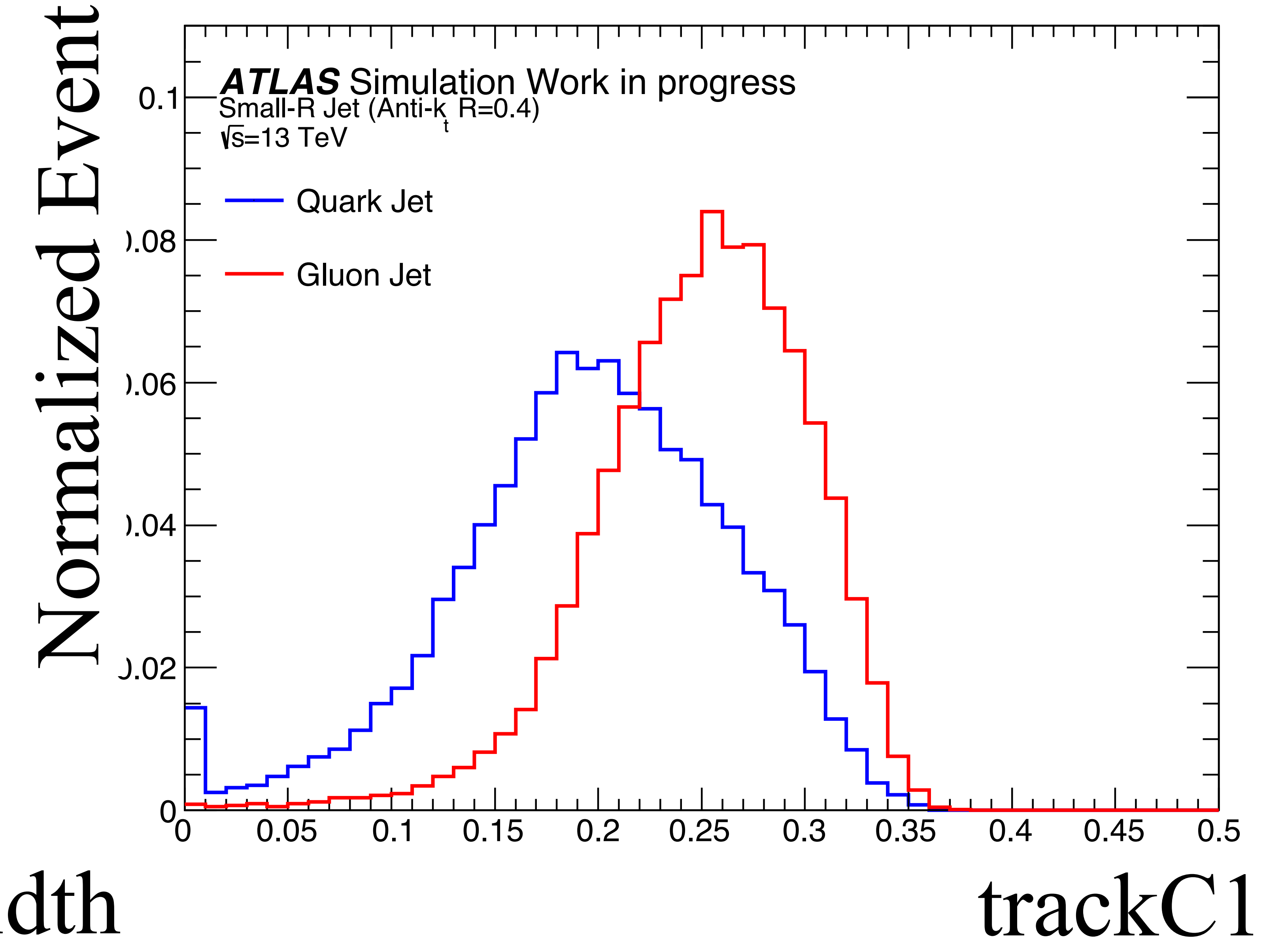




High-level Input







The distance between the tracks and the associated jet

$$\mathrm{trackwidth} = \frac{\sum_i (p_{T,i}^{trk} \times \Delta R(trk_i, Jet))}{\sum_i p_{T,i}^{trk}}$$

where $\Delta R(x,y) = \sqrt{(\eta_x - \eta_y)^2 + (\phi_x - \phi_y)^2}$;

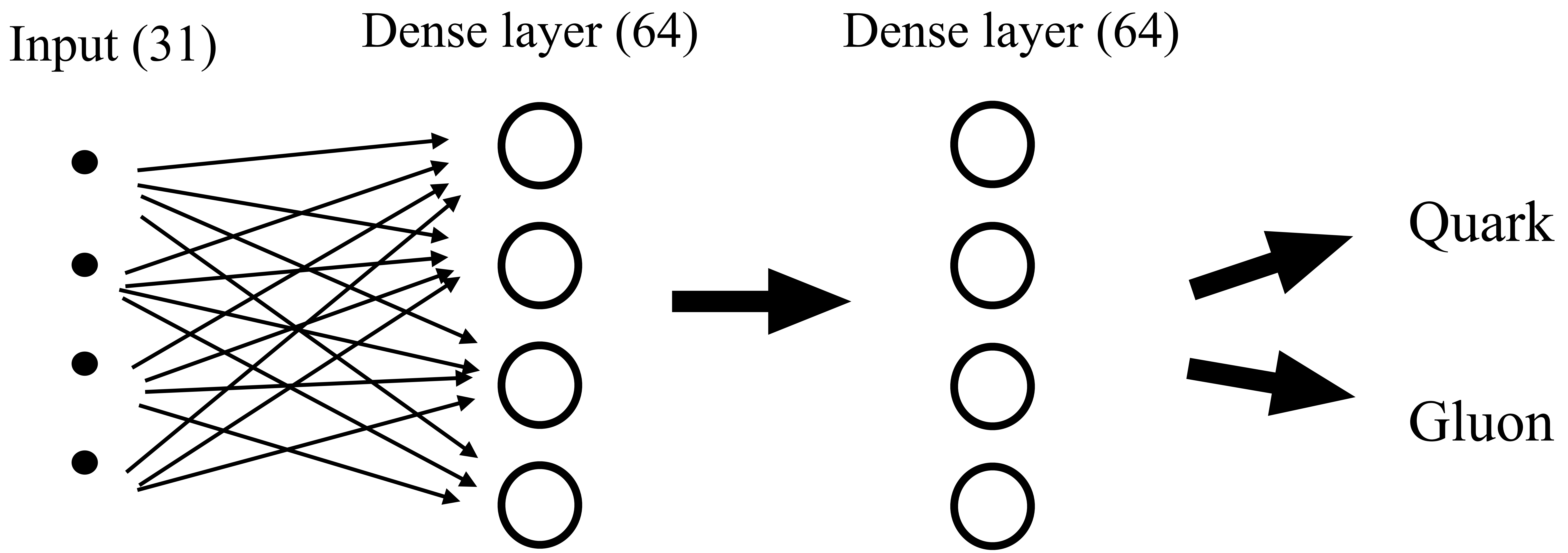
The summation of the distance weighted by the momentum of each track.

$$\mathrm{trackC1} = \frac{\sum_i \sum_j (p_{T,i}^{trk} p_{T,j}^{trk} \times \Delta R^{\beta}(trk_i, trk_j))}{(\sum_i p_{T,i}^{trk})^2}$$

where $\beta = 0.2$

- Using the variables individually is challenging to separate the quark and gluon jets.
- Therefore, the high-level variables are inputted into the TMVA BDT and MLP to have better performance.

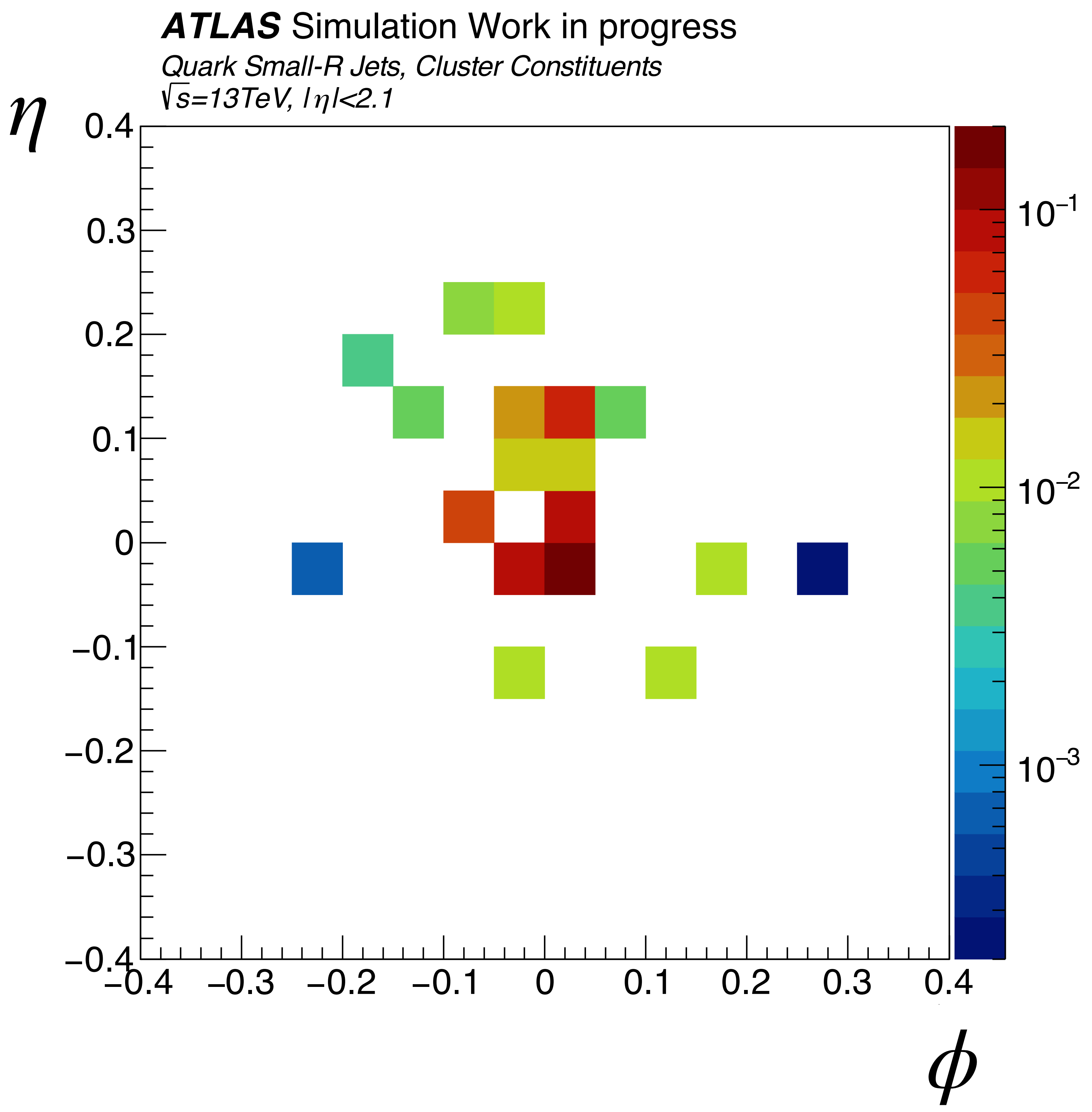
MLP Architecture

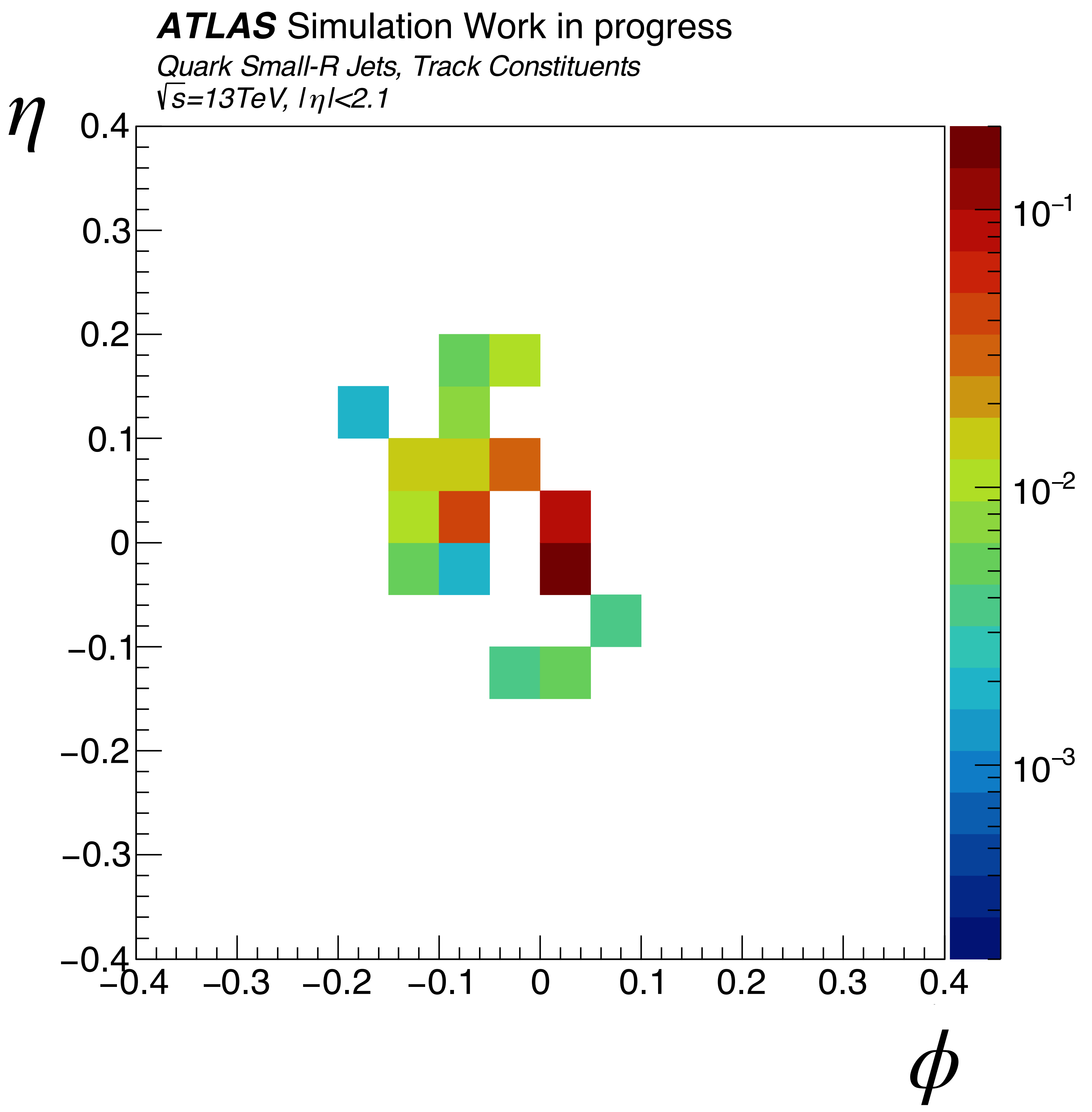


- The weight (Arrow mark) is upgraded by the gradient descent method. $\Delta\omega_{\mathrm{ji}}(\mathrm{n}) = -\lambda \frac{\partial\varepsilon(\mathrm{n})}{\partial\nu_{\mathrm{j}}(\mathrm{n})} \mathrm{y}_{\mathrm{i}}(\mathrm{n})$
- λ is the learning rate, y_{i} is the output of the previous neuron, ε is the square sum of error, and ν is the variable in the kernel function.

- The high-level inputs imply that the amount of information will decrease during the calculation.

Image Input

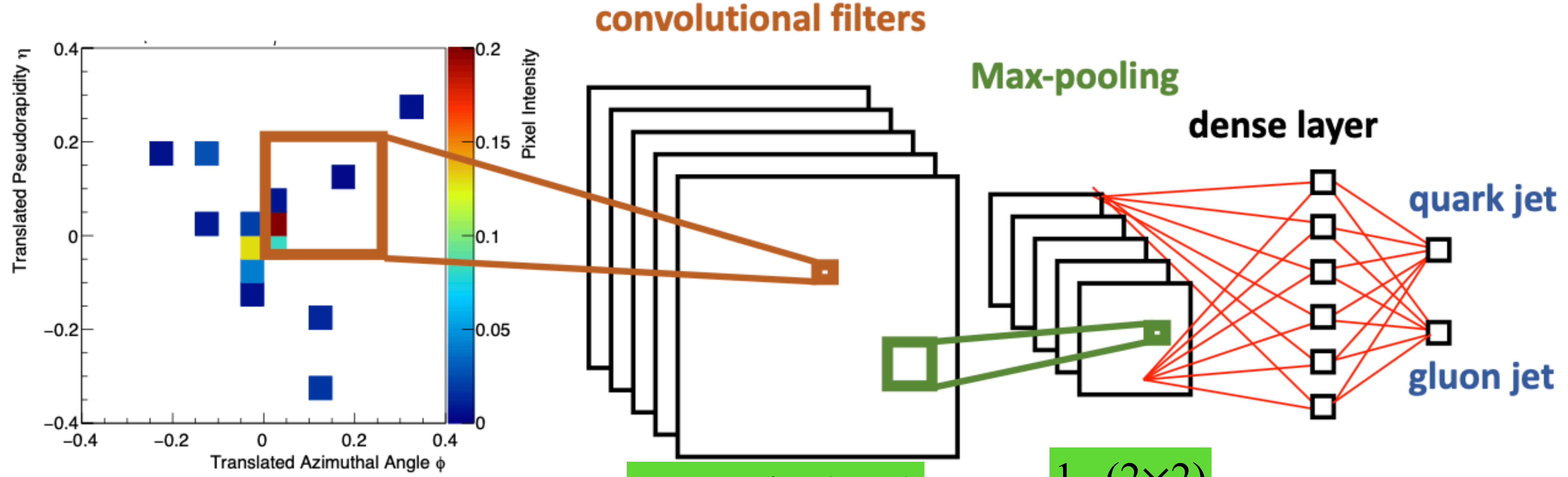




- The image input provides another way to utilize low-level information to separate the quark and gluon jets.
- Both the topo-cluster constituents and the track constituents images have 16×16 pixels in the $|\eta\&\phi| < 0.4$.
- The η and ϕ are rotated to $\eta_{jet} = \phi_{jet} = 0$.
- z axis is the p_T of partons in the jet and p_T is normalized to one for better training performance.

CNN Architecture

Input : $16 \times 16 \times 2$ channel (track + cluster)



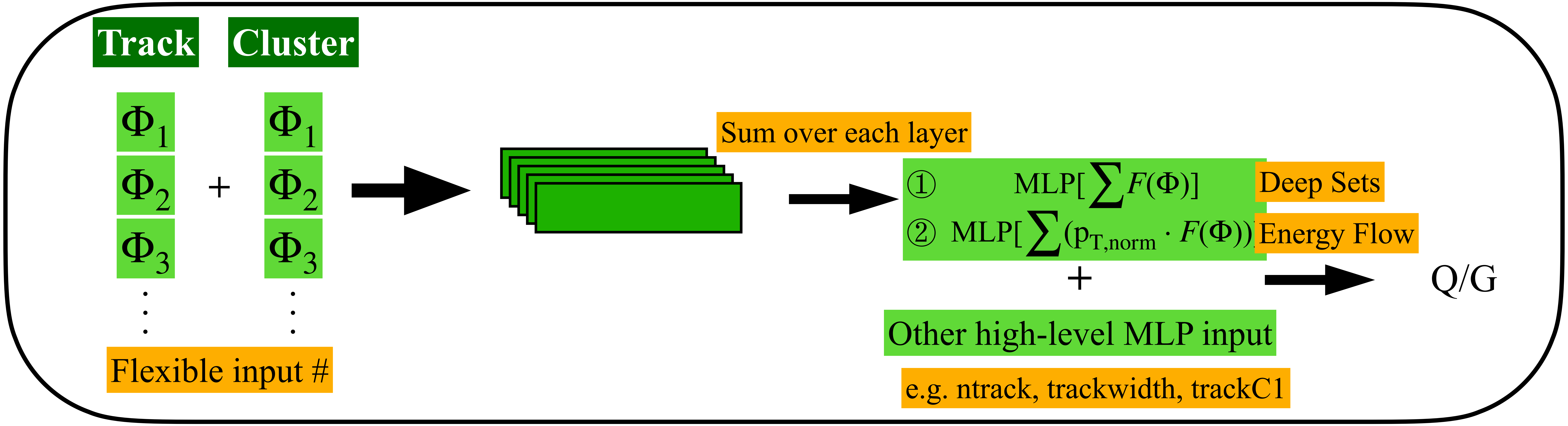
1. 256ch×(7×7)
2. 128ch×(5×5)
3. 64ch×(3×3)

1. (2×2)

ATL-PHYS-PUB-2017-017

- The Convolutional Neural Network model (CNN) is able to use more low-level inputs than the MLP and BDT models.
- Using the CNN model, we could utilize the information of positions with cluster and track inputs.

Pointwise Convolution Architecture

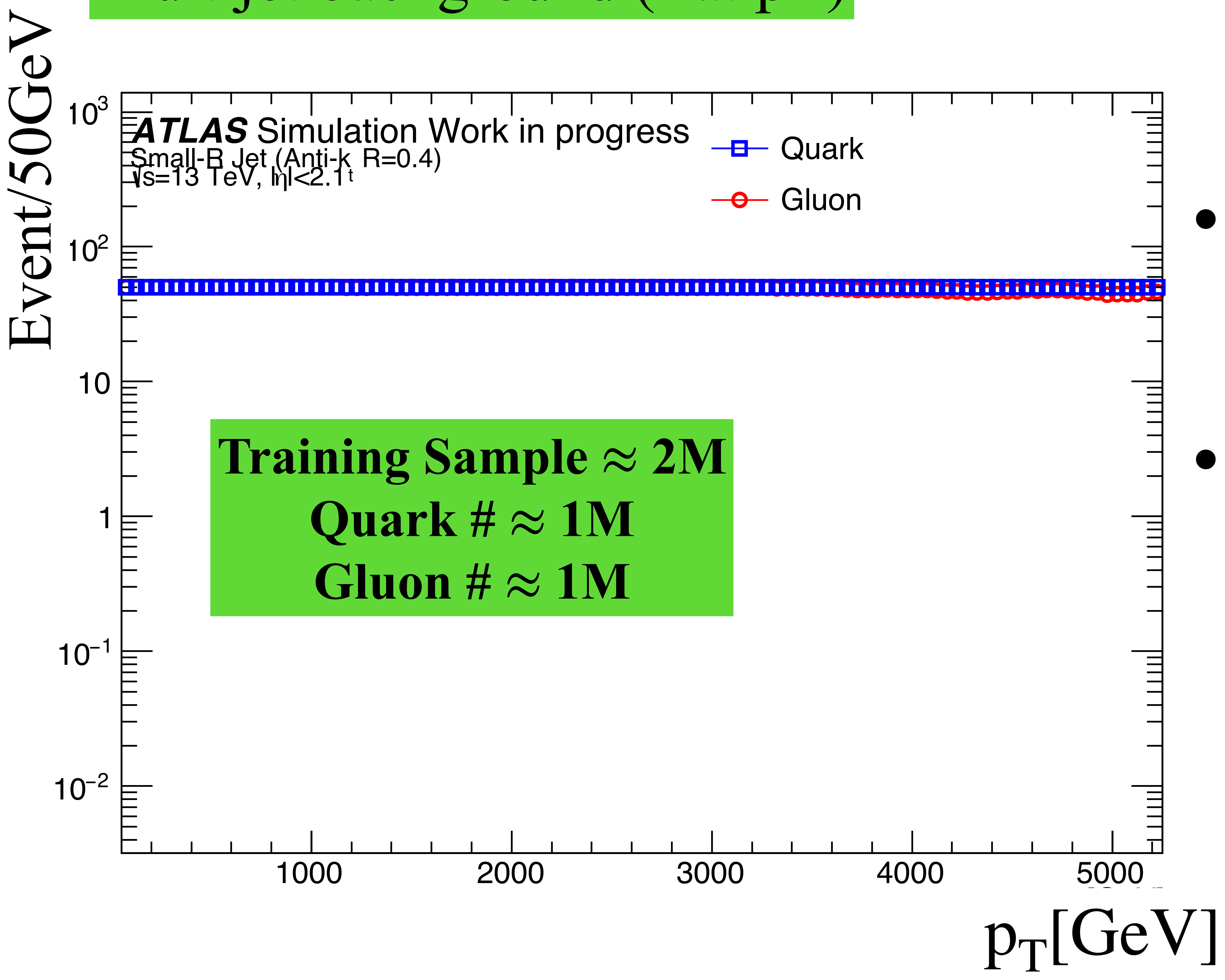


- Since the input used in the CNN model are pixels, it implies we have some limitation on pixel resolution which equates to $\left(\frac{\eta \text{ range}}{\text{pixel \#}} \times \frac{\phi \text{ range}}{\text{pixel \#}} = \frac{0.8}{16} \times \frac{0.8}{16}\right)$
- The PW Conv model can input ϕ and η directly rather than in pixel gives better performance.
- The Energy Flow model satisfies Infrared and collinear safety (IRC safe).

arXiv:0906.1833

MC Sample

Multijet background (flat pT)

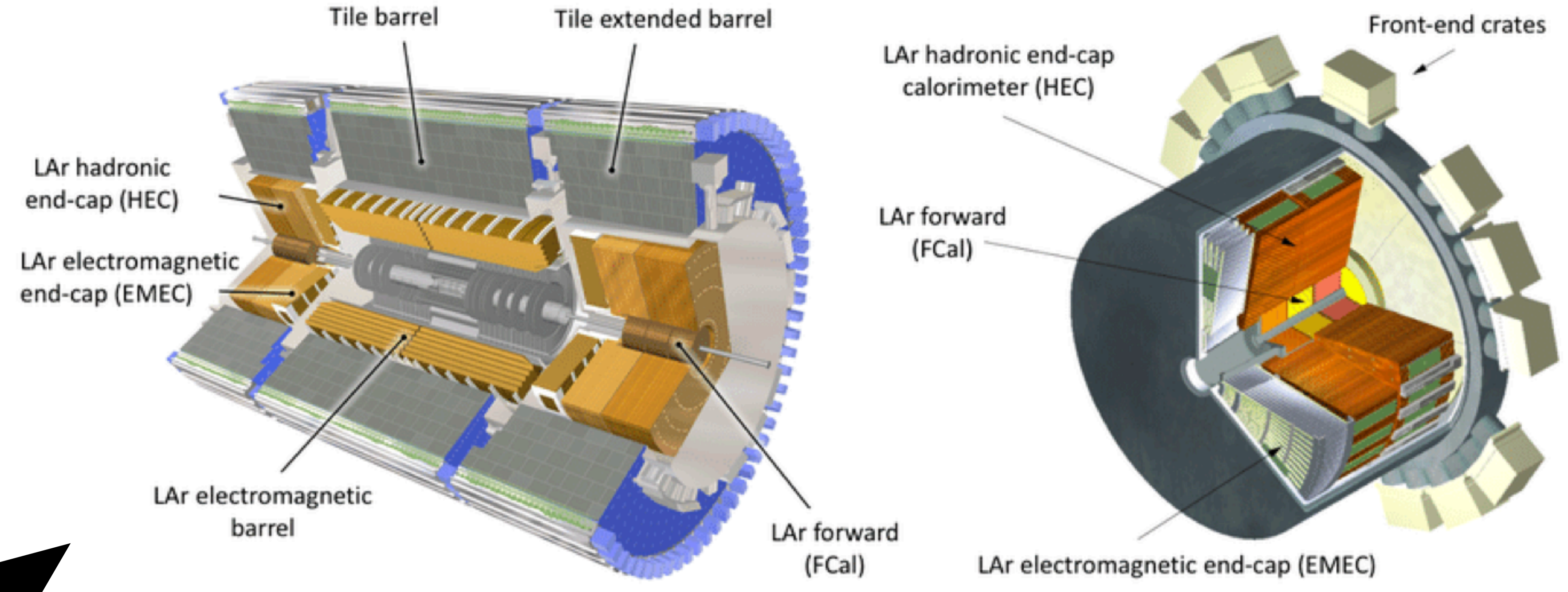


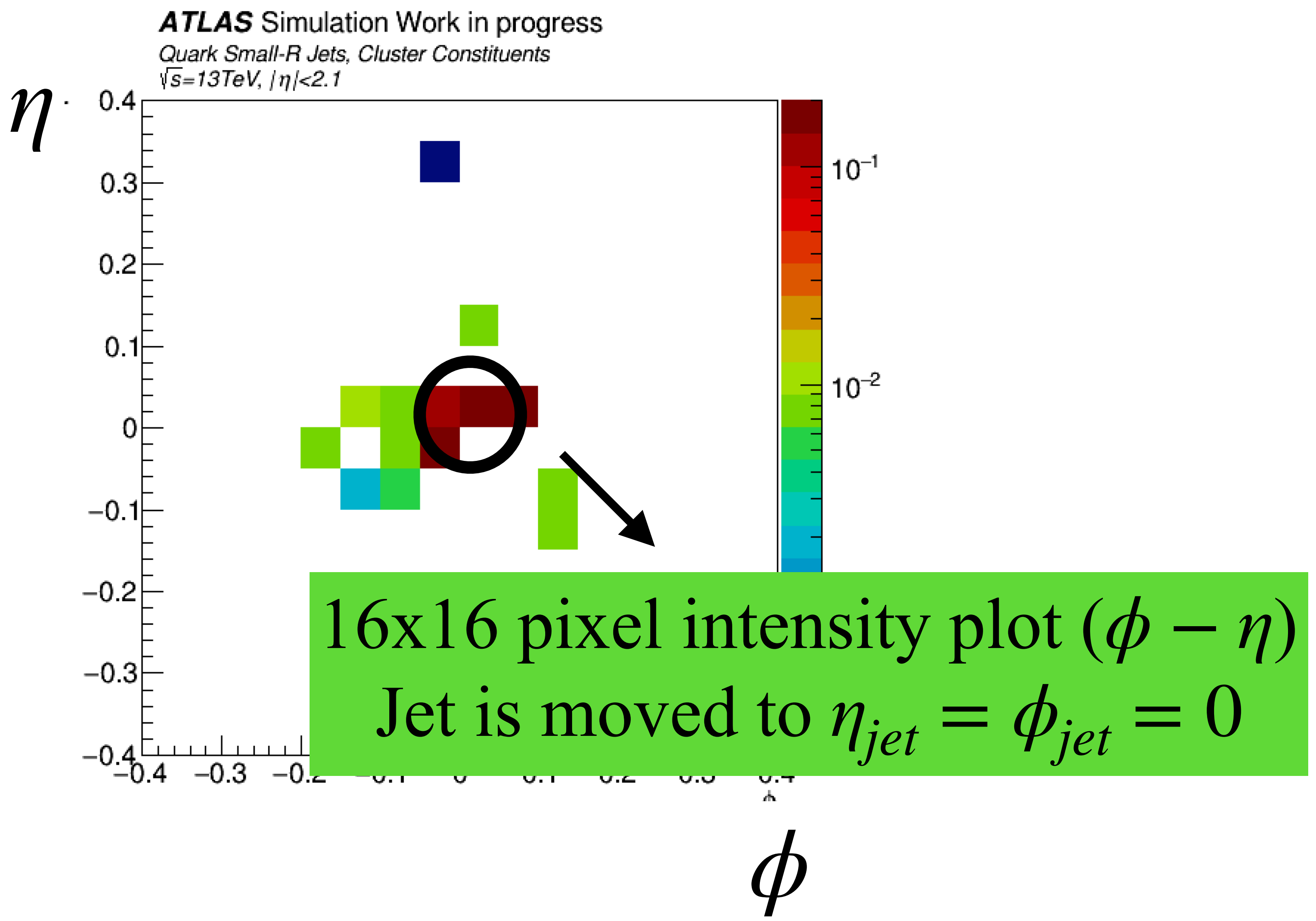
Pileup and noise are considered

- The dijet process@leading order (multijet) MC samples evolved with Pythia 8 have been used to train the q/g tagging models.
- The quark and gluon jets are defined by the labels generated in the MC sample. (Three labels (d, u, s) and one label (g) for quark and gluon jets respectively.)
- In order to have better learning for the q/g tagging model, other steps are taken.
  - The p_T of quark jet and gluon jets is flatted to the same distribution and normalized to one to make sure models are not affected by p_T bias.
  - The quark and gluon inputs are collected in the same numbers since the training performs better with the same amount of quark and gluon inputs in neural network models.

Inputs

	Input
○ BDT	ntrack trackwidth trackC1
○ MLP	ntrack trackwidth trackC1 calorimeter layer energy
○ CNN	$\eta, \phi, \dfrac{p_T(\phi,\eta)}{\text{all } p_T} = p_{T,norm}$
○ PW Conv	$\eta, \phi, p_{T,norm}$ calorimeter layer energy ntrack trackwidth trackC1

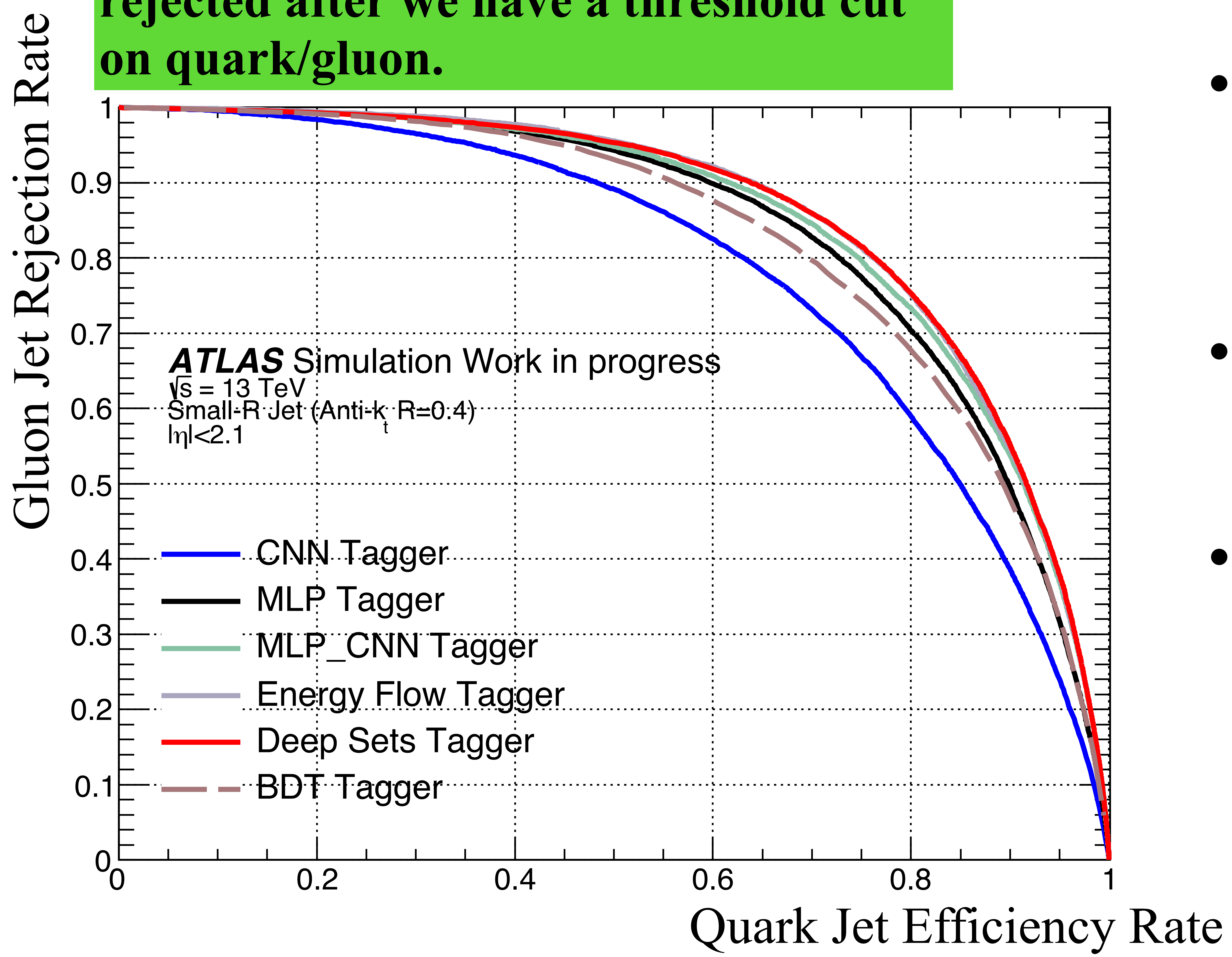




ROC Curve (Receiver Operating Characteristic)

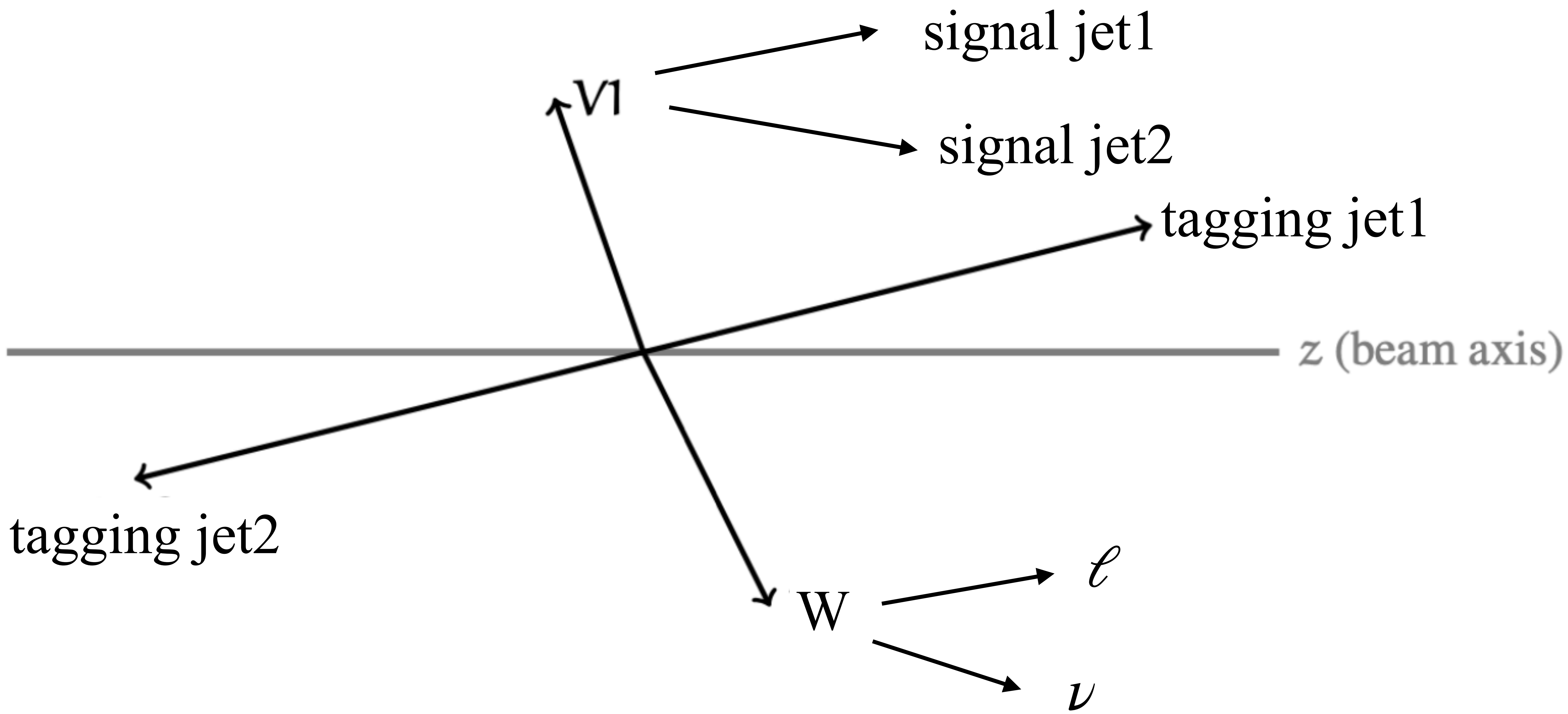
- **Quark Efficiency/Gluon Rejection : The percentage of quark left/gluon rejected after we have a threshold cut on quark/gluon.**

$$\text{Quark Efficiency} = \frac{N_{\text{true postive}}}{N_{\text{positive}}}, \quad \text{Gluon Rejection} = 1 - \frac{N_{\text{true negative}}}{N_{\text{negative}}}$$



- Deep Sets (Red Curve) and Energy Flow (Gray Curve) models have almost the same performance and achieve the best performance among all models.
- Since the pointwise models can utilize the low-level inputs, the improvement is more significant.
- Compared to the BDT model, the gluon rejection rate of the Pointwise models improved approximately **10% better at the 80% quark efficiency rate**, Pointwise and BDT models are around **75% and 68%** respectively.

1-lepton semileptonic VBS Process



- This thesis only focuses on the 1-lepton semileptonic VBS channel for convenience. The signal candidate events were selected by four hadron jets, in addition to exactly one lepton in the event.
- The main sources of the background process are W+jets and $t\bar{t}$. About 32%, 12%, and 1% of W+4jets events have 2, 3 and 4 gluon-induced jets in the selected jets.

Improving Results

	Selection stratgy	Bkg systematic uncertainty	Signal	W+jet	$t\bar{t}$	Significance	Compared to baseline significance, improving by
Baseline	2leading	10%	44	42±14	128±38	0.78	
	minmass	10%	45	57±18	103±39	0.73	
With q/g tagger	neural 2leading	10%	48	46±13	148±41	0.81	3.8%
	neural minmass	10%	65	83±22	242±58	0.77	5.5%

- **2leading strategy** :
  - It has a larger probability of selecting the jets from $t\bar{t}$ samples but with fewer background events.
  - There are 1059 VBS signal events, 297845 W+jet events, and 235519 $t\bar{t}$ events.
- **minmass strategy** :
  - This makes sure more jets decay from W/Z vector bosons, but the significance is worse.
  - There are 1419 VBS signal events, 454939 W+jet events, and 409285 $t\bar{t}$ events.

Conclusion

- Because of the BDT q/g tagger limitation of only using high-level input variables, neural network q/g tagging models introduced in this study are considered to solve this issue.
- The gluon rejection rate of the Energy Flow and Deep Sets models improved approximately **10%** than the conventional BDT q/g tagging model at the 80% quark efficiency rate.

- The Energy Flow q/g tagging models are examined in the 1-lepton semileptonic VBS process.
- The improving amplitude of the neural 2leading and neural minmass is approximately **3.8% and 5.5%** respectively, where the neural 2leading strategy has the **best significance of 0.81.**

- Try to use other kinds of neural network model (such as the GNN model.).
- Adding other input variables, such as particle charge.

Backup

東京大学
素粒子物理国際研究センター
International Center for Elementary Particle Physics
The University of Tokyo

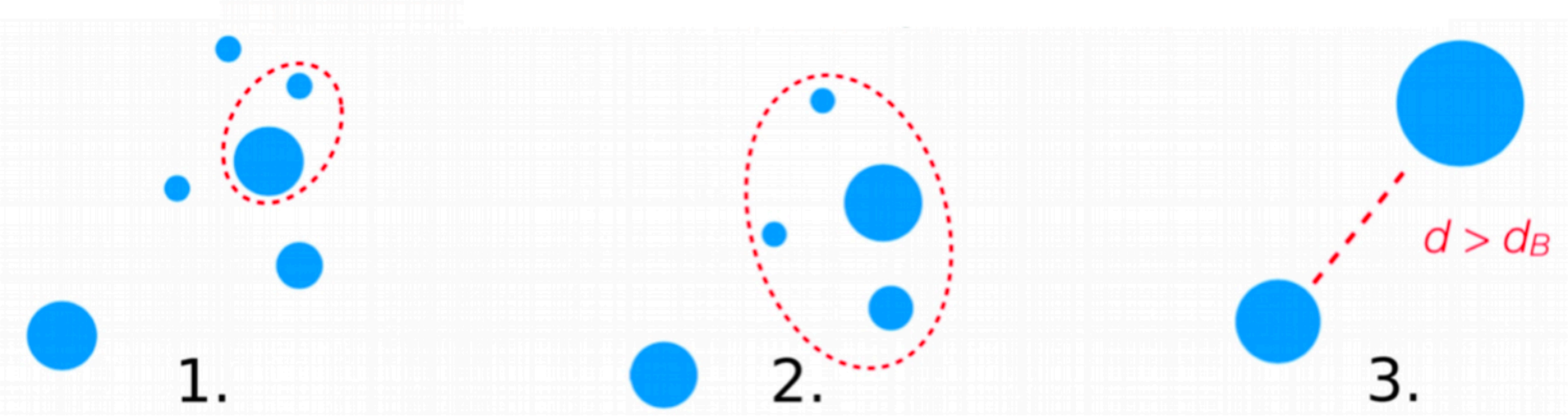
ATLAS EXPERIMENT

Baseline Analysis and Improving Limit

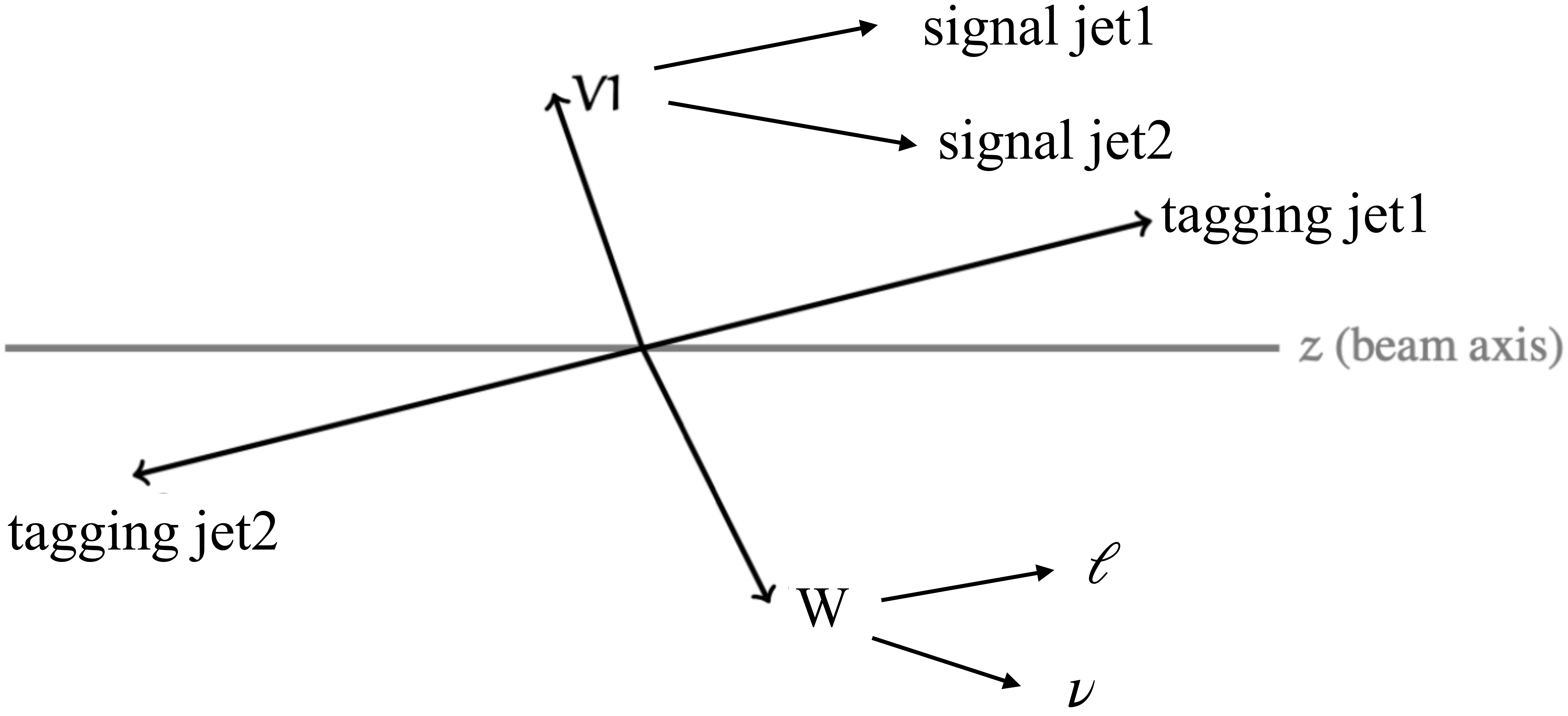
- **Sensitivity of the baseline analysis**
  - The BDT discriminant output threshold is scanned to obtain the best significance. Therefore, The numbers of the signal and background events are different with different systematic uncertainty assumptions.

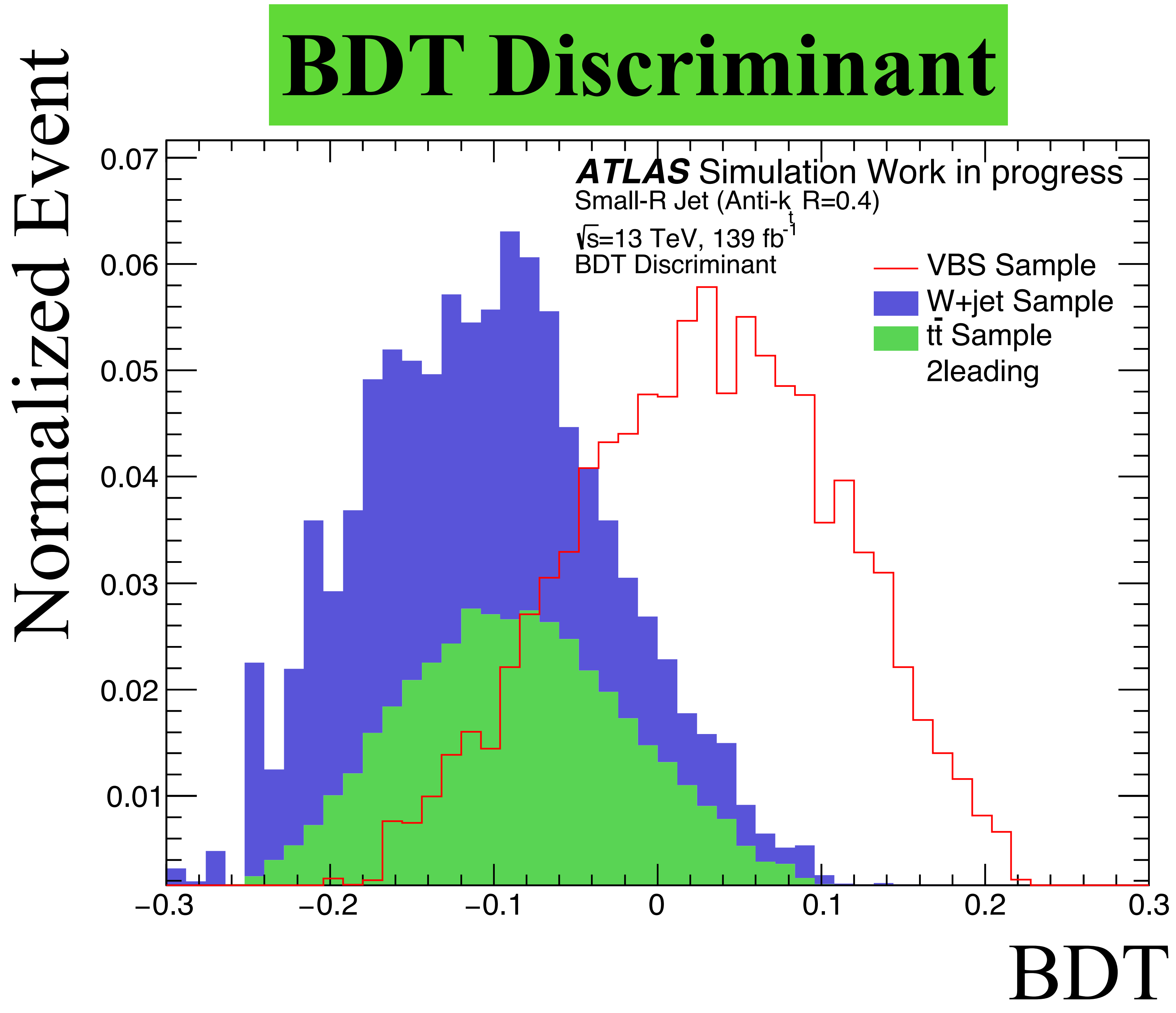
- ***q/g* tagging improving limit**
  - Consider there is a perfect q/g tagging model that can separate quark and gluon jets 100%.
  - The result shows that the significance of perfect q/g tagging **improves by approximately 60%** compared to the significance of baseline analysis.

Reconstruction of Jet



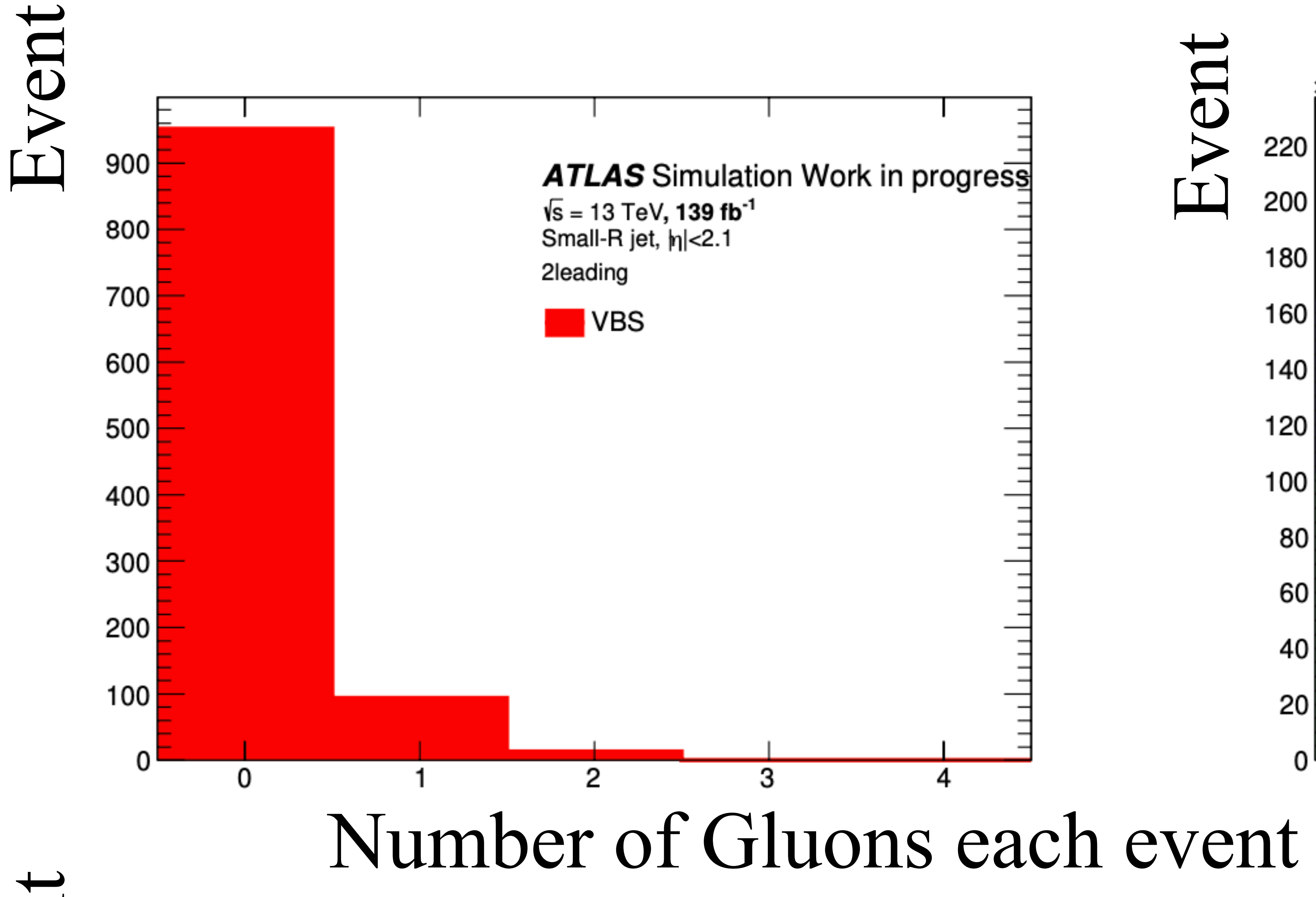
1-lepton semileptonic VBS Process

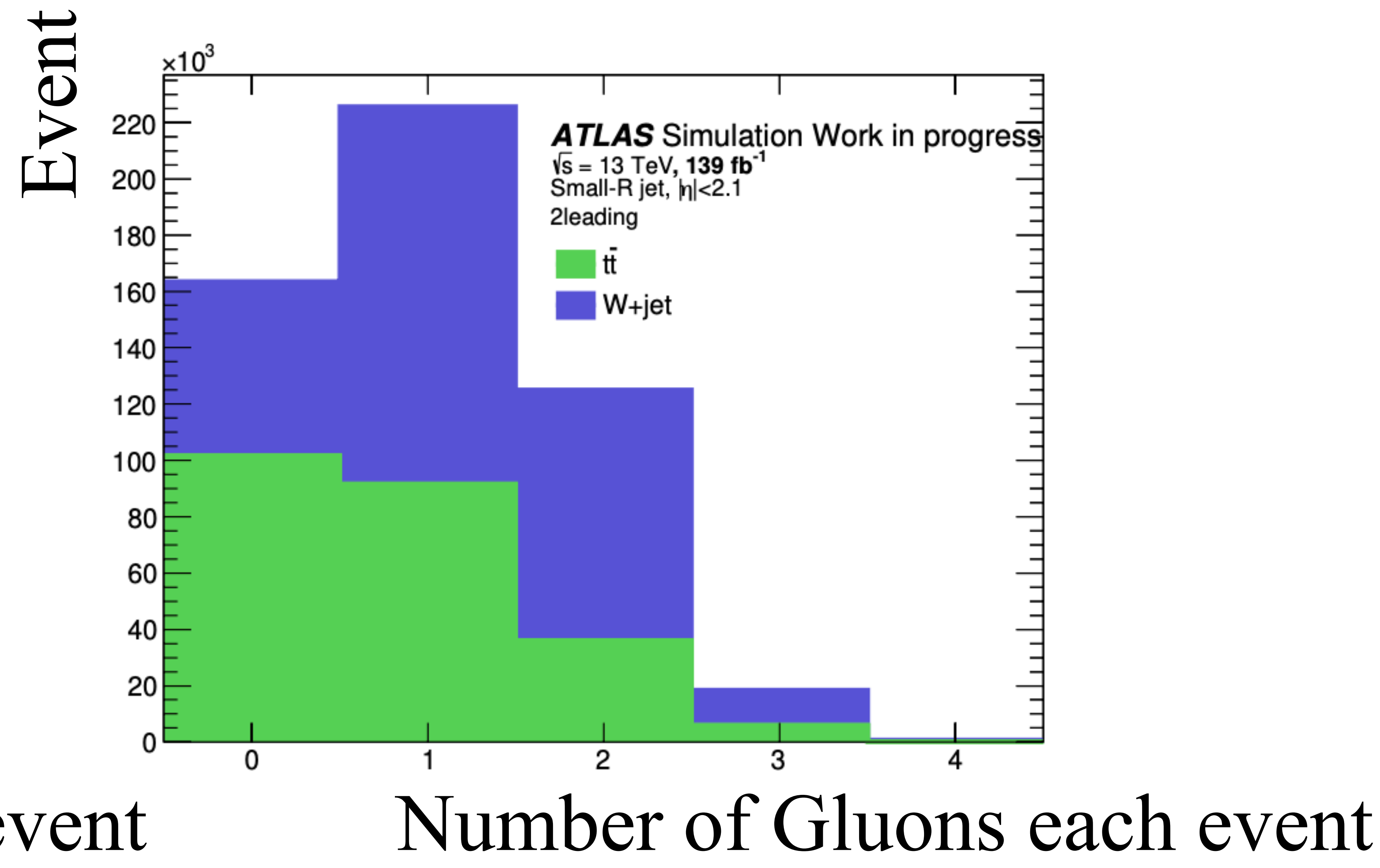


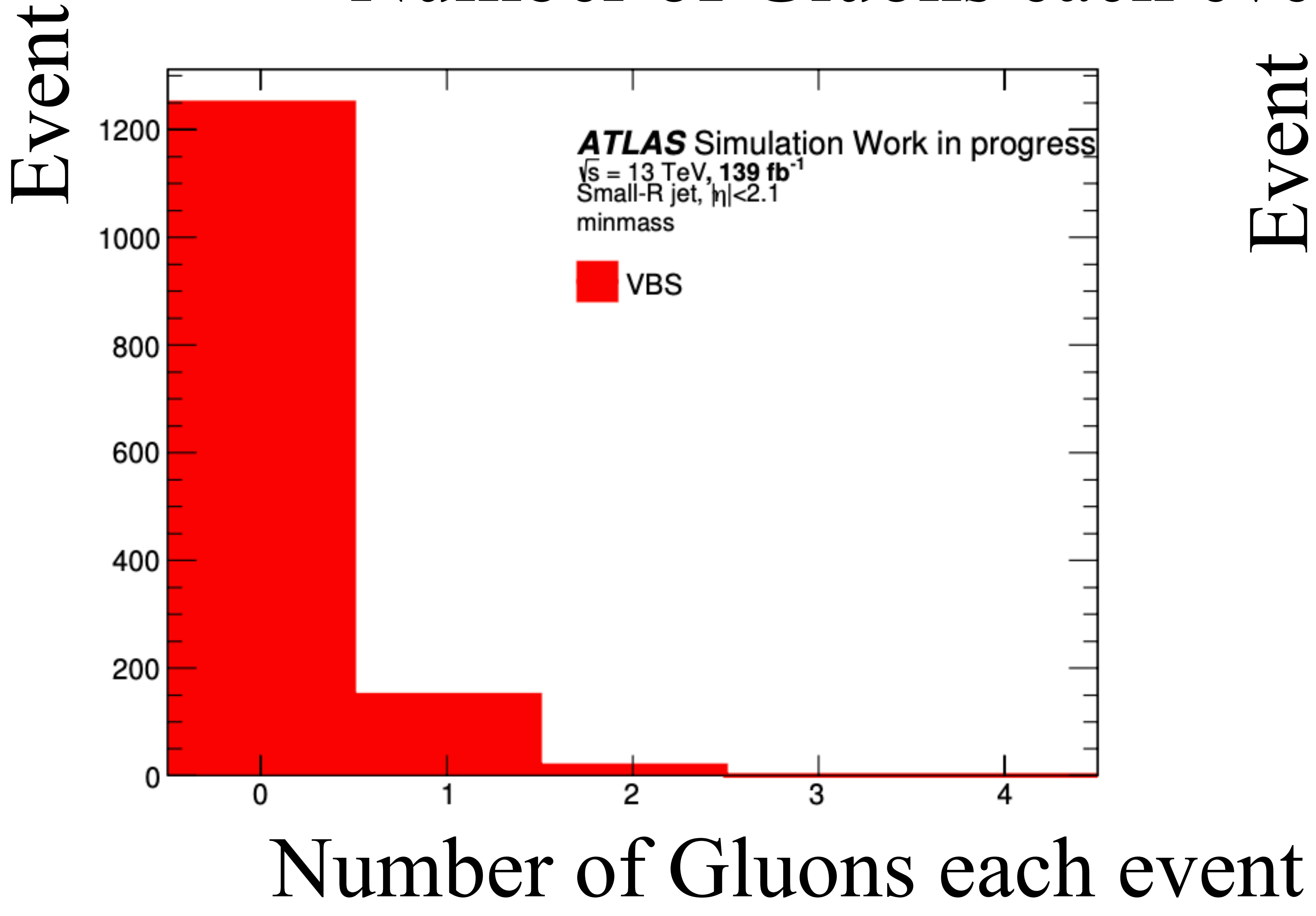


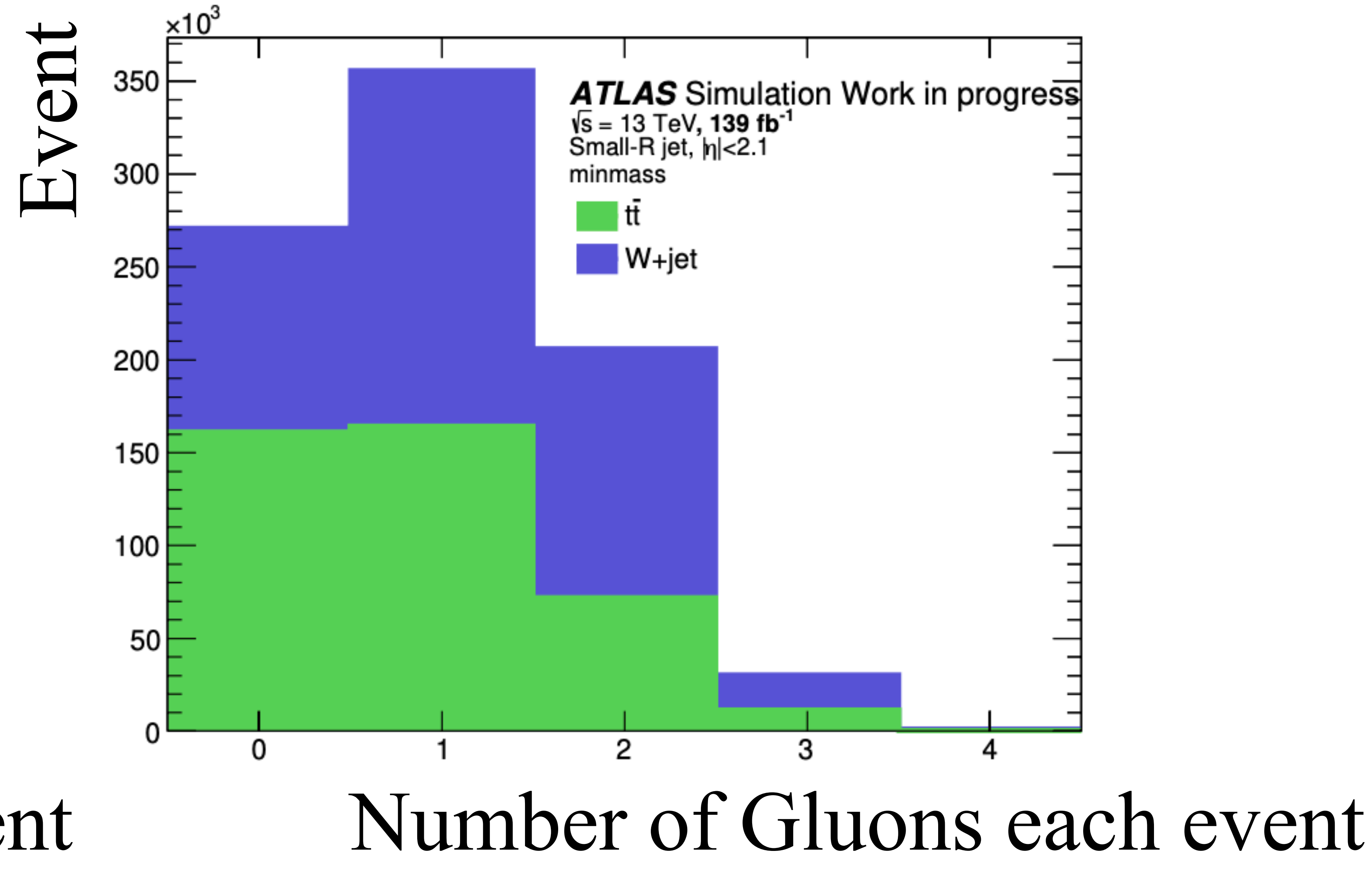
- This thesis only focuses on the 1-lepton semileptonic VBS channel for convenience. The signal candidate events were selected by four hadron jets, in addition to exactly one lepton in the event.
- The BDT Discriminant method is considered to enhance the purity of signal events.
- The main sources of the background process are W+jets and $t\bar{t}$. About 32%, 12%, and 1% of W+4jets events have 2, 3 and 4 gluon-induced jets in the selected jets.

Number of Gluons each event









BDT Architecture

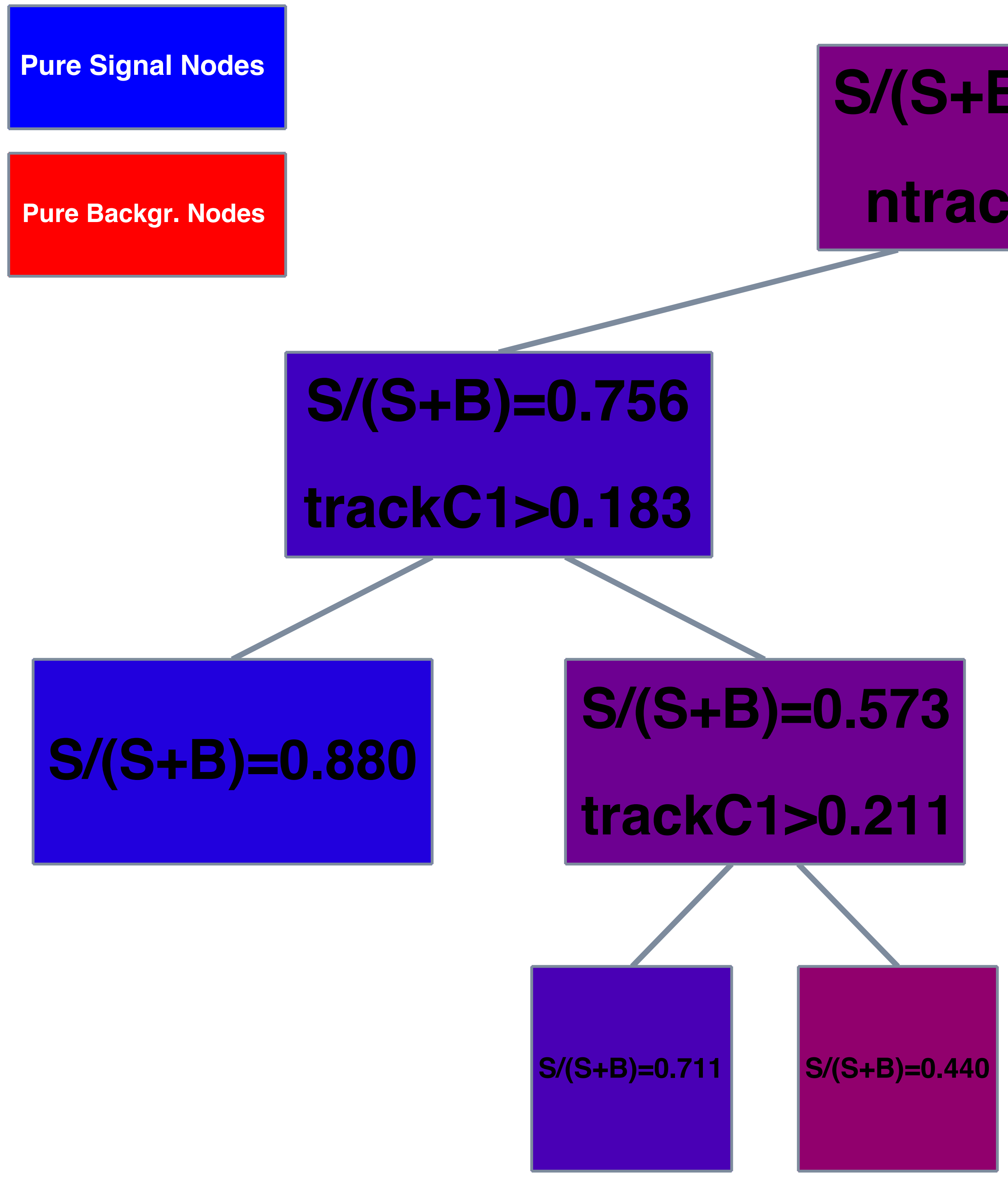


Table 4.1: The setup of BDT model.

AnalysisType=Classification
NTrees=850
MinNodeSize=2.5%
MaxDepth=3
AdaBoostBeta=0.5
BaggedSampleFraction=0.5
SeperationType=GiniIndex
nCuts=20

- First, give a cut on one of the variables and calculate the remaining signal and background percentage.
- Secondly, choose the better one and give another cut on another variable.

- The BDT q/g tagger is considered as a baseline model compared to the neural network models.

Review of CNN model

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

6 x 6 image with
3 x 3 kernel

1	0	-1
0	1	3
-2	0	1

3 x 3 kernel

		5	1		

Layer i

Layer i+1

Calorimeter input

Table 4.2: The calorimeter layers used in this study.

LAr barrel	LAr EM endcap	Hadronic endcap	Tile barrel
PreSamplerB	PreSamplerE	HEC0	TileBar0
EMB1	EME1	HEC1	TileBar1
EMB2	EME2	HEC2	TileBar2
EMB3	EME3	HEC3	
Tile gap	**Tile extended barrel**	**Forward EM endcap**	**Mini FCAL**
TileGap1	TileExt0	FCAL0	MINIFCAL0
TileGap2	TileExt1	FCAL1	MINIFCAL1
TileGap3	TileExt2	FCAL2	MINIFCAL2
			MINIFCAL3

Pointwise Convolution Architecture

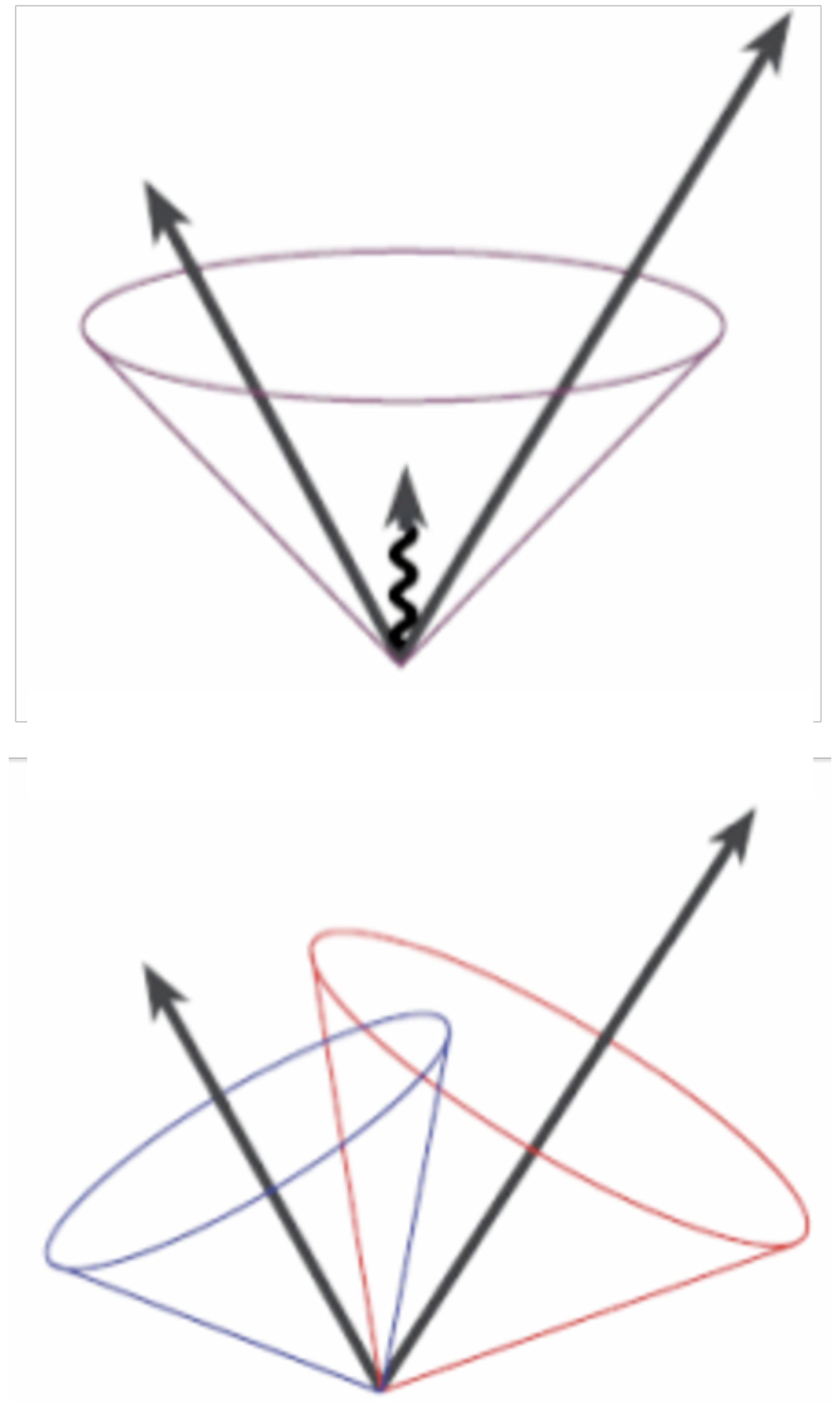
- We construct two kinds of Pointwise Convolution Architecture here.
  - Deep Sets
    - Input $\Phi = [p_{T,norm}, \eta, \phi]$
    - Layer = $\mathrm{MLP}[\sum F(\Phi)]$
  - Energy Flow
    - Input $\Phi = [\eta, \phi]$
    - Layer = $\mathrm{MLP}[\sum (p_{T,norm} \cdot F(\Phi))]$

The Energy Flow model satisfies Infrared and collinear safety (IRC safe)

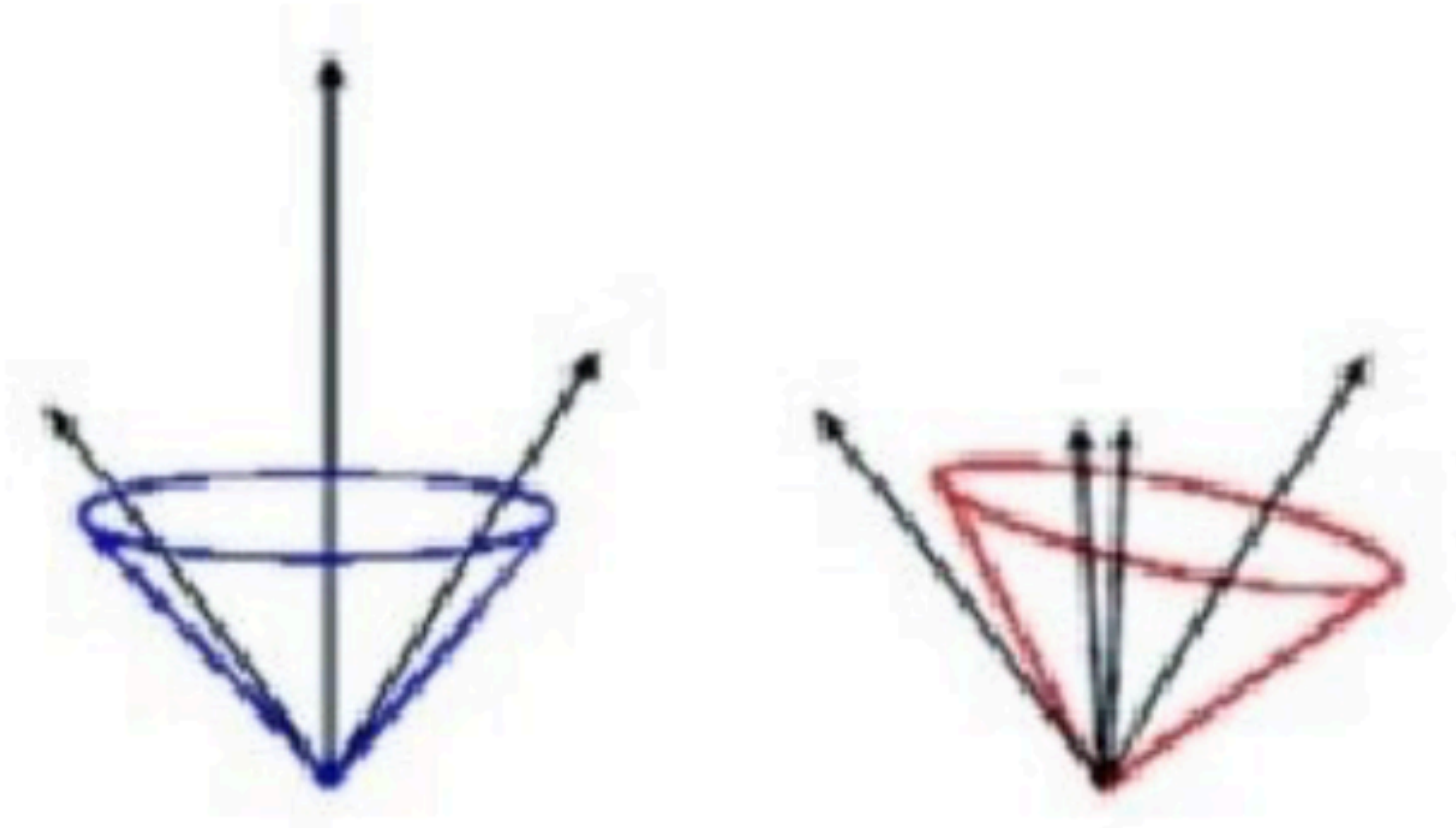
IRC safe means the model remains unchanged if we have addition of collinear splittings and soft emission effect.

arXiv:0906.1833

IRC Safe



- Infrared Safety
  - Soft emission doesn't change the classification



- Collinear Safety
  - Particle splits into two with same energy doesn't change the result.

Conventional Strategy V.S Neural Network

- The QCD interaction generates many gluons as background in the LHC. Therefore, it is essential to develop a tool to separate the quarks and gluons, which is known as Quark/Gluon tagging (q/g tagging).
- Neural network models have played a pivotal role in machine learning recently since some of them are able to make use of more low-level variables.

Conventional Strategy : BDT

- Use Boosted Decision Tree (BDT)
  - nTrack
  - Jet width related variables
    - trackwidth
    - trackC1
- Hard to utilize more low-level inputs.
- The BDT model is considered as a baseline compared with the neural network models.

Neural Network

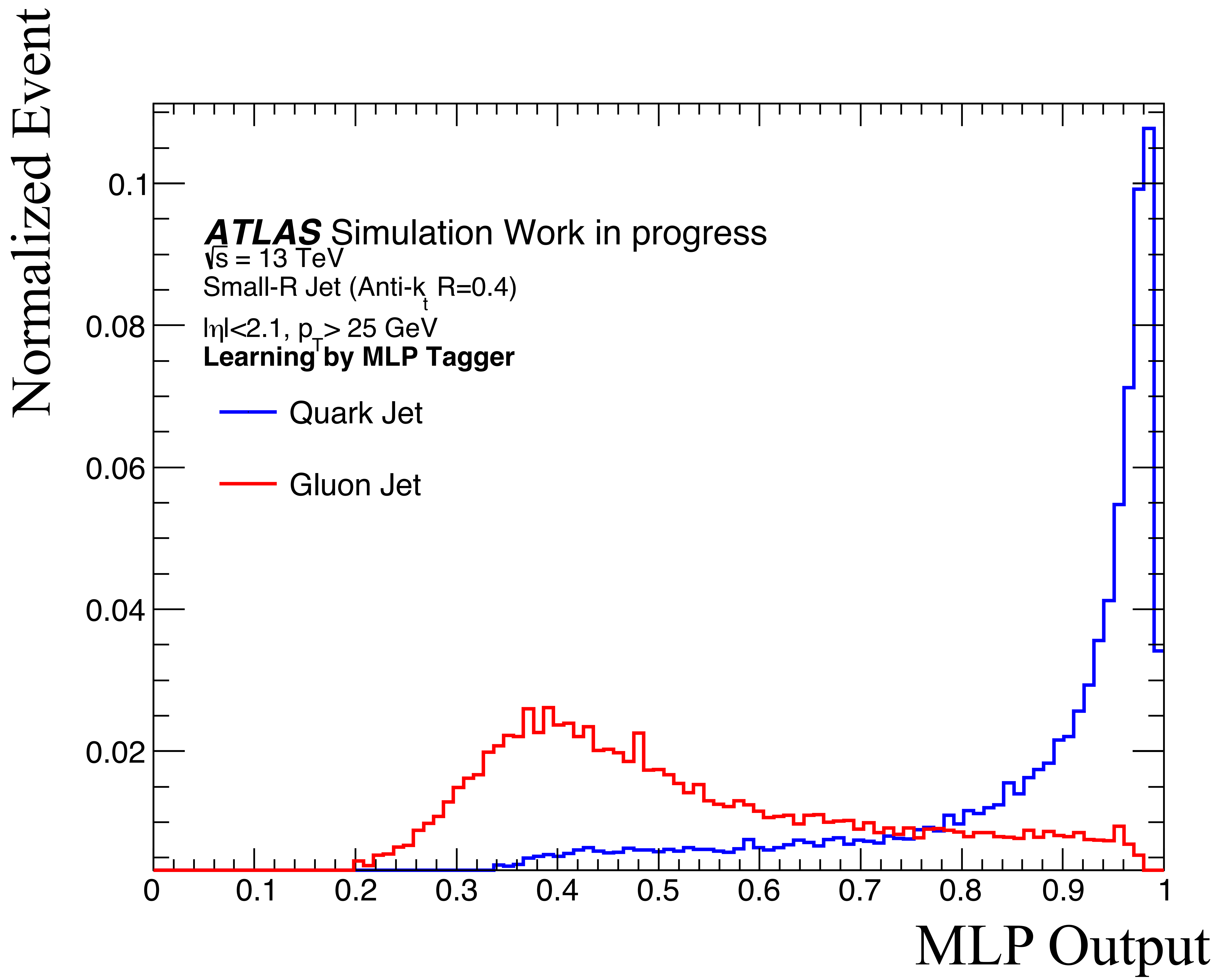
- There are many kinds of neural network
  - MultiLayer Perceptron (MLP)
  - Convolutional Neural Network (CNN)
  - MLP_CNN
  - Pointwise Convolution (PW Conv)
    - Deep Sets
    - Energy flow
- Some models can use low-level inputs that means we can utilize much more information.

Comparison of all models

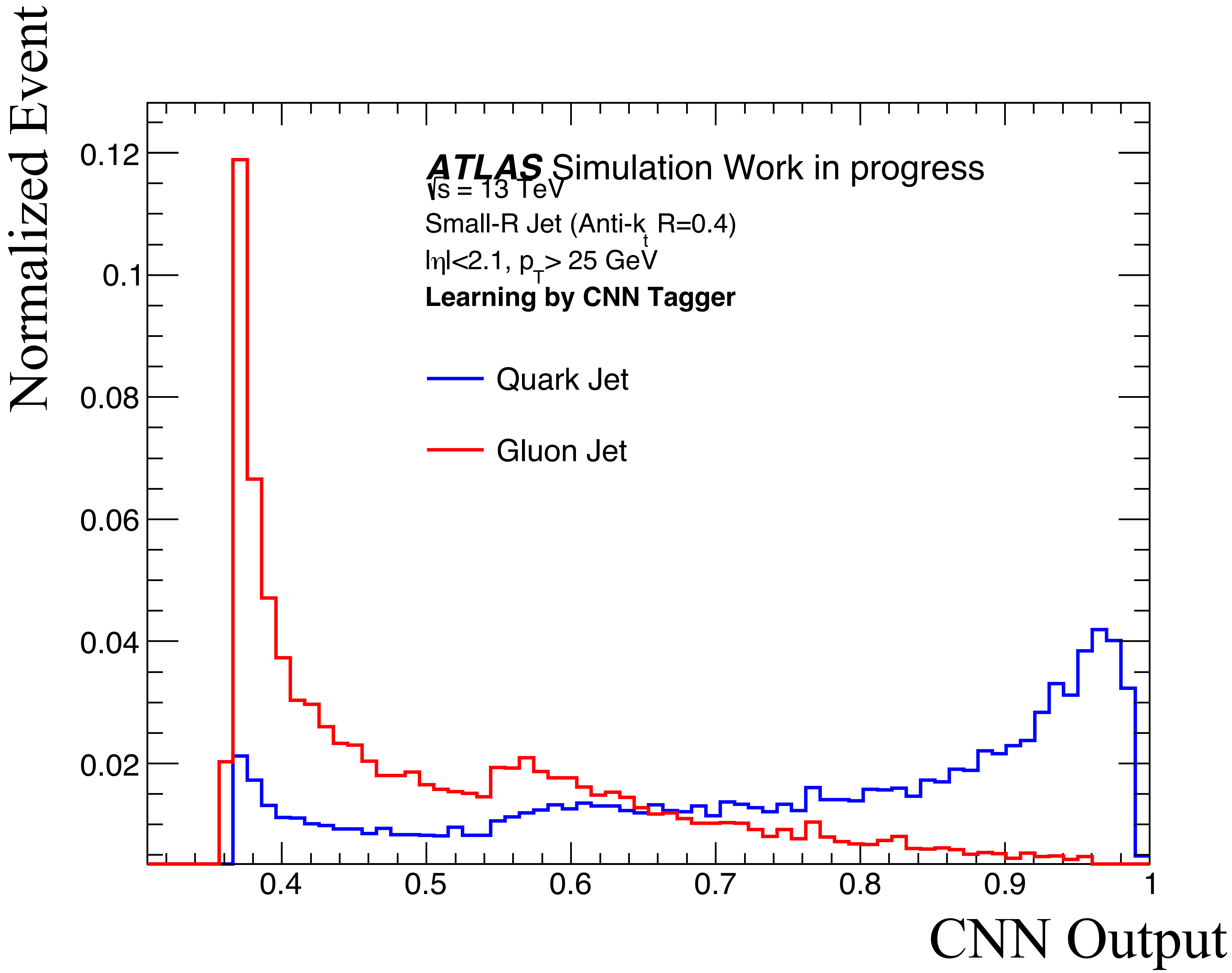
- BDT and MLP models :
  - These two models conveniently input the high-level variables, such as the number of track, trackwidth and trackC1.

- CNN model :
  - The input of the CNN model is an image. Therefore, it is helpful to utilize the variables related to the position.
  - In this paper, the track and cluster p_T images are inputted, giving a way to utilize the low-level variables.

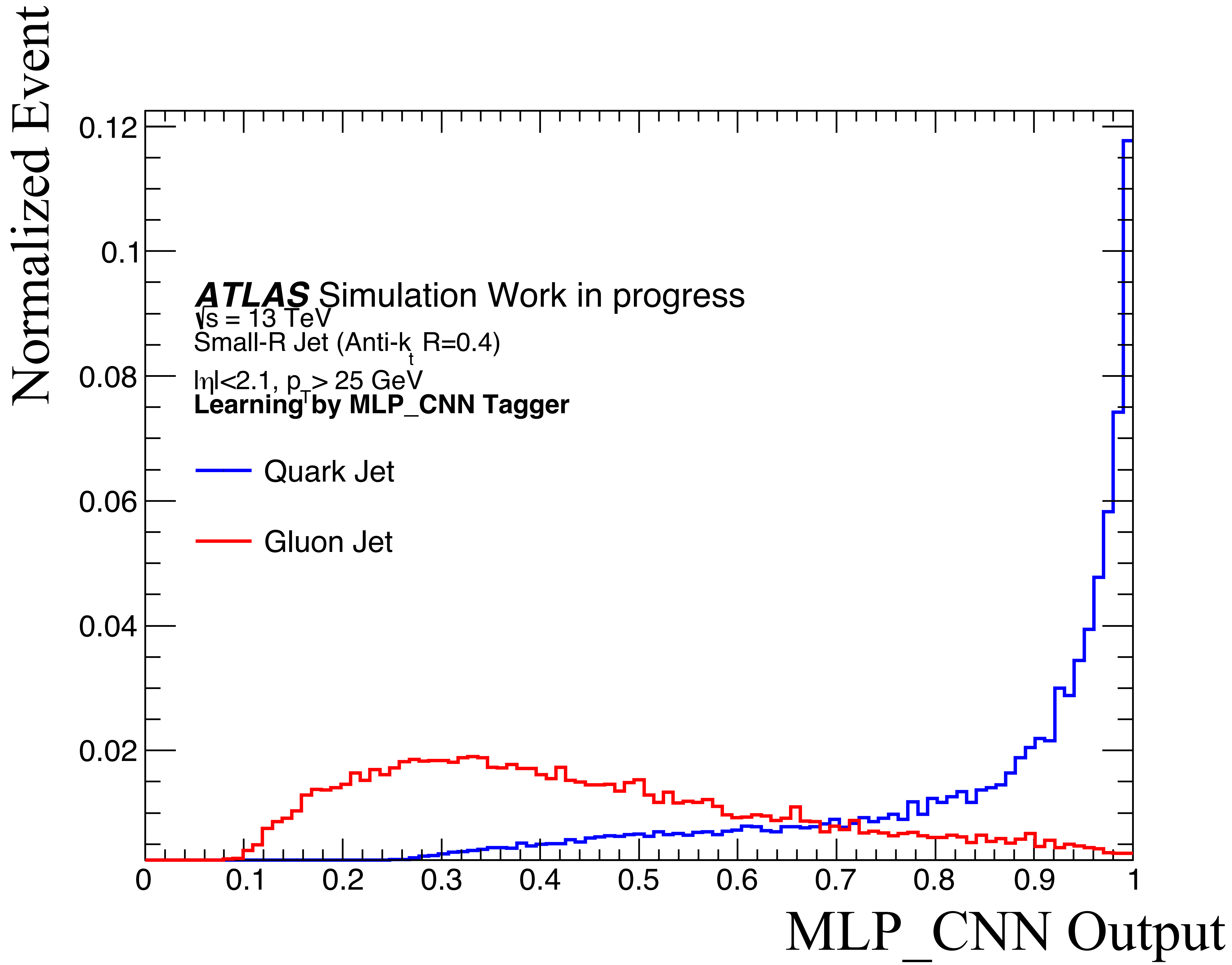
- Pointwise Convolution :
  - The PW Conv model inputting ϕ and η directly rather than in pixel avoids losing other information.
  - It is also convenient to have additional inputs, whether high or low-level variables, such as track charge.



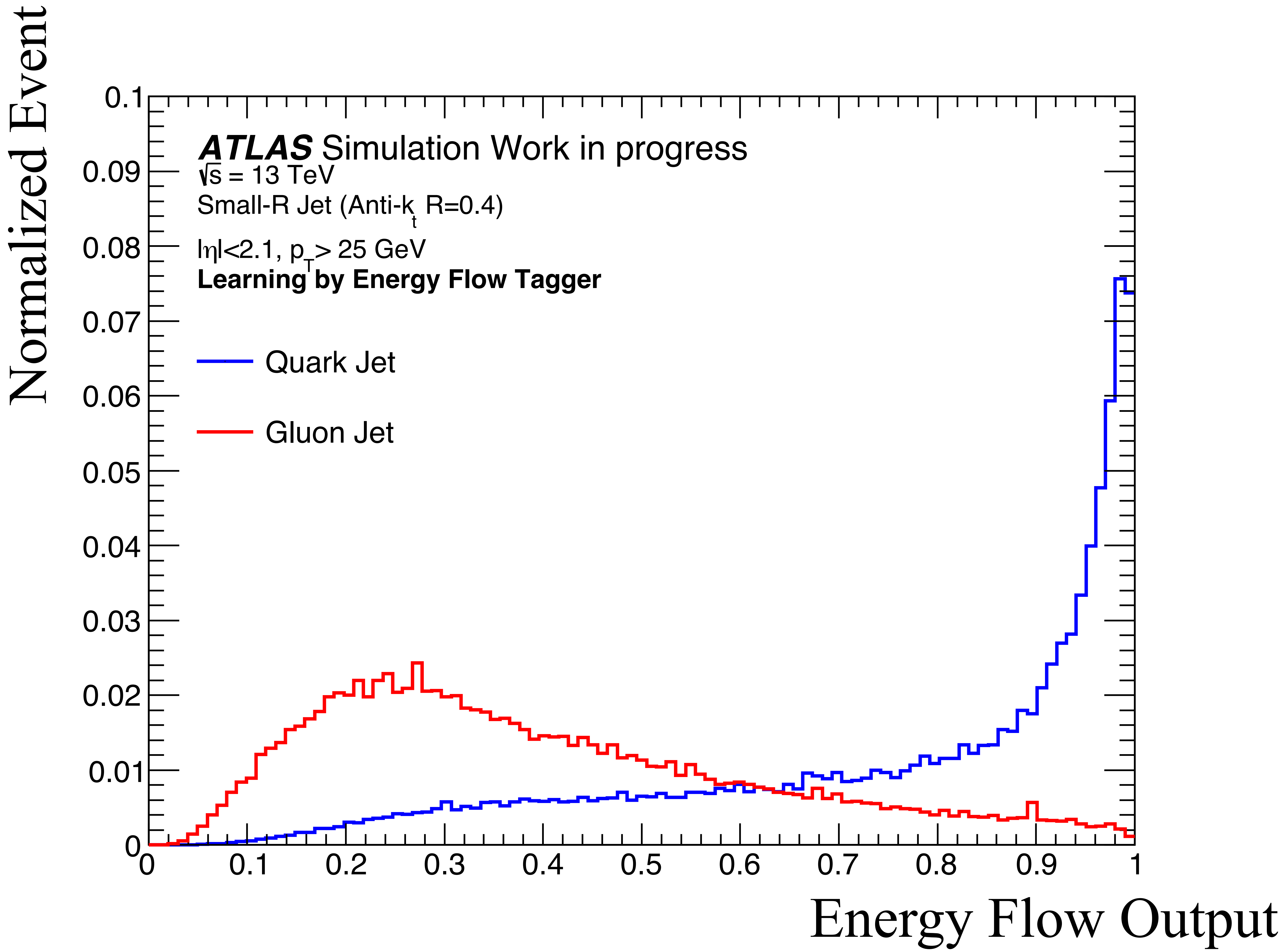
Normalized Event
ATLAS Simulation Work in progress
√s = 13 TeV
Small-R Jet (Anti-kt R=0.4)
|η|<2.1, pT > 25 GeV
Learning by MLP Tagger
Quark Jet
Gluon Jet
0.1
0.08
0.06
0.04
0.02
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
MLP Output



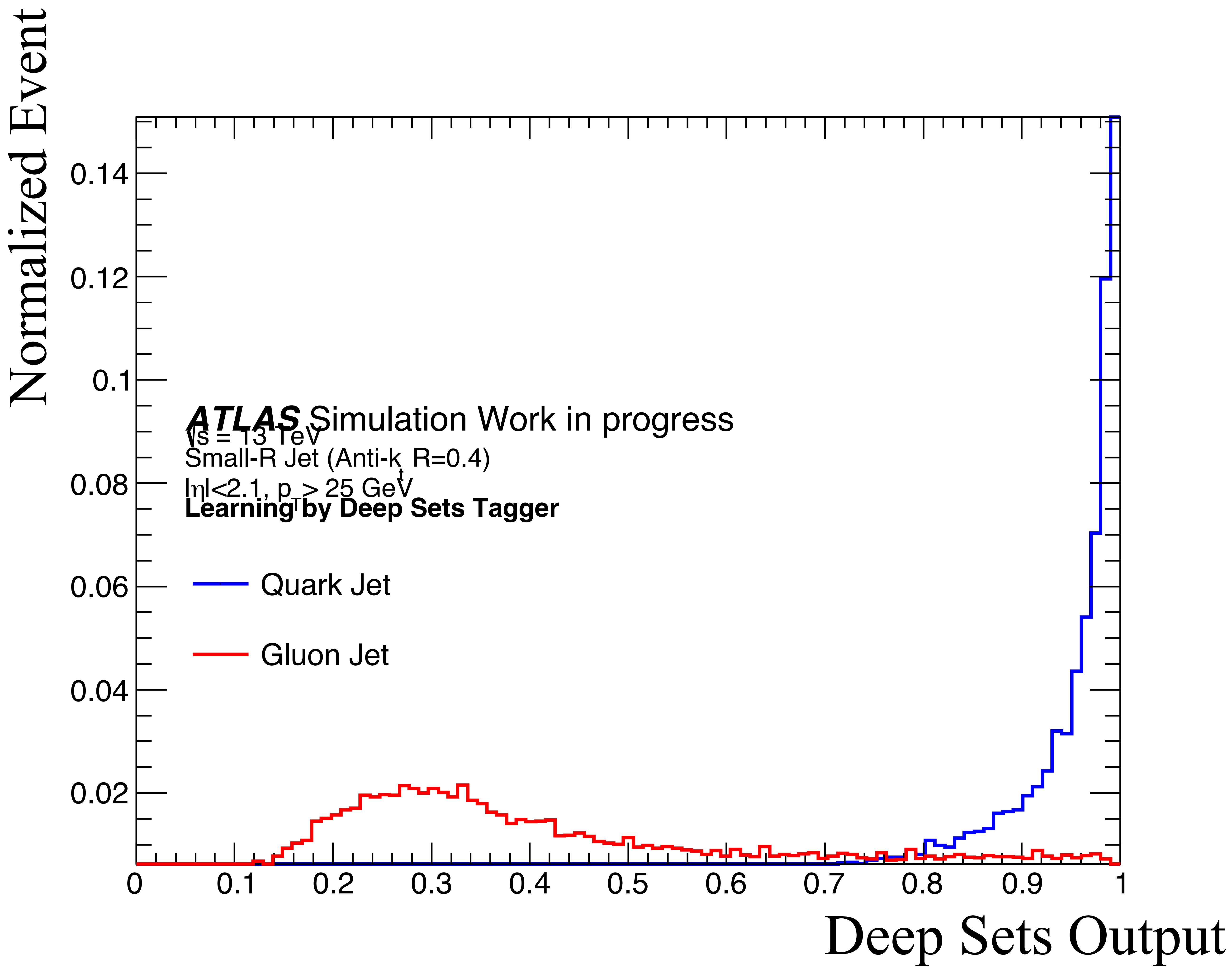
ATLAS Simulation Work in progress
√s = 13 TeV
Small-R Jet (Anti-k_t R=0.4)
|η|<2.1, p_T > 25 GeV
Learning by CNN Tagger
Quark Jet
Gluon Jet
Normalized Event
0.12
0.1
0.08
0.06
0.04
0.02
0.4
0.5
0.6
0.7
0.8
0.9
1
CNN Output



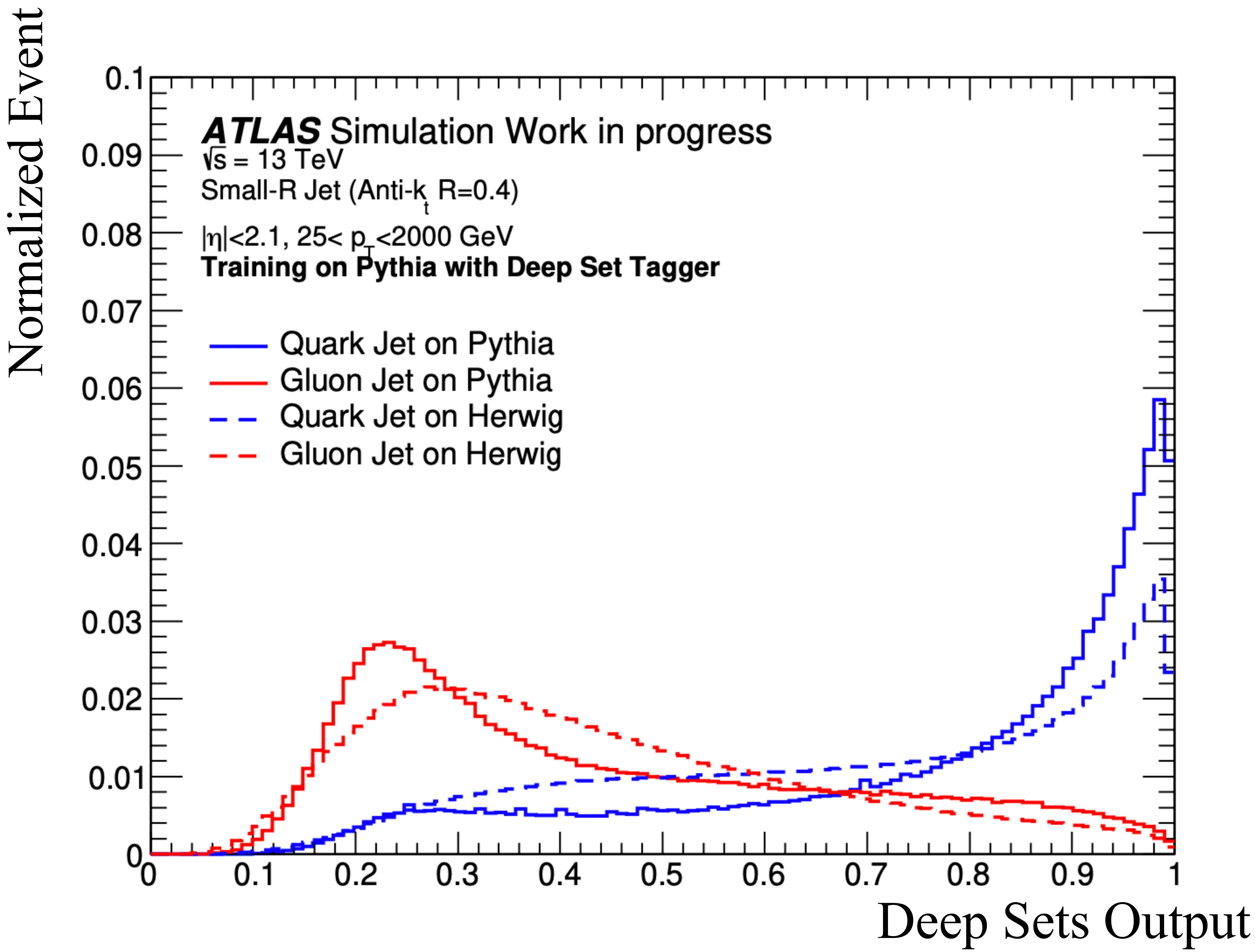
Normalized Event
0.12
0.1
0.08
0.06
0.04
0.02
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
MLP_CNN Output
ATLAS Simulation Work in progress
√s = 13 TeV
Small-R Jet (Anti-k_t R=0.4)
|η|<2.1, p_T > 25 GeV
Learning by MLP_CNN Tagger
Quark Jet
Gluon Jet



ATLAS Simulation Work in progress
√s = 13 TeV
Small-R Jet (Anti-k_t R=0.4)
|η|<2.1, p_T > 25 GeV
Learning by Energy Flow Tagger
Quark Jet
Gluon Jet
Normalized Event
Energy Flow Output


Normalized Event
0.14
0.12
0.1
0.08
0.06
0.04
0.02
ATLAS Simulation Work in progress
√s = 13 TeV
Small-R Jet (Anti-k_t R=0.4)
|η|<2.1, p_T > 25 GeV
Learning by Deep Sets Tagger
Quark Jet
Gluon Jet
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Deep Sets Output

Pythia v.s Herwig Distribution

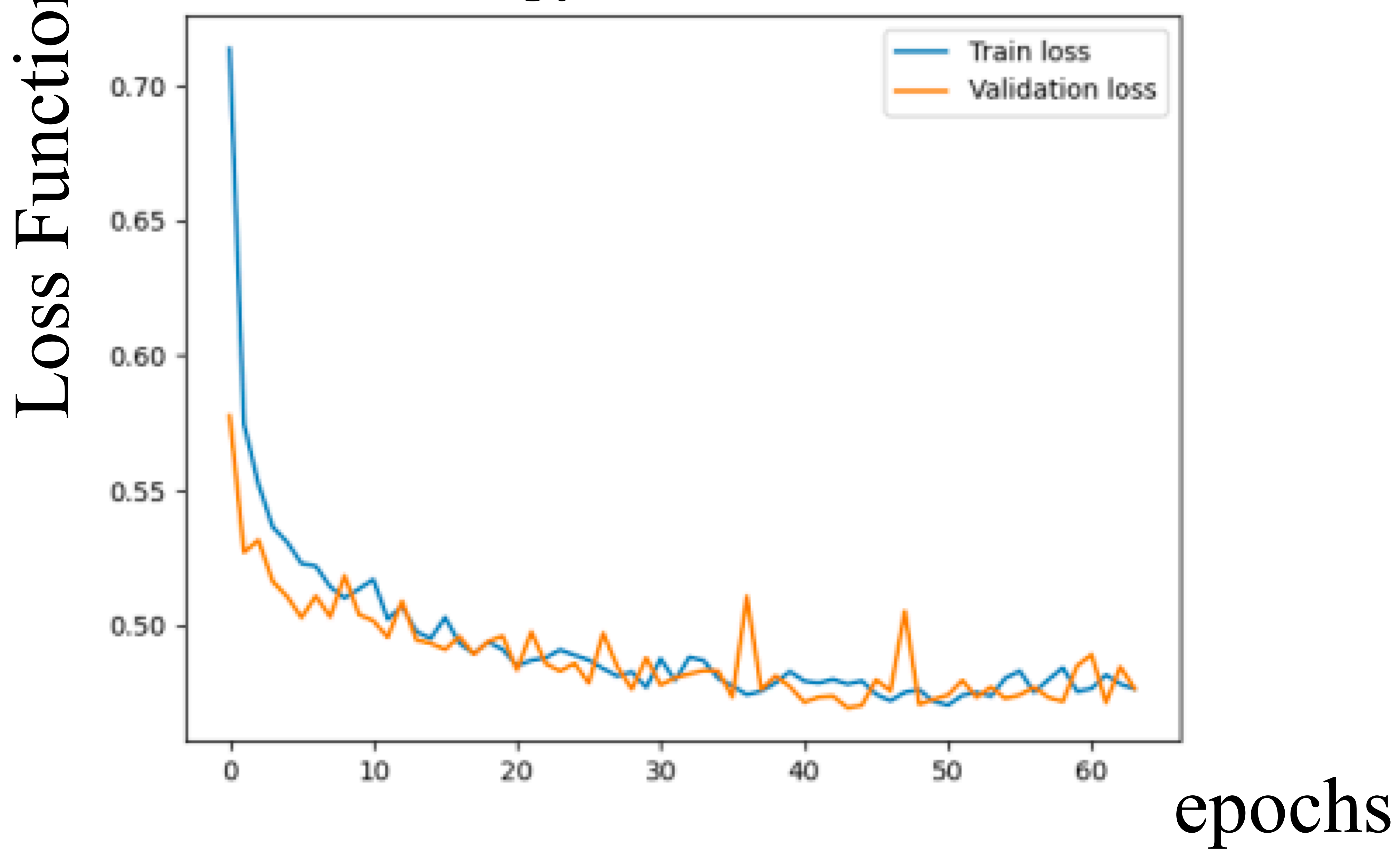


Overtraining

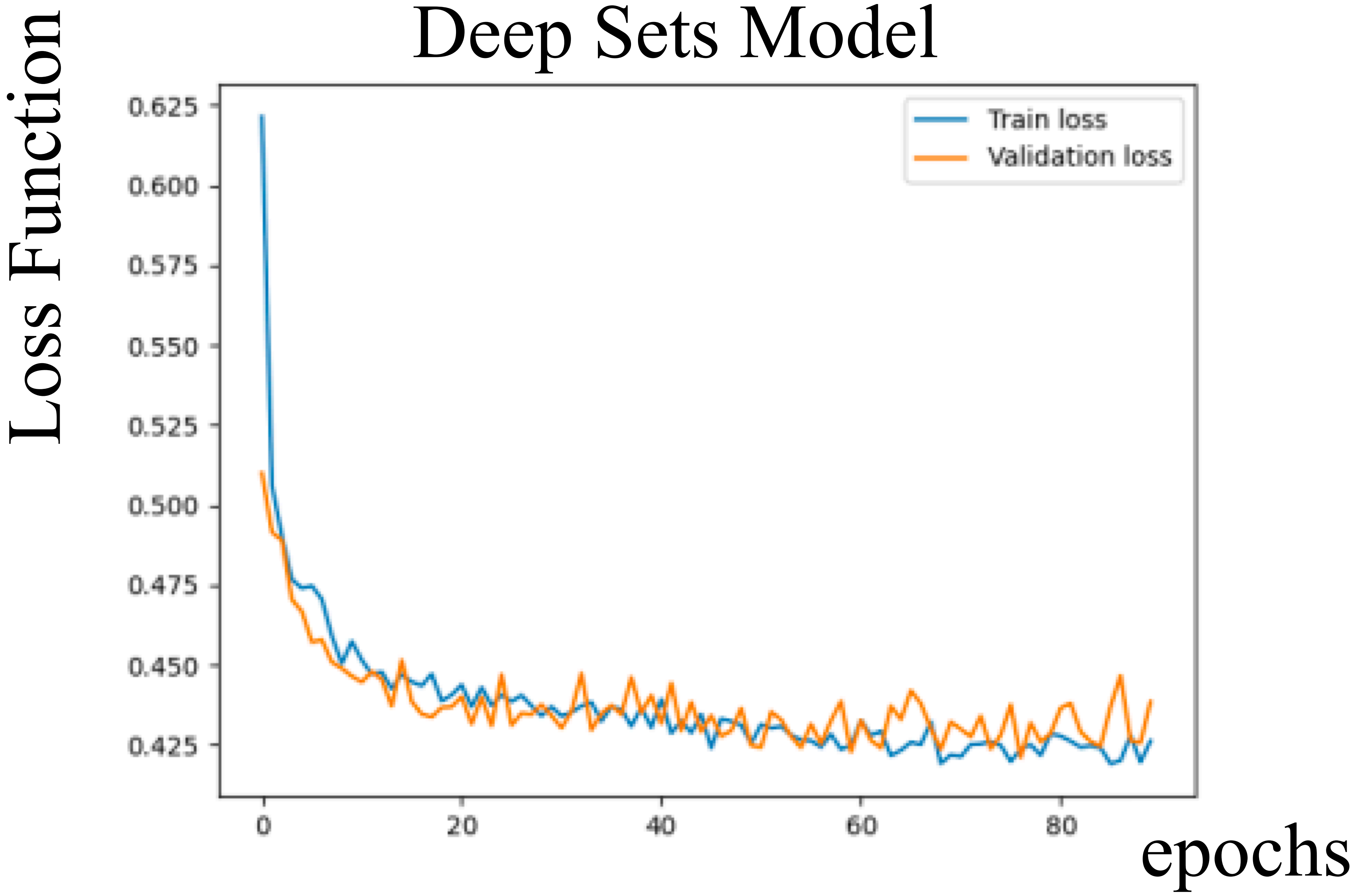
Table 4.3: The setup of neural network models in this thesis.

Activation Function=ReLU
Dropout=0.3
Regularization=L2
Optimizer=Adam
Learning Rate=10^{-5}
Output=Softmax function
Loss function=SparseCategoricalCrossentropy

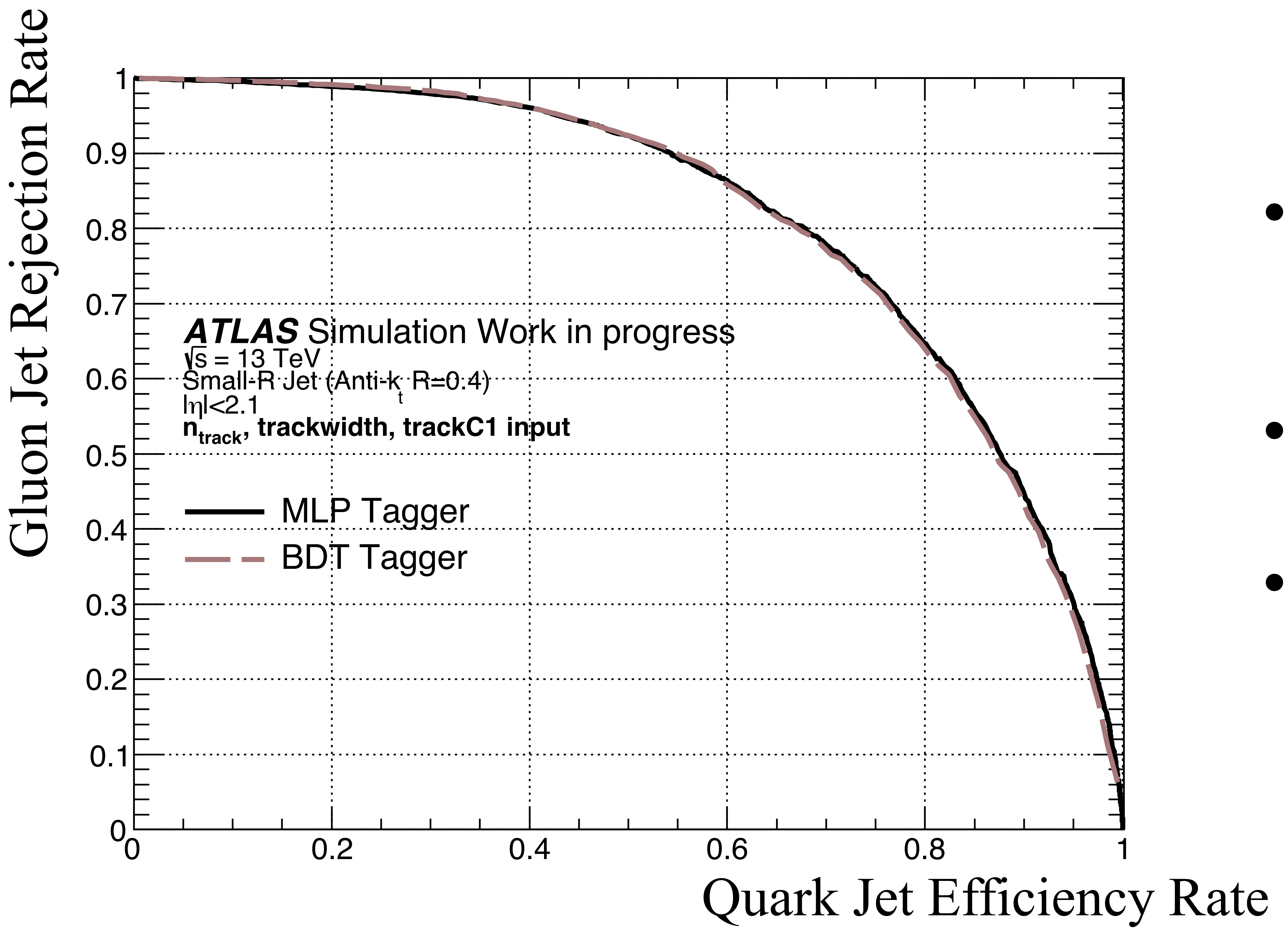
Energy Flow Model



Deep Sets Model

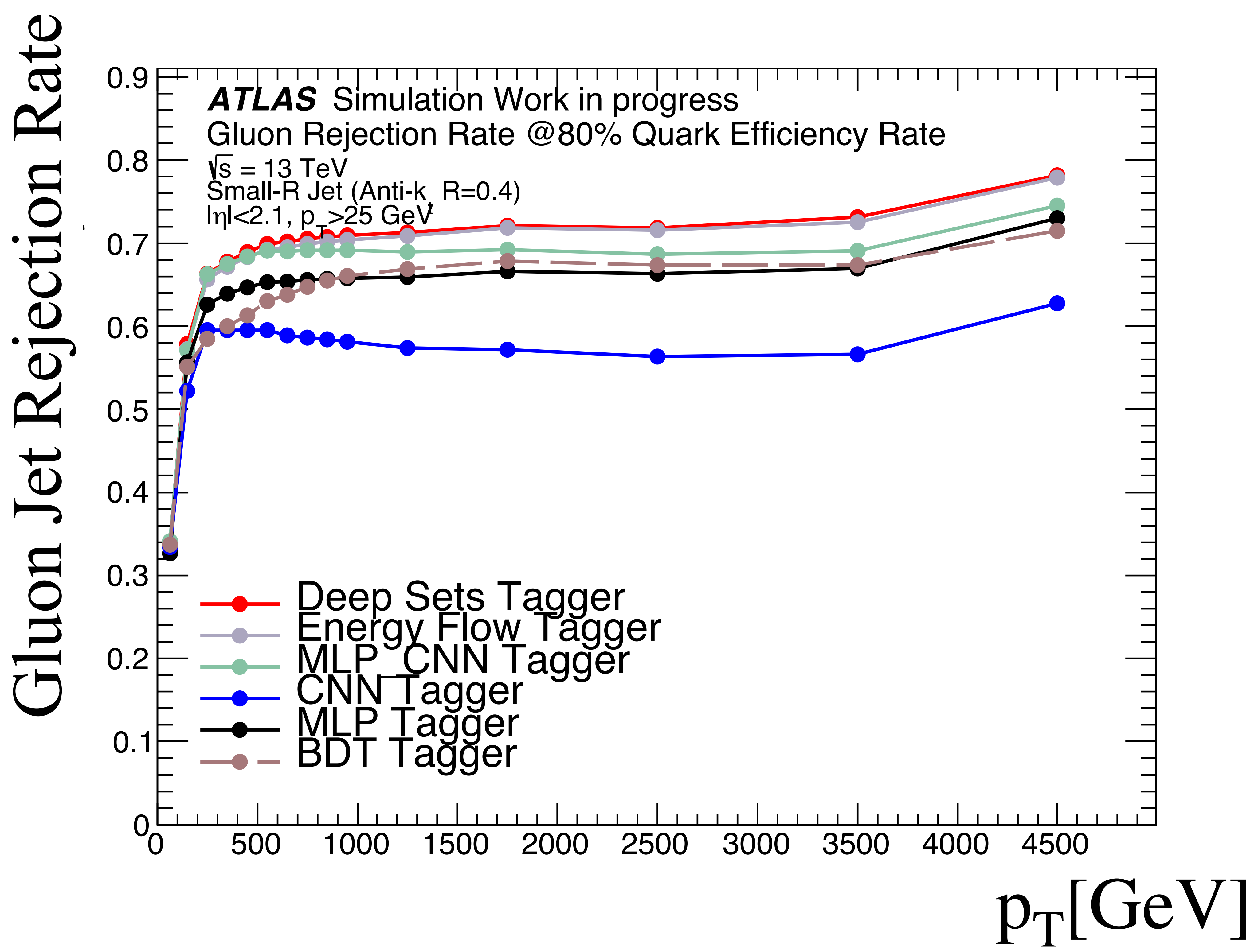


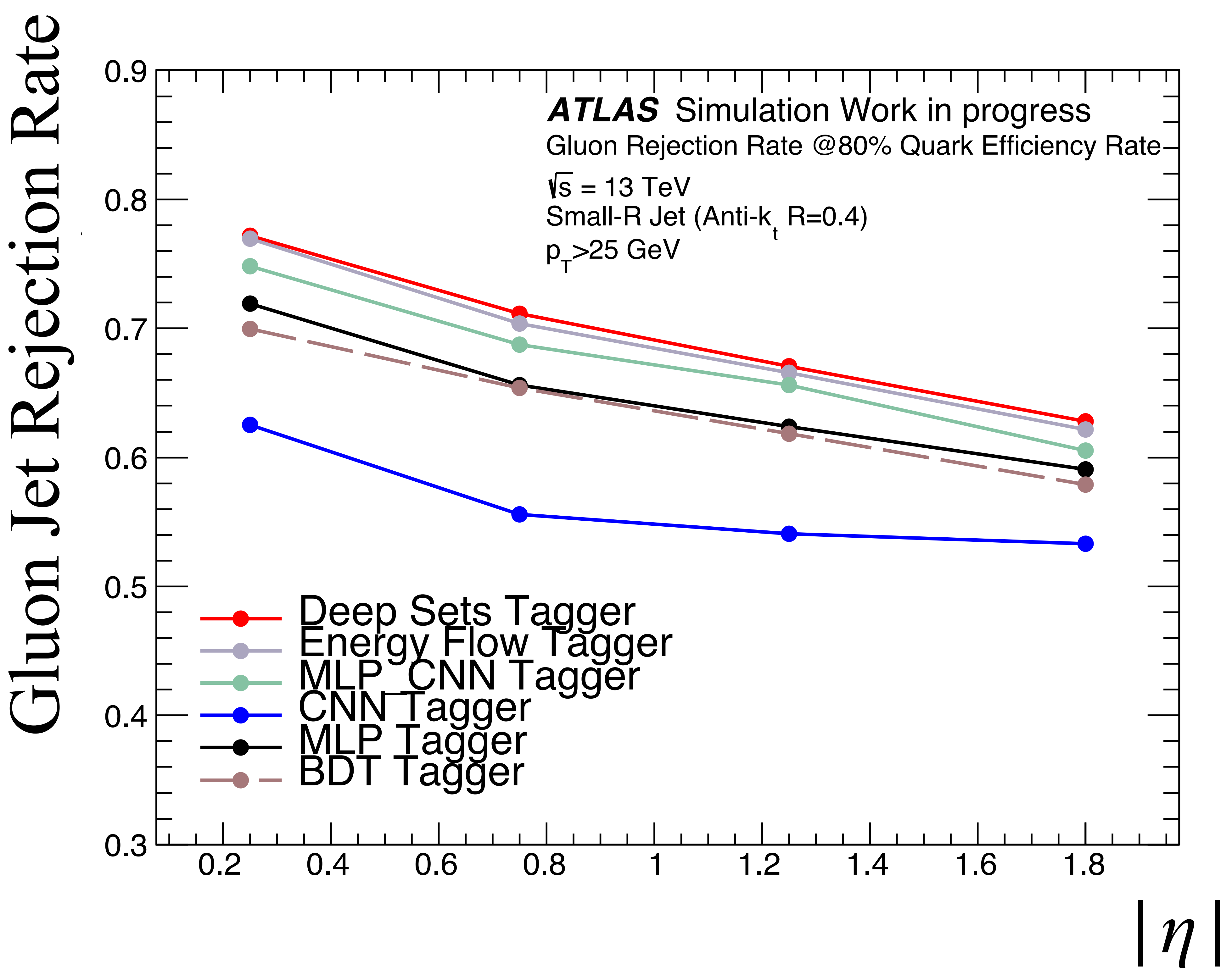
ROC Curve (Receiver Operating Characteristic)



- For comparing the difference between different architecture, the MLP uses the same input variables as BDT.
- The result is shown in the left plot and the difference between them is tiny.
- From this, it is known that the tagger performance depends more on the input variables and the way to tackle the inputs, and not so much on the model architecture.

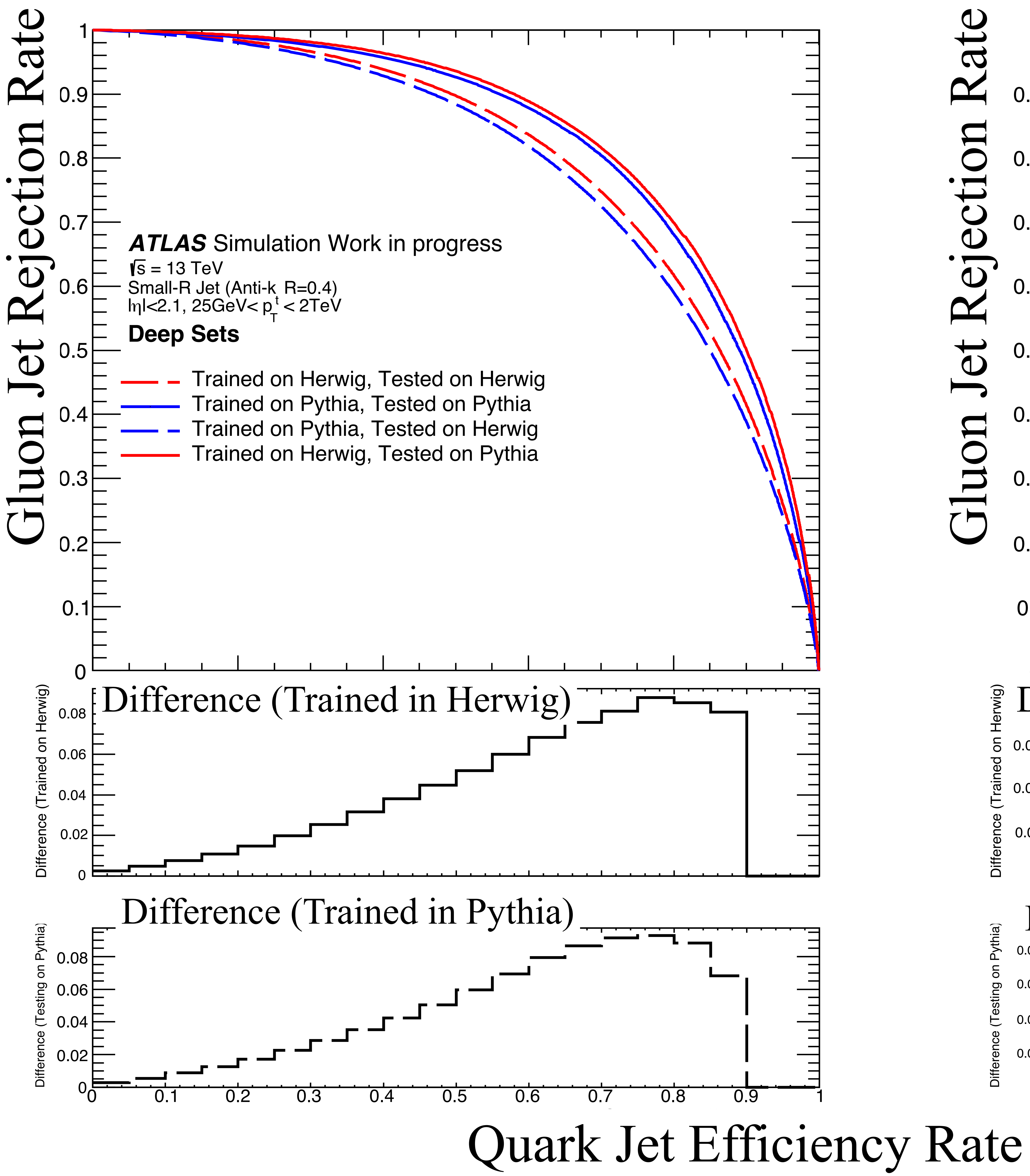
Gluon Rejection Rate @ 80% Quark efficiency Rate

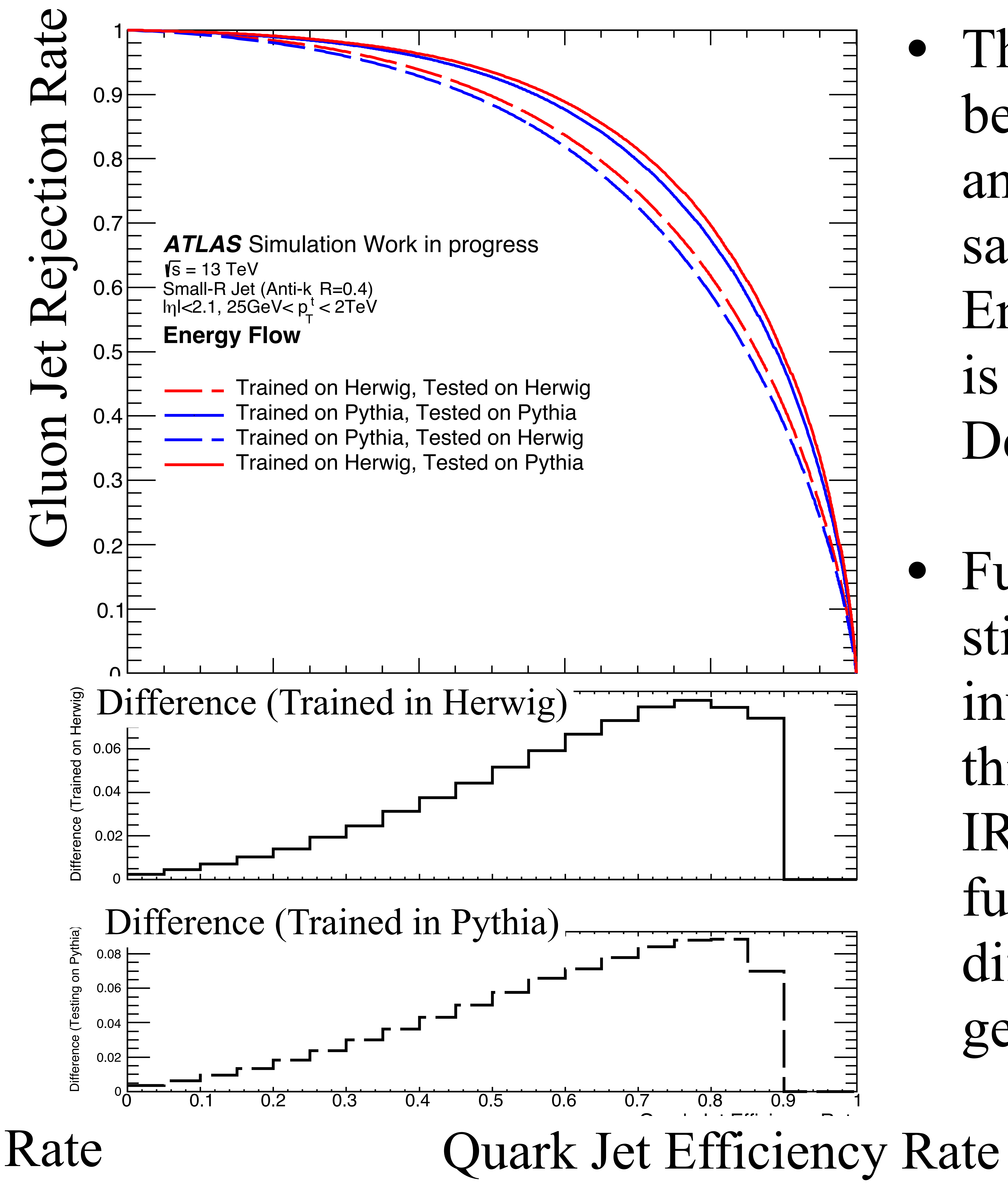




- Worse performance at low p_T region results from the similar track and topo-cluster distribution of low energy Quark/Gluon.
- At lower η, the number of associated jet tracks increases and gives more information, bringing better accuracy and better performance.
- At $p_T < 200$ GeV region, Deep Sets, Energy Flow and MLP_CNN still have better performance than the BDT model.

Pythia vs Herwig





- The difference between the Pythia and Herwig MC samples of the Energy Flow model is smaller than the Deep Sets.
- Further research is still needed to investigate whether this originates from IRC-safe or the fundamental difference between generators.

VBS Process MC Sample

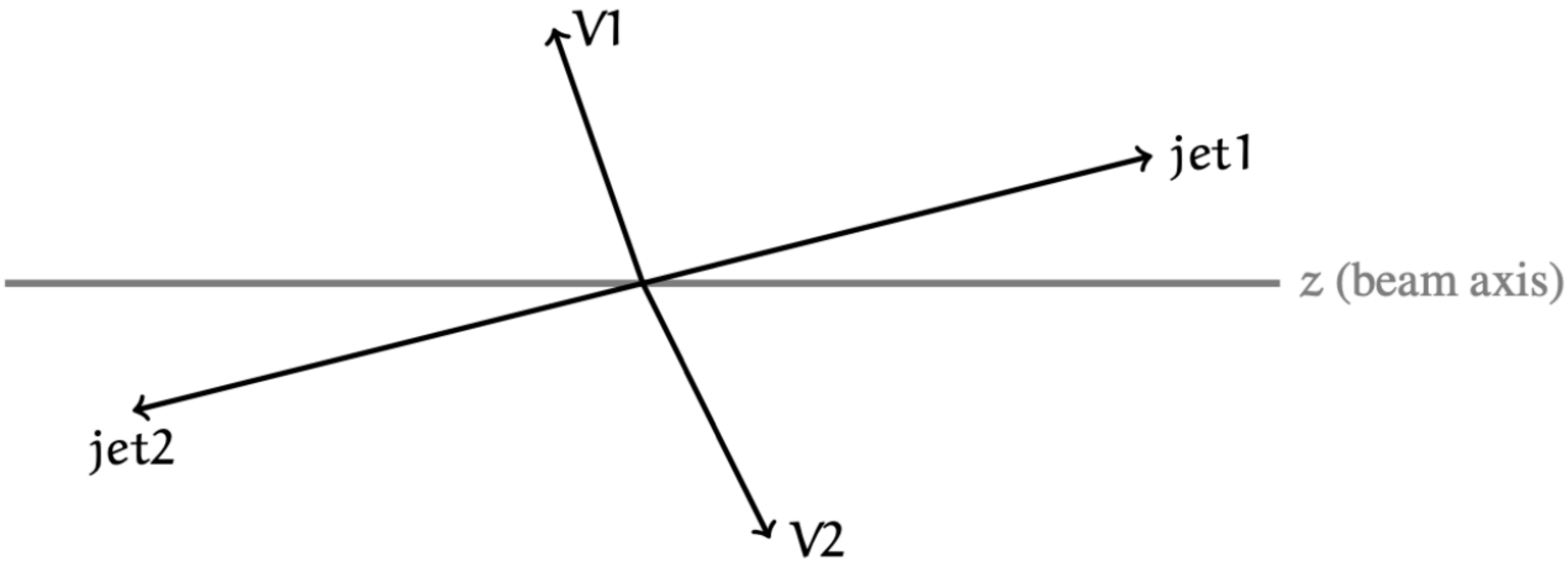


Table 5.1: The MC samples used in Chapter 5.

Process	Generator	Cross section[pb]
VWjj→ $\ell\nu qq$+jj	MadGraph + Pythia8 +EvtGen	2.25
W+jet → $\ell\nu$ + jet	SHERPA 2.2.1	6.16×10^4
$t\bar{t} \to \ell\nu qq$bb	POWHEG + Pythia8 + EveGen	396.89

- The validation focuses on the 1-lepton channel of the Semileptonic Vector Boson Scattering MC sample.
- There are two signal events considered, $WW/WZ \to \ell\nu qq$+jj.
- The background events are $W + \text{jet} \to \ell\nu jj + jj$ and $t\bar{t} \to \ell\nu qq bb$.

Physics Object

Table 5.2: The electron object definition used in this validation.

Cut	Selection
p_T	$p_T > 30$ GeV
η	$\|\eta\| < 2.47$
Track to Vertex Association	$\|d_0^{BL}(\sigma)\| < 5$, $\|\Delta z_0^{BL} \sin\theta\| < 0.5$ mm
Identification	ElectronID = TightLH

Table 5.3: The Muon object definition used for the MC sample in this chapter.

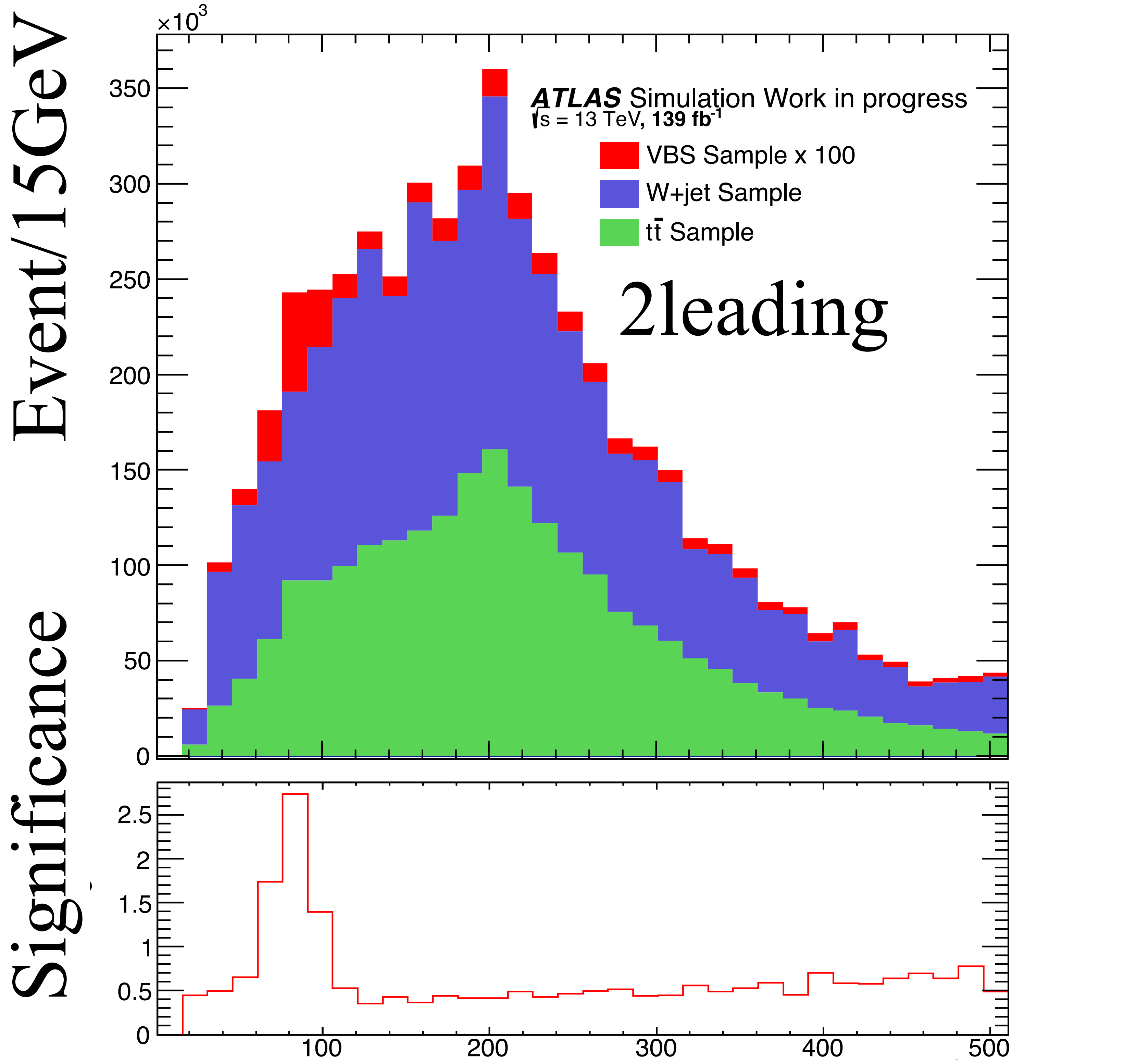
Cut	Selection
p_T	$p_T > 30$ GeV
η	$\|\eta\| < 2.5$
Track to Vertex Association	$\|d_0^{BL}(\sigma)\| < 3$, $\|\Delta z_0^{BL} \sin\theta\| < 0.5$ mm
Identification	MuonQuality = Medium

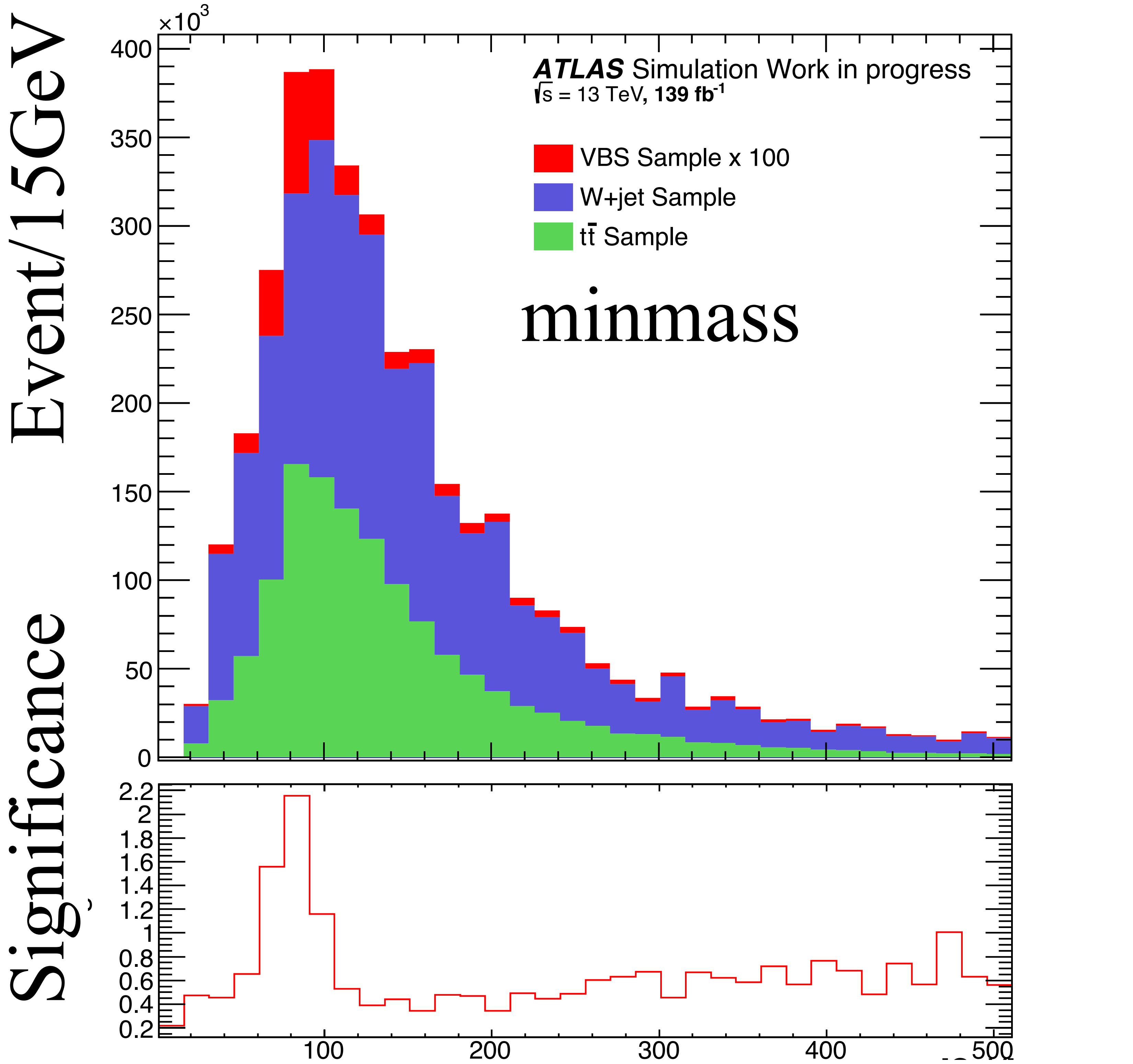
Physics Object

Table 5.4: The Small-R Jet object definition used for the MC sample in this chapter. EMPFlow represents the particle flows (small-R jets here) are reconstructed with electromagnetic scale topo-cluster.

Cut	Selection
Algorithm	anti-k_T (R=0.4)
Input Constituent	EMPFlow
p_T	p_T >20 GeV
η	$\|\eta\| < 4.5$
JVT	> 0.95 for $p_T < 120$ GeV and $\|\eta\| < 2.4$ > 0.11 for $p_T < 120$ GeV and $2.4 < \|\eta\| < 2.5$ (Medium Working Point)

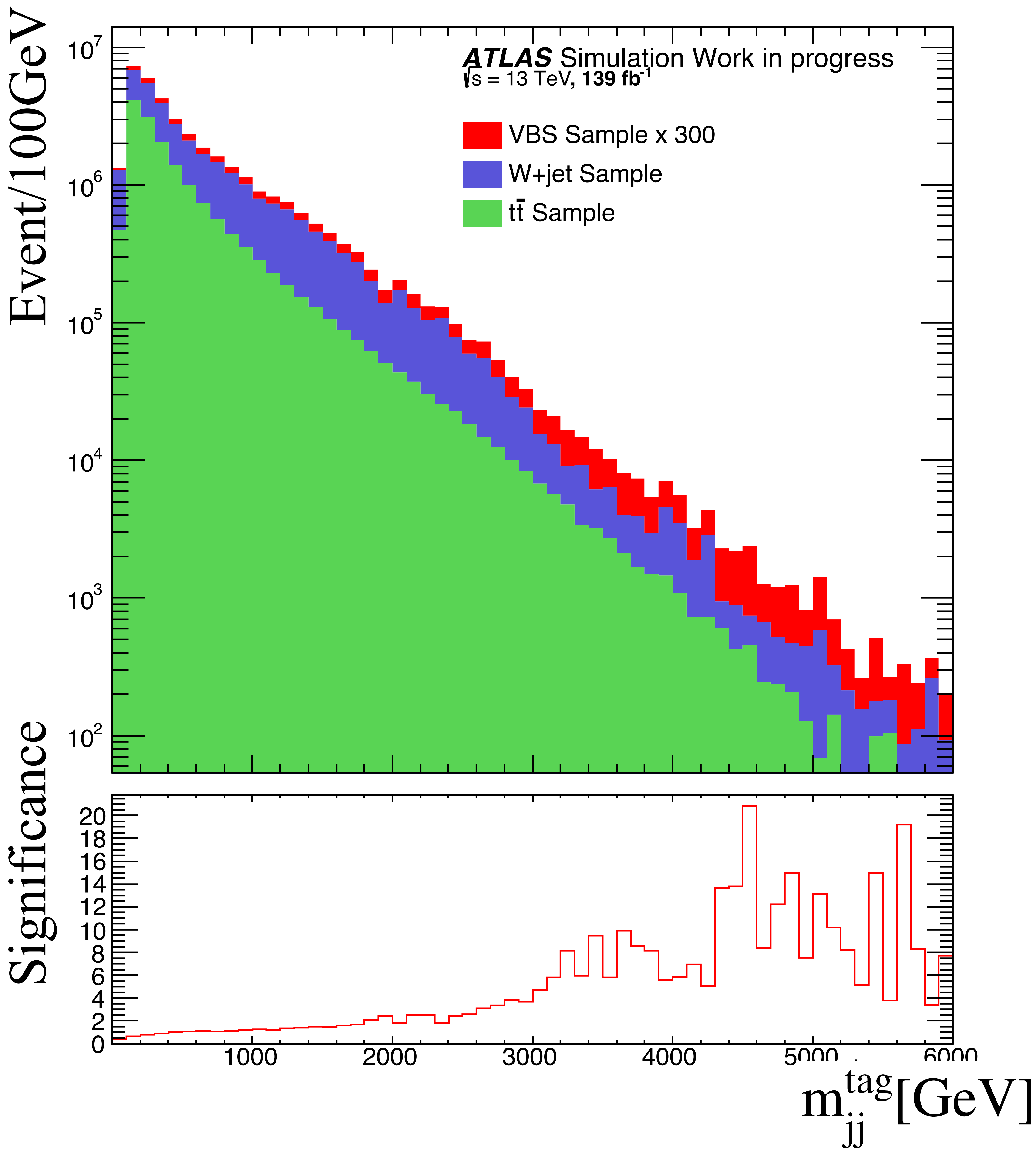
Signal jets pair



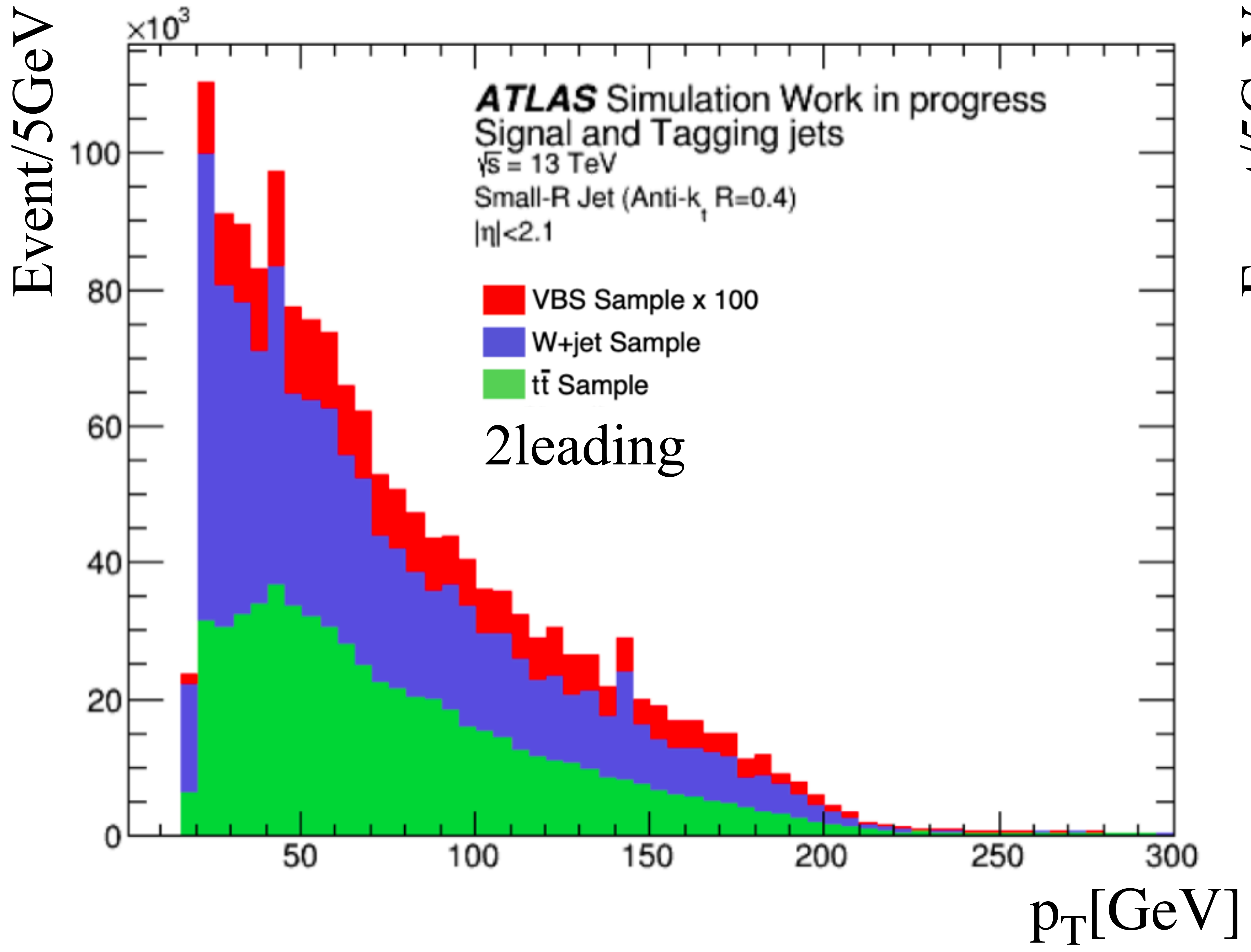


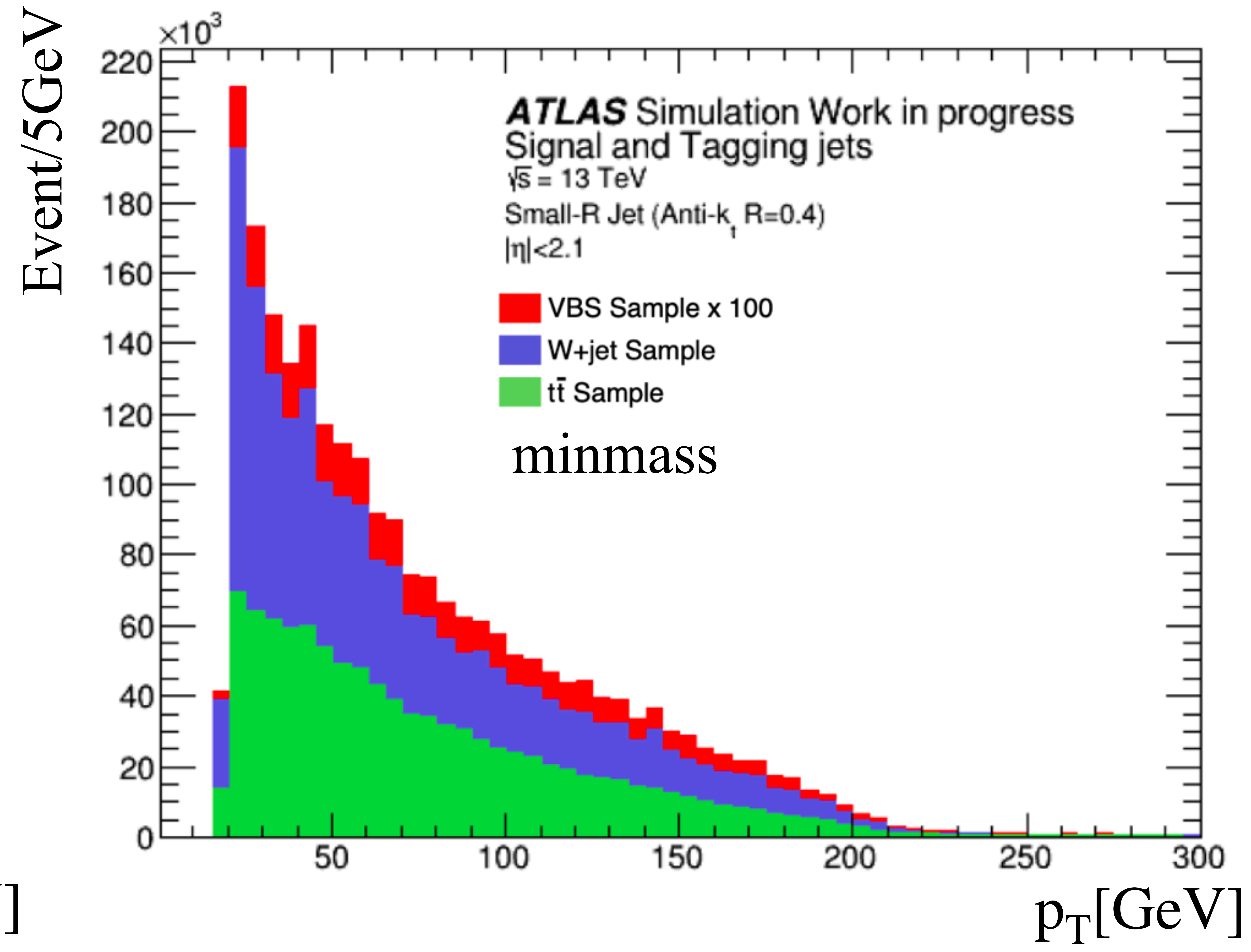
- 2leading strategy :
  - It has a larger probability of selecting the jets from $t\bar{t}$ samples but with fewer background events.
  - There are 1059 VBS signal events, 297845 W+jet events, and 235519 $t\bar{t}$ events.
- minmass strategy :
  - This makes sure more jets decay from W/Z vector bosons, but the significance is worse.
  - There are 1419 VBS signal events, 454939 W+jet events, and 409285 $t\bar{t}$ events.

m_{jj}^{tag} Distribution

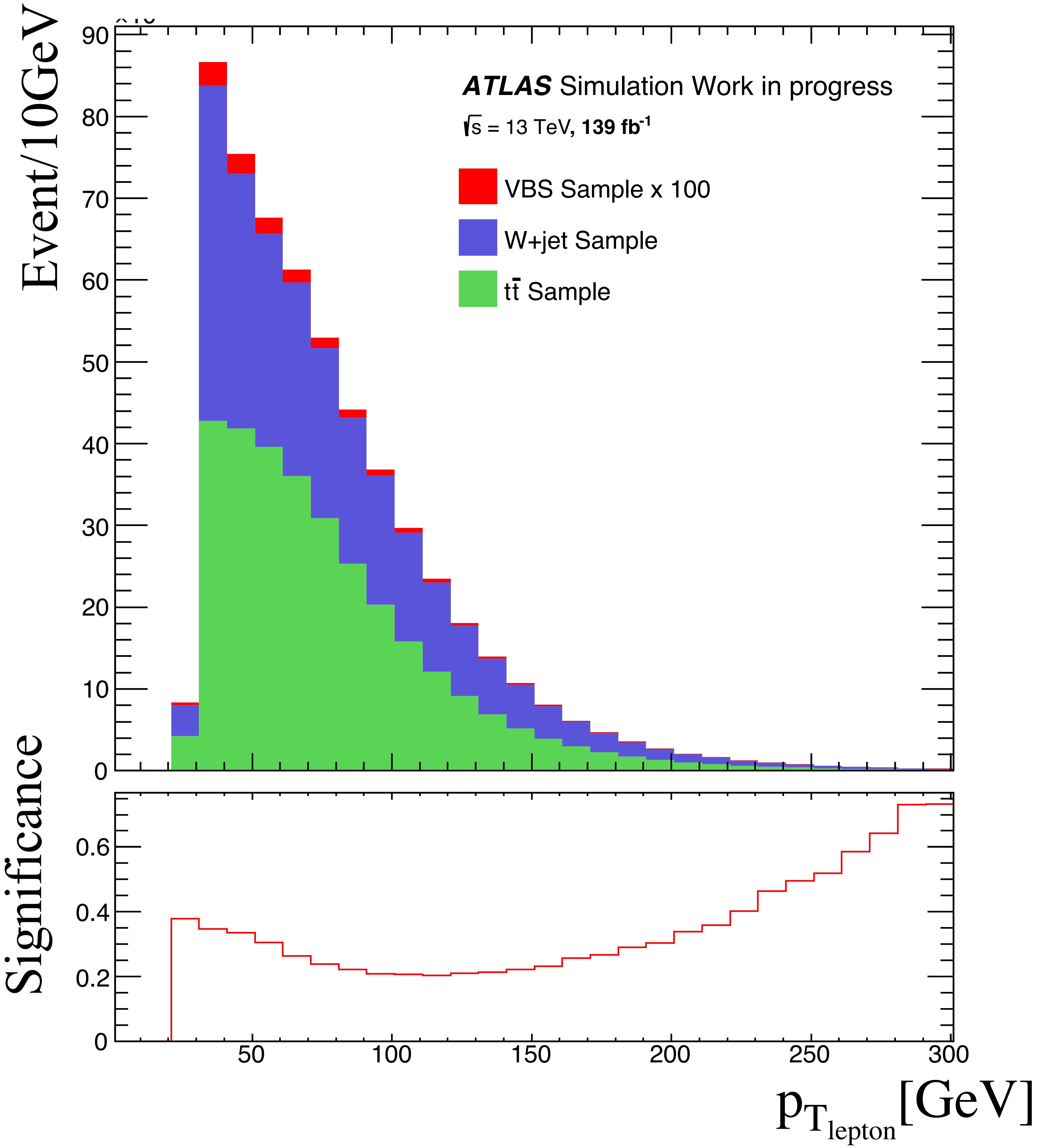


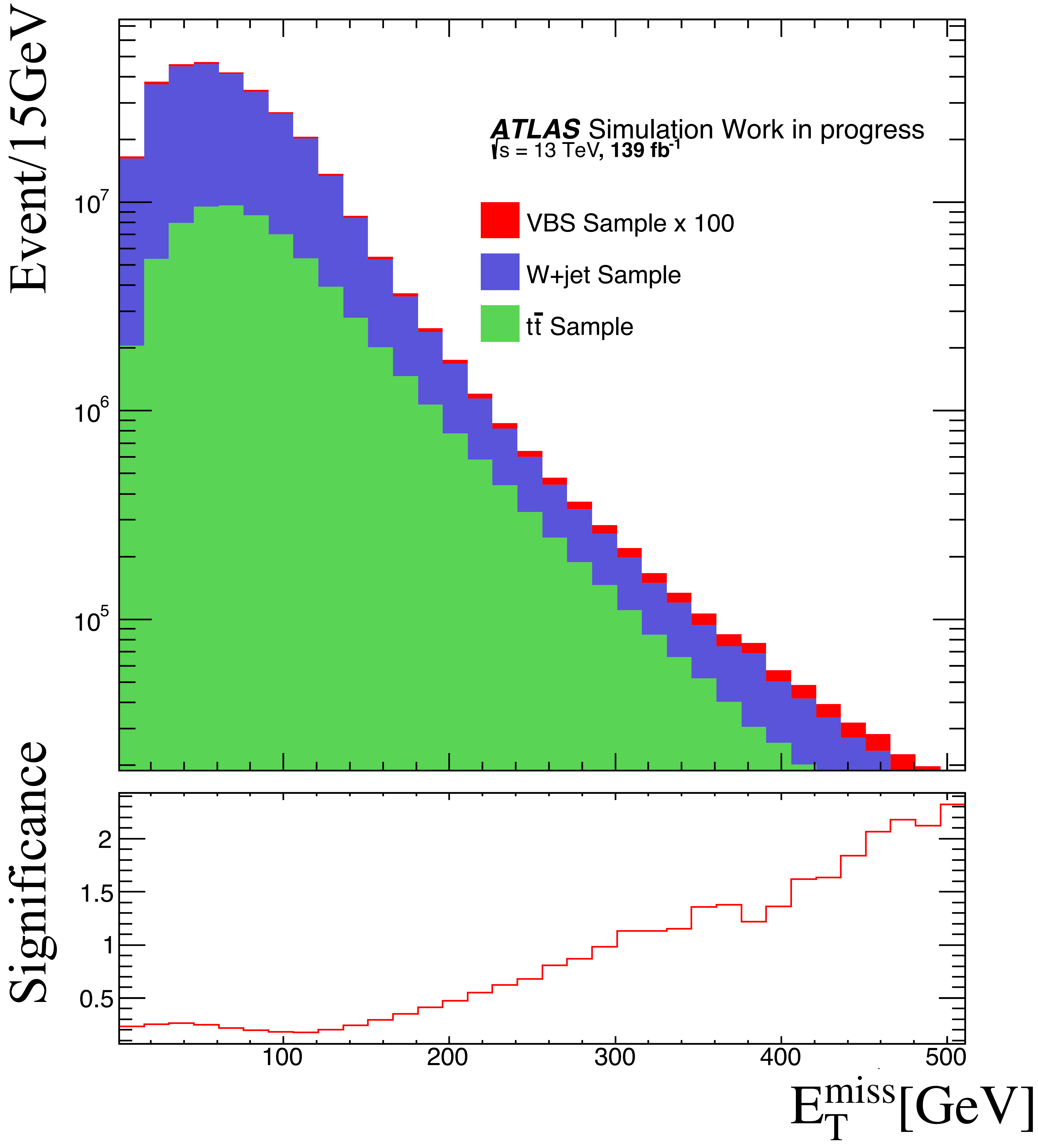
Jet p_T Distribution



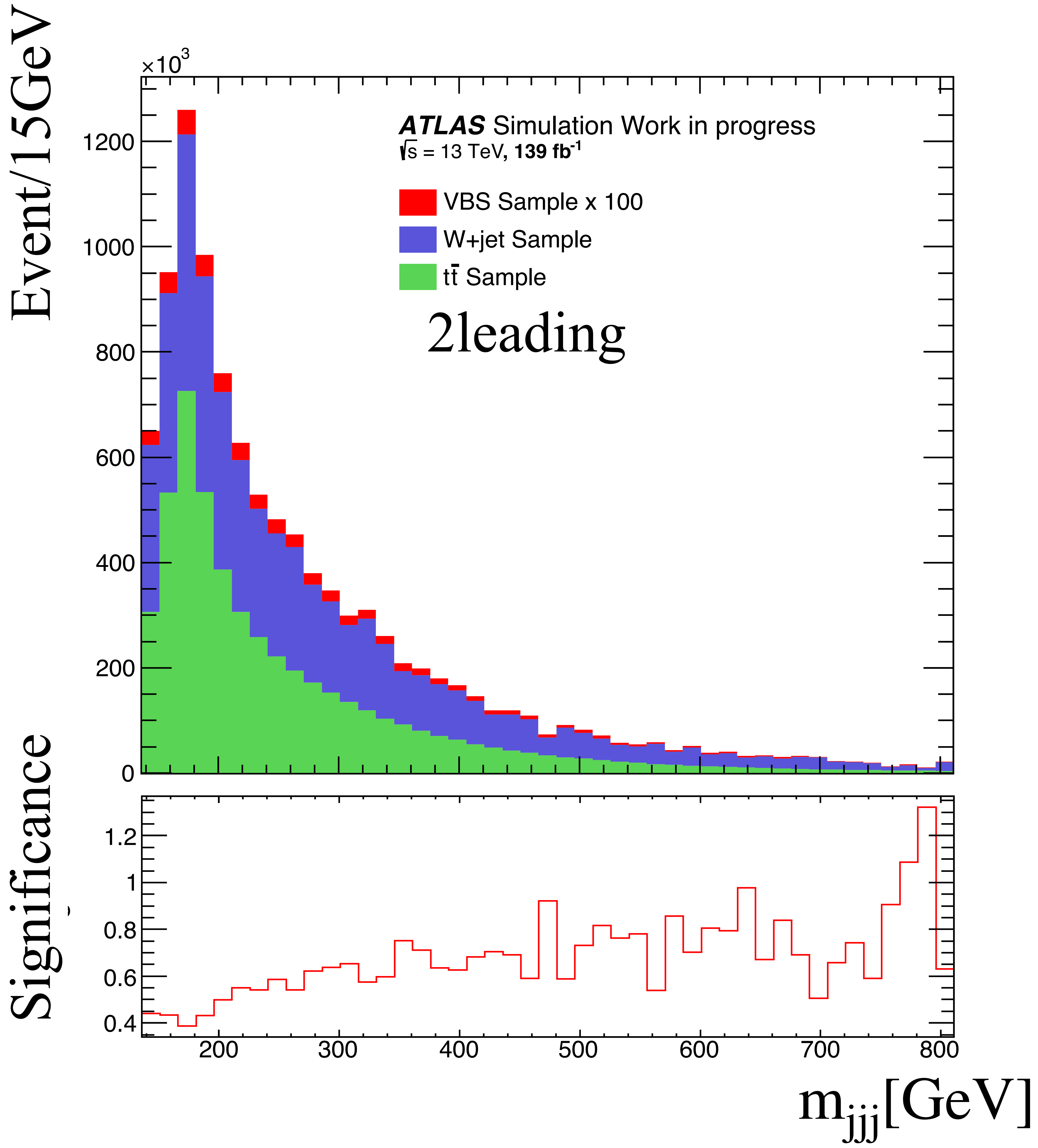


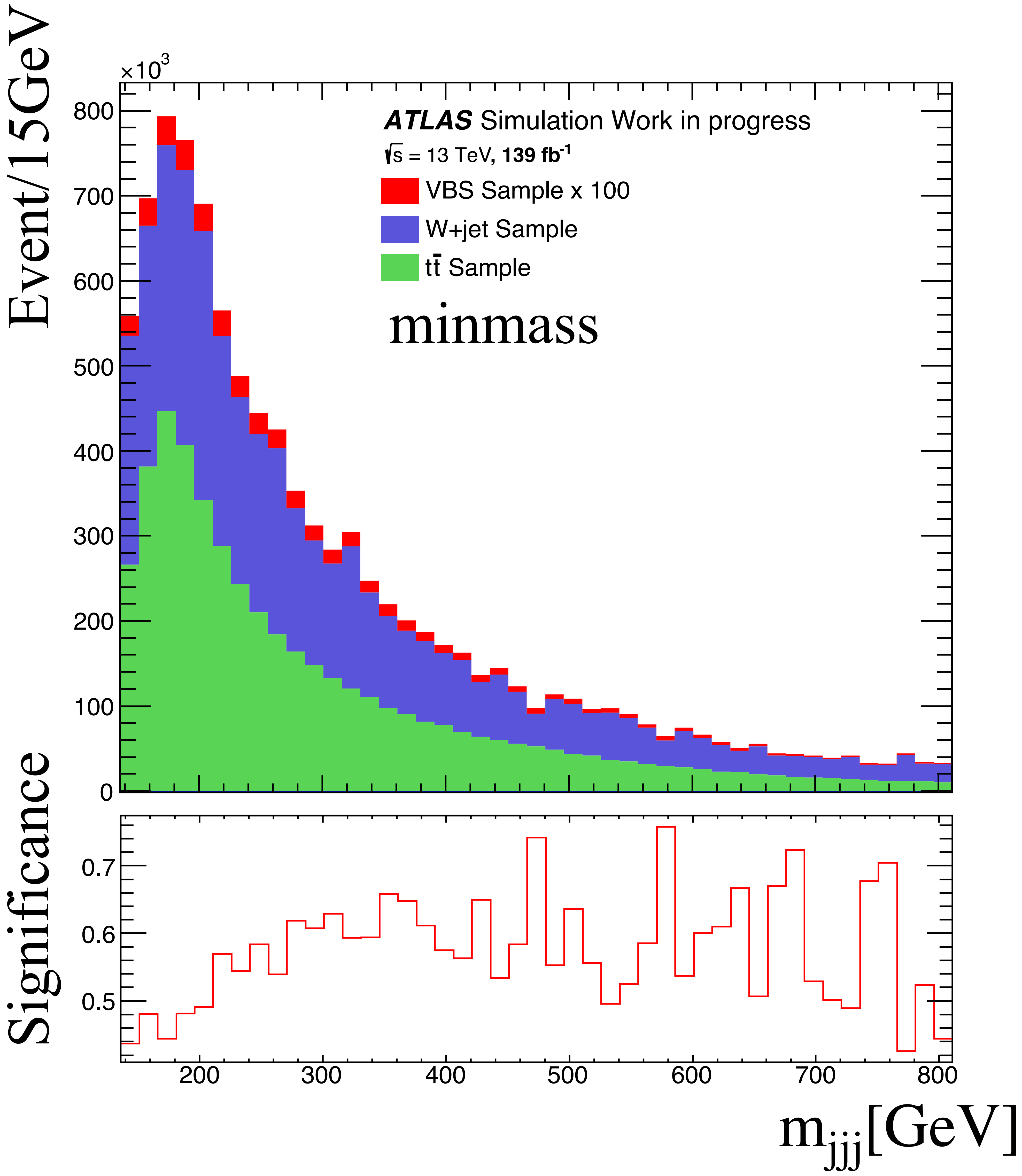
Lepton p_T and E_T^{miss} Distribution



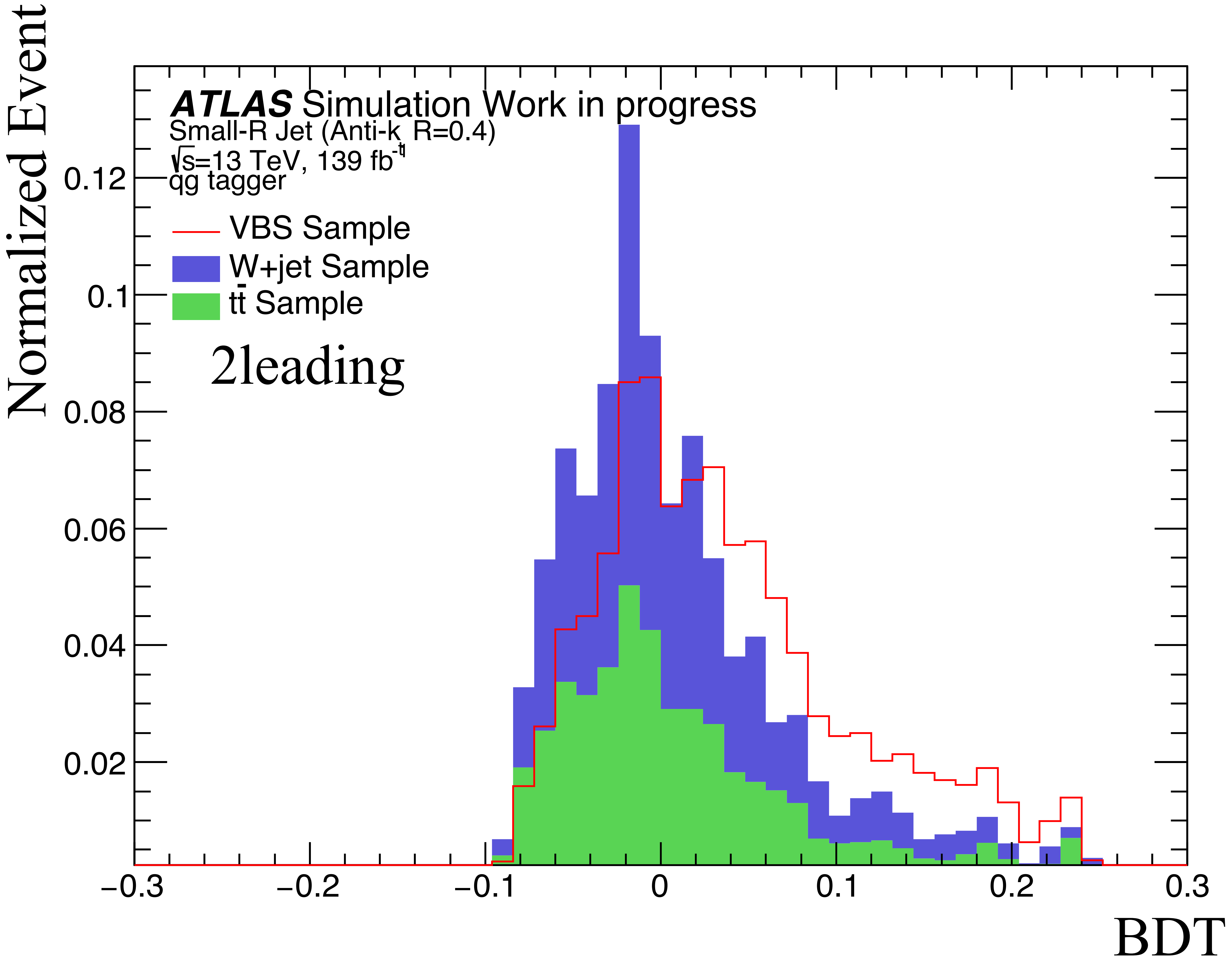


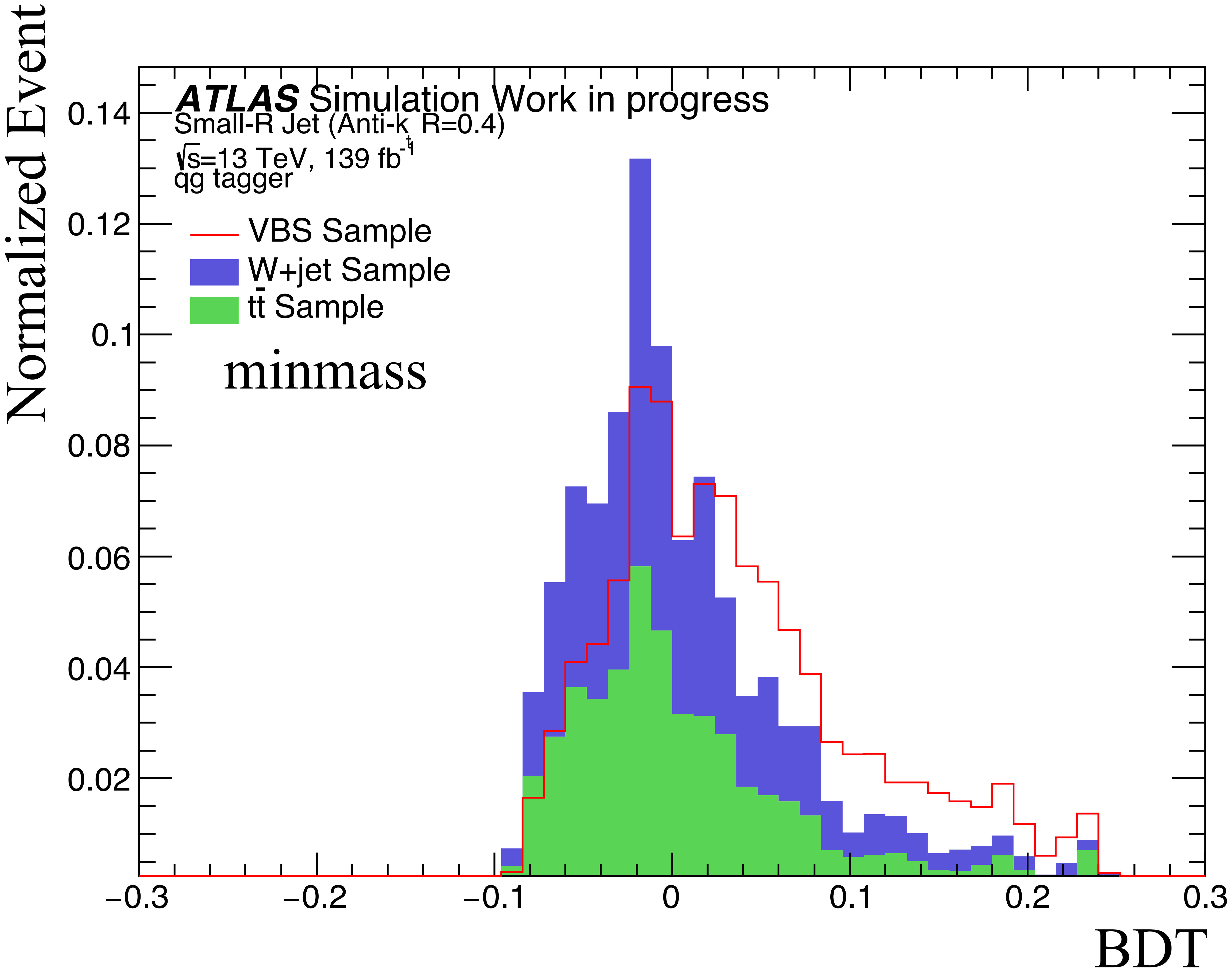
m_{jjj} Distribution





BDT q/g tagger Distribution





Selection

Objects	Cuts	threshold
$W \to \ell\nu$	Number of Tight leptons	1
	Number of Loose leptons	0
	E_T^{miss}	> 80 GeV
	$p_T(\ell)$	> 30 GeV
tagging jets pair	tagging jets p_T	> 30 GeV
	m_{jj}^{tag}	> 400 GeV
Signal jets pair	Number of signal jets	≥ 2
	p_T(signal jet)	> 20 GeV for $\lvert\eta\rvert < 2.1$ > 30 GeV for $2.1 < \lvert\eta\rvert < 4.5$
	Leading jet p_T	> 40 GeV
Others	Signal Region	$64 < m_{jj} < 106$ GeV
	Number of additional b-tagged jets	0
	m_{jjj}	> 220 GeV

- A tagging represents the jet coming from partonic quarks inside the collision protons.
- A signal jet stands for the jet originating from the vector boson decay.
  - 2leading strategy: The signal jets pair is selected by choosing the two highest p_T jets other than tagging jets.
  - minmass strategy: The signal jets pair is selected by choosing two jets other than tagging jets that have the closest invariant mass from W/Z boson mass.
- The invariant mass of two signal jets and a specific jet is required as $m_{jjj} > 220$ GeV, where the specific jet is selected by calculating the closest invariant mass m_{jjj} from the top quark mass.

Discovery Signicance

- Discovery significance is an estimating value to describe the opportunity that the observation is not from the background fluctuation and is believable.
- The significance in this study was calculated by **BinomialExpZ** in **RooStats**.

$$Z(N_s, N_b, \delta_b) = \sqrt{2}\,\mathrm{erf}^{-1}(1 - 2p),$$

where the p-value is $$p = \int_0^\infty \mathrm{d}b\; N(b; N_b, \delta_b N_b) \sum_{i=N_s+N_b}^{\infty} P(i; b),$$

where N and P are Gaussian and Poisson distribution respectively.

Improving Results

Selection stratgy	Bkg systematic uncertainty	Signal	W+jet	$t\bar{t}$	Significance
2leading	10%	44	42±14	128±38	0.78
	20%	44	42±14	128±38	0.58
	30%	37	30±14	99±43	0.40
minmass	10%	45	57±18	103±39	0.73
	30%	45	57±20	103±43	0.57
	30%	45	57±24	103±49	0.39
2leading with perfect q/g tagger	10%	65	14±6	134±42	1.11
	20%	65	14±7	134±48	0.91
	30%	60	13±7	114±51	0.70
minmass with perfect q/g tagger	10%	71	13±7	168±42	1.19
	20%	53	3±4	99±37	0.97
	30%	53	3±4	99±43	0.76
neural 2leading	10%	48	46±13	148±41	0.81
	20%	48	46±15	148±48	0.60
	30%	34	18±10	89±40	0.42
neural minmass	10%	65	83±22	242±58	0.77
	20%	35	33±15	79±34	0.59
	30%	22	1±1	40±23	0.44

Compared to baseline significance, improving by

56.9%

76.4%

3.9%

8.2%

- The better improving rates indicate that the minmass strategy has more space to improve.
- Compared to baseline analysis, the q/g tagger introduced in this study improves by about 5%.

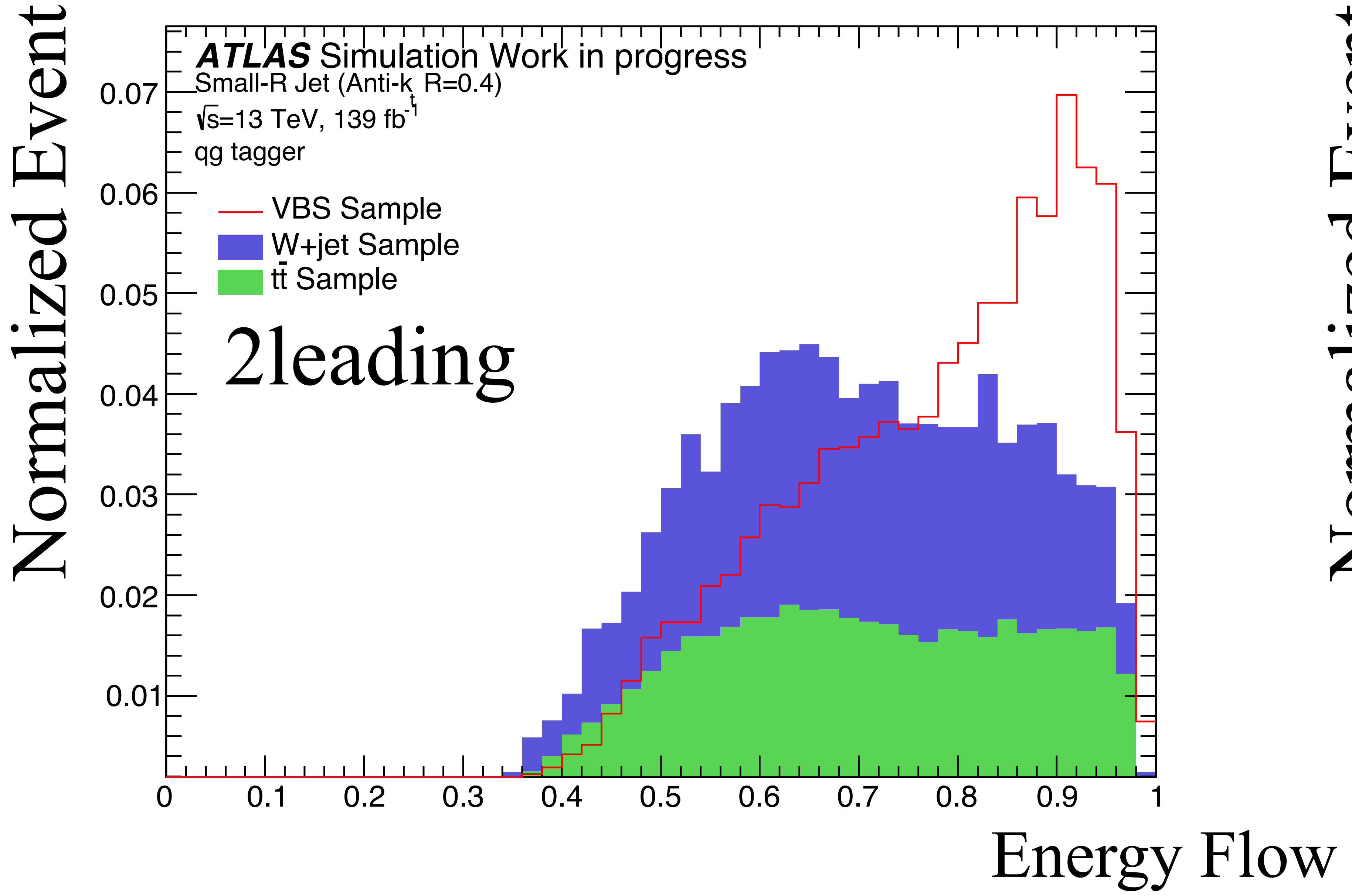
Order Strategy

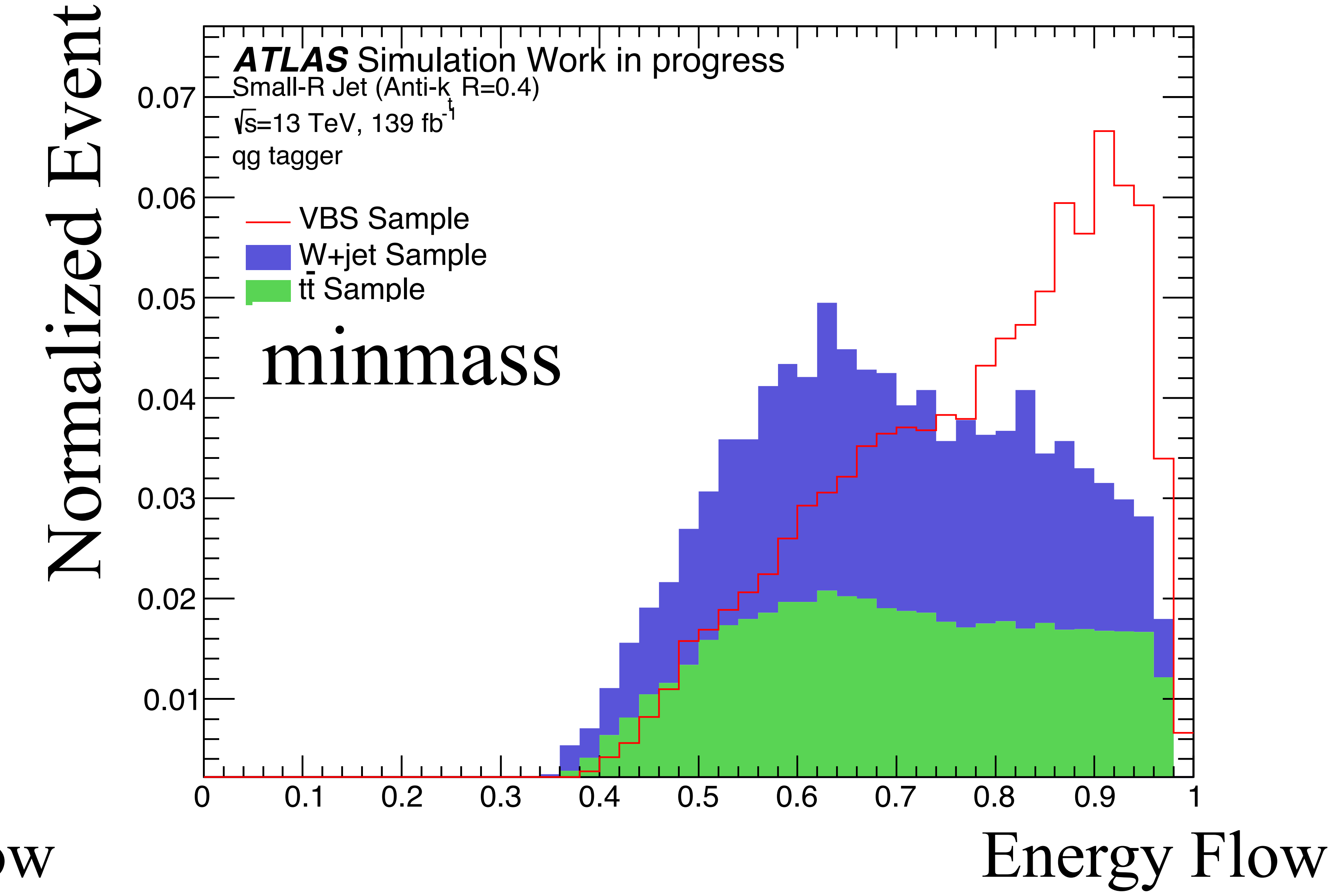
- Unlike 2leading and minmass selection strategies, order strategy selects tag- ging jet and signal jet pairs by the magnitude of Energy Flow tagger output.
- Compared to the significance baseline in Section 5.3.2, it is found that the neural order selection strategy performs worse than the 2leading with only the BDT discriminant strategy.

Table E.1: A summary of the order strategy significance.

Selection Stratgy	Bkg Systematic Uncertainty	Signal	W+jet	$t\bar{t}$	Significance
order	10%	59	68±16	208±51	0.754
	20%	59	68±16	208±51	0.529
	30%	24	8±6	59±24	0.363

Application to 1-lepton semileptonic VBS

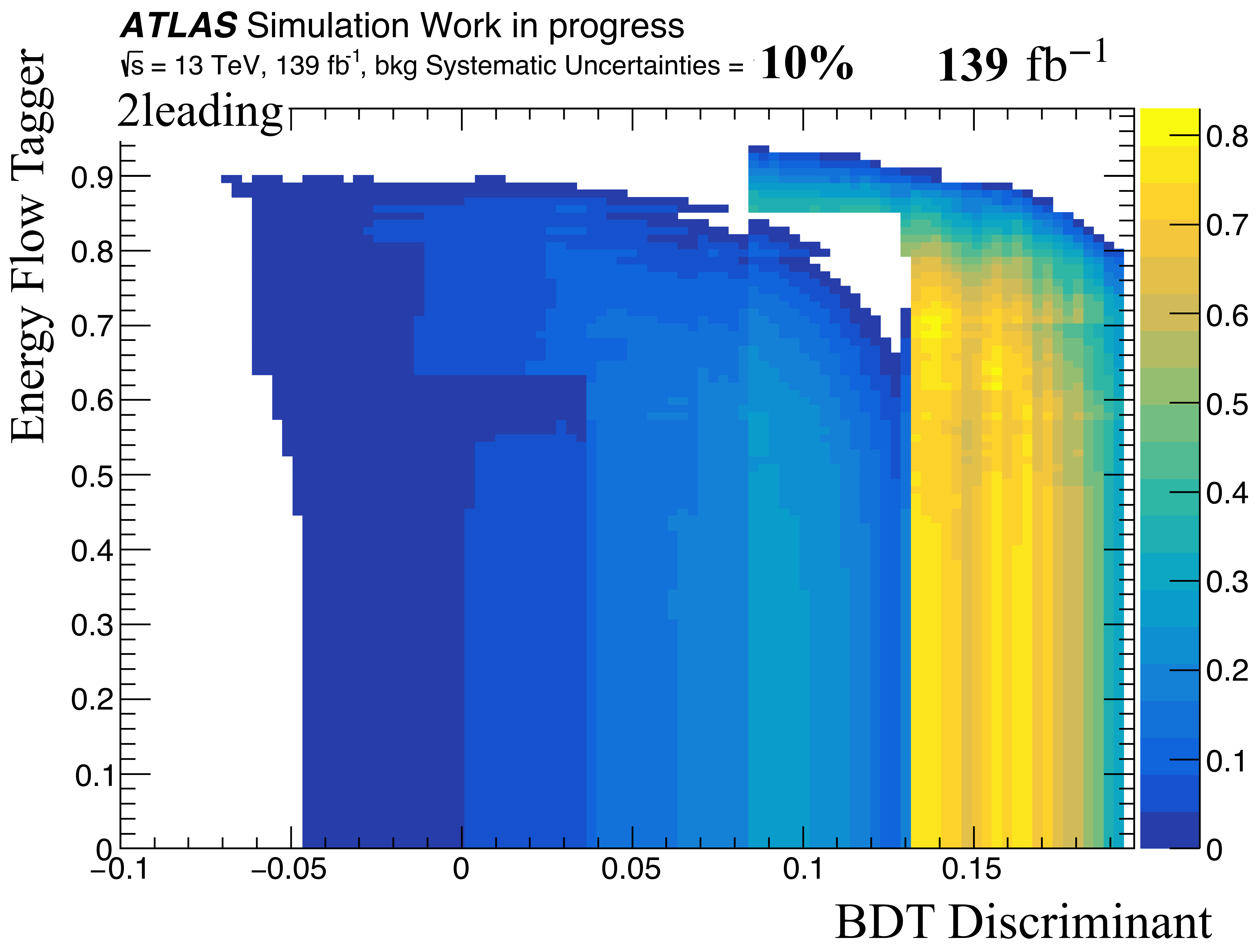




- The **2leading** strategy has better significance because the background event is less if the q/g tagging models are not applied.
- Although the **minmass** strategy has the worse significance compared to the 2leading strategy, it has more statistics of signal and background events, indicating the minmass strategy might have more space to improve by applying the q/g tagging models.
- Finally, the Energy Flow q/g tagging and the BDT Discriminant outputs are applied to obtain the best significance.

	2leading with BDT Discriminant	minmass with BDT Discriminant
Energy Flow q/g tagging	Neural 2leading	Neural minmass

q/g tagging and improvement



- x : BDT to enhance the signal from the background.
- y : q/g tagger threshold to define quark jets.
- The z axis represents the significance after the thresholds are applied.
- The q/g tagger and the BDT discriminant threshold are scanned to search for the best significance.