

Next steps

- Explore new variables (high and low level)
- Add definition
- 2d plot eta and pt vs vars
- Central and forward
- Extend to new samples Zee, Zmumu, ggH
- Feature importance

High stats, new samples, PU=0

<https://its.cern.ch/jira/browse/ATLMCPROD-10818>

Simulation: Use 25.0.0 with fixed timing of hard scatter (i.e. including fix from <https://its.cern.ch/jira/browse/ATLINFR-5081>) and latest geometry of ATLAS-P2-RUN4-03-00-01.

✓	e8481 s4290 r15420 T: done finished Produced events: 4999000 submitted edit (saved)
✓	41 + 801168/mc.Py8EG_A14NNPDF23LO_jj_JZ3.py (FullSim)Py8 LO dijet JZ3 (mu=0), new conditions mc21_14TeV.801168.Py8EG_A14NNPDF23LO_jj_JZ3.evgen.EVNT.e8481 events: 5000000 e8481 s4290 r15420 T: finished finished Produced events: 4998000 submitted edit (saved)
✓	42 + 801169/mc.Py8EG_A14NNPDF23LO_jj_JZ4.py (FullSim)Py8 LO dijet JZ4 (mu=0), new conditions mc21_14TeV.801169.Py8EG_A14NNPDF23LO_jj_JZ4.evgen.EVNT.e8481 events: 5000000 e8481 s4290 r15420 T: finished finished Produced events: 4989000 submitted edit (saved)
✓	43 + 801170/mc.Py8EG_A14NNPDF23LO_jj_JZ5.py (FullSim)Py8 LO dijet JZ5 (mu=0), new conditions mc21_14TeV.801170.Py8EG_A14NNPDF23LO_jj_JZ5.evgen.EVNT.e8481 events: 5000000 e8481 s4290 r15420 T: finished finished Produced events: 4992000 submitted edit (saved)

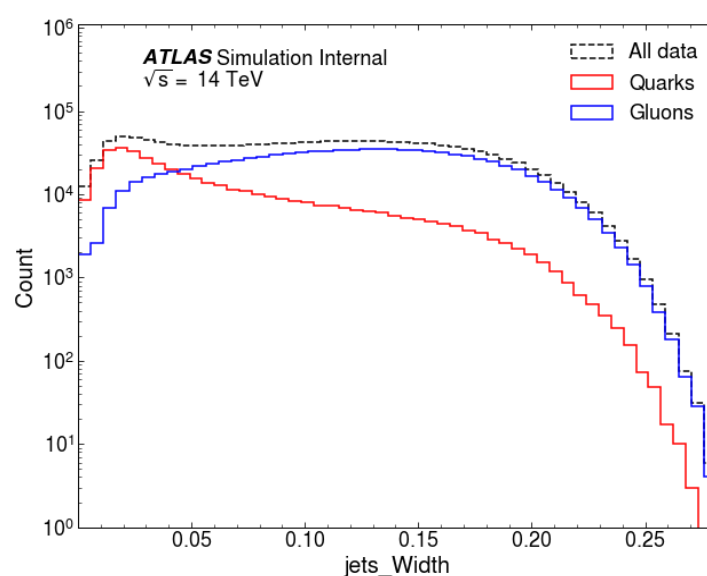
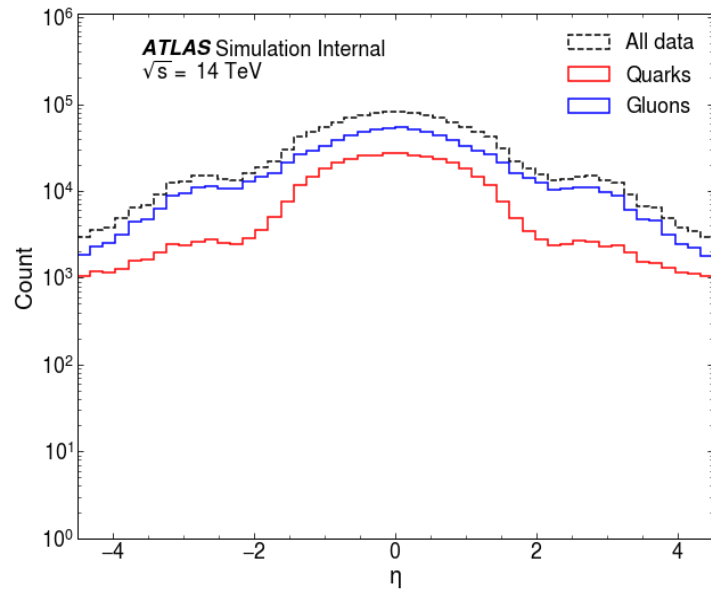
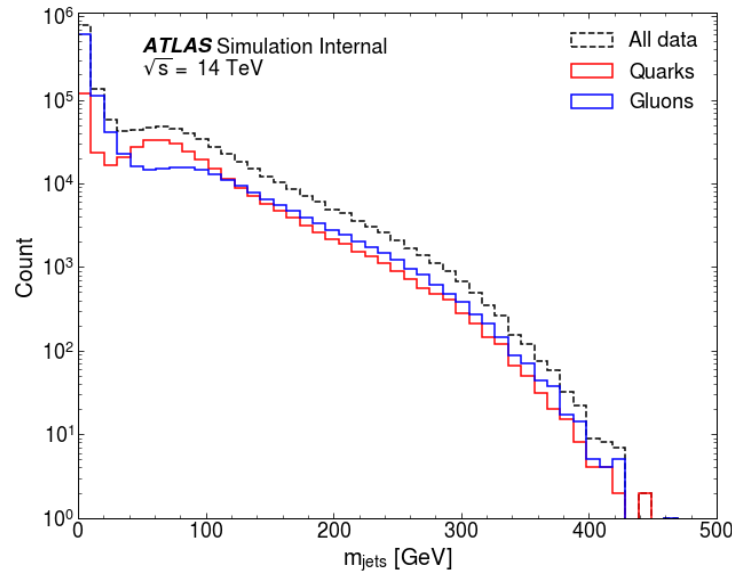
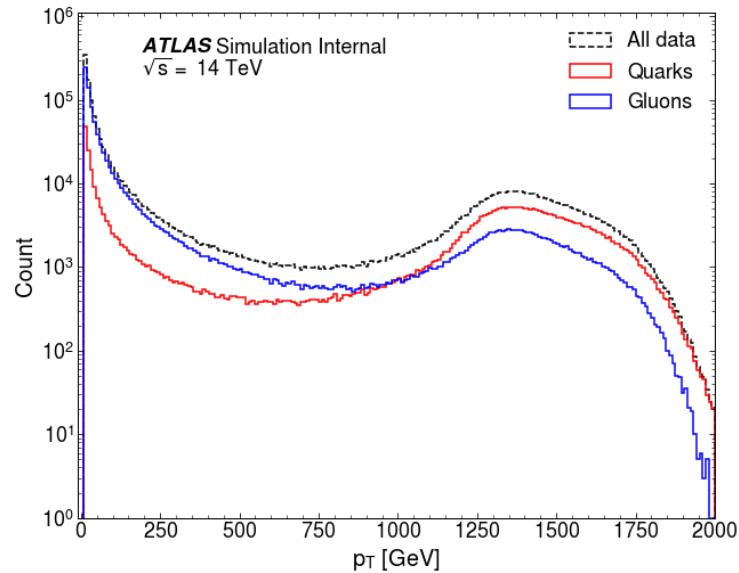
Label	leading p_T^{jet} range, [GeV]
JZ1	20–60
JZ2	60–160
JZ3	160–400
JZ4	400–800
JZ5	800–1300
JZ6	1300–1800
JZ7	1800–2500
JZ8	2500–3200
JZ9	3200–3900
JZ10	3900–4600

Samples available

- Red flag -> producing derivation (rucio)
- Derivations finished stored in lyon clusters

- mc21_14TeV.601229.PyPy8EG_A14_ttbar_hdamp258p75_SingleLep.recon.AOD.e8481_s4290_r15420
- mc21_14TeV.601230.PyPy8EG_A14_ttbar_hdamp258p75_dil.recon.AOD.e8481_s4290_r15420
- mc21_14TeV.601237.PyPy8EG_A14_ttbar_hdamp258p75_allhad.recon.AOD.e8481_s4290_r15420
- mc21_14TeV.601189.PyPy8EG_AZNLO_Zee.recon.AOD.e8481_s4290_r15420
- mc21_14TeV.601190.PyPy8EG_AZNLO_Zmumu.recon.AOD.e8481_s4290_r15420
- mc21_14TeV.601191.PyPy8EG_AZNLO_Ztautau.recon.AOD.e8481_s4290_r15420
- mc21_14TeV.801165.Py8EG_A14NNPDF23LO_jj_JZ0.recon.AOD.e8481_s4290_r15420
- mc21_14TeV.801166.Py8EG_A14NNPDF23LO_jj_JZ1.recon.AOD.e8481_s4290_r15420
- mc21_14TeV.801167.Py8EG_A14NNPDF23LO_jj_JZ2.recon.AOD.e8481_s4290_r15420
- mc21_14TeV.801168.Py8EG_A14NNPDF23LO_jj_JZ3.recon.AOD.e8481_s4290_r15420
- mc21_14TeV.801169.Py8EG_A14NNPDF23LO_jj_JZ4.recon.AOD.e8481_s4290_r15420
- mc21_14TeV.801170.Py8EG_A14NNPDF23LO_jj_JZ5.recon.AOD.e8481_s4290_r15420 -> DOING A TEST WITH THIS SAMPLE
- mc21_14TeV.801171.Py8EG_A14NNPDF23LO_jj_JZ6.recon.AOD.e8481_s4290_r15420
- mc21_14TeV.801172.Py8EG_A14NNPDF23LO_jj_JZ7.recon.AOD.e8481_s4290_r15420
- mc21_14TeV.801173.Py8EG_A14NNPDF23LO_jj_JZ8.recon.AOD.e8481_s4290_r15420
- mc21_14TeV.600026.PyH7EG_NNPDF3_AZNLO_VBFH125_ZZ4nu_MET75.recon.AOD.e8481_s4290_r15420

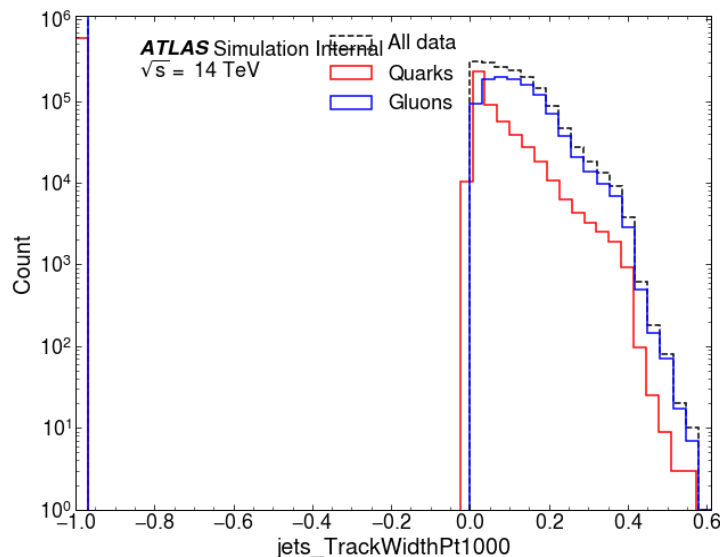
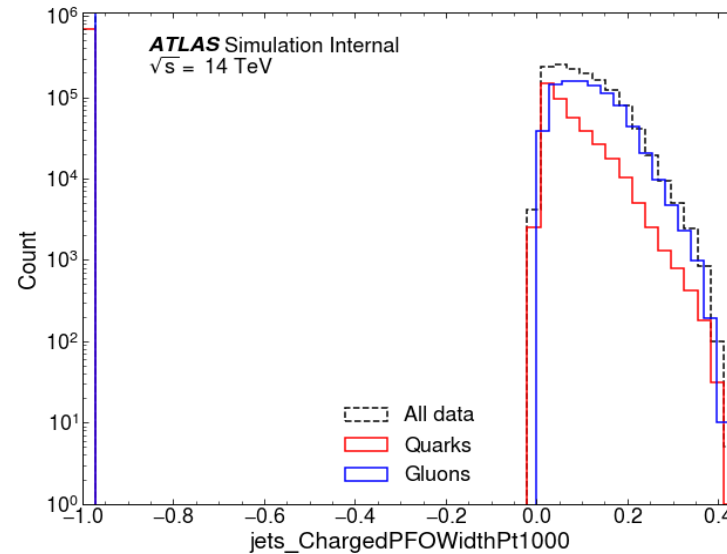
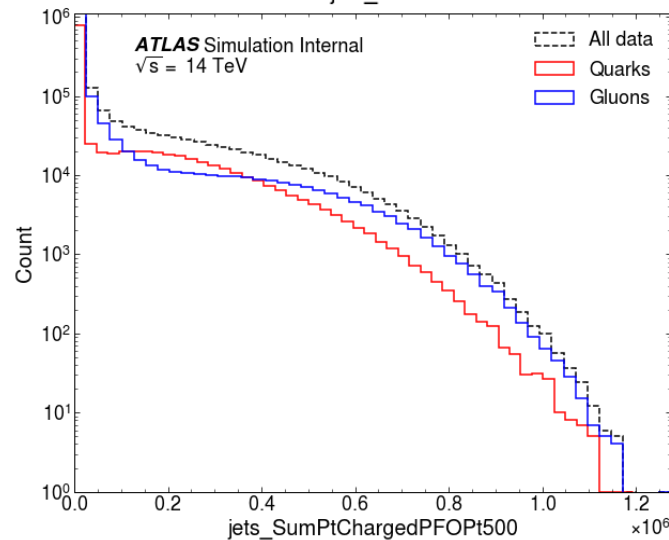
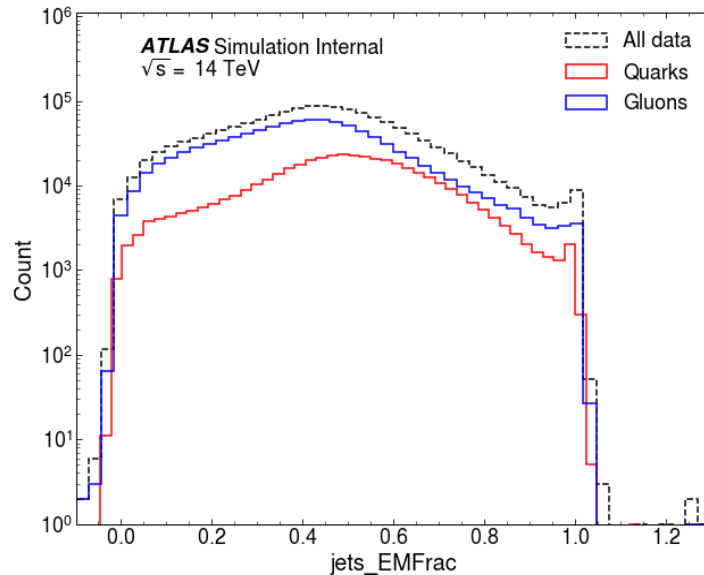
Variables



- **Jets** are reconstructed with the anti- k_t with $R = 0.4$ as PFOs as constituents.
- PFOs are a collection of **topo-clusters** formed from energy deposits in calorimeter cells and an algorithmic combination of charged-particle **tracks** with those topo-clusters.
- **Jet width** calculated using PFOs with $p_T > 1000$ MeV

$$\frac{\sum_{\text{trk} \in \text{jet}} p_T^{\text{track}} \Delta R_{\text{trk}, \text{jet}}}{\sum_{\text{trk} \in \text{jet}} p_T^{\text{track}}}$$

Variables



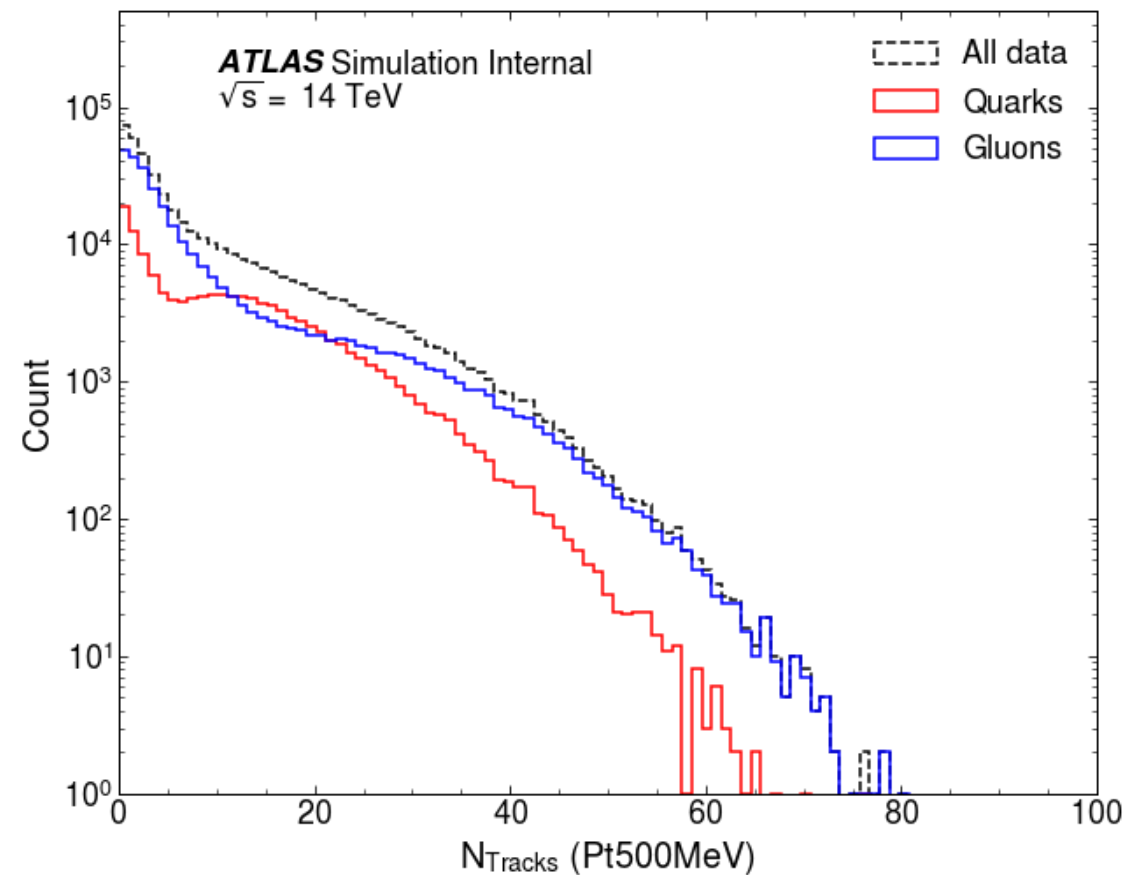
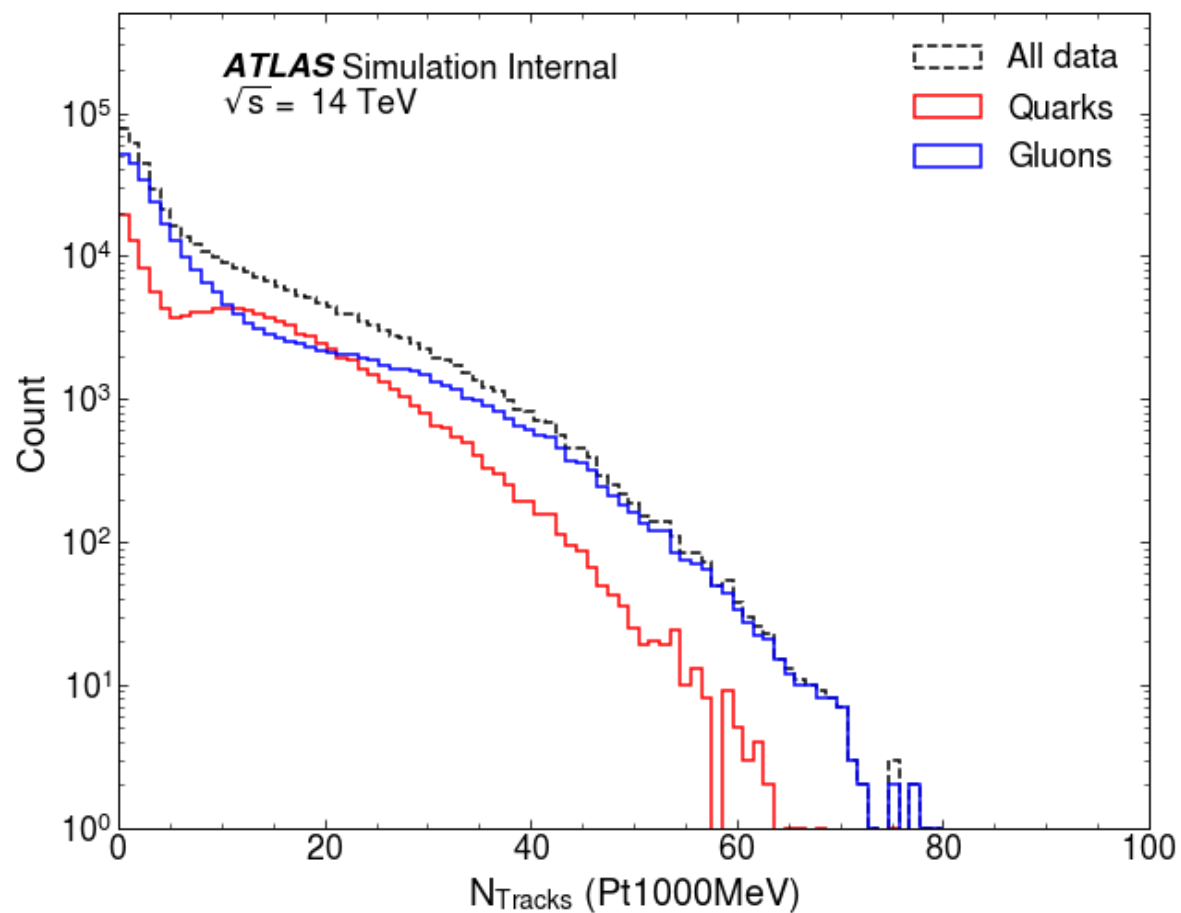
EMFrac: Fraction of the jet energy deposited in the EM calorimeter. Fraction of the jet energy carried by charged particles.

ChargedPFOWidthPt1000: Numbers of charged PFOs passing two different pT thresholds 1000 MeV

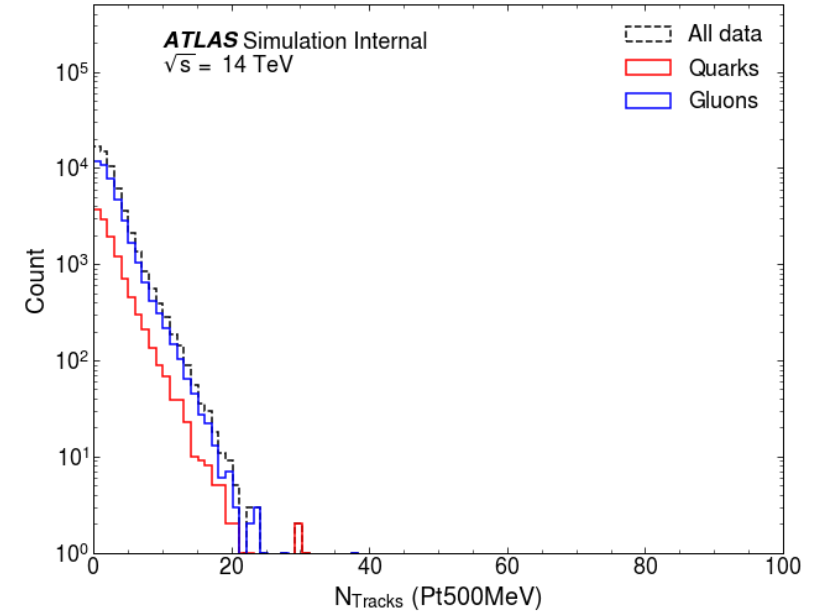
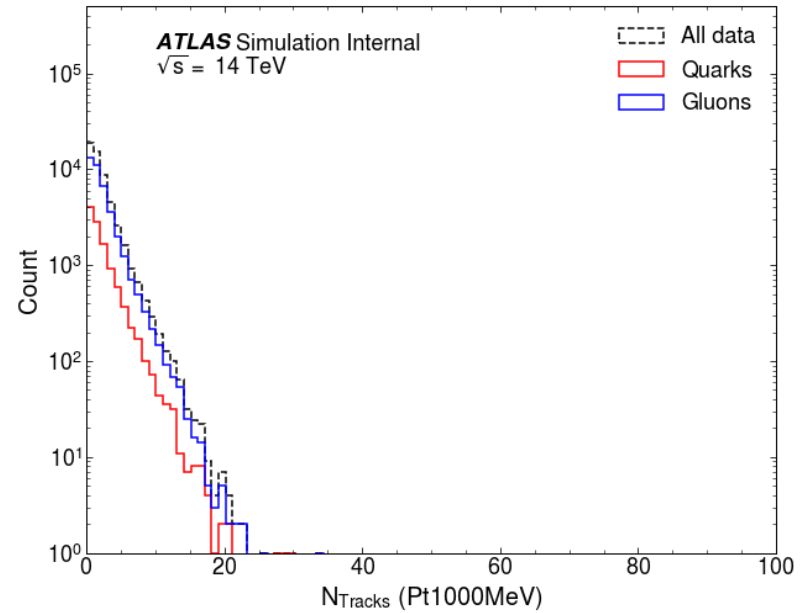
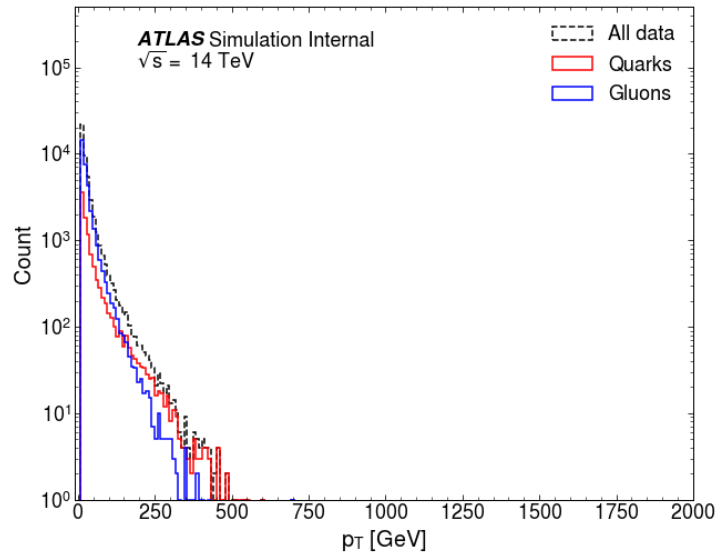
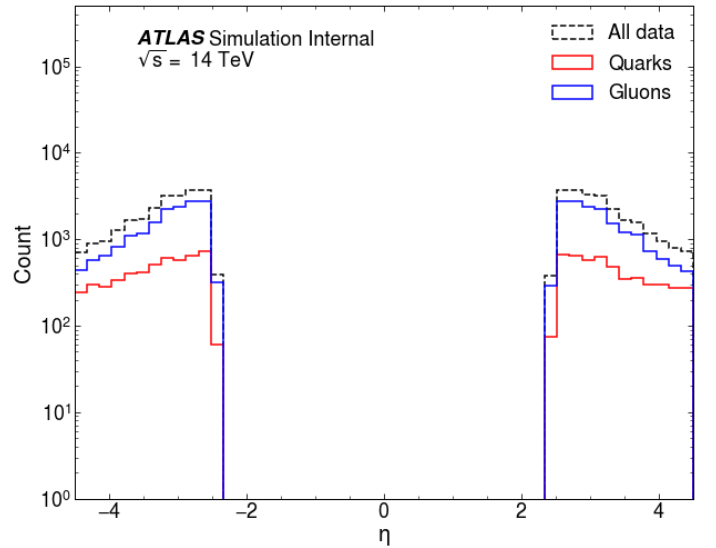
SumPtChargedPFOpt500: the sum of the transverse momenta of all charged particle flow objects that have a transverse momentum greater than 500 MeV.

TrackWidthPt1000: Jet width calculated using tracks with pT > 1000 MeV

Variables



Number of tracks (forward)



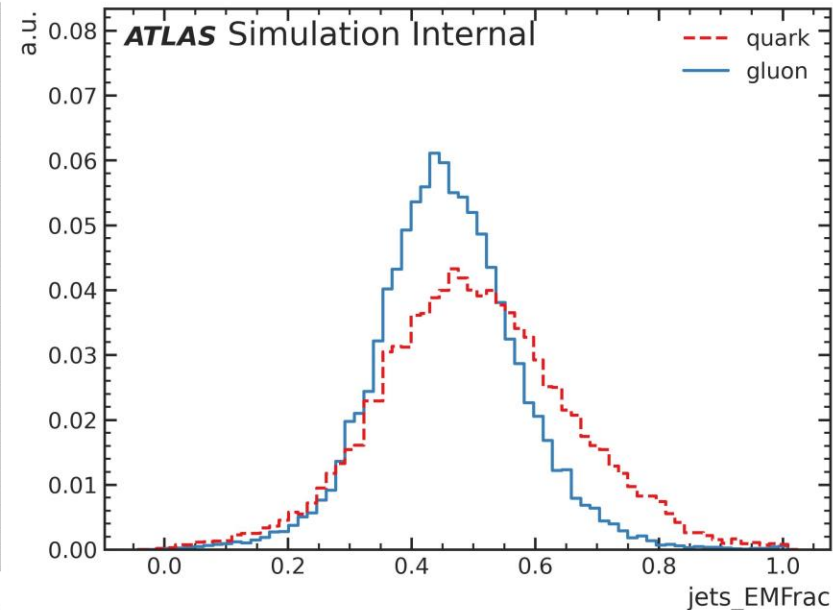
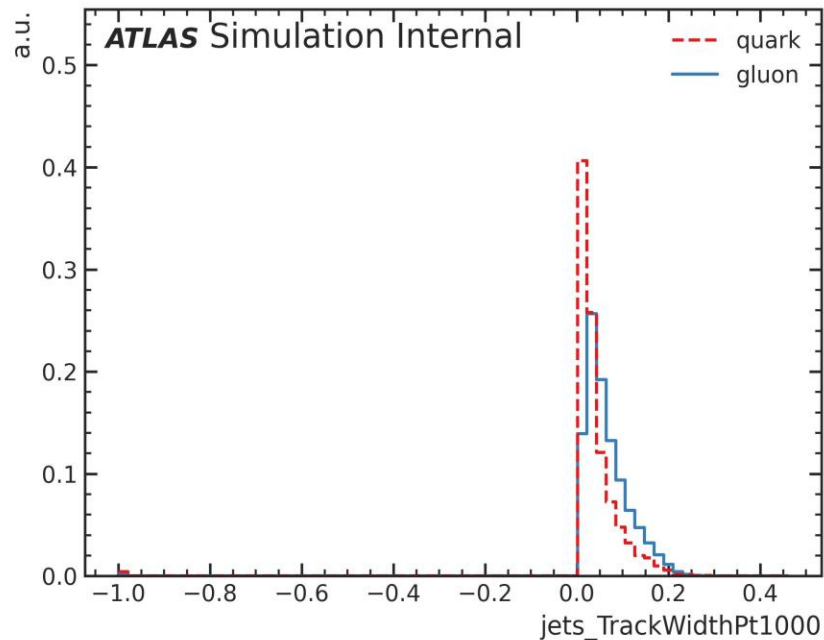
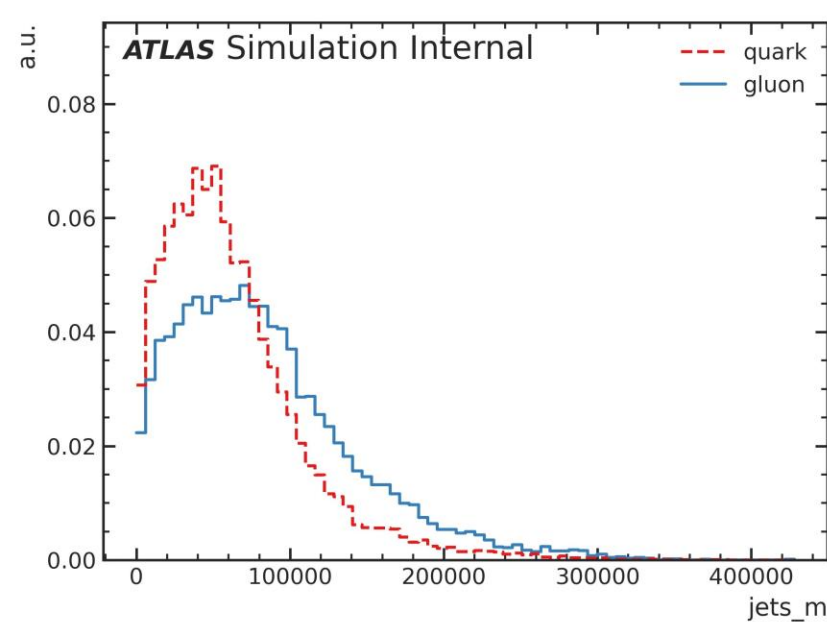
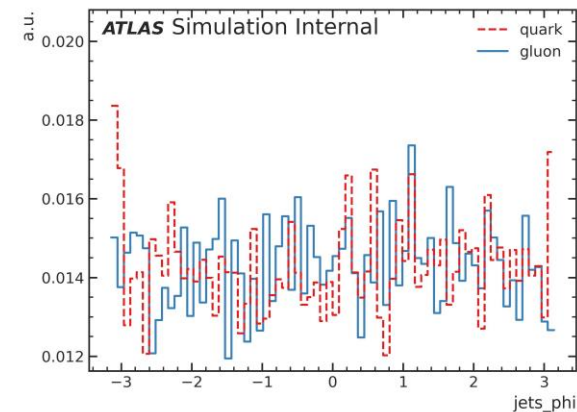
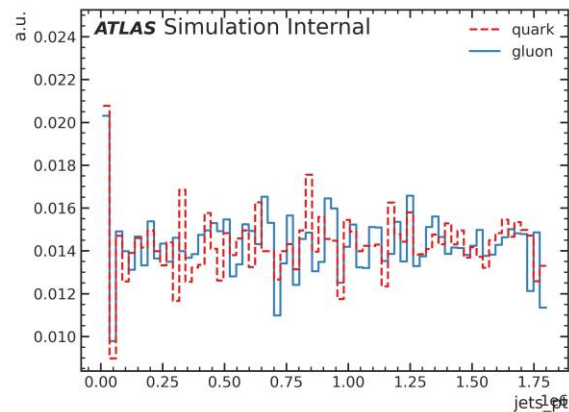
Train with 801171

1M events processed

Train: 4M jets

Test: 0.5M jets

Val: 0.5M jets



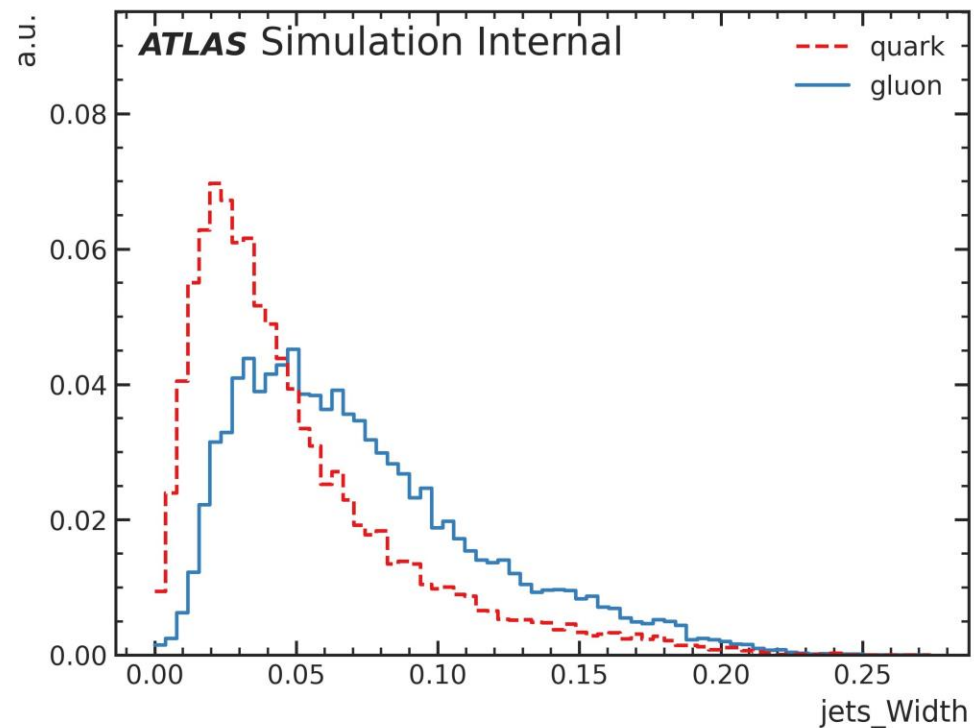
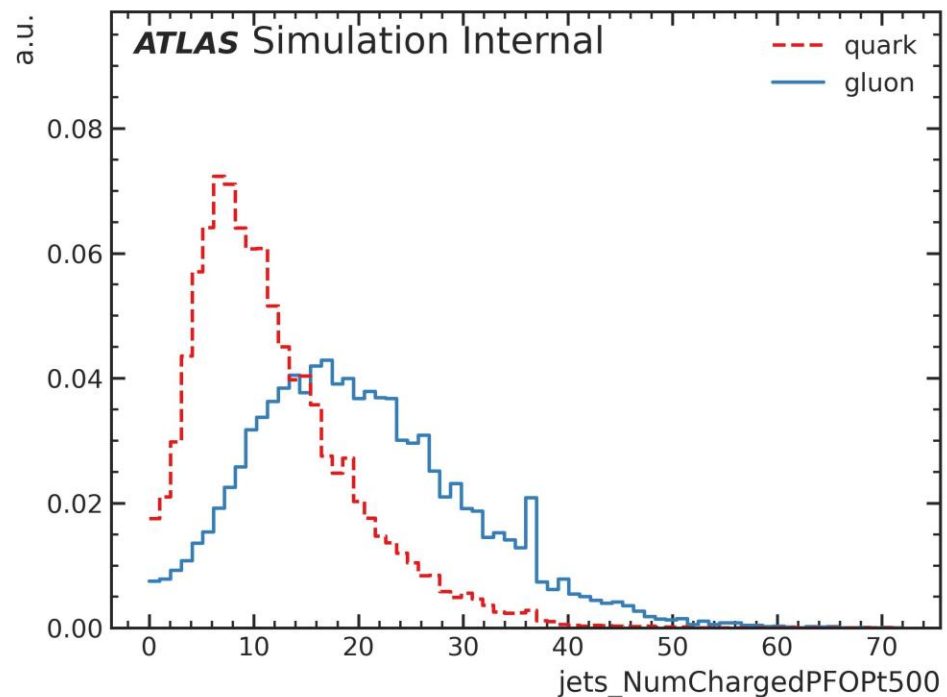
Train with 801171

1M events processed

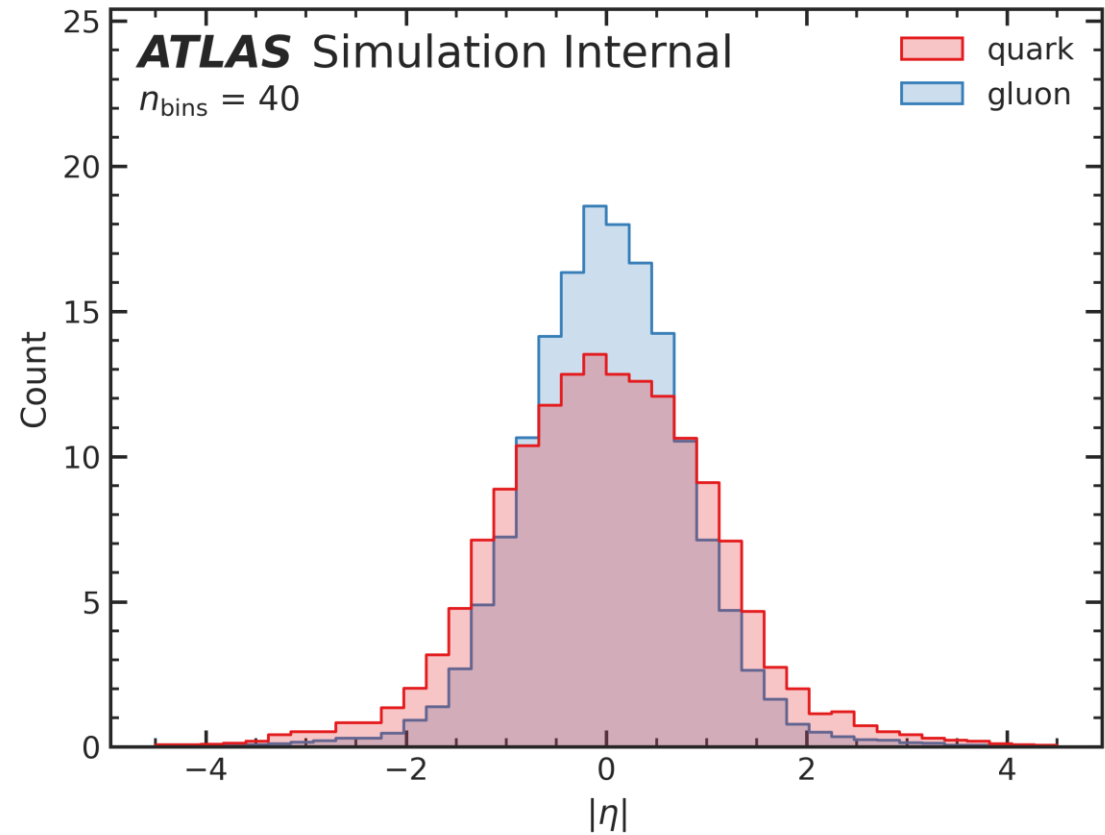
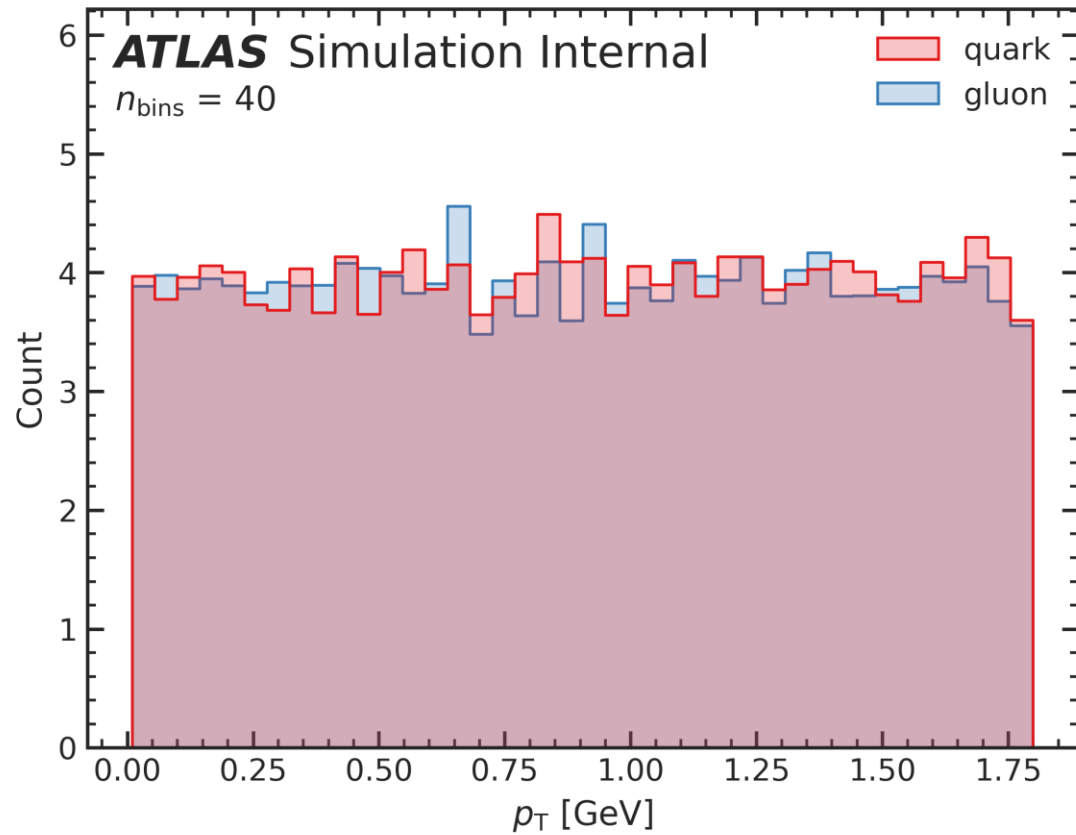
Train: 4M jets

Test: 0.5M jets

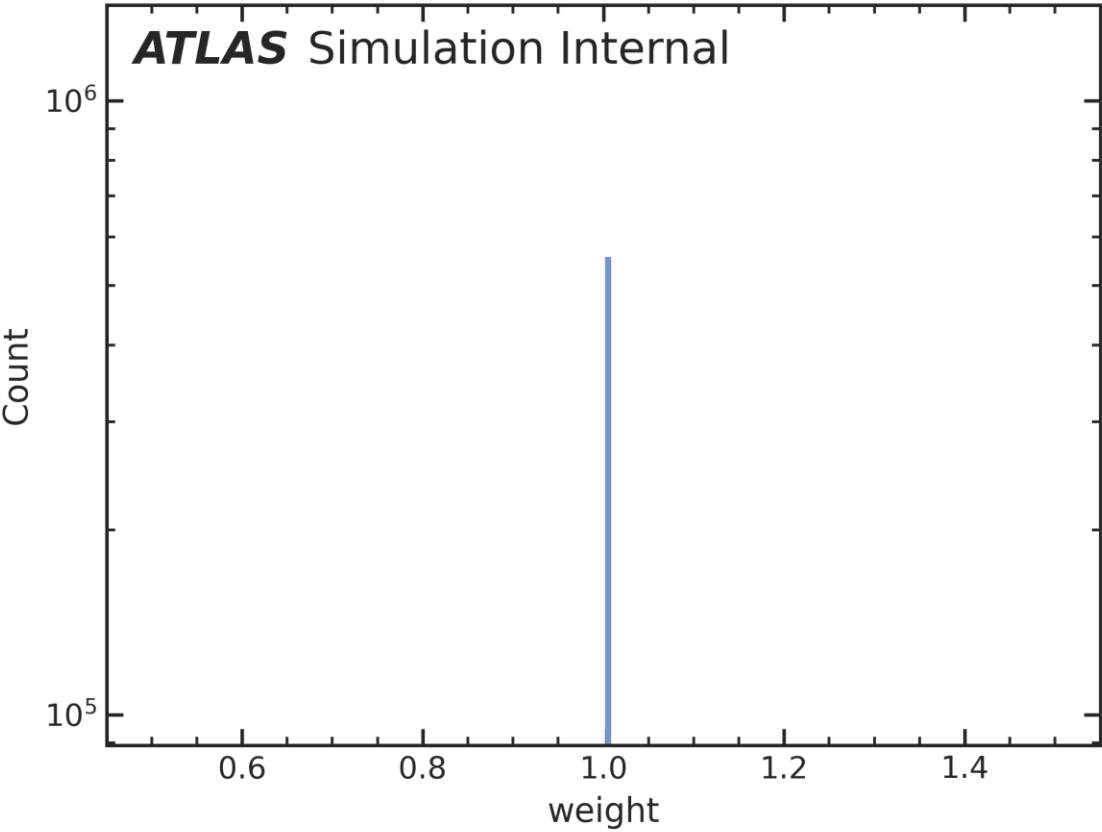
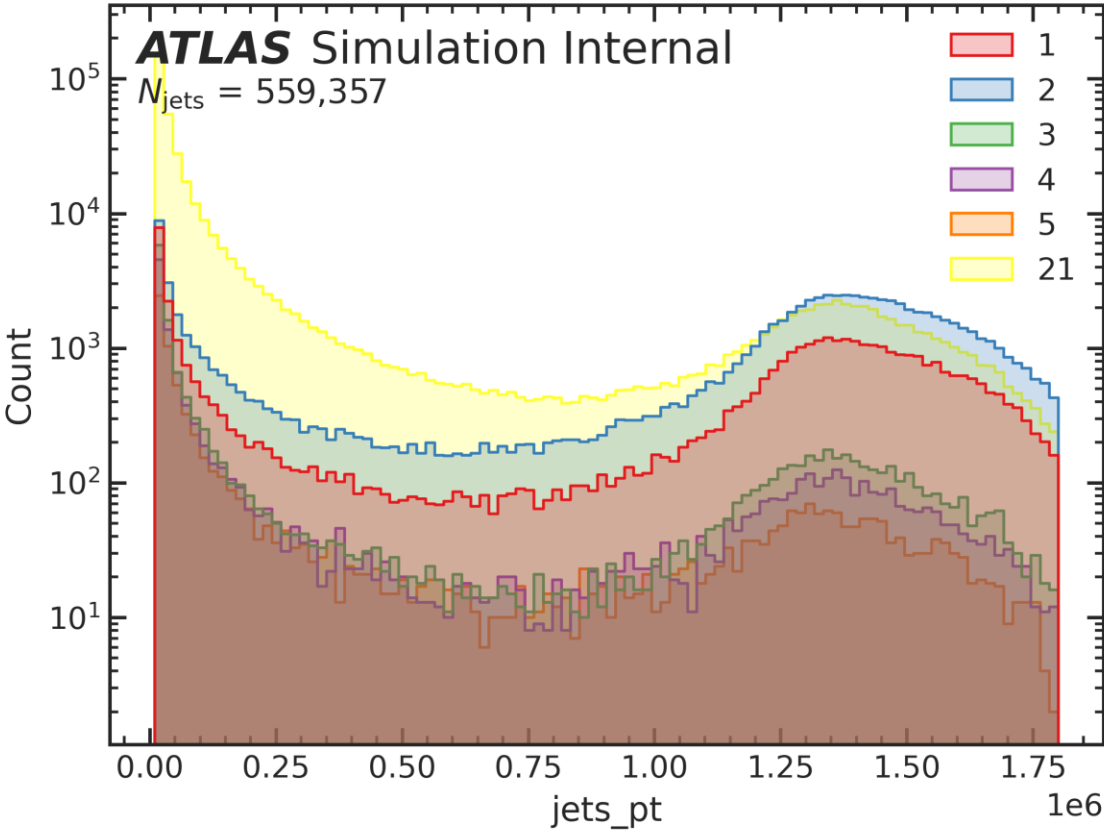
Val: 0.5M jets



Train with 801171



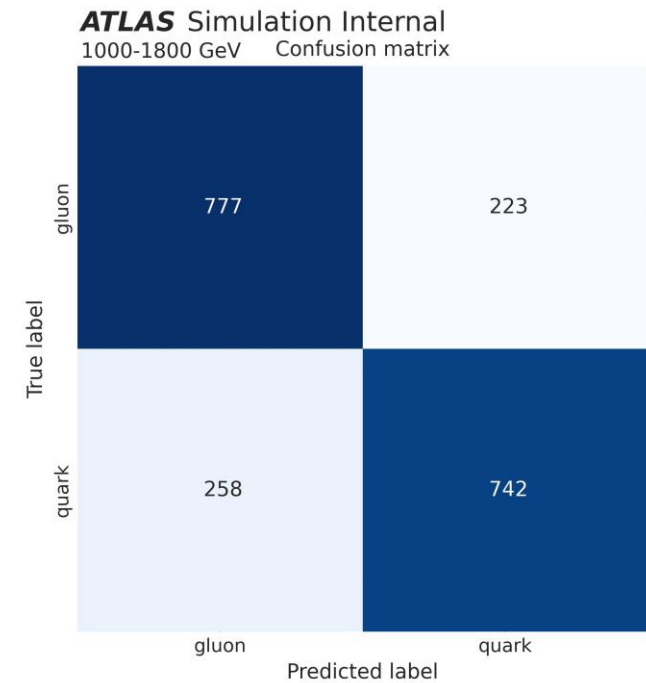
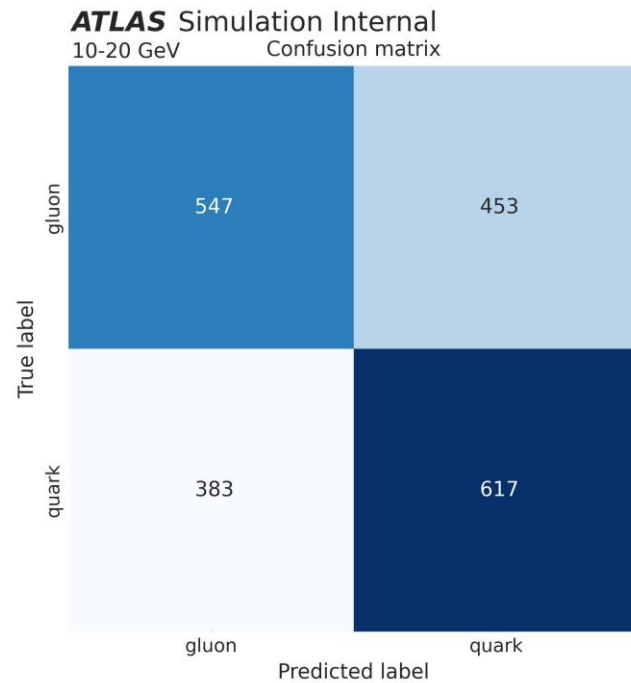
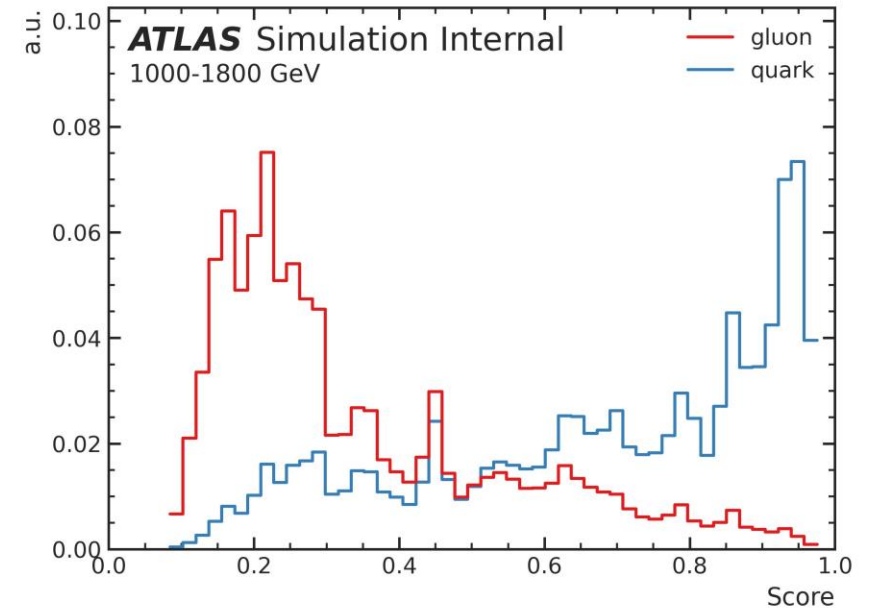
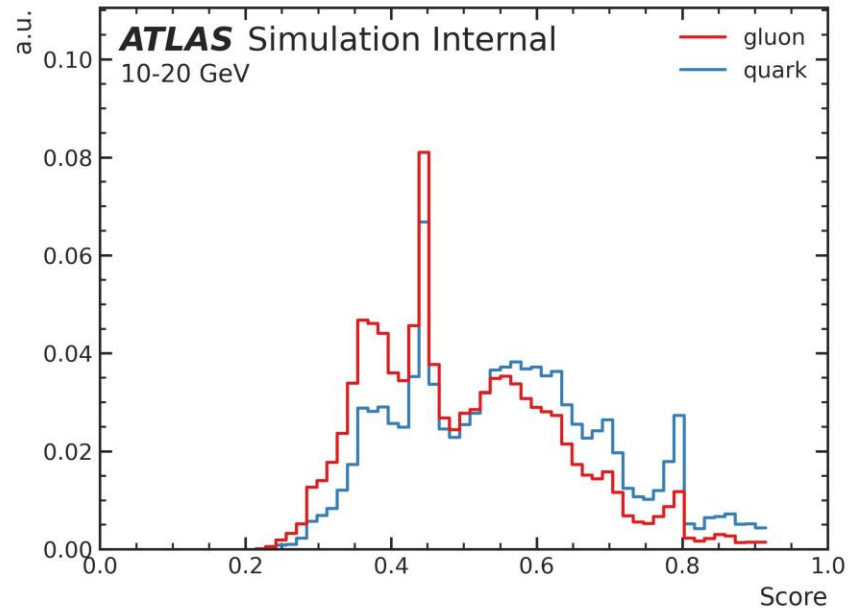
Test with 801171



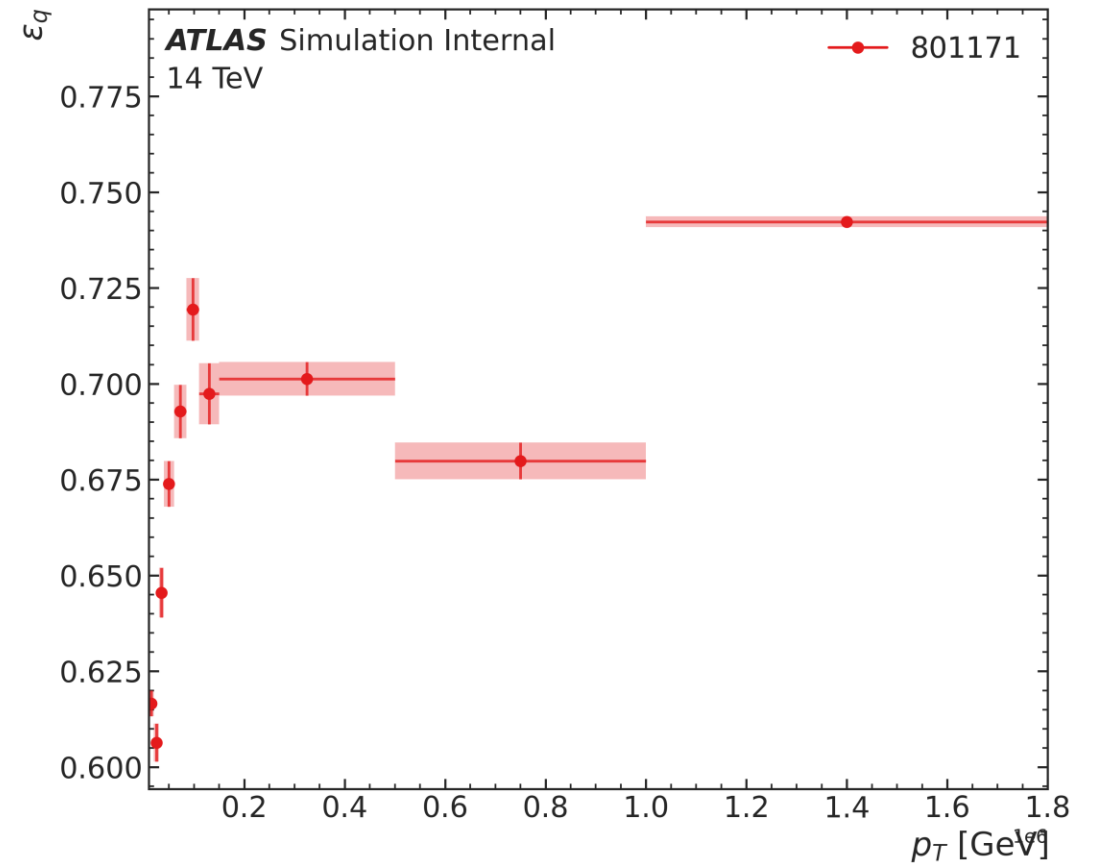
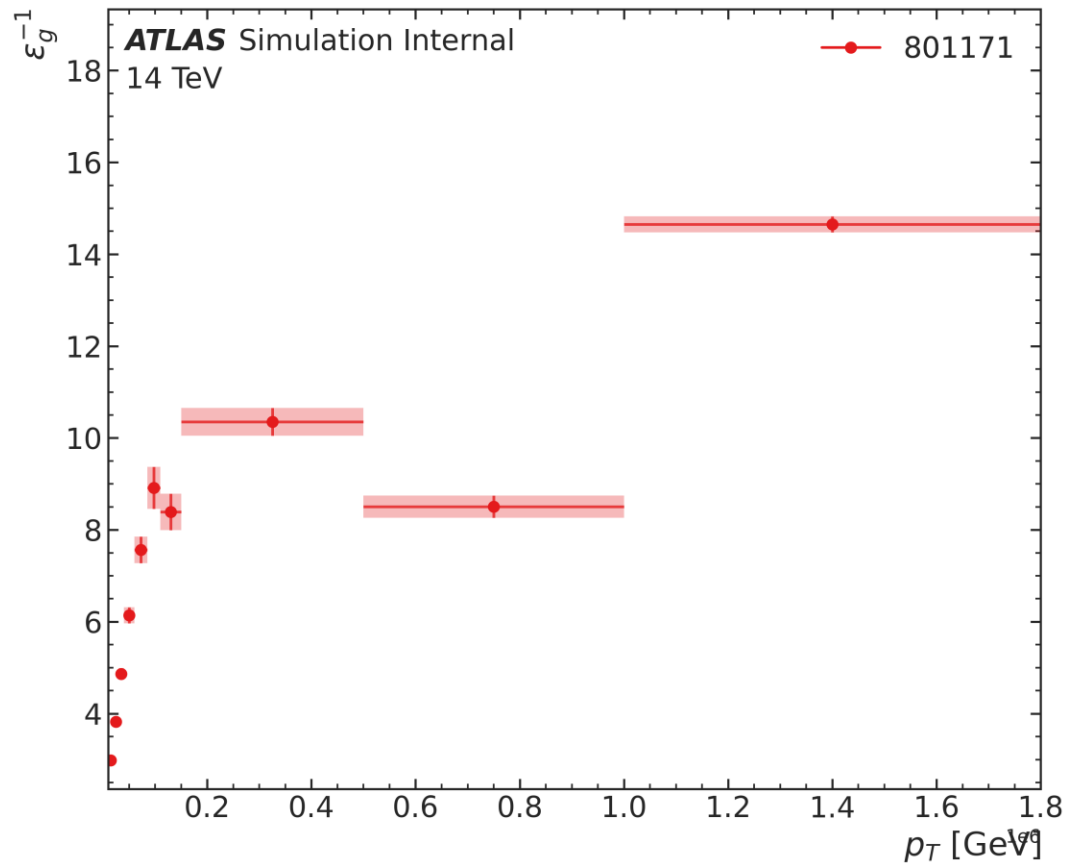
Evaluation

bins: [10000, 20000, 30000, 40000, 60000, 85000, 110000, 150000, 500000, 1000000, 1800000] MeV

Py8EG_A14NNPDF23LO_jj_
JZ6



Evaluation: Py8EG_A14NNPDF23LO_jj_JZ6



Next steps

- Include all datasets --> create a baseline!
- Explore new variables (high and low level)
- Add definition --> need to create a table
- 2d plot eta and pt vs vars
- Central and forward
- ~~Extend to new samples Zee, Zmumu, ggH~~ (not added to the new production)
- Feature importance
- Present the baseline in ATLAS Meeting!

Bonus!

