THE OHIO STATE UNIVERSITY

ATLAS EXPERIMENT

Recent developments in flavour tagging and b-jet triggers in the ATLAS collaboration

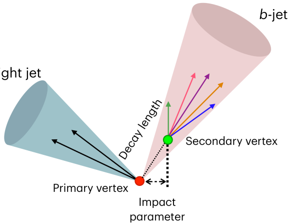
Roshan Joshi

On behalf of the ATLAS Collaboration

September 30, 2025

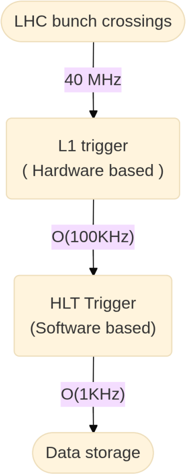
Introduction: flavour tagging

- Heavy flavour (b/c) tagging is important for Higgs physics and BSM physics at the LHC [eg: $HH \to 4b$, $VH(bb/cc)$].
- Exploits decay properties such as:
    - longer lifetime of b-/c-hadrons, O(ps), leading to displaced tracks and secondary vertex,
    - high track multiplicity,
    - larger mass compared to light quarks/gluon,
    - semi-leptonic decays of b-/c-hadrons.



Introduction: trigger

- Not feasible to record every collision - only select those events that are "interesting" → trigger.
- b-jet triggers select events that are likely to contain jets coming from b-hadron decay. Important for searches like $HH \to 4b, bb\tau\tau, HHH \to 6b$, etc.
- Updates on this talk:
    - Transformer based b-tagging, and $b + \tau$-tagging in ATLAS for small-R jets.
    - Muon-in-jet trigger.
    - Transformer based $H \to bb$ tagging in ATLAS b-jet trigger for large-R jets.



LHC bunch crossings

40 MHz

L1 trigger
(Hardware based)

O(100KHz)

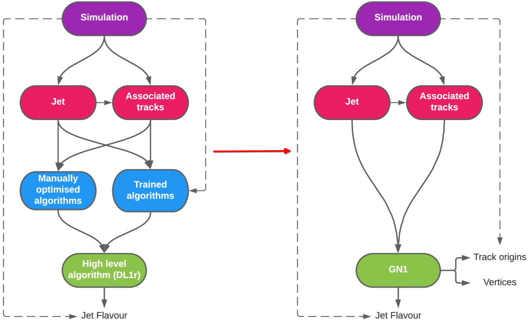
HLT Trigger
(Software based)

O(1KHz)

Data storage

History of ML-enabled flavour tagging in ATLAS

- Run2 taggers used output of low-level algorithms (which constructed discriminating variables exploiting large impact parameters of tracks, reconstructing displaced vertices, etc.).
- New taggers use jet & associated track features directly.

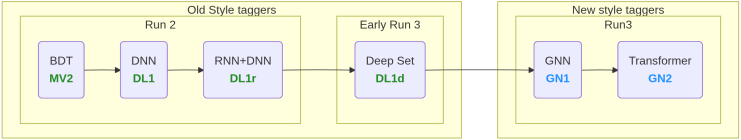


Simulation
Jet
Associated tracks
Manually optimised algorithms
Trained algorithms
High level algorithm (DL1r)
Jet Flavour

Simulation
Jet
Associated tracks
GN1
Track origins
Vertices
Jet Flavour

ATL-PHYS-PUB-2022-027

† Only offline small-R jet tagging covered in this talk. For offline large-R jet tagging see talk by Jurjan.



Old Style taggers

Run 2

BDT **MV2**

DNN **DL1**

RNN+DNN **DL1r**

Early Run 3

Deep Set **DL1d**

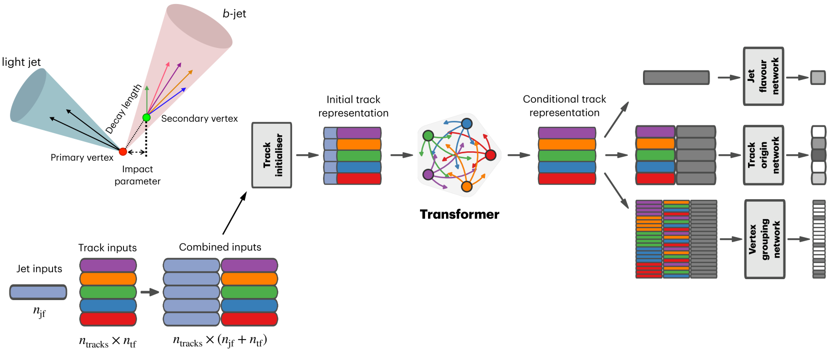
New style taggers

Run3

GNN **GN1**

Transformer **GN2**

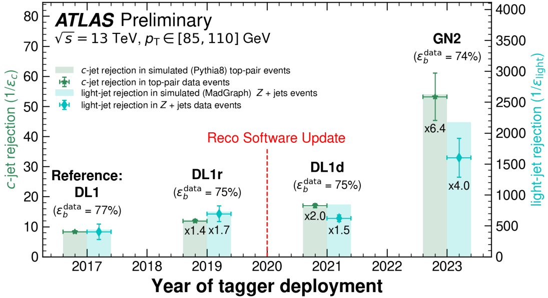
Transformer based tagger: GN2



ATLAS-FTAG-2023-05

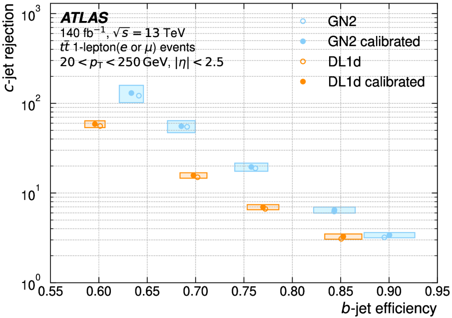
- Inputs: Jet p_T, η; track parameters, their associated uncertainties, and impact parameters; info on inner detector hits.
- Also incorporates auxiliary tasks like track origin determination, track-pair vertexing.

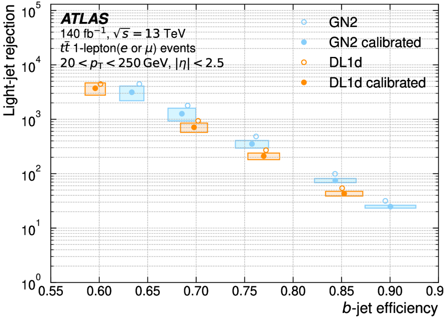
Transformer based tagger: GN2



ATLAS-FTAG-2023-05, ATLAS-PHOTO-2025-044-1

$$D_b = \log\left(\frac{p_b}{f_c p_c + f_\tau p_\tau + (1 - f_c - f_\tau) p_u}\right)$$





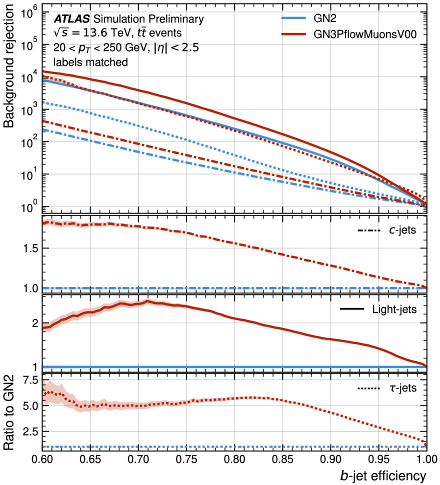
Upcoming tagger: GN3

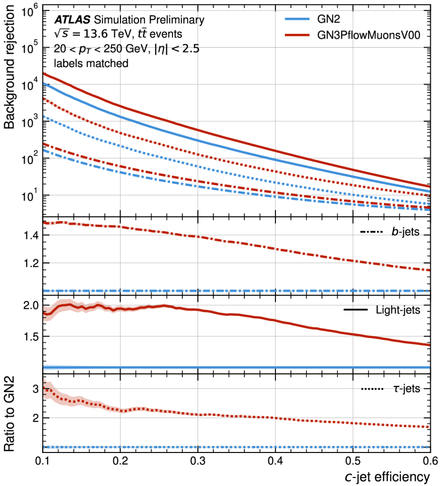
- Improves on GN2 by (ATLAS-FTAG-2025-01):
  - using looser track cuts,
  - using neutral particle flow information along with the usual charged tracks,
  - using lepton identification information,
  - using improved transformer architecture.
- Has more output classes (light-jet splitting), and more auxiliary tasks than GN2.

$$D_b(\text{GN2}) = \log\left(\frac{p_b}{f_c \cdot p_c + f_\tau \cdot p_\tau + (1 - f_c - f_\tau) \cdot p_{\text{light}}}\right)$$

$$D_b(\text{GN3}) = \log\left(\frac{p_b}{f_c \cdot p_c + f_\tau \cdot p_\tau + (1 - f_c - f_\tau) \cdot (p_{ud} + p_g + p_s)}\right)$$

GN3 vs GN2 performance





ATLAS-FTAG-2025-01

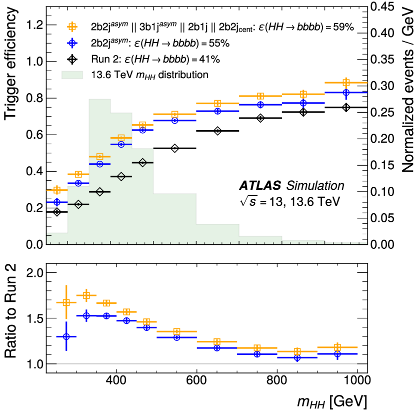
History of online (trigger) taggers

- The development in online taggers follows the offline taggers very closely in terms of architecture being used.
- In Run 2 and early Run3 the trigger taggers used output of low-level algorithms that were fed to high level algorithms.
- In Run 3, following offline development, ATLAS b-jet trigger also moved to using graph neural networks (GN1) and then to transformer (GN2).
- GN2 currently being used in b-jet trigger (small-R jets). No plans to move to GN3 for remainder of Run 3.

Novel b-jet trigger strategy for Run3 with focus on HH signal, but potentially useful in HHH analyses

b-jet triggers in $HH \to 4b$

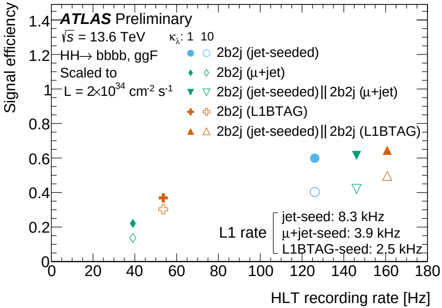
- Thresholds in Run2:
  - 2b1j: two b-tagged with $p_T > 55$ GeV and one additional jet with $p_T > 100 - 150$ GeV
  - 2b2j: two b-tagged jets and two additional jets with $p_T > 35$ GeV
- Asymmetric thresholds in Run3 in 2022(2023) with DL1d(GN1):
  - 2b2j(asym): ≥ 4 jets with $p_T > 80(75), 55(50), 28(25)$, 20 GeV, two b-tagged



ATLAS-TRIG-2022-02

muon-in-jet b-jet triggers in $HH \to 4b$

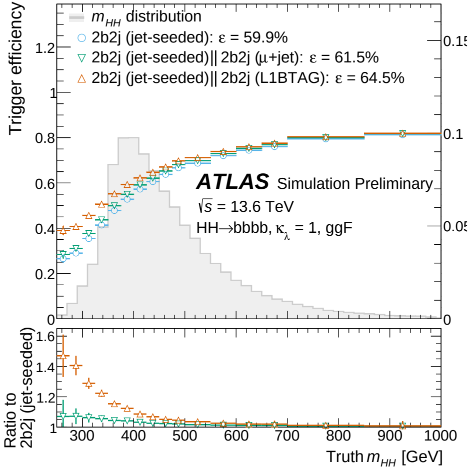
- 20% of b-hadron decay chain lead to muons. Can exploit the presence of muons in b-jet triggers:
  - A muon (in L1) +jet seeded chain running since start of Run3 called $2b2j(\mu + \text{jet})$ with muon $p_T > 8$ GeV (no ΔR matching between jet and muon)
- New design: use ΔR matching between μ and jet to do b-tagging at L1 (L1BTAG) → **NEW IN 2025!**:

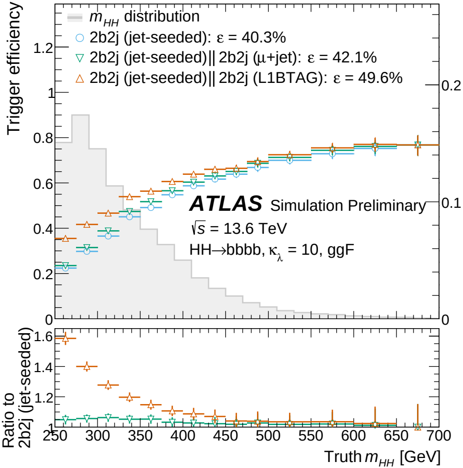


BJet Trigger Public Results

- two L1 jets with $p_T > 50, 40$ GeV, a L1 muon with $p_T > 5$ GeV geometrically matched to a jet above 20 GeV by L1Topo system, and four HLT jets with $p_T > 40, 35, 25, 20$ GeV.
- Helps reduce L1 rate for μ+jet chain.

muon-in-jet b-jet triggers in $HH \to 4b$

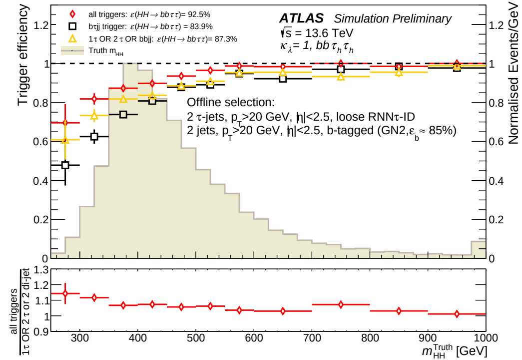


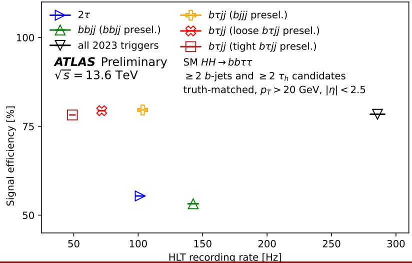


BJet Trigger Public Results

b-jet triggers in $HH \to bb\tau_{\text{had}}\tau_{\text{had}}$

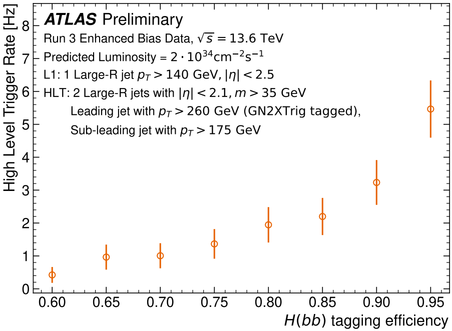
- Single tau: 1τ with $p_T > 160$ GeV
- Di-tau: 2τ with $p_T > 30, 20$ GeV
- $2b + 2j$(asym)
- $1b + 1\tau + 2j$: 4 jets with $p_T > 65, 40, 25, 20$ GeV, requiring 1 b-tagged jet, and 1τ with $p_T > 25$ GeV (**NEW SINCE 2024**)
- TauTrigger public results

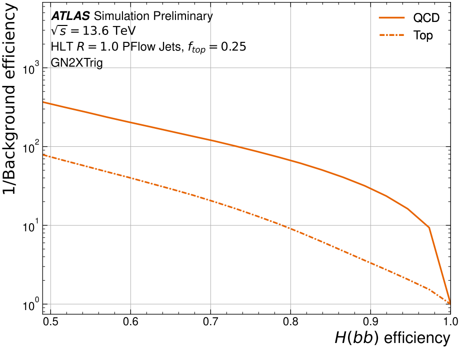




Large-R jet $H \to bb$ tagging

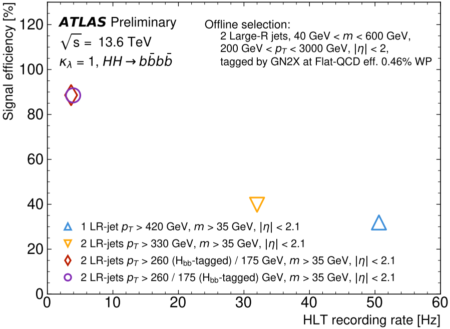
- Current large-R jet primary has high threshold (j420 or 2j330). Can use $H(bb)$ tagging to lower thresholds & improve efficiency in medium pT range [200, 400] GeV.
- Model uses transformer architecture following offline GN2X tagger.
- Uses properties of tracks associated with large-R jet, and jet p_T and η as input features.

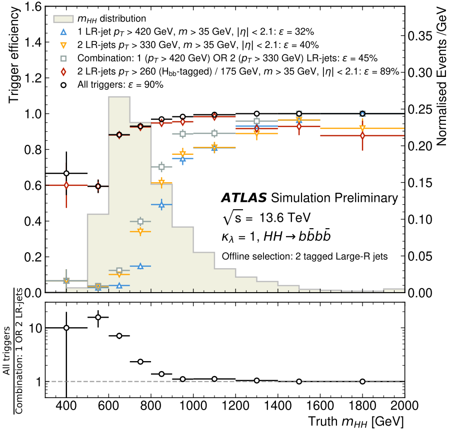




BJet Trigger Public Results

Large-R jet $H \to bb$ tagging: $HH \to 4b$





BJet Trigger Public Results

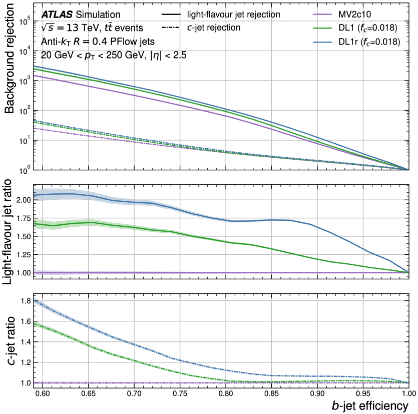
Summary

- New machine learning models bring considerable improvement in flavour-tagging. Further improvement can be accomplished by supplying more information to the networks.
- These models can then be used in trigger to improve background rejection thereby decreasing rates.
- Novel methods in trigger can bring efficiency gains in di-higgs (and tri-higgs) analyses - $b + \tau$ triggers bring efficiency gain of 8.6% in $HH \to bb\tau\tau$ and boosted jet $H(bb)$ trigger may bring efficiency gain of up to 40% in boosted $HH \to 4b$, etc.
- The same triggers can potentially benefit HHH analyses.

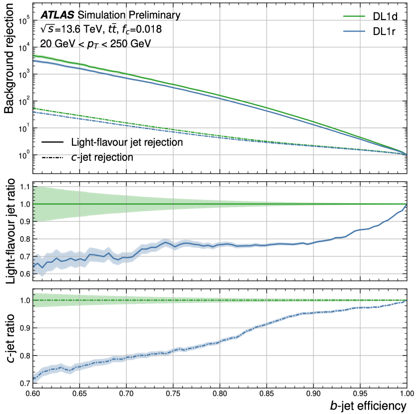
BACKUP SLIDES

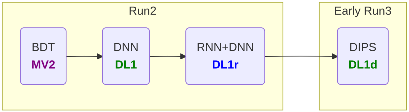
ML-enabled flavor tagging: early days

CERN-EP-2022-226



ATLAS-FTAG-2022-004





Run2

BDT **MV2**

DNN **DL1**

RNN+DNN **DL1r**

Early Run3

DIPS **DL1d**

Explanation of terms in formula of discriminants

$$D_b(\mathrm{GN2}) = \log\left(\frac{p_b}{f_c \cdot p_c + f_\tau \cdot p_\tau + (1 - f_c - f_\tau) \cdot p_{\mathrm{light}}}\right)$$

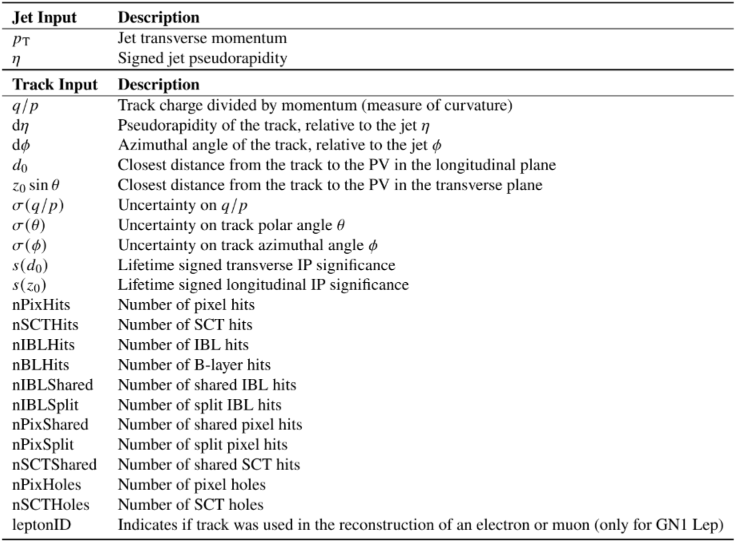
$$D_b(\mathrm{GN3}) = \log\left(\frac{p_b}{f_c \cdot p_c + f_\tau \cdot p_\tau + (1 - f_c - f_\tau) \cdot (p_{ud} + p_g + p_s)}\right)$$

- p_b: output score of the model corresponding to the b-jet class.
- p_c: output score of the model corresponding to the c-jet class.
- p_τ: output score of the model corresponding to the tau-lepton class.
- p_{light}: output score of the model corresponding to the light-jet class.
- p_{ud}: output score of the model corresponding to the up/down jet class.
- p_g: output score of the model corresponding to the gluon jet class.
- p_s: output score of the model corresponding to the strange jet class.
- f_c, f_τ: fractions controlling contribution of p_c and p_τ to the discriminant respectively.
- Note: $p_{\mathrm{light}} = p_{ud} + p_g + p_s$

Values of f_c, f_τ

- GN2 (b-tagging discriminant): $f_c = 0.2$, $f_\tau = 0.01$ (f_c tuned to reach a much higher c-jet rejection, while still achieving a better light-jet rejection, compared with DL1d)
- DL1d (b-tagging discriminant): $f_c = 0.018$ (no tau class)
- GN2 (c-tagging discriminant): $f_b = 0.3$, $f_\tau = 0.01$
- DL1d (c-tagging discriminant): $f_b = 0.1$

Input variables to GN1



Jet Input	**Description**
p_{T}	Jet transverse momentum
η	Signed jet pseudorapidity

Track Input	**Description**
q/p	Track charge divided by momentum (measure of curvature)
$\mathrm{d}\eta$	Pseudorapidity of the track, relative to the jet η
$\mathrm{d}\phi$	Azimuthal angle of the track, relative to the jet ϕ
d_0	Closest distance from the track to the PV in the longitudinal plane
$z_0 \sin\theta$	Closest distance from the track to the PV in the transverse plane
$\sigma(q/p)$	Uncertainty on q/p
$\sigma(\theta)$	Uncertainty on track polar angle θ
$\sigma(\phi)$	Uncertainty on track azimuthal angle ϕ
$s(d_0)$	Lifetime signed transverse IP significance
$s(z_0)$	Lifetime signed longitudinal IP significance
nPixHits	Number of pixel hits
nSCTHits	Number of SCT hits
nIBLHits	Number of IBL hits
nBLHits	Number of B-layer hits
nIBLShared	Number of shared IBL hits
nIBLSplit	Number of split IBL hits
nPixShared	Number of shared pixel hits
nPixSplit	Number of split pixel hits
nSCTShared	Number of shared SCT hits
nPixHoles	Number of pixel holes
nSCTHoles	Number of SCT holes
leptonID	Indicates if track was used in the reconstruction of an electron or muon (only for GN1 Lep)

- Same input variables used in GN2 except for the variables relating to holes in silicon tracker (no impact on performance).