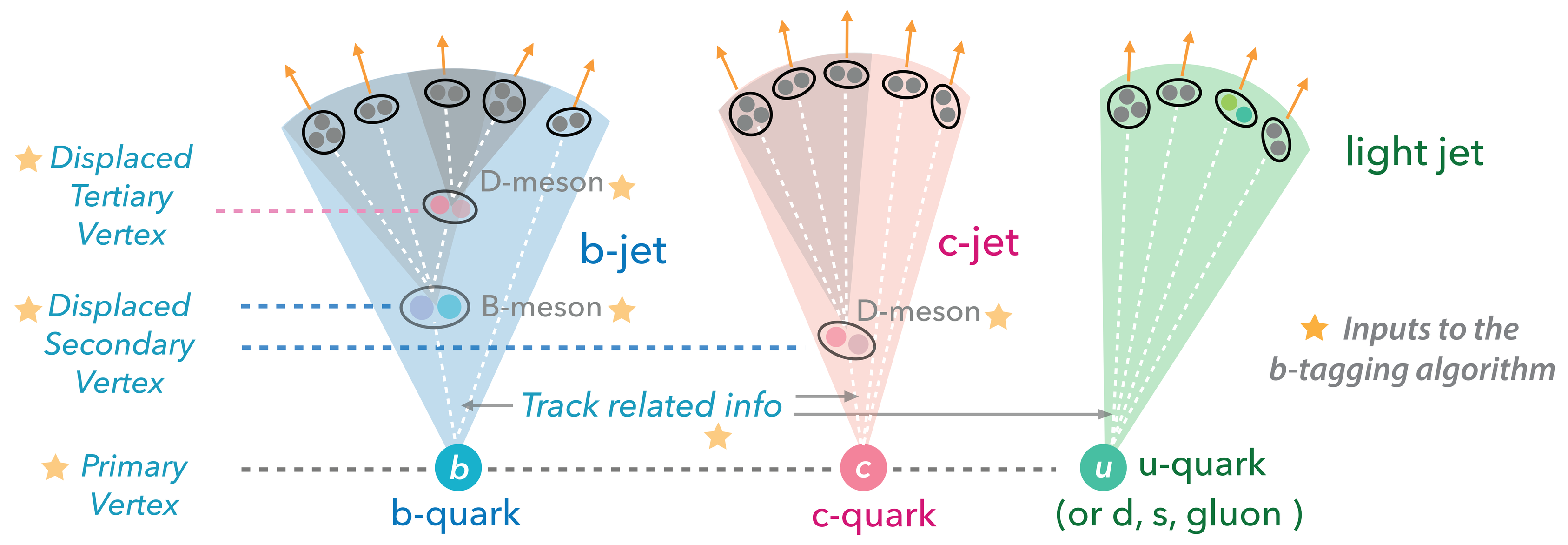
Graph Neural Network (GNN) based Truth-tagging Tool in ATLAS

BOOST 2022 (15^{th}-19^{th} August)

Lakmin Wickremasinghe (Osaka University), on behalf of the **ATLAS Collaboration**

1 Tagging Jets

b-tagging is done using a set of machine learning based algorithms (such as the ***DL1r algorithm***) which exploit B-hadron decay features to indentify jets. *Reference (2)

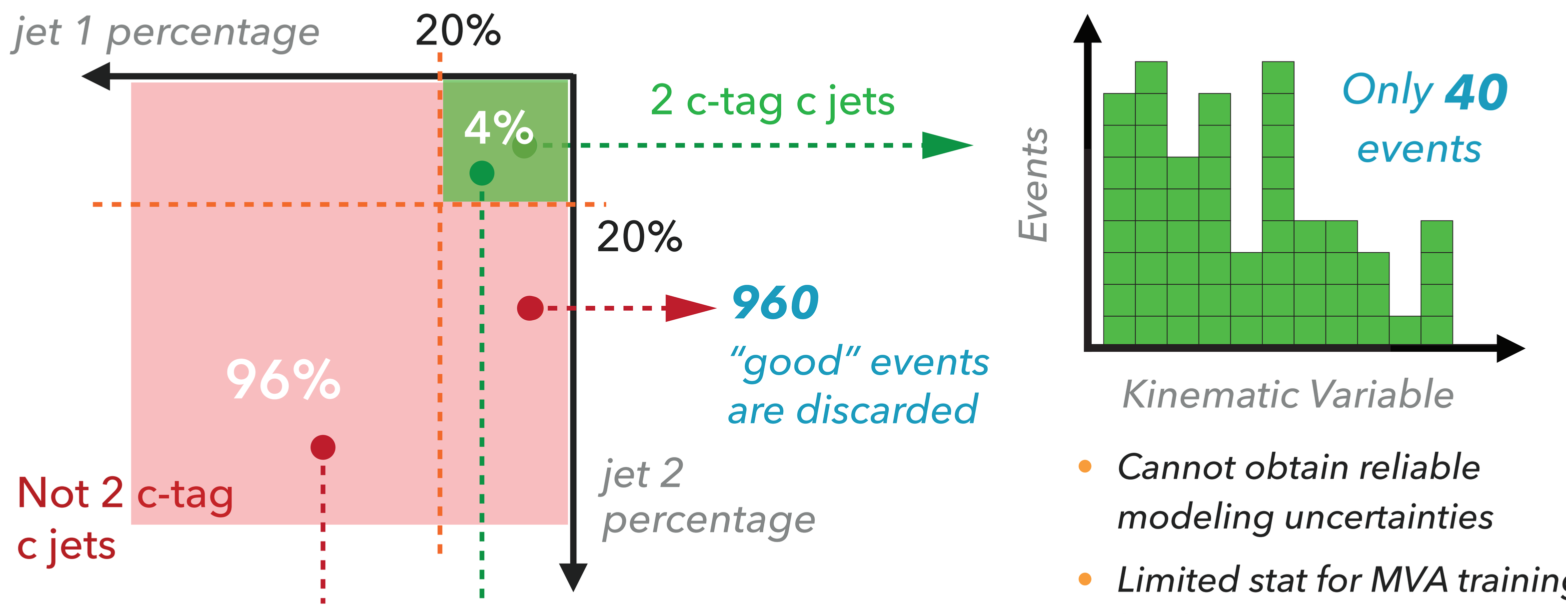


A jet is identified as tagged if its b-tagging score (given by the ***DL1r discriminant***) is higher than a certain threshold (***operating point***)

$$\text{Jet Tagging Efficiency } (\epsilon_{jet}) = \frac{\text{Number of tagged jets of a flavor}}{\text{Total number of jets of the same flavor}}$$

2 Direct Tagging vs Truth Tagging

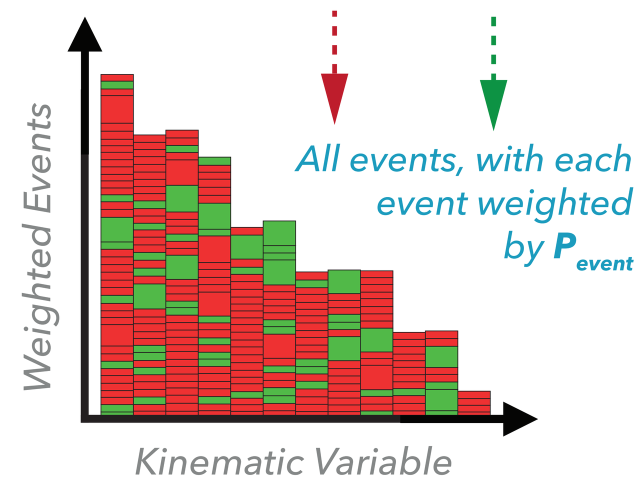
*Let's say we have **1000** events with c-jets, and ϵ_{jet} is **20%***



- *Cannot obtain reliable modeling uncertainties*
- *Limited stat for MVA training leads to overtraining*

What if we can use all the events?

Same problem, we look in a different way:
*What's the **probability P_{event}** of getting a 2 c-tag event?*



MC stat error is much smaller

This probability is,

$P(\text{event is tagged}|\theta)$

$= P(\text{jet}_1 \text{ is tagged}|\theta) \cdot P(\text{jet}_2 \text{ is tagged}|\theta) \cdot (1- P(\text{jet}_3 \text{ is tagged}|\theta)) \cdot \ldots$

$= \varepsilon_{j1}(x|\theta) \cdot \varepsilon_{j2}(x|\theta) \cdot (1 - \varepsilon_{j3}(x|\theta)) \cdot \ldots$

where εj(x|θ) the tagging efficiency of the jet j parametrised as a function of the set of variables θ

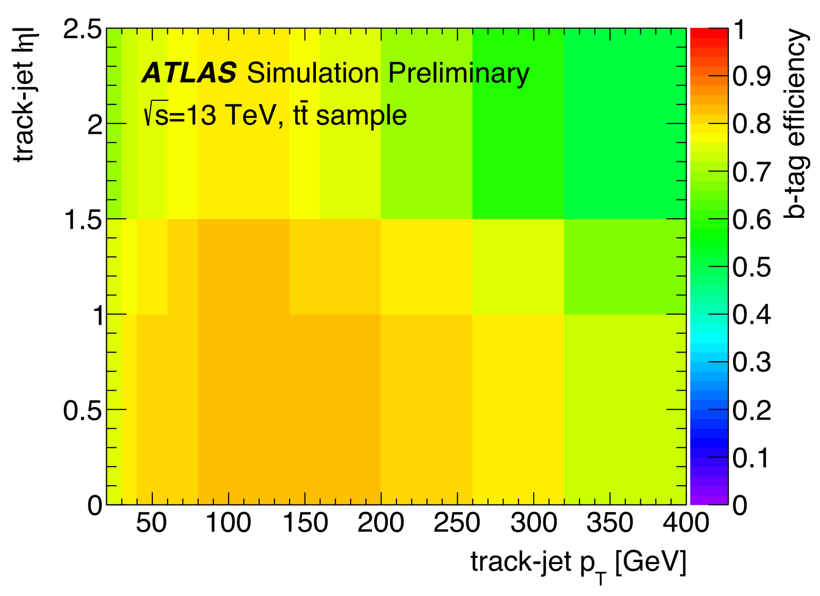
Hence, parametrization of $\boldsymbol{\varepsilon_{jet}}$ is important for this method to work!

The *DL1r discriminant bins*

B 85 - 100	B 77 - 85	B 70 - 77	B 60 - 70	B 60

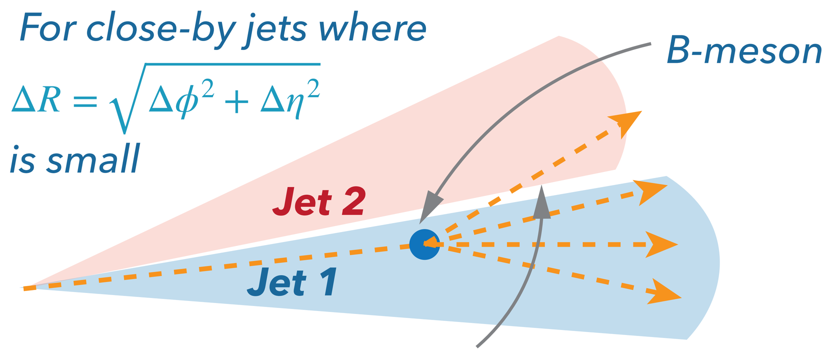
DL1r bin edges: 1.31, 2.06, 2.71, 3.88 (DL1r)

3 Effiency Maps vs GNN's



- Maps are defined from the two dominant parameters affecting ϵ_{jet} ; p_T and η
- *But ϵ_{jet} depends on multiple parameters, and is affected by close-by-jet effects*

For close-by jets where $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ is small



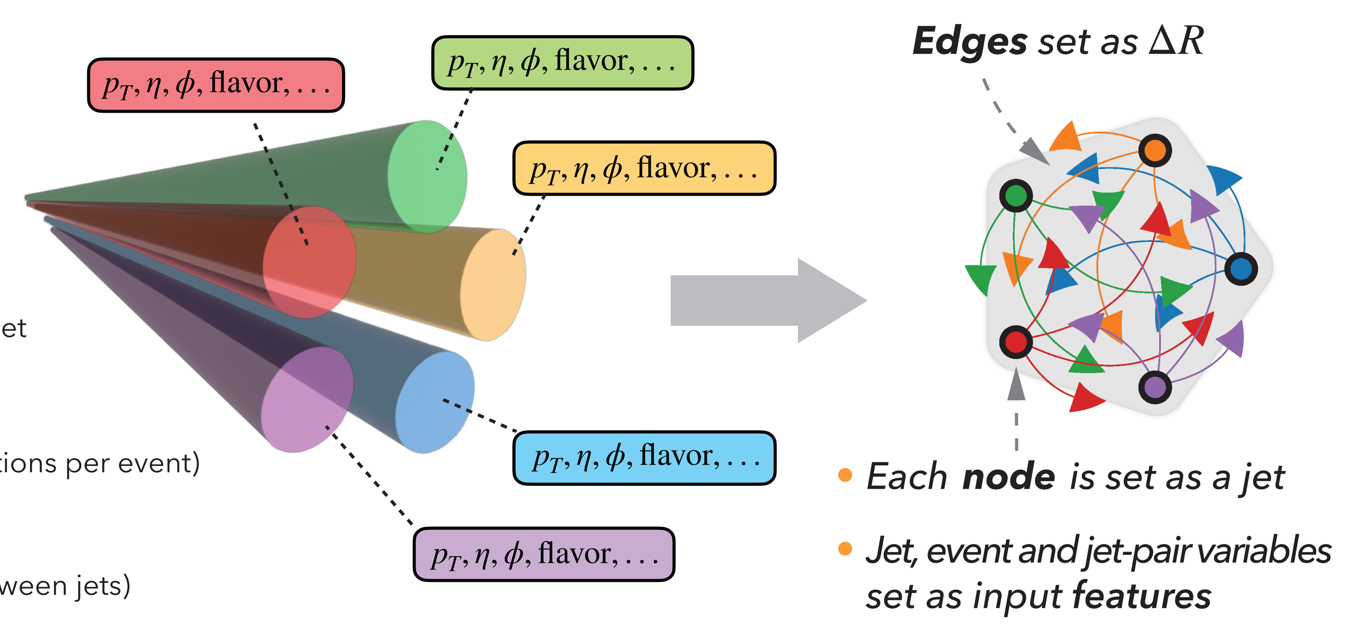
*Tracks from nearby **Jet 1** entering **Jet 2** affects the tagging efficiency of **Jet 2***

GNN Model

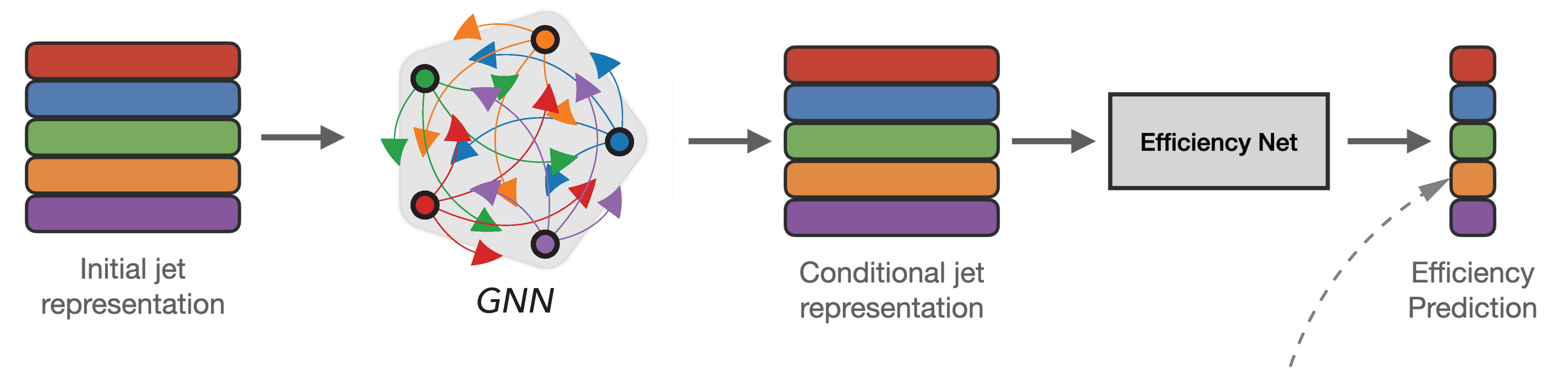
12 *Jet Variables*
Jet p_T, η, ϕ and *flavor*-label
mass, p_T, η, ϕ of the p_T leading *b* or *c* hadron in the jet

1 *Event Variable*
$\langle\mu\rangle$ (Avg. number of interactions per event)

1 *Jet-Pair Variables*
ΔR (Angular seperation between jets)



- Each **node** is set as a jet
- Jet, event and jet-pair variables set as input ***features***



Loss function: *Binary Cross Entropy*

$$\text{B.C.E} \rightarrow \frac{p(\theta)}{p(\theta)+q(\theta)} = \varepsilon(\theta) = \text{efficiency as a function of } \theta$$

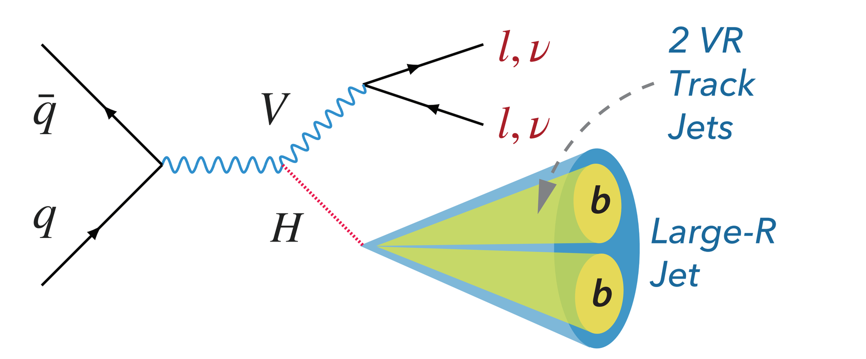
Where $p(\theta)$, $q(\theta)$ is the distribution of tagged, not-tagged jets

A vector of efficiencies (per DL1r bin) for each jet

Training (e.g):	0	0	0	1	0
To predict:	ε_1	ε_2	ε_3	ε_4	ε_5

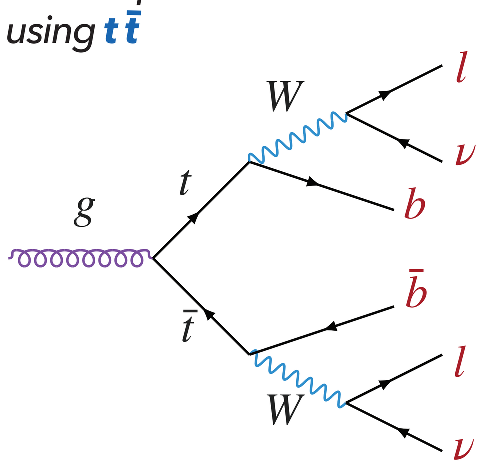
4 Performance (test case)

Analysis Scenario: Boosted VHbb



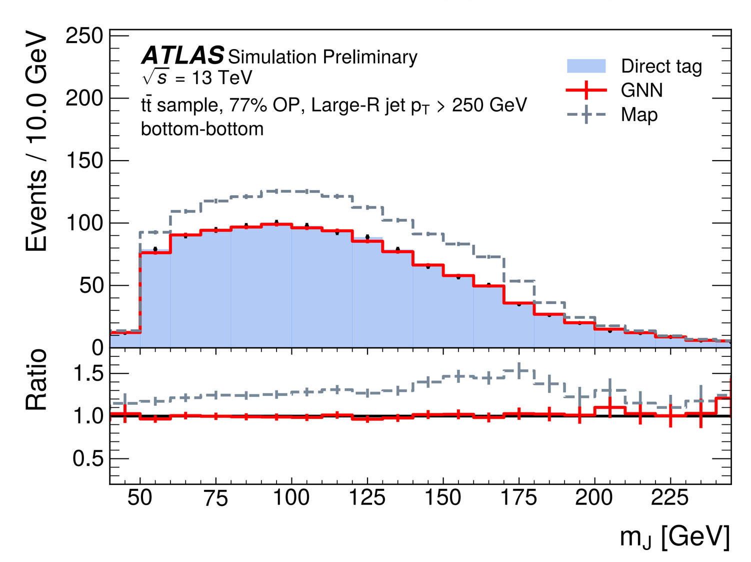
*Tagging is applied to the two leading **Variable Radius (VR) Track Jets** inside the Large-R Jet*

$\boldsymbol{t\bar{t}}$ is a major background in VHbb. The GNN models performance evaluation is done using $\boldsymbol{t\bar{t}}$

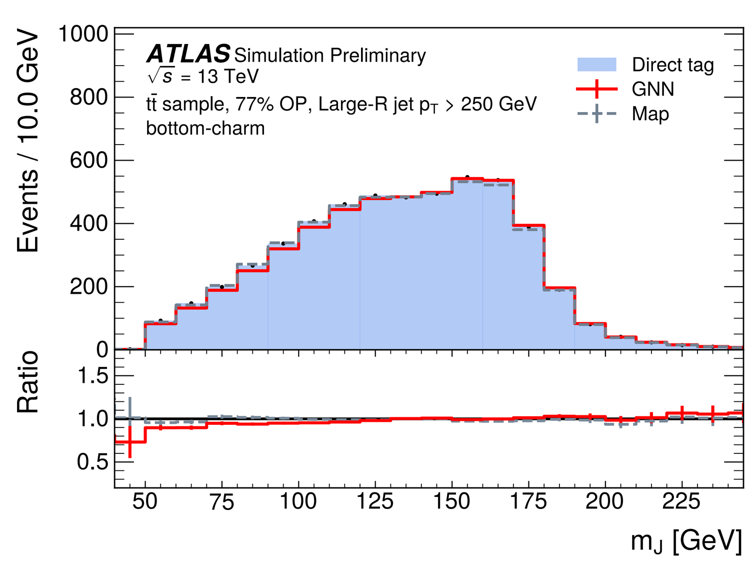


DL1r discriminant 77% Operating Point :
$\varepsilon_{b\text{-}jet} = 77\%$
$1/\varepsilon_{c\text{-}jet} = 5$
$1/\varepsilon_{light\text{-}jet} = 148$

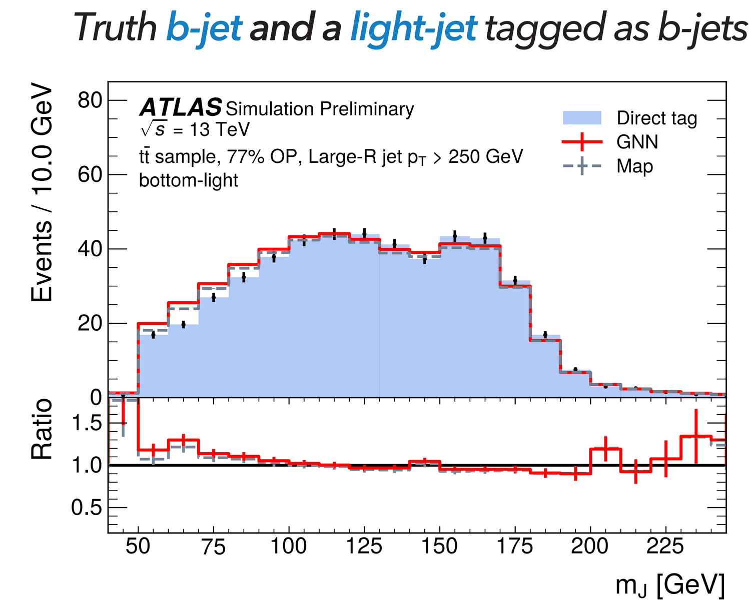
Two truth b-jets tagged as b-jets



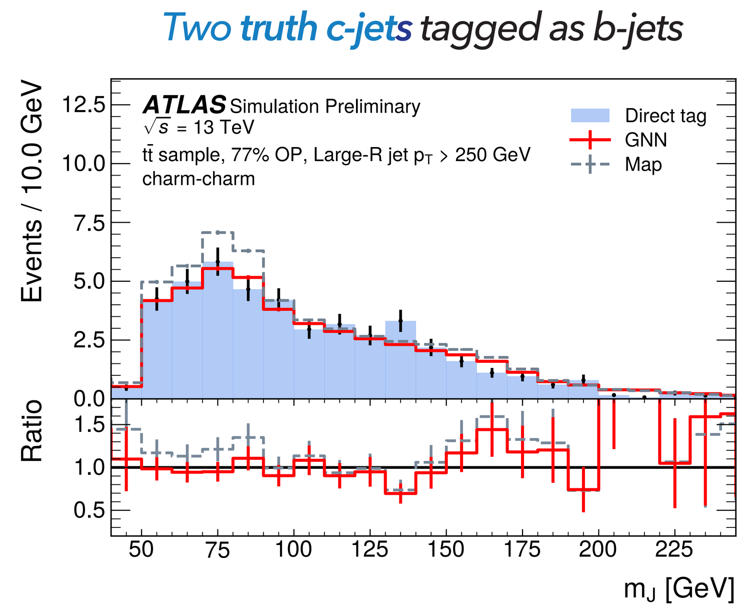
Truth b-jet and a c-jet tagged as b-jets



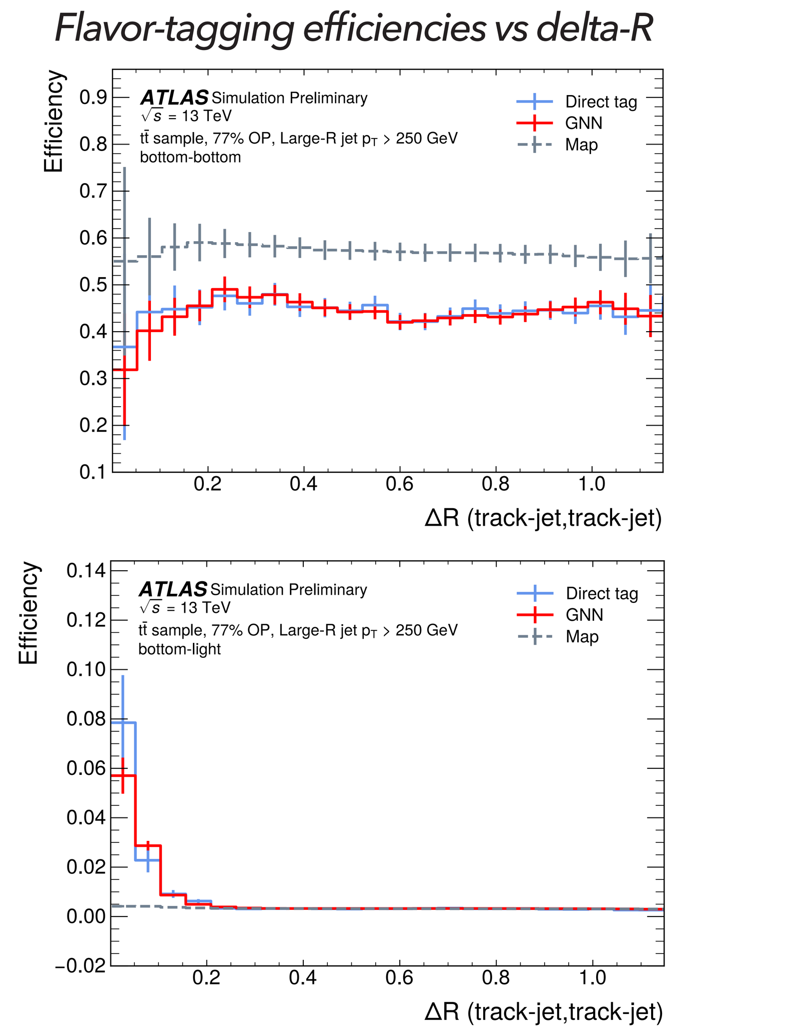
Truth b-jet and a light-jet tagged as b-jets



Two truth c-jets tagged as b-jets



Flavor-tagging efficiencies vs delta-R



Overall, the closure to direct tagging is better with GNN truth tagging, which confirms the superior performance of the GNN to parametrise the jet tagging efficiency, compared to maps.

References:

(1) ATLAS Collaboration, Flavour Tagging Efficiency Parametrisations with Graph Neural Networks, ATL-PHYS-PUB-2022-041

(2) ATLAS Collaboration, ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s}$ = 13 TeV, Eur. Phys. J. C 79 (2019) 970

(3) F. A. Di Bello et al., Efficiency Parameterization with Neural Networks, Comput. Softw. Big Sci. 5 (2021) 14