PROCEEDINGS OF SCIENCE

New ATLAS *b*-tagging Algorithm for Run-3

Martino Tanasini, on behalf of the ATLAS Collaboration*

Università di Genova,
Via Dodecaneso 33, Genova, Italia

E-mail: martino.tanasini@cern.ch

The ability to identify jets containing b-hadrons (b-jets) is of essential importance for the scientific programme of the ATLAS experiment at the Large Hadron Collider, underpinning the observation of the Higgs boson decay into a pair of bottom quarks, Standard Model precision measurements, and searches for new phenomena. The ATLAS flavour tagging algorithms rely on powerful multivariate and deep machine learning techniques. These algorithms exploit tracking information and secondary vertex reconstruction in jets to establish the jet's flavour. Both specifically designed observables sensitive to the distinct properties of b-jets and neural networks operating directly on the charged-particle tracks within the jet are used. In this proceeding, we review the state-of-the-art in flavour tagging algorithms developed by the ATLAS collaboration and of their expected performance using simulated data.

41st International Conference on High Energy physics - ICHEP2022
6-13 July, 2022
Bologna, Italy

*Speaker

https://pos.sissa.it/

PoS(ICHEP2022)1090

1. Flavour tagging in a nutshell

Flavour tagging is the identification of jets containing b-hadrons (b-jets) or c-hadrons (c-jets) and discriminating them from those containing neither c nor b-hadrons (light-flavour jets). Due to their long lifetime, of the order of 1.5 ps, b-hadrons have a significant mean flight length before decaying. This peculiar feature, in addition to the high masses and decay multiplicities of the b-hadrons, allows for the identification of b-jets. In particular, tracks associated to b-jets are expected to have larger impact parameters (Fig. 1), the impact parameter being defined as the minimum distance between the track and the primary interaction point. Furthermore, the displaced decays of b-hadrons typically are reconstructed as secondary vertices distinguished from the primary interaction point.

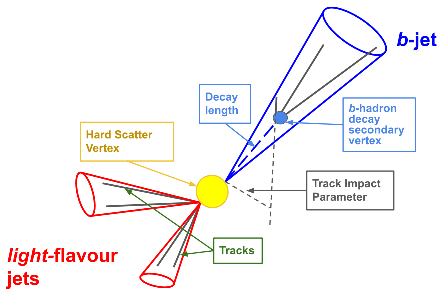


Figure 1: Visualisation of the features that distinguish b-jets (in blue) from *light*-flavour jets (in red). The displaced decays of b-hadrons lead to the presence of secondary vertices in the b-jets. Tracks associated to the charged particles produced in the decay can have large impact parameters.

2. The algorithms

2.1 DL1 series

The baseline b-tagging algorithms used in ATLAS [1] Run-2 analyses belong to the DL1 family [2]. They consist of Deep Neural Networks (DNN) combining the information produced by a set of low-level algorithms [3] about the kinematics of the charged particle tracks and about the presence and features of secondary vertices associated to the jets.

Low-level algorithms are divided into two broad categories. The first consists of track-based taggers that exploit the large impact parameter of tracks originated by b-hadron decays. IP2D and IP3D [3] assign per-track probabilities of coming from a jet of a given flavor and combine these probabilities in log-likelihood ratio (LLR) discriminants. RNNIP [4] is a Recurrent Neural Network (RNN) trained to identify the flavour of a jet from a set of track and jet features. It is capable of dealing with multidimensional representation of tracks and of exploiting the correlations between them. The order by which tracks in jets are given as input to RNNIP during the training affects its performance. Since there is no natural preferred way of ordering tracks in jets, a conventional ordering by decreasing impact parameter significance was chosen. DIPS (Fig. 2)[5] treats tracks as unordered sets. It has been shown to surpass RNNIP both in terms of training time and performance.

Recently, an improved version of the Dips algorithm trained on a dataset containing tracks selected with looser requirements compared to previous algorithms (DIPS loose [6]) has been developed.

The second category consists of algorithms that explicitly reconstruct displaced vertices: SV1 and JetFitter . SV1 reconstructs a single secondary vertex in the jet. It then provides features of the reconstructed vertex as inputs for DL1 algorithms. JetFitter aims at reconstructing the full b-hadron decay chain by exploiting the topological structure of weak b and c-hadron decays.

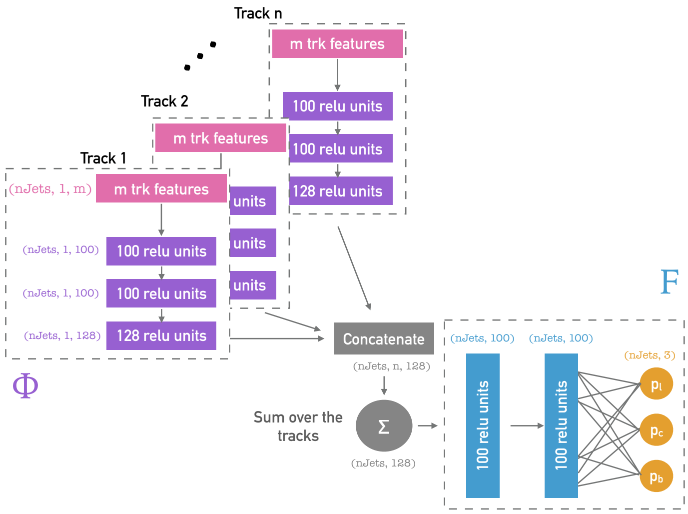


Figure 2: Architecture for the DIPS algorithm [5].

2.2 GN1

The current state-of-the-art flavour tagging algorithm GN1 (Fig. 3) [7] is based on a Graph Neural Network (GNN) which operates on the tracks associated to a jet as a fully connected graph. Differently than DL1 algorithms, GN1 is an algorithm with a monolithic structure. It directly operates on tracks to perform jet flavour-tagging. At the same time, it performs vertexing and track classification, which are defined as auxiliary tasks that guide the GNN toward an understanding of the underlying physics, hence removing the need for low-level algorithms.

3. *b*-tagging performance

The performance of a b-tagging algorithm is quantified by its power to reject c- and light-flavour jets for a given b-jet tagging efficiency. It can be visualized by means of Receiver Operating Characteristics (ROC) curves. They display the fraction of b-jets correctly b-tagged vs the fraction of non b-jets correctly rejected for any working point of a classifier. The larger the area under the ROC curve, the better the performance.

Fig. 4 shows the ROC curves of DL1r (a member of the DL1 series using RNNIP and not DIPS), GN1 and GN1Lep (a version of GN1 which includes an additional track-level input, indicating if the track was used in the reconstruction of an electron, a muon or neither). For a b-jet tagging efficiency of 70% GN1 outperforms DL1r light-flavour (c)-jet rejection by a factor of $\sim 1.8(\sim 2.1)$ for simulated jets coming from $t\bar{t}$ decays with transverse momentum $20 < p_T < 250$ GeV. For

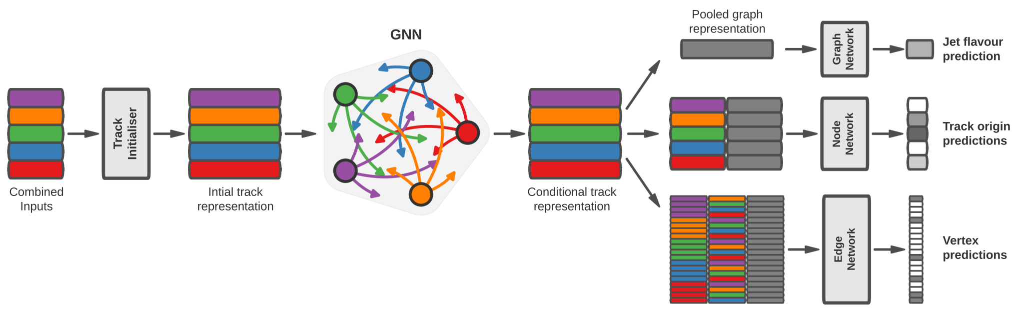
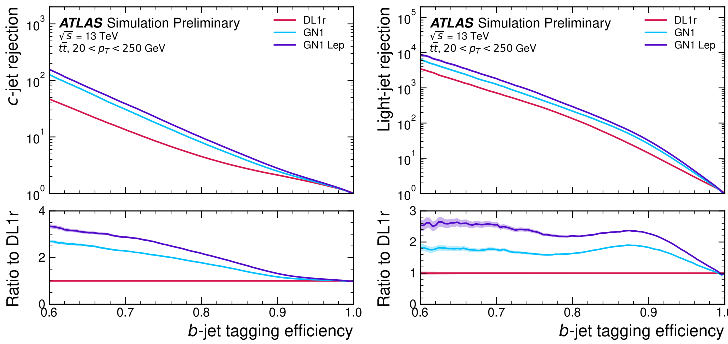
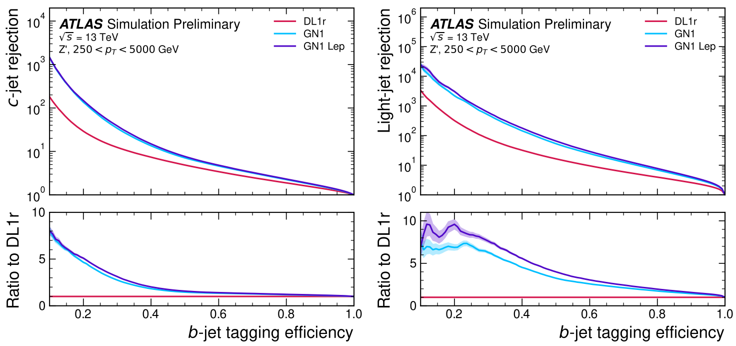


Figure 3: The network architecture of GN1 [7]. Inputs are fed into a per-track initialisation network, which outputs an initial latent representation of each track. These representations are then used to populate the node features of a fully connected graph network. After the graph network, the resulting node representations are used to predict the jet flavour, the track origins, and the probability of each track pair to belong to a common vertex.

simulated jets coming from Z' decays with transverse momentum $250 < p_T < 5000$ GeV the light-flavour (c)-jet rejection improves by a factor $\sim 6(\sim 2.8)$ for a comparative 30% b-jet efficiency.



(a) $t\bar{t}$



(b) Z'

Figure 4: The c-jet and light-jet rejections as a function of the b-jet tagging efficiency for jets in a simulated $t\bar{t}$ sample, with transverse momentum $20 < p_T < 250$ GeV (4a); and for jets in a simulated Z' sample, with transverse momentum $250 < p_T < 5000$ GeV (4b) [7].

References

[1] ATLAS Collaboration, *JINST* **3** (2008) S08003

[2] ATLAS Collaboration *JHEP* **03** (2020) 145

[3] ATLAS Collaboration *Eur. Phys. J. C* **79** (2019) 970

[4] ATLAS Collaboration *ATL-PHYS-PUB-2017-003*, https://cds.cern.ch/record/2255226

[5] ATLAS Collaboration *ATL-PHYS-PUB-2020-014*, https://cds.cern.ch/record/2718948

[6] ATLAS Collaboration, http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2021-004/

[7] ATLAS Collaboration, *ATL-PHYS-PUB-2022-027*, https://cds.cern.ch/record/2811135