IL NUOVO CIMENTO **45 C** (2022) 69
DOI 10.1393/ncc/i2022-22069-8

COMMUNICATIONS: SIF Congress 2021

Developing novel jet flavour tagging techniques at high transverse momentum for the ATLAS experiment at the LHC

M. TANASINI

Dipartimento di Fisica, Università di Genova - Genova, Italy

received 30 January 2022

Summary. — The kinematics of b-hadrons decay and the associated time of flight at high scales of momentum challenge conventional approaches to identify the jets that contain them. In the context of the ATLAS experiment, current identification algorithms are based on the reconstruction of charged tracks and of secondary decay vertices inside jets. A new experimental technique to identify the flavour of jets at the TeV scale is presented. I focus on the spatial distribution of clusters of silicon pixels activated by the passage of charged particles in the tracking system near to the identified jet axis. Exploiting the discriminant power of the patterns that they display could enhance the capability of the algorithms to identify the flavour of jets at an energy scale interesting for the searches of new physics at the LHC.

1. – Introduction

Due to colour confinement, QCD colour charges are never observed as isolated free-states. When one of such charges is produced in a collision at a particle collider like the Large Hadron Collider (LHC) [1], it undergoes a cascade particle production process, called hadronisation, that leads to a colourless final state. Given their collimated topologies, such states are called *jets*. Jets are thus the experimental signature that indicates the production of quarks or gluons in collision events.

Flavour tagging is the process of identifying jets containing b-hadrons (b-jets), discriminating them from jets containing c-hadrons and no b-hadrons (c-jets) and from those containing neither c nor b-hadrons (*light*-flavour jets). Efficient flavour tagging algorithms are of great importance for many aspects of the scientific programme of high-energy physics experiments at the LHC, spanning from the famous observations of the Higgs boson decay into a pair of bottom quarks made by ATLAS [2] and CMS [3] in 2018 [4, 5] to precision measurements of the Standard Model (SM), to searches for new phenomena.

Flavour tagging in ATLAS is done by means of an algorithm called DL1r [6], a Deep Neural Network (DNN) that combines into three scalar outputs the information from the

kinematics of the charged particle tracks and about the presence and features of secondary vertices associated to the jets. These outputs represent the probabilities assigned by the algorithm to each jet of being a b, c or *light*-flavour jet. Section **2** is dedicated to describe DL1r in greater detail.

DL1r performance decreases when it is used to select jets with high transverse momentum (*i.e.*, with a high value of the projection of the momentum in the plane perpendicular to the beam-line) p_T. This behaviour can be understood by considering how the jet environment changes when more energy and momentum are at stake, and its explanation is the object of sect. **3**.

It has been shown in [7] that making a direct use of the clusters of silicon pixels activated by the passage of charged particles in the inner parts of a detector can restore some of the discriminating power which is lost when dealing with high-p_T jets. This happens because the displaced decays of high-p_T b-hadrons produce patterns in the geometrical disposition of clusters in jets which are recognizable even when the jet environment is too extreme to properly reconstruct high quality tracks and secondary vertices. In sect. **4**, I describe the simplest possible discriminating variable that can be obtained by using directly the cluster information.

In sect. **5**, I describe the Boosted Decision Tree (BDT) that I trained on ATLAS Monte Carlo (MC) simulated samples by making use of the cluster information in addition to a track-based flavour tagging algorithm. Preliminary results suggest that the additional cluster information enhances the algorithm's capability of separating b-jets from *light*-flavour jets in the 0.1–1.5 TeV p_T range with peaks of 30% around 0.5 TeV.

The fact that very simple cluster-based discriminating variables are capable of boosting the performance of a much more sophisticated track-based algorithm is promising and encourages to think about refined ways of exploiting the cluster topology for the flavour-tagging classification task. Section **6** summarizes this paper and indicates some conceivable developments that will be the object of future studies.

2. – Flavour tagging in ATLAS

Flavour tagging in ATLAS is largely based on the tracks produced by charged particles in the Inner Detector (ID) and associated to the jets. The ATLAS ID system [8] is the first encountered by particles generated in the collisions, and consists of three subsystems: a silicon pixel detector, made of four layers of active modules placed at increasing radii with respect to the beam pipe, a silicon micro-strip detector (SCT) with four layers, and a transition radiation tracker (TRT) made of straw tubes. The ID is placed inside a solenoid magnet providing a 2 T axial field and covers the $|\eta| < 2.5$ region of acceptance.

Due to their long lifetime, of the order of 1.5 ps ($\langle c\tau \rangle \approx 450$ μm), b-hadrons have a significant mean flight length $\langle l \rangle = \beta\gamma c\tau = p\tau/m_0$ in the ID before decaying. This peculiar feature, in addition to the high masses and decay multiplicities of the b-hadrons, allows identifying the b-jets. In particular, tracks associated to jets containing b-hadrons are expected to have higher impact parameters (fig. 1). Furthermore, the displaced decays of b-hadrons generally lead to the presence of at least a secondary vertex in the b-jets, distinguished from that of the primary interaction point.

A set of b-tagging algorithms has been developed by the ATLAS Collaboration [9]. They take advantage of the various features that distinguish the signatures of b-hadrons in the detector to select b-jets. First, low-level algorithms use features of tracks and secondary vertices associated to the jets to provide educated guesses about their identity.

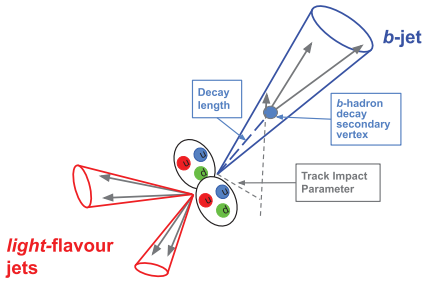

Fig. 1. – Visualisation of the features that distinguish jets containing b-hadrons (b-jets, in blue) from jets not containing either b or c-hadrons (*light*-flavour jets, in red). The displaced decays of b-hadrons lead to the presence of secondary vertices in the b-jets. Tracks associated to the charged particles produced in the decay can have high impact parameters.

Then, their outputs are collectively used as inputs of DL1r, a multivariate high-level algorithm that produces the discriminant variable which is used for the final selection.

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. They are called IP2D, IP3D, RNNIP and DIPS. IP2D and IP3D assign per-track probabilities of coming from a jet of a given flavor and combine these probabilities in log-likelihood ratio (LLR) discriminants. The former uses the significance of the transverse impact parameters of tracks, the latter uses also the significance of their longitudinal impact parameter. The LLR technique cannot deal with higher dimensional representations of tracks and it does not allow to account for the correlations between them. RNNIP [10], a Recurrent Neural Network (RNN) that represents tracks via multiple variables including the significance of their transverse and longitudinal impact parameters, their share of the total jet p_T and their angular separation with respect to the jet axis, is capable of exploiting these correlations instead. The order by which tracks in jets are given as input to RNNIP during the training affects its performance. This fact is a source of ambiguity, since there is no natural preferred way of ordering tracks in jets. DIPS [11] is a DeepSet based NN that solves this ambiguity by treating the tracks as unordered.

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 uses the features of the reconstructed vertex to provide eight discriminant variables to pass to DL1r. They include the number of tracks associated with the SV1 vertex, the invariant mass of the secondary vertex, its energy share with respect to all the tracks associated with the jet, and the three-dimensional decay length significance.

JetFitter aims at reconstructing the full b-hadron decay chain by exploiting the topological structure of weak b and c-hadron decays. Hence, it can isolate also tertiary vertices when present. Similarly to SV1, it provides eight discriminating variables, including the track multiplicity at the JetFitter displaced vertices, the invariant mass of tracks associated with these vertices, their energy share and their average three-dimensional decay length significance, which are then used as DL1r inputs.

With the exception of DIPS, all outputs of the low-level taggers are combined in DL1r, a version of the DL1 algorithm [12] upgraded by adding the RNNIP outputs to its list of inputs. DL1r is a fully connected feed-forward NN with three scalar outputs, p_b, p_c and p_u, that represent the probabilities it assigns to each jet of being a b, c or *light*-flavour

jet. p_b, p_c and p_u are combined into a single discriminant variable D_{DL1r}, defined as

$$D_{\mathrm{DL1r}} = \ln\left(\frac{p_b}{f_c p_c + (1 - f_c), p_u}\right) \tag{1}$$

where f_c is a tunable parameter representing the effective c-jet fraction in the background training sample. Selections on the values of D_{DL1r} are used to define regions in which enter jets that are b-tagged with increasing confidence.

3. – b-jet identification performance at high p_T

DL1r belongs to a category of algorithms called *classifiers*, whose aim is to predict the class to which the items of a dataset belong according to some identifying features. Classifiers are trained on labelled data: knowing the belonging class of each item of a dataset permits to tune the parameters of a classifier to model the correlations between the identifying variables and the belonging class. Their performance is then evaluated on an independent labelled dataset by considering the fraction of items correctly and incorrectly classified in the different classes.

DL1r is trained and evaluated on MC simulated $t\bar{t}$ and Z' samples produced by the ATLAS Collaboration. Since they have very different kinematic distributions (the second being enriched in high p_T jets), training the model on both allows using it in many different types of analyses. Its performance in separating b from c and *light*-flavour jets is depicted in Receiver Operating Characteristic (ROC) curves. They are obtained by moving the selection cut on D_{DL1r} and plotting the resulting fraction of correctly identified b-jets (b-tagging efficiency) against the rate of *light*-flavour or c-jets misidentified as b-jets (fake rate) [13].

DL1r performance is correlated with various features of the jets it classifies and of the events in which they are produced. In particular, it gets worse when dealing with jets with very high p_T (fig. 2). This loss can be attributed to multiple concurrent factors, that are illustrated in the following.

The complex process that starts from the production of a free quark or gluon in a collision, evolves via its QCD radiation and ends in the colourless jet final state resulting from the hadronization of the radiated coloured particle is called fragmentation. The multiplicity of charged particles produced in a jet during the fragmentation process is proportional to the p_T of the initiating quark or gluon [14]. This correlation is clearly shown in fig. 3, in which the fraction of tracks associated to charged particles produced in the fragmentation process in b-jets is shown as a function of the jet p_T. It can be seen that this fraction goes from 0.1 at 50 GeV to more than 0.3 at 1 TeV. In the same p_T range, the fraction of reconstructed tracks associated to charged particles belonging to the b-hadron decay chain peaks at 0.3 at 250 GeV and is reduced to 0.15 at 1 TeV. As I explained in detail in sect. **2**, low-level algorithms take advantage of this second category of tracks to identify b-jets. Hence, a reduction of its relative fraction with respect to that of the tracks from fragmentation blurs the differences that are exploited by the low-level algorithms to discriminate between the different flavours of jets, contributing to the decrease of DL1r performance.

Figure 3 also shows an increase of fake tracks, *i.e.*, tracks with less than 50% of the hits produced by a single simulated particle, as a function of the jet p_T. This increase, as well as the reduction in the number of reconstructed tracks associated to the b-hadron decay chain, arises from the difficulties raised by tracking in dense environments [15].

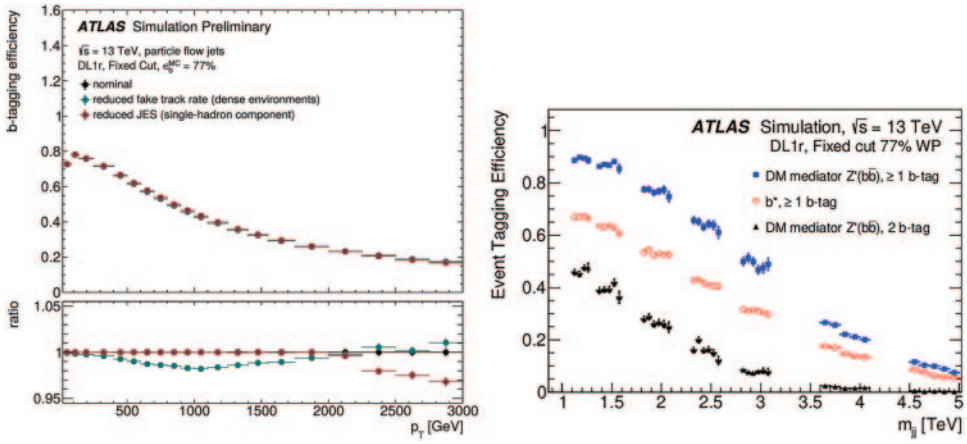

Fig. 2. – (Left) The DL1r b-tagging efficiency for a fixed value of the selection cut on D_{DL1r} (corresponding to an amount of 77% of the total number of b-jets correctly identified) as a function of the jet p_T in simulation [16]. The blue and red markers correspond to the efficiency in the presence of uncertainty components affecting the track reconstruction and the jet momentum calibration, respectively. (Right) In the search for new dijet resonances described in [6], dijet events originated by the decay of a hypothetical Z' boson from a leptophobic Z' dark-matter mediator model (blue and black dots) and of an excited b-quark from a compositness model (red dots) are required to contain at least one (1 b-tag) or exactly two (2-btag) b-tagged jets to be selected. The fraction of those correctly selected by the b-tagging cut decreases as a function of the mass of the dijet resonance m_{jj}. This happens because high mass dijet resonances decay in high p_T b-jets, more difficult to identify.

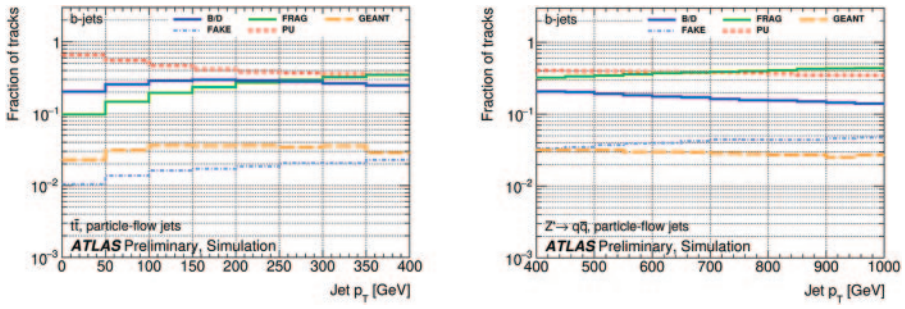

Fig. 3. – Fraction of reconstructed tracks from different categories associated to b-jets as a function of the jet transverse momentum p_T, depicted for a sample of simulated $t\bar{t}$ events (left) and $Z' \to q\bar{q}$ (right). The categories are defined according to the origin of tracks: tracks belonging to the decay chains of b or c-hadrons (B/D), tracks associated to the other charged particles produced in the fragmentation process (FRAG), secondary charged particles from interactions with the detector materials or from decays of long-lived hadrons (GEANT), and tracks from pileup (PU). Tracks with less than 50% of the hits produced by a single simulated particle are classified as fake tracks (FAKE). The relative fraction of fake and fragmentation tracks with respect to the other categories increases at high p_T [17].

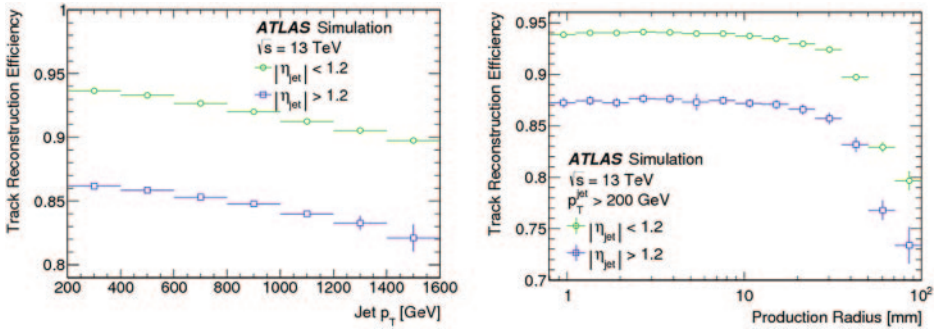

Fig. 4. – The track reconstruction efficiency is compared for charged particles produced before the IBL in jets with $|\eta| < 1.2$ ($|\eta| > 1.2$) as a function of the jet p_T (left) and of the production radius of the charged particles for simulated dijet MC events, where charged particles are not required to be created before the IBL (right) [15].

Dense environments are regions of the detector with many charged particles divided by low angular separations. Being sprays of collimated particles, hadronic jets are typical examples. The density of the environment in jets is correlated to their p_T both because of the increase in multiplicity of particles from fragmentation and because Lorentz boosts shrink their cones proportionally to their momentum. These combined effects result in the drop of efficiency of reconstructing tracks in high p_T jets, shown in fig. 4 (left) for particles produced before the Insertable B-Layer (IBL), the innermost layer of the ATLAS pixel detector.

The situation is even more complicated when dealing with the reconstruction of tracks associated to charged particles belonging to the chain of b-hadrons displaced decays. If the p_T of a b-hadron is high enough, it may decay beyond the first layers of the ID. When this happens, track reconstruction efficiency is affected in two ways. Particles produced in decays happening beyond the first active layers of the ID create fewer clusters, since they do not interact with those situated upstream to the decay point. Furthermore, since they are in general closer to the next active layer with respect to those generated in the primary interaction point, the average separation between the clusters they produce is smaller. The combination of these two factors leads to the drop in track reconstruction efficiency for particles with high production radius (*i.e.*, produced distant from the beam line in the plane perpendicular to it) shown in fig. 4 (right).

An effect of the combined action of all these features of high p_T jets is depicted in fig. 5. There, it is shown that the amount of b-jets with a reconstructed secondary vertex decreases at high p_T. Instead, it increases in c and *light*-flavour jets: the latter are to be considered as fake. The depicted trend is exacerbated for jets with p_T greater than 800 GeV, such as those produced in the decay of the dijet resonances with masses at the TeV scale shown in fig. 2 (right).

4. – Patterns in the distribution of clusters in high p_T b-jets

It has been observed in [7] that the distributions of clusters in b-jets containing b-hadrons that decay beyond the first layers of the ID present peculiar patterns. In fig. 6, I represented schematically such distributions for idealized high p_T and low p_T b-jets, and for a *light*-flavour jet. When a b-hadron decays between two pixel layers, the multiple charged particles it produces do not interact with the layers upstream to the decay point.

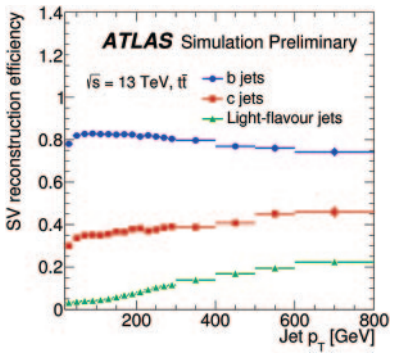

Fig. 5. – The SV1 reconstruction efficiencies as a function of jet p_T [18].

Thus, an increase in the number of clusters is registered in the first layer beyond the decay point. The increase is not present in the idealized b-jet whose b-hadron decays before the first pixel layer, nor in the idealized *light*-flavour jet.

Some simple calculations allow predicting the momentum scale above which b-hadron decays beyond the first layer of the ID are to be expected. Given b-hadron masses of the order of 5 GeV, imposing an average flight length l of 33.25 mm (the radial distance of the first pixel layer to the beam-line) in the formula $l = \gamma\beta c\tau = p\tau/m_0$, and inverting it to find the momentum p leads to a momentum of the b-hadron of about 370 GeV for the average value $< c\tau > \approx 450\ \mu$m. Accounting for the fact that b-hadrons momenta carry an average of 85% of the total momenta of the jets in which they are contained [19], allows to conclude that decays between two layers of the ID become likely for jets with p_T greater than 400 GeV. This value is close to that for which DL1r performance starts to decrease (fig. 2). Hence, being able to exploit the increase in the number of clusters per pixel layer as an additional feature to identify b-jets would be especially beneficial where track-based and secondary-vertexing algorithms start to falter.

The visual representations of jets in fig. 6 are idealized in the sense that, modulo the displaced b-hadron decay in the high p_T scenario, they display a constant number of clusters per pixel layer. This is not what happens with real jets. Multiple charged particles can be created in the interactions of particles with the detector material. Others may be absorbed by the pixel layers. Long Lived Particles (LLP) that are not b or c-hadrons may be present in the jets and decay between two layers of the ID. Local

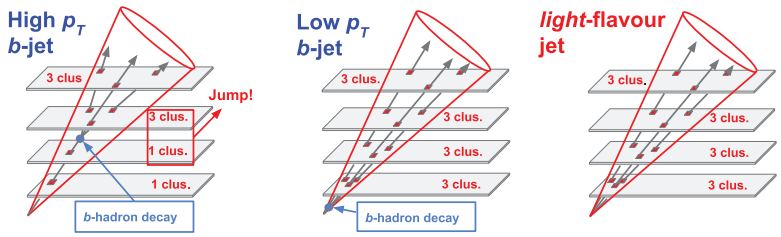

Fig. 6. – Distribution of clusters in idealized high and low p_T b-jets and in a *light*-flavour jet. The number of clusters increases in the first pixel layer beyond the displaced decay of the b-hadron in the high p_T scenario. This identifying feature can be used to boost the performance of b-tagging algorithms in the p_T region in which it decreases the most.

variations in the number of particles produced in pile-up proton collisions erroneously associated to the jets can affect the number of clusters registered inside them. Particles could exit the reconstructed jet cone and the clusters they activated in the outer layers would then be lost to the jet.

Notwithstanding these sources of background, in [7] it is indicated that the patterns in the distributions of clusters inside jets can be used to identify those containing b-hadrons. I will now describe the algorithm that I developed to test the same ideas with the proper tools and MC simulations developed for the ATLAS experiment, and how I quantified the improvement they can provide to the performance of the track-based IP3D low level tagger when dealing with jets with high p_T.

5. – A Boosted Decision Tree using cluster-based information and the IP3D outputs to identify high p_T b-jets

I used TMVA [20], the toolkit for multivariate analysis integrated in the ROOT analysis framework [21] to train a Boosted Decision Tree (BDT) to classify b-jets and *light*-flavour jets belonging to the MC simulated Z' sample described in detail in [12]. The sample is designed to exhibit a flat p_T spectrum up to 6 TeV, allowing for very high p_T studies.

I imposed some additional cuts to the sample. For the sake of simplicity, jets are required to have a value $|\eta| < 1$, hence to be reconstructed in the barrel. b-jets containing more than one b-hadron are discarded, as well as those whose b-hadron decays outside the silicon pixel detector.

As an input for the BDT, I represented each jet by its p_T and $|\eta|$, by the likelihood assigned by the IP3D algorithm to the the b and *light*-flavour jet hypotheses, by the LLR between these two, by the four numbers of clusters per pixel layer in a cone with angular opening $\mathrm{dR} = 0.04$ around the jet axis, and by the three differences between subsequent pairs of these four numbers.

The last three variables proved to be redundant: training the BDT with the four numbers of clusters per pixel layer or adding the three differences as well resulted in identical performance of the algorithm.

I used only clusters activated inside cones with an angular opening of $\mathrm{dR} = 0.04$ around the jet axis because 90% of those associated to the b-hadrons decay chain fall in that region. All the clusters not associated to such decays, instead, are evenly distributed in the jet cone. Loosening or tightening the value of the cut on dR resulted in a decrease in the performance of the algorithm.

I used a training sample of approximately 10^5 b-jets and 4×10^5 *light*-flavour jets. To avoid biases in the model due to the over-representation of one class, I renormalized the number of *light*-flavour jets to the number of b-jets. p_T and $|\eta|$ are included among the input variables to exploit their eventual correlations with the others. Since p_T is distributed differently in the two classes of jets, its spectrum for the b-jets is reweighted to match that of the *light*-flavour jets. This avoids biases in the model that may cause it to assign the flavour of the jets based on the absolute values of their kinematic properties.

The BDT training hyperparameters are listed in table I. They were chosen adapting the values optimised for the MV2 b-tagging algorithm [9] and oscillating them while monitoring the performance of the obtained model. The BDT proved to perform similarly when moving its training hyperparameters around the chosen values. This can be explained by the limited number of input variables used by the model, that make its convergence rather simple.

TABLE I. – *List of the hyperparameters used to train the BDT. Each one is defined in* [20].

Parameter	Tuned value
Number of trees	850
Maximum depth	Number of input variables
Minimum node size	0.5%
Separation type	Gini Index
Number of cuts	20
Boosting type	AdaBoost
β	0.5
Bagged sample fraction	0.3

5˙1. *Results*. – I compared the performance of the BDT when trained using p_T,$|\eta|$, and the IP3D variables and when adding also the four numbers of clusters per pixel layer. The comparison allows quantifying the boost given by the additional information on the distributions of clusters in jets to the performance of the track-based IP3D algorithm.

The evaluation sample consisted of approximately 4×10^5 *light*-flavour and 10^5 b-jets, to which no global or kinematic reweighting procedure was applied. I split the sample in five bins according to the p_T of the jets. The bins divided the p_T range of 100–2500 GeV in five intervals, the first being 400 GeV wide, the others 500 GeV each.

In each bin, I computed the ratio between the ROC curves of the BDT trained with and without including the cluster-based variables: the performance is increased everywhere. Considering b-tagging efficiencies above 60% (lower working points are not considered as interesting and have never been used for physics analyses by the ATLAS Collaboration), the ratio of the *light*-flavour jet rejections between the two models is maximal for the 100–500 and 500–1000 GeV p_T bins. There, the ratios differ only by slight fluctuations, and show a maximum increase of about 35% at 60% working point. The increase in the performance is less pronounced for the 1000-1500 GeV p_T bin, where, at the b-tagging efficiency of 60%, it is of 20%; and it is minimal in the 1500–2000 and 2000–2500 GeV p_T bins, where the very similar ratio curves show a maximal increase of 10% at the b-tagging efficiency of 60%.

I then computed the ratio of the *light*-flavour jet rejections between the two models as a function of the p_T of the jets for a value of the b-tagging efficiency fixed at 77%. In accord with the previous observations, the ratio shows a maximal increase of 30% around 500 GeV, and settles down to values close to 1 for jets with p_T greater than 1500 GeV. Considerations analogous to those put forward in sect. **4** indicate that for such b-jets the average decay lengths of the b-hadrons become greater than the radial distance of the outermost layer of the pixel detector to the beam line. Given the cut that I imposed on b-jets whose b-hadrons decay outside the silicon pixel detector, it is guaranteed that the decay products of all the b-hadrons in the sample cross at least one pixel layer in the ID. Nevertheless, when the decay happens very close to the outermost pixel layer, the charged particles produced may not have the time to separate enough to be resolved by the detector before crossing it. Such cases are increasingly likely for b-jets with p_T greater than 1500 GeV. When they occur, the distribution of clusters in the associated b-jets do not display the identifying feature described in sect. **4**. Hence, a value of the ratio close to 1 for the BDT evaluated on these jets is to be expected.

6. – Conclusion

Identifying jets containing b-hadrons at high scales of momentum challenge conventional approaches. In this study, a BDT was trained in the context of the ATLAS experiment to classify b-jets and *light*-flavour jets belonging to a MC simulated sample enriched in high p_T jets. The training was done both by representing jets by their p_T, $|\eta|$ and by simple track-based variables based on the LLR-inspired IP3D algorithm, and by adding innovative cluster-based variables in the vicinity of the jet axis for capturing discriminant information which is at least partially lost when traditional flavour-tagging techniques are applied to jets with high momenta. The addition proved to enhance the performance of the model when evaluated on jets in the 0.1–2.5 TeV p_T spectrum.

This result paves the way to think about refined ways to exploit the identifying potential contained in the distributions of clusters in jets. Algorithms related to reconstructing the jet substructure may succeed in increasing the relative fraction of clusters associated to the b-hadron decay chain by discarding some of the others. The use of NNs such as Graph Neural Networks (GNN) could be tested to see whether they are capable of exploiting unanticipated correlations in the cluster distributions to identify the b-jets. These ideas will be the object of future studies.

REFERENCES

[1] Evans L. and Bryant P., *JINST*, **3** (2008) S08001.
[2] ATLAS Collaboration, *JINST*, **3** (2008) S08003.
[3] CMS Collaboration, *JINST*, **3** (2008) S08004.
[4] ATLAS Collaboration, *Phys. Lett. B*, **786** (2018) 59.
[5] CMS Collaboration, *Phys. Rev. Lett.*, **121** (2018) 121801.
[6] ATLAS Collaboration, *JHEP*, **03** (2020) 145.
[7] Huffmann B. T., Jackson C. and Tseng J., *J. Phys. G*, **43** (2016) 085001.
[8] ATLAS Collaboration, `https://cds.cern.ch/record/331063`.
[9] ATLAS Collaboration, *Eur. Phys. J. C*, **79** (2019) 970.
[10] ATLAS Collaboration, ATL-PHYS-PUB-2017-003.
[11] ATLAS Collaboration, ATL-PHYS-PUB-2020-014.
[12] ATLAS Collaboration, ATL-PHYS-PUB-2017-013.
[13] ATLAS Collaboration, `http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2019-005/`.
[14] ATLAS Collaboration, *Phys. Rev. D*, **100** (2019) 052011.
[15] ATLAS Collaboration, *Eur. Phys. J. C.*, **77** (2017) 673.
[16] ATLAS Collaboration, ATL-PHYS-PUB-2021-003.
[17] ATLAS Collaboration, `http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2021-003/`.
[18] ATLAS Collaboration, ATL-PHYS-PUB-2017-011.
[19] Peterson C., Schlatter D., Schmitt I. and Zerwas P. M., *Phys. Rev. D.*, **27** (1983) 105.
[20] Hoecker A. *et al.*, arXiv:physics/0703039.
[21] Brun R. and Rademakers F., *Nucl. Instrum. Methods A.*, **389** (1997) 81.