Draft version 1.1

ATLAS NOTE

ATLAS-CONF-2011-XXX

June 6, 2012

Identification and Tagging of Double *B*-hadron jets from Gluon-Splitting to $b\bar{b}$ with the ATLAS Detector

The ATLAS Collaboration

Abstract

A method is presented allowing the identification of merged b-quark jets originating from gluon-splitting. The technique exploits the kinematic differences between merged and single B-hadron jets using track-based jet shape and jet substructure variables combined in a multivariate likelihood discriminant. The ability to reject b-jets from gluon splitting is important in order to reduce and to improve the estimation of the b-tag background in Standard Model analyses and in new physics searches. In the simulation, the algorithm rejects 95% (50%) of merged B-hadron jets while retaining 50% (90%) of the single B-hadron jets, although the exact values depend on the jet p_{T}.

1 Introduction

The ability to identify jets containing B-hadrons is important for the high-p_{T} physics program of the ATLAS experiment. Two robust b-tagging algorithms taking advantage of either the impact parameter of tracks [1] or the reconstruction of secondary vertices [2] were developped and successfully used for several analyses of the 2010 and 2011 data. Building on this experience, more advanced and performing b-tagging algorithms have been recently commissioned with 2011 data [3]. All b-tagging algorithms rely on the relatively long decay length of B-hadrons that gives rise to large impact parameter children tracks and displaced decay secondary vertices. The more advanced taggers use multivariate techniques to further increase the discrimination between b-jets and light jets. These algorithms, however, are not sensitive to the number of B-hadrons within the jet. In particular they would equally tag gluon jets if they give rise to a close-by B-hadron pair via gluon splitting, as depicted in Fig. 1. We will henceforth call "merged" b-jets or $b\bar{b}$ jets the b-tagged jets containing two B-hadrons. The ability to single out b-tagged jets from gluon splitting has several applications in different lines of analysis: measurement of QCD beauty production, $t\bar{t}$ and single top production, reduction of background in searches with b-quarks in the final state, and the study of substructure in fat jets, where $g \to b\bar{b}$ jets compete with boosted $Z \to b\bar{b}$ and $H \to b\bar{b}$ jets.

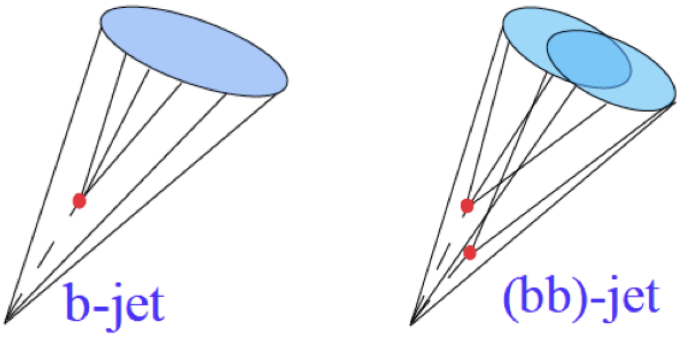

Figure 1: b-tagging algorithms select jets originating both from the fragmentation of a single b quark ("single" b-jets, left image) or from the splitting of a gluon into a pair of close-by $b\bar{b}$ quarks ("merged" b-jets, right image).

There are two possible strategies to attempt the identification of b-jets containing two B-hadrons. One of them relies on the direct reconstruction of the two B-decay secondary vertices [4]. This has the further advantage of allowing the measurement of the angular separation between the B-hadrons, but suffers from the low efficiency of a double b-tag requirement plus additional reconstruction inefficiencies at small angular separation between the two Bs. In this paper we develop an alternative method that does not rely on explicitly finding vertices, but exploits the substructure differences between single and merged b-jets, combining them in a multivariate analysis.

The note is organized as follows. In section 2 we review the physics cases where this tool finds natural applications. Sections 3 and 4 review the Data and Monte Carlo samples and their reconstruction and section 5, the criteria for selection of events, jets and tracks. The kinematic variables that differentiate between single and merged b-jets are discussed in section 6 and the validation of their MC distributions with QCD data in section 7. The construction of the multi-variate discriminator is presented in section 8 and the discussion of the systematic uncertainties in section 9. Section 10 investigates the performance of the tagger with other Monte Carlo generators and, finally, section 11 summarizes the results and discusses future improvements and new ideas.

2 Motivation

Within the Standard Model (SM) a range of production channels exists for heavy-quark jets, e.g. pure QCD production or production in association with heavy bosons (W, Z, H). Furthermore, b-quarks enter in many collider searches, notably because they are produced in the decays of various SM particles, e.g. top quarks and the Higgs boson (if light), and of numerous particles appearing in proposed extensions of the SM. The ability to distinguish genuine b-quark jets from those produced via gluon splitting is thus of wide application. Here we briefly discuss three cases, the measurement of QCD b-quark production, the reduction of background in SM and BSM analyses, and studies of jet substructure.

The measurement of the inclusive b-jet spectrum

Studies of QCD bottom production are important because of the correspondence between parton level production and the observed hadron level, and their potential to provide information on the b-quark parton distribution function, the only component of the proton structure thought to be generated entirely perturbatively from the DGLAP evolution of the other flavours. The theoretical calculation of the inclusive b-jet spectrum presents rather important uncertainties ($\sim 50\%$), considerably larger than those for the light jet inclusive spectrum ($\sim 10-20\%$) [5]. These arise from poor convergence of the perturbative series, as some NLO terms are considerably larger than LO ones. Fig. 2 shows examples of diagrams at LO ("flavor creation"), where the b and $\bar{b}$ quarks fly in opposite directions, and at NLO ("gluon splitting", GSP), where a gluon splits collinearly into a close-by $b\bar{b}$-pair that the clustering algorithm classifies within the same jet. A jet containing both b and $\bar{b}$ is considered to be a plain b-jet in standard definitions.

The largest uncertainties are associated to the GSP contribution, that receives a strong enhancement from collinear logarithms. This channel however does not even correspond to one's physical idea of a b-jet, i.e. the one induced by a hard b-quark, and it seems somehow unnatural to include it at all as part of one's b-jet spectrum. Reference [6] proposes a solution for improving the accuracy of the prediction of the b-jet spectrum through a new jet reconstruction scheme which maintains the correspondence between partonic flavour and jet flavour: the flavour-k_t algorithm [7]. Specifically, a jet containing equal number of b-quarks and b-antiquarks is labeled in an IR-safe way as a light jet, so that jets that contain a b and $\bar{b}$ from the gluon splitting channel do not contribute to the b-jet spectrum. The measurement proposed in referene [6] thus requires that jets identified as gluon splitting are removed from the b-jet spectrum.

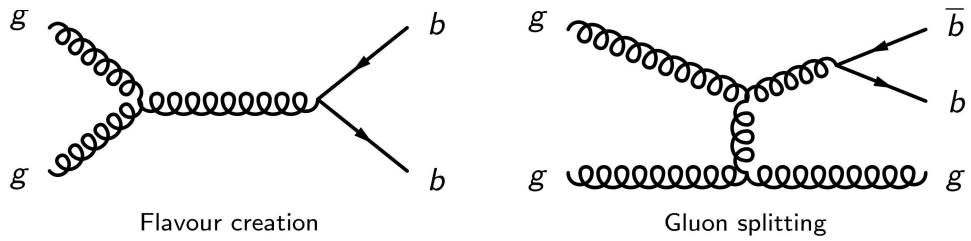

Figure 2: Two examples of Feynman diagrams contributing up to next-to-leading order to QCD bottom production: Flavor creation (LO) and Gluon splitting (NLO).

Rejection of background in Standard Model analyses and beyond-SM searches

Efficient tagging of merged b-jets from gluon splitting can provide an important handle to understand, estimate and/or reject b-tagged backgrounds to SM and new physics searches at the LHC.

SM physics analyses that rely on the presence of b-quarks in the final state, such as top quark

physics (either in the $t\bar{t}$ or the single top channels) or associated Higgs production ($WH \to \ell\nu b\bar{b}$ and $ZH \to \nu\nu b\bar{b}$), suffer from backgrounds that can be partially removed by a merged b-jet tagger. These are the reducible background from QCD, that can produce b-jets from gluon splitting, and the irreducible background due to W bosons plus b-jets. One of the leading order diagrams in W + b-jets corresponds to $W+g$ followed by $g \to b\bar{b}$ with the b-pair produced at small angles and reconstructed as a merged jet. The relevance of $b\bar{b}$ jets in this channel is supported by NLO calculations of the production of W bosons and two jets with at least one b-quark at the LHC [8]. For jets with $p_{\mathrm{T}} > 25$ GeV, and $|\eta| < 2.5$, they indicate that the cross section for $W(b\bar{b})j$ is almost a factor of two higher than $Wb\bar{b}$, where $(b\bar{b})$ denotes the case in which the two b quarks are merged into the same jet.

New physics searches with b-quarks in the final state also benefit from rejection of QCD and $W+b$ backgrounds which have b-jets arising from gluon splitting. An example is the search for supersymmetry in the framework of generic R-parity conserving models [9]. The superpartners of quarks and gluons are expected to be copiously produced via the strong interaction at the LHC. The partners of the right- and left-handed quarks, $\tilde{\mathrm{q}}_L$ and $\tilde{\mathrm{q}}_R$, can mix to form two mass eigenstates and, since mixing is proportional to the corresponding fermion masses, it becomes most important for the third generation producing sbottom and stop with masses significantly lighter than those of the other squarks. In this model, thus, sbottom and stop production is expected to dominate. As they chain decay to b-quarks and the lightest supersymmetric particle, the signature for this channel is $E_{\mathrm{T}}^{\mathrm{miss}}$ + (single) b-jets. Identification of merged b-jets is thus a potential tool to reduce SM backgrounds.

Jet substructure and boosted objects

At the LHC, many of the particles considered to be heavy at previous accelerators will be frequently produced with a transverse momentum greatly exceeding their rest mass, like the electro-weak gauge bosons $W^{\pm}$ and Z^0, the top quark, the Higgs boson (or bosons) and possibly other new particles in the same mass range. These boosted objects, produced either by recoil against other energetic objects or from decays of even heavier BSM particles, upon decay can give rise to a highly collimated topology too close to be resolved by jet algorithms. For theses cases, sophisticated tools have been developed in the last years [10, 11] to analyse the substructure of the ensuing jet and reveal its heavy-particle origin.

The analysis of $b\bar{b}$ jets from gluon splitting is an ideal testbed for studying jet substructure in data, as it provides a large supply of boosted, merged jets. Furthermore, understanding $g \to b\bar{b}$ jets is important as they are themselves the background to boosted object searches, like $Z \to b\bar{b}$ or $H \to b\bar{b}$. In particular, it has recently been suggested [12] that WH and ZH production can become potential discovery and analysis channels by restricting one's attention to the $\sim 5\%$ of events in which the vector and Higgs bosons have large transverse momentum, $p_{\mathrm{T}H} \gtrsim 200$ GeV. Understanding the much more common QCD events with $b\bar{b}$ jets will be essential before attempting to measure these rare final states.

3 Data and Monte Carlo samples

The method presented in this note relies on Monte Carlo predictions for the signal (single b) or background (merged b) hypotheses. The accuracy of the simulation is validated with data by comparing the distributions of the different variables explored in this analysis.

Samples of jet events from proton-proton collision processes are simulated with PYTHIA8 [13] using a $2 \to 2$ matrix element at leading order in the strong coupling to model the hard subprocess, p_{T}-ordered parton showers to model additional radiation [14], underlying event and multiple parton interactions [15], and fragmentation and hadronisation based on the Lund string model [16]. The ATLAS MC11 tune of the soft model parameters was used [17]. In order to have sufficient statistics over the entire p_{T} spectrum, eight samples were generated with different thresholds of the hard-scattering partonic transverse momentum $\hat{p}_T$. Events from different samples were mixed taking into account their respective

production cross sections.

The GEANT4 [18] software toolkit within the ATLAS simulation framework [19] propagates the generated particles through the ATLAS detector and simulates their interactions with the detector material. The energy deposited by particles in the active detector material is converted into detector signals in the same format as the detector read-out. Finally the Monte Carlo generated events are processed through the trigger simulation package of the experiment, and are reconstructed and analyzed with the same software as for the real data. The simulated data sample used for the analysis gives an accurate description of the pile-up content and detector conditions for the full 2011 data-taking period.

The data samples employed correspond to proton-proton collisions at $\sqrt{s} = 7$ TeV delivered by the LHC and recorded by ATLAS between May and November 2011, with the LHC running with 50 ns bunch spacing, and bunches organized in bunch trains. Only data collected during stable beam periods in which all sub-detectors were fully operational are used. After the application of the data quality selection, the surviving data corresponds to an integrated luminosity of 4.7 fb^{-1}. The LHC performance steadily improved during 2011. In particular the average number of minimum-bias pile-up events, originating from collisions of additional protons in the same bunch as the signal collision, grew from from 3 to 20. This fact will be of importance when discussing the selection of discriminating variables.

For the study of systematic effects and for result comparison, other Monte Carlo samples were utilised. Results were produced with the HERWIG++ generator [20] and with PYTHIA8 using the Perugia tune [21].

4 Reconstruction

Experimental data and simulated events were both reconstructed using the latest version available of the ATLAS software. In this section we brielfy describe the reconstruction of the two key objects used in this analysis, namely jets and tracks.

Jets are reconstructed using the anti-k_t jet algorithm [22] with a distance parameter $R = 0.4$, using calorimeter topological clusters [23] as input. Several quality criteria are applied to eliminate "fake" jets that are caused by noise bursts in the calorimeters and energy depositions belonging to a previous bunch crossing [24].

The jet energies are corrected for inhomogeneities and for the non-compensating nature of the calorimeter by using p_{T}- and η-dependent calibration factors determined from Monte Carlo simulation [25]. This calibration is referred to as the EM+JES scale. Using test beam results, in-situ track and calorimeter measurements, estimations of pile-up energy depositions, and detailed Monte Carlo comparisons, an uncertainty on the absolute jet energy scale was estabished. This uncertainty is smaller than $\pm 10\%$ for $\eta < 2.8$ and $p_{\mathrm{T}} > 20$ GeV. More sophisticated techniques undergoing commissioning, such as local cluster weighting, are expected to considerably improve the jet energy uncertainty and resolution [26].

The tracks of charged particles with a pseudorapidity $|\eta| < 2.5$ are reconstructed in the the Inner Detector. It is composed of a barrel, consisting of 3 Pixel layers, 4 double layers of single-sided silicon strip sensors, and 73 layers of Transition Radiation Tracker straws concentric with the beam, plus a system of disks on each end of the barrel, occupying in total a cylindrical volume around the interaction point of radius of 1.15 m and length of 7.024 m. The Pixel detector's innermost layer is located at a radius of 5 cm from the beam axis, has a position resolution of approximately 10 μm in the $r-\phi$ plane and 115 μm along the beam axis (z). Tracks with $p_{\mathrm{T}}^{\mathrm{track}} > 400$ MeV and consistent with the beamspot are associated in primary vertices via a finding/fitting sequential algorithm. Several primary vertices can be reconstructed per event due to the presence of in-time pile-up. The one with at least five associated tracks, a z position within 100 mm of the ATLAS geometrical center, and the largest $\sum_{\mathrm{trk}} p_{\mathrm{T}}^2$, is selected as the one associated to the hard interaction.

Jets are classified as b-quark candidates by the ATLAS MV1 b-tagging algorithm, based on a neural network that combines the information from three high-performance taggers: IP3D, SV1 and JetFitter [3]. These three tagging algorithms use a likelihood ratio technique in which input variables are compared to smoothed normalized distributions for both the b- and background (light- or in some cases c-jet) hypotheses, obtained from Monte Carlo simulation. The IP3D tagger takes advantage of the signed transverse and longitudinal impact parameter significances. The SV1 tagger reconstructs an inclusive vertex formed by the decay products of the b-hadron and relies on the invariant mass of all tracks associated to the vertex, the ratio of the sum of the energies of the tracks in the vertex to the sum of the energies of all tracks in the jet and the number of two-track vertices. The JetFitter tagger exploits the topology of the primary, b- and c-vertices and combines vertex variables with the flight length significance. The b-tagging performance is determined using a simulated $t\bar{t}$ sample and is calibrated using experimental data with jets containing muons and with a sample of $t\bar{t}$ events [27].

5 Event and object selection

The event sample for this analysis was collected using a logical OR of single jet triggers which select events with at least one jet with transverse energy above a given threshold at the highest trigger level. The ATLAS Trigger system uses three consecutive trigger levels. At the hardware Level 1 and local software Level 2, cluster-based jet triggers are used to select events. The Level 3, the so-called Event Filter, runs the offline anti-k_t jet finding algorithm with $R = 0.4$ on topological clusters over the complete calorimeter. At this stage, the transverse energy thresholds, expressed in GeV, are: 20, 30, 40, 55, 75, 100, 135, 180. These triggers reach an efficiency of 99% for events having the leading jet with an offline energy higher than the corresponding trigger thresholds by a factor ranging between 1.5 and 2. The triggers with the lowest p_{T} thresholds were prescaled by up to five orders of magnitude, and typically the same jet trigger is prescaled ten times more in the later data taking periods compared to the early ones.

The offline event selection requires at least one primary vertex candidate with 5 or more tracks. All jets, with transverse momentum between 40 and 480 GeV, were required to be in a region with full tracking coverage, $|\eta_{jet}| < 2.1$, and they were classified in eight p_{T} bins chosen such as to match the jet trigger 99% efficiency thresholds (in GeV): 40, 60, 80, 110, 150, 200, 270, 360. Only jets tagged as b-jets using the MV1 b-tagging algorithm at the 60% efficiency working point were considered. b-tagged jets with close-by jets ($\Delta R < 0.8$) with p_{T} higher than 7 GeV at electromagnetic scale were not included in the analysis. Unless otherwise indicated, performance plots are shown for one medium p_{T} bin (80 to 110 GeV) and one high p_{T} bin (200 to 270 GeV).

In the case of MC the reconstructed b-tagged jets were further classified into single and merged b-jets based on truth Monte Carlo information. A B-hadron is considered to be associated to a jet if the ΔR distance in $\eta - \phi$ space between the direction of the hadron and the jet axis is smaller than 0.4. Jets were labeled as merged (single) b-jets if they contain two (only one) B-hadron.

It is important to select genuine tracks belonging to jets. Only tracks located within a cone of radius $\Delta R(jet^{\mathrm{reco}}, \mathrm{track}) \leq 0.4$ around the jet axis were considered. Cuts on $p_{\mathrm{T}}^{\mathrm{trk}} > 1.0$ GeV and the χ^2 of the track fit, $\chi^2/ndf < 3$, are applied. In addition, tracks are required to have a total of at least seven precision hits (pixel or micro-strip) in order to guarantee at least 3 z-measurements. Tracks are also required to fulfill cuts on the transverse and longitudinal impact parameters at the perigee to ensure that they arise from the primary vertex. As cutting on impact parameter (IP) significance might be detrimental for b-jets, where large IP values are expected, the relaxed cuts were used, $|IP_{xy}| < 2$ mm, and $|IP_z \sin\theta| < 2$ mm, with θ being the polar angle measured with respect to the beam axis.

6 Kinematic differences between single and double B-hadron jets

The differences between genuine b-quark jets and $b\bar{b}$ jets are expected to arise from the two-subjet (two B-hadrons) substructure of merged jets. They are thus expected, for the same jet p_{T}, to have higher track-multiplicity and be wider than single b-jets. Based on these characteristics simulated QCD samples of b-tagged jets were used to study the following properties, discussed in the next paragraphs, built from jet constituents either at calorimeter level (topological clusters) or tracks associated to the jet:

- Jet multiplicity (number of constituents)
- Jet width, p_{T} weighted
- Jet Mass
- Maximum ΔR between pairs of constituents
- ΔR between 2 k_t subjets within the b-jet
- τ_2: 2-subjettiness
- τ_2/τ_1
- ΔR of leading constituents
- Eccentricity

Figures 3 to 12 show the distributions and correlations of some of these variables in selected bins of b-tagged jet p_{T}.

Figure 3 shows the distribution of jet track multiplicity. This variable is simple to calculate and carries important information of the jet inner structure. It was observed that merged b-jets contain on average around two more tracks than single b-jets at low jet p_{T}, with a larger difference at higher p_{T} values. The jet track multiplicity corresponds to tracks with p_{T} above 1 GeV, satisfying the quality cuts described in section 5. The effect of using a minimum track p_{T} of 0.5 GeV was also examined. This was motivated by the fact that it could lead to an improvement in discrimination if it captured more information about the fragmentation process. On ther other hand, a lower minimum track p_{T} can make the method more sensitive to pile-up with the addition of soft tracks incorrectly associated to the jets. What it was observed is that reducing the p_{T} cut only widens the distributions without increasing the separation between single and merged jets.

Figure 4 shows that, as expected, merged b-jets are wider than single b-jets. The track-jet width was computed as the track p_{T} weighted average of the ΔR distance between the associated tracks and the jet axis:

$$\textit{track-jet width} = \frac{\sum_{i=1}^{N} p_{\mathrm{T}}^{trk_i} \Delta R(trk_i, jet)}{\sum_{i=1}^{N} p_{\mathrm{T}}^{trk_i}} \quad (1)$$

where N is the total number of tracks in the jet.

In Fig. 5 the correlation between the track-jet width and the jet track multiplicity is shown for single and merged b-jets. These two variables alone provide a good discrimination for tagging $b\bar{b}$ jets.

The calorimeter jet width (as in Equation 1 but using topological clusters) gives also good separation. However, this variable is more sensitive to the amount of pile-up in the event than its track-based counterpart. In Fig. 6 the distributions of calorimeter width for single and merged b-jets can be seen for events with low and high Number of Primary Vertices (NPV), in a low p_{T} region where the effect of pile-up is more important. In Fig. 7 the same distributions are shown for the track-jet width. Calorimeter jet width varies with NPV and due to this behavior the track-based version is more suitable as a more robust discriminator. For similar reasons, the jet topological cluster multiplicity and the jet mass were discarded as discriminating variables.

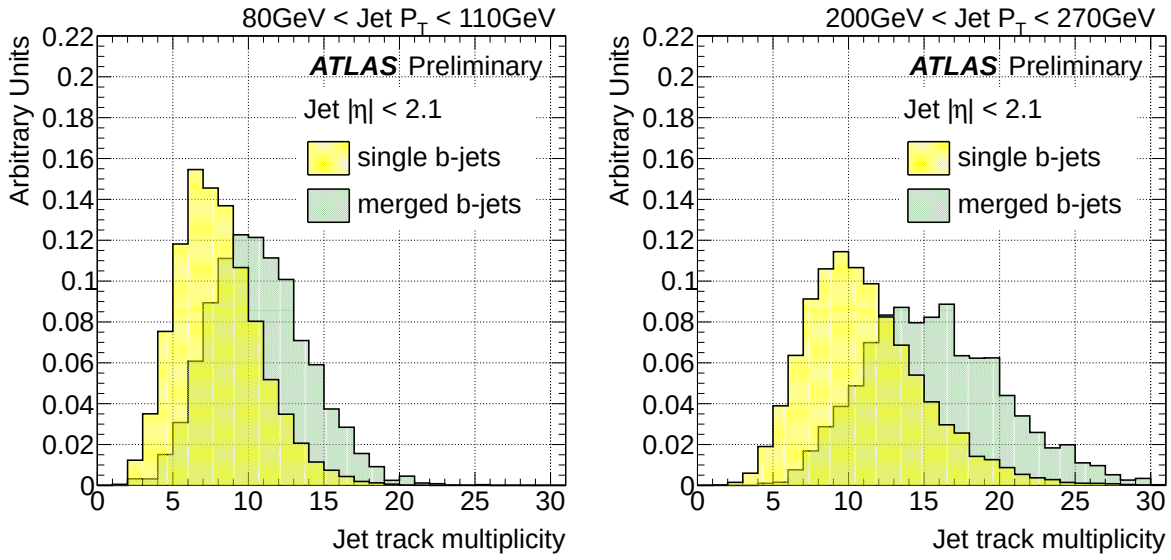

Figure 3: Distribution of the track multiplicity in jets for single and merged *b*-jets between 80 GeV to 110 GeV (left) and 200 GeV to 270 GeV (right).

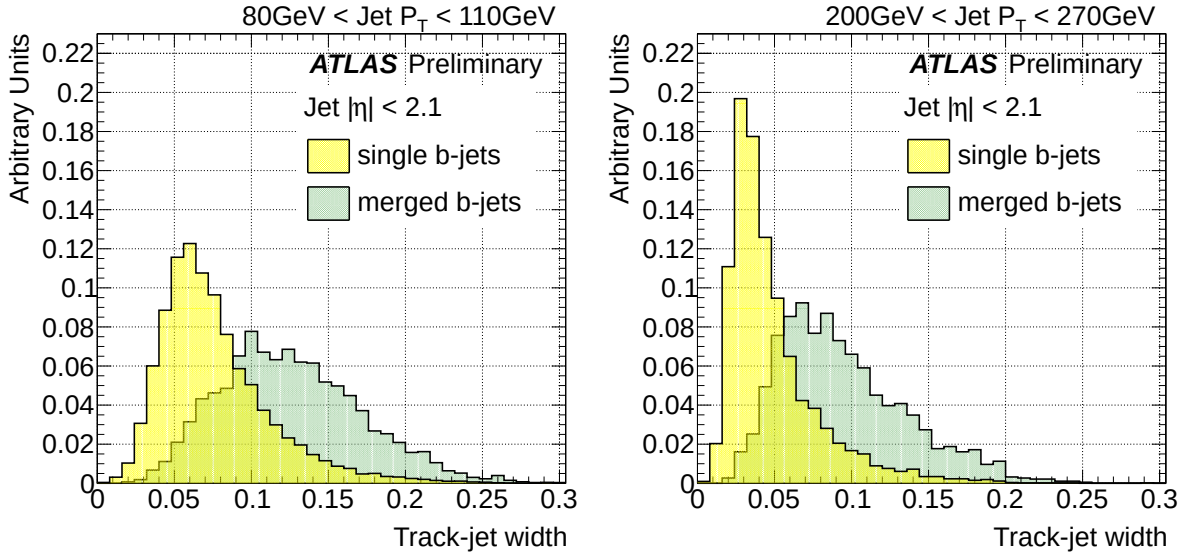

Figure 4: Distribution of track-jet width in jets for single and merged *b*-jets between 80 GeV to 110 GeV (left) and 200 GeV to 270 GeV (right).

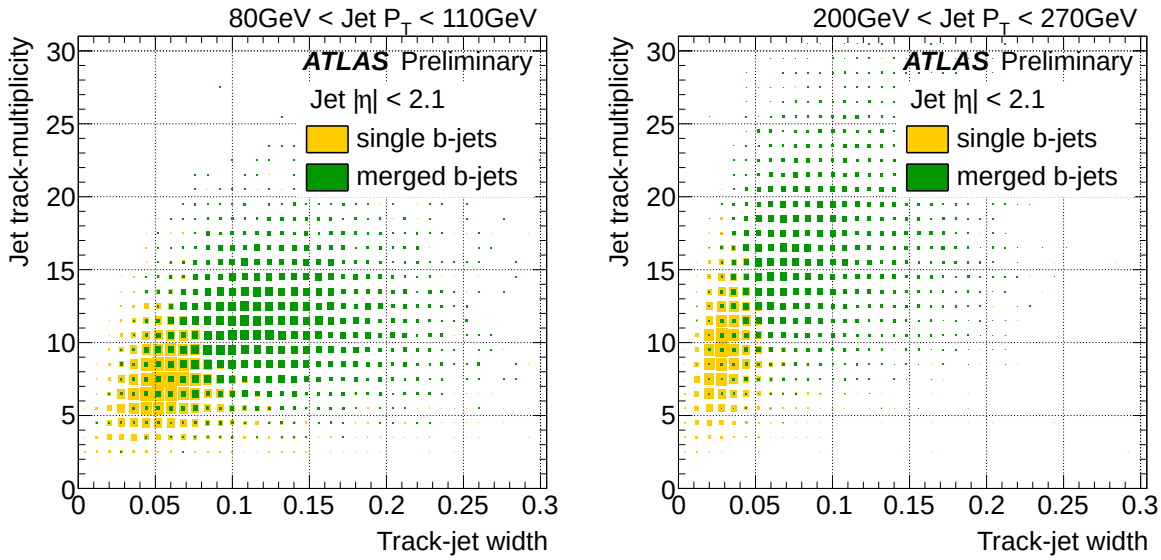

Figure 5: Correlation between jet track multiplicity and track-jet width for single and merged *b*-jets between 80 GeV to 110 GeV (left) and 200 GeV to 270 GeV (right).

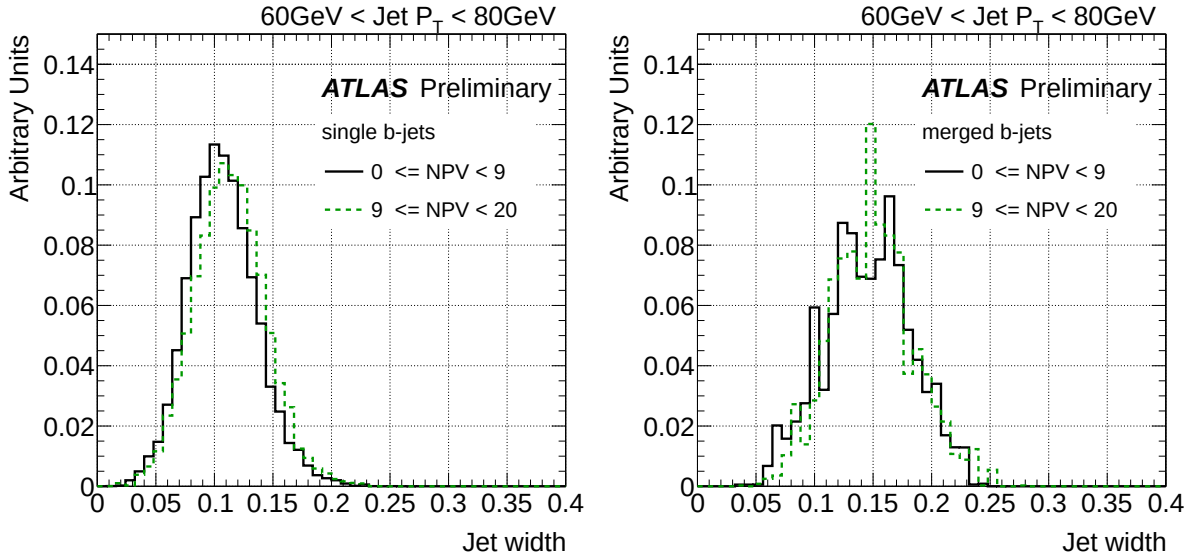

Figure 6: Distribution of calorimeter jet width (using topological clusters) for single (left) and merged (right) *b*-jets in two bins of Number of Primary Vertices for jets between 60 GeV to 80 GeV.

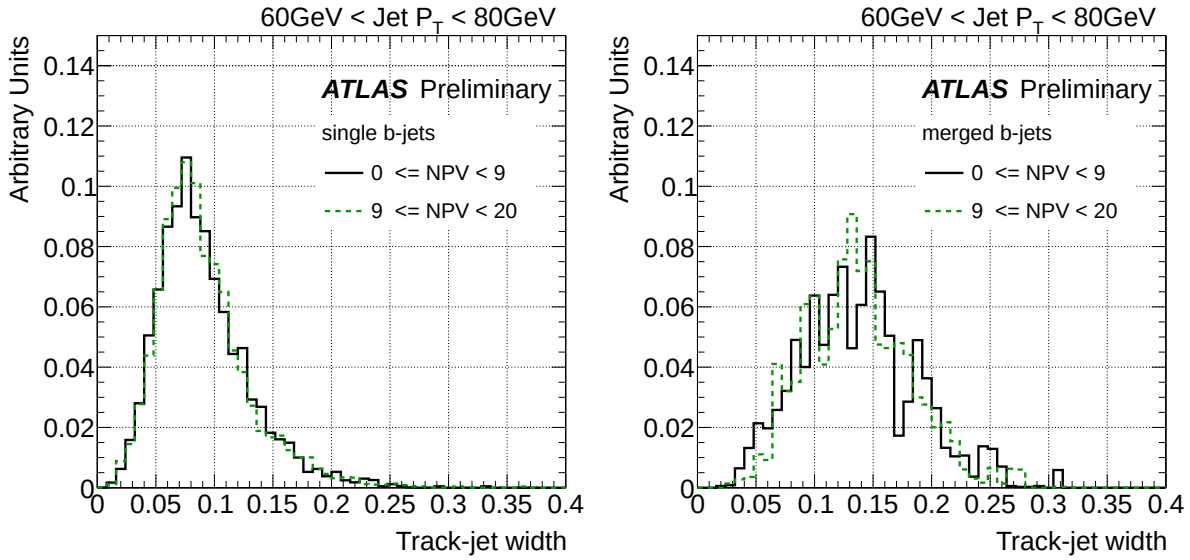

Figure 7: Distribution of track-jet width for single (left) and merged (right) b-jets in two bins of Number of Primary Vertices for jets between 60 GeV to 80 GeV.

Figure 8 shows the distribution of the maximum ΔR between track pairs in the jets ($\mathrm{Max}\{\Delta R(trk,trk)\}$). Merged b-jets show significantly higher values for this variable over a broad range of jet p_{T}. The distinct characteristic of this variable is that the separation between single b-jets and merged does not depend on jet p_{T}. In spite of its good discrimination power, we have looked for alternatives to $\mathrm{Max}\{\Delta R(trk,trk)\}$ as it is not an infrared safe observable and is sensitive to soft tracks originating from pile-up.

The distribution of the ΔR between the axes of the two exclusive k_t subjets in the jet is shown in Fig. 9 for single and merged b-jets. In order to build this variable the k_t algorithm [28] is applied to all the tracks associated to the jet using a large k_t distance parameter to ensure that all of them get clustered. The clustering is stopped once it reaches exactly two jets. We observe that this variable also provides good separation, with the advantage of infrared safeness and insensitivity to pile-up.

N-subjettiness variables, as described in Ref. [29], were originally designed to identify boosted objects, like electroweak bosons and top quarks, decaying into collimated shower of hadrons which a standard jet algorithm would reconstruct as single jets. It is defined as:

$$\tau_N = \frac{1}{\sum_k p_{\mathrm{T}k} R_0} \sum_k p_{\mathrm{T}k} \min\{\Delta R_{S_1,k}, \Delta R_{S_2,k}, ..., \Delta R_{S_N,k}\} \quad (2)$$

where R_0 is the jet radius used in the jet clustering algorithm and the sum runs over the constituents of the jet. To avoid dependence on pile-up we consider the track-based n-subjettiness, where the sum is over the tracks in the b-tagged jet. $\Delta R_{S_j,k}$ is the distance in the rapidity-azimuth plane between the axis of subjet j and constituent track k. This jet shape variable quantifies to what degree a jet can be regarded as composed of N subjets. For instance, a jet with a two pronged structure, with all tracks clustered along two directions, is expected to have a smaller τ_2 value than a jet with tracks uniformly distributed in $\eta - \phi$ space.

Plots of τ_2 are shown in Fig. 10. In spite of its expected 2-prong substructure, merged b-jets have higher values of τ_2 than single b-jets. The explanation of this behavior can be found in Fig. 11, where its correlation with track-jet width ($\sim \tau_1$) is shown for single and merged b-jets. The two variables are highly correlated and for this reason wider jets have a larger τ_2. This suggests to switch from an absolute to a width-normalized τ_2. Fig. 12 thus shows the distributions of τ_2/τ_1. This ratio is often used but,

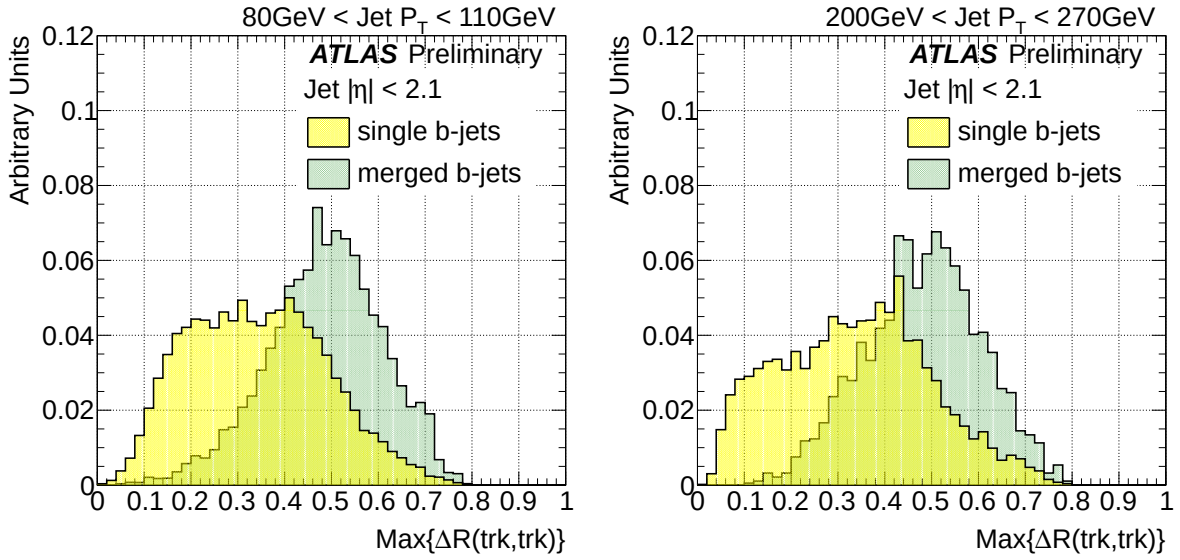

Figure 8: Distribution of the maximum ΔR between pairs of tracks in jets for single and merged b-jets between 80 GeV to 110 GeV (left) and 200 GeV to 270 GeV (right).

although as expected somewhat larger values are obtained for single than for merged b-jets, especially at high p_{T}, we decided not to use this variable as it offers only marginal discrimination. Variables such as the ΔR between the two leading constituents of the jet (those with highest transverse momentum) and the jet eccentricity (the ratio of the principal axes of the jet area) did not show good discrimination either and were not considered for the multivariate study.

7 Validation of discriminating variables with data

In order to study the extent to which the simulation reproduces the distributions observed in data for the different variables explored a set of comparison plots is presented. Fig. 13 shows the distributions of jet track multiplicity, track-jet width and ΔR between the axes of the two k_t subjets, in two different jet p_{T} bins in dijet Monte Carlo and data events collected by ATLAS during 2011. The distributions are normalized to unit area to allow for shape comparisons. There is a good agreement between data and simulation. It should be remarked that the observed agreement is actually not a direct validation of the description in the MC of the relevant variables, but its convolution with the simulated relative fractions of light-, c-, b- and bb-jets in the b-tagged generated jet sample. To some extent, some level of compensation can take place between these two effects.

8 $g \to b\bar{b}$ likelihood tagger

A discriminant between single b-jets and merged b-jets was built by training a simple likelihood estimator in the context of the Toolkit for Multivariate Data Analysis, TMVA [30].

A sub-set of the dijet Monte Carlo sample was used for training. After the event and jet selections were performed, the b-tagged jets with $|\eta| < 2.1$ were classified as signal (single b-jets) or background (merged b). The likelihood training was done in bins of calorimeter jet p_{T}. Signal and background jets were not weighted by the dijet samples cross-sections to allow the contribution of subleading lower p_{T} jets from high p_{T} events, and thus increase the statistics of merged jets in the low p_{T} bins. For the

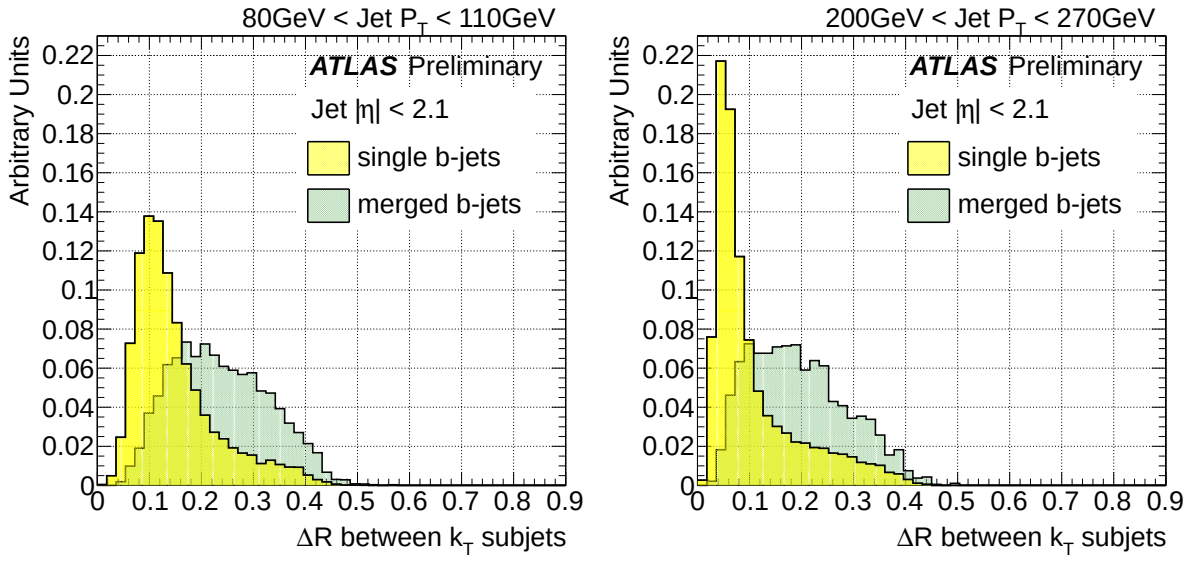

Figure 9: Distribution of the ΔR between the axes of the two k_t subjets in the jet for single and merged b-jets between 80 GeV to 110 GeV (left) and 200 GeV to 270 GeV (right).

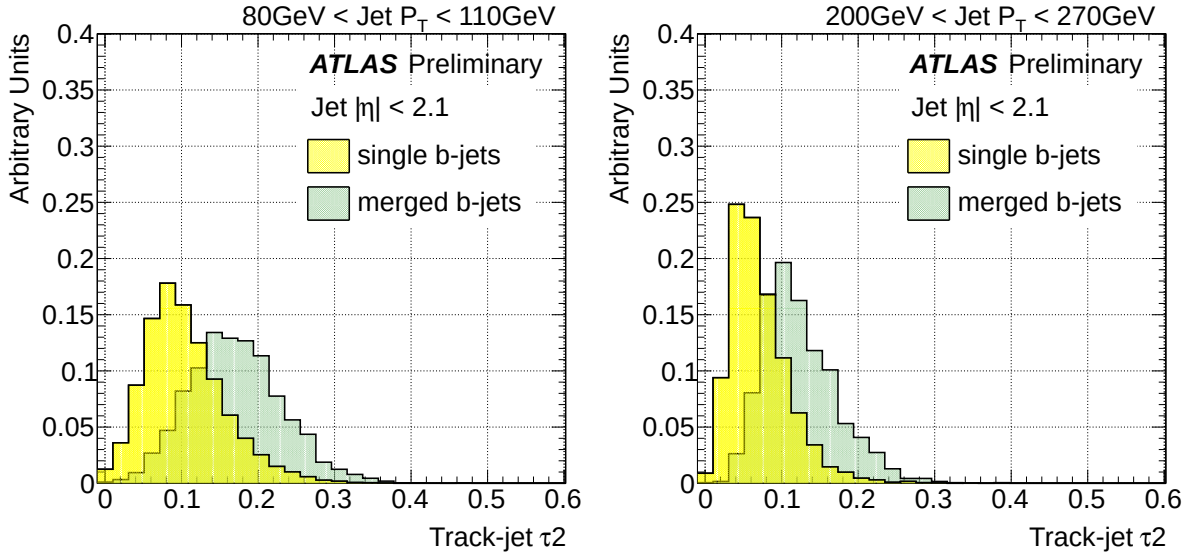

Figure 10: Distribution of τ_2 in jets for single and merged b-jets between 80 GeV to 110 GeV (left) and 200 GeV to 270 GeV (right).

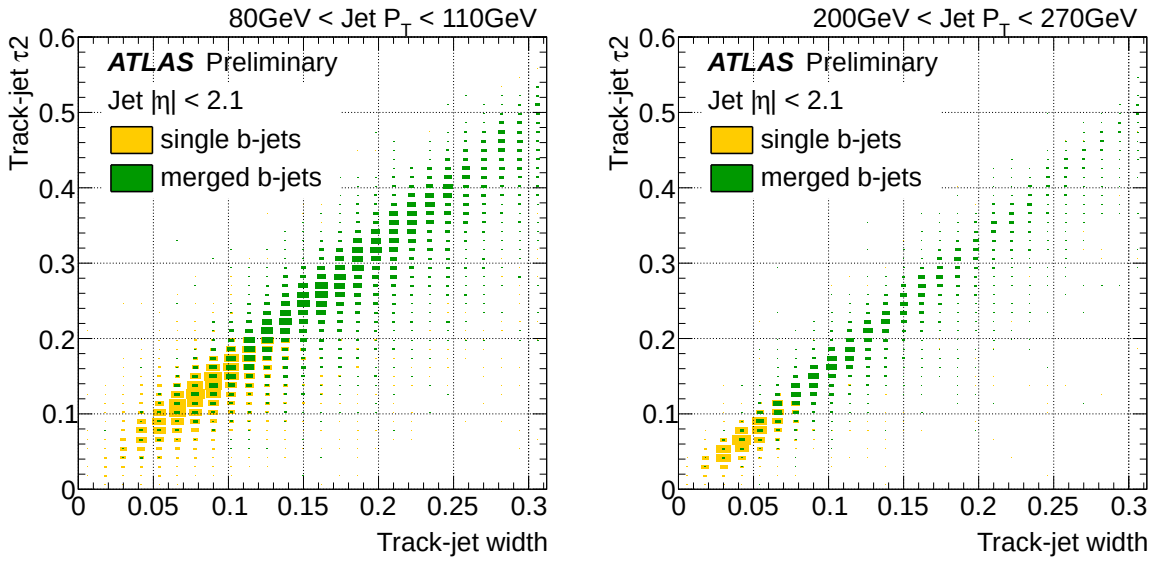

Figure 11: Correlation between τ_2 and track-jet width for single and merged b-jets between 80 GeV to 110 GeV (left) and 200 GeV to 270 GeV (right).

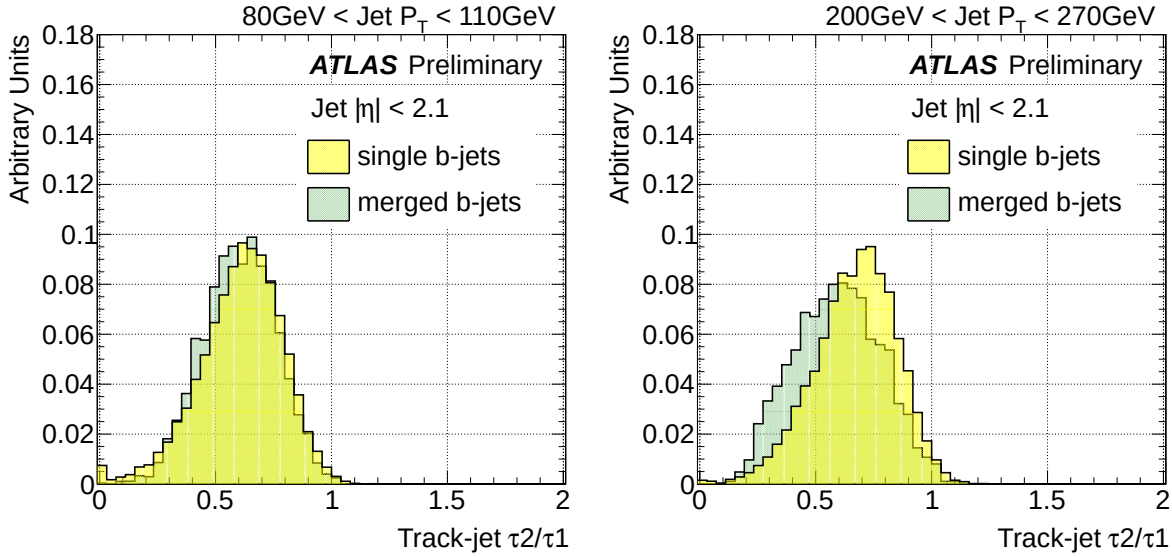

Figure 12: Distribution of τ_2/τ_1 in jets for single and merged b-jets between 80 GeV to 110 GeV (left) and 200 GeV to 270 GeV (right).

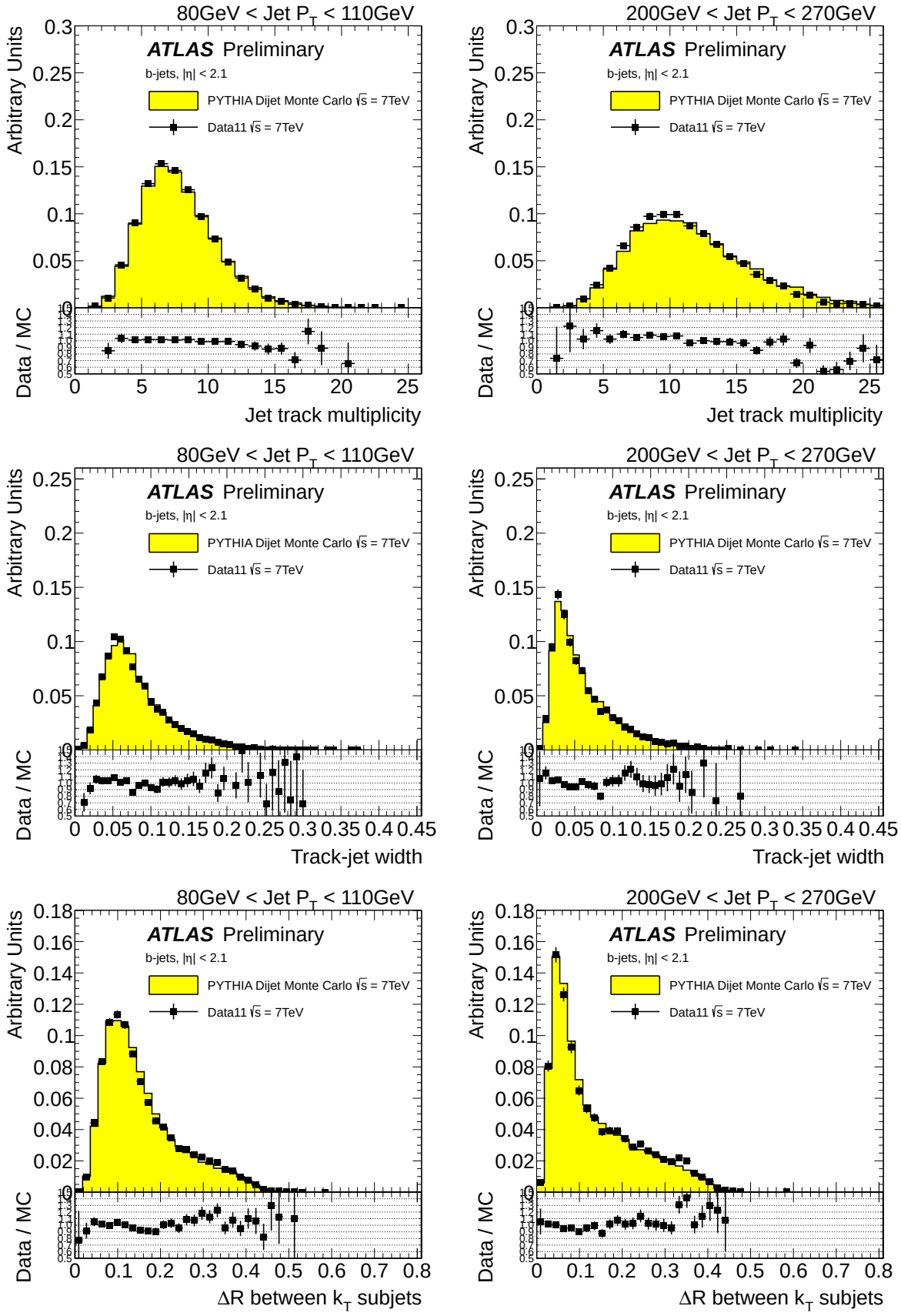

Figure 13: Distribution of three tracking variables in 2 different jet p_{T} bins, for experimental data collected by ATLAS during 2011 (solid black points), and simulated data (filled histograms). The ratio data over simulation is shown at the bottom of each plot.

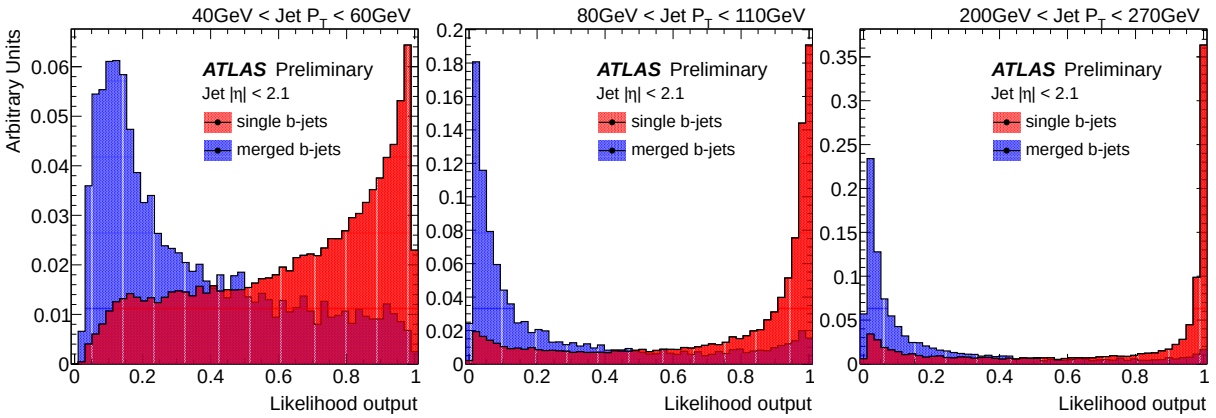

Figure 14: Distribution of the $g \to b\bar{b}$ likelihood output for single and merged b-jets for low, medium and high p_{T} jets.

evaluation of the method the same procedure was followed.

Several combinations of the tracking and jet shape variables studied in the previous section were tested as input variables. We found that the following three offer the best performance:

1. Jet track multiplicity
2. Track-jet width
3. ΔR between the axes of 2 k_t subjets within the jet

A requirement of at least two matching tracks was imposed to all b-tagged jets in order to build the third variable listed. This cut was applied in both training and testing samples.

The distribution of the likelihood output for single and merged b-jets is shown in Fig. 14 for low, medium and high transverse momentum jets.

The performance of the $g \to b\bar{b}$ tagger in the simulation can be displayed in a plot of rejection $(1/\varepsilon_{bkg})$ of merged b-jets as a function of single b-jet efficiency, where ε_{bkg} is the probability that a $b\bar{b}$-jet passes the tagger. This is shown in Fig. 15 for the eight bins of jet p_{T} mentioned in section 5. The performance improves with p_{T}:

- $p_{\mathrm{T}} > 40$ GeV: rejection above 8 at 50% eff.
- $p_{\mathrm{T}} > 60$ GeV: rejection above 10 at 50% eff.
- $p_{\mathrm{T}} > 200$ GeV: rejection above 30 at 50% eff.

The likelihood was trained with jets that had been first tagged by the MV1 algorithm. In order to use the $g \to b\bar{b}$ classifier for jets tagged by another tagger a new training is required.

The rejection of merged jets attained as a function of p_{T} for the 50% and 60% efficiency working points are summarized in Table 1, together with their relative statistical error. These are propagated from the Poisson fluctuations of the number of events in the merged and single $b\bar{b}$ distributions. The error is slightly lower for the 60% efficiency working point because a higher efficiency allows for a greater number of Monte Carlo events to measure the performance.

9 Systematic studies

The development, training and performance determination of the tagger is based on simulated events. Although the agreement between simulation and data explored in section 7 is a necessary validation

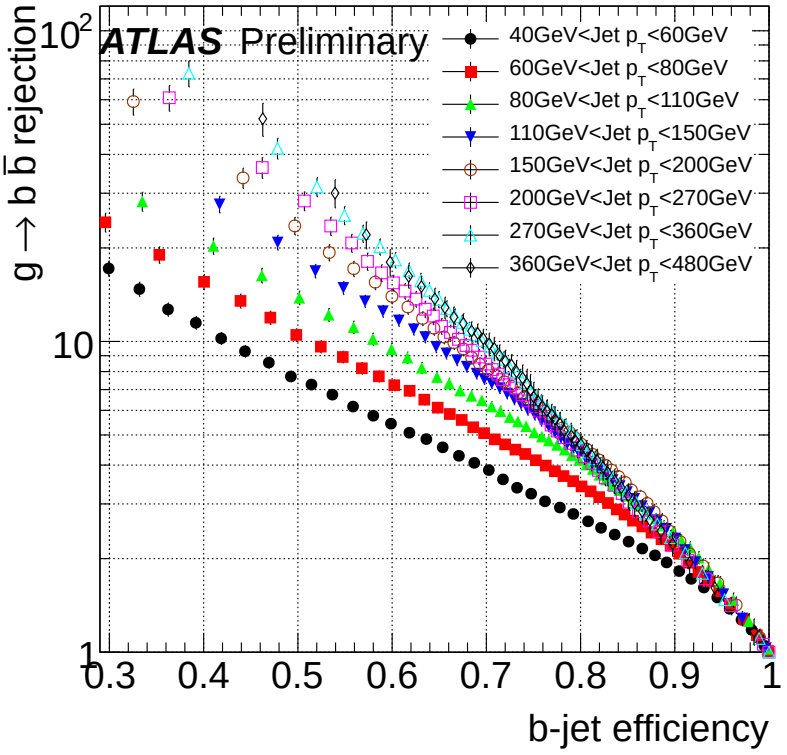

Figure 15: Rejection of $g \to b\bar{b}$ merged b-jets as a function of b-jet efficiency for dijet events in 8 jet p_{T} bins.

Jet p_{T}	single b-jet efficiency 50%		single b-jet efficiency 60%	
(GeV)	Rejection	stat.err.	Rejection	stat.err.
40 - 60	8	4%	5	3%
60 - 80	10	4%	7	4%
80 - 110	14	5%	9	4%
110 - 150	19	5%	12	4%
150 - 200	23	5%	14	5%
200 - 270	30	7%	16	6%
270 - 360	36	7%	19	6%
360 - 480	41	8%	18	8%

Table 1: The merged b-jet rejection for the 50% and 60% efficiency working points in bins of p_{T}.

condition, it is also important to investigate how the tagger performance depends on systematics relevant in the data. In particular we have considered:

- presence of additional interactions (pile-up)
- uncertainty in the b-jet tagging efficiency
- uncertainty in the track reconstruction efficiency
- uncertainty in the track transverse momentum resolution
- uncertainty in the jet transverse momentum resolution

I. Pile-up

The size of this effect was studied by comparing the performance of the likelihood discriminant with b-jets in events with small (1-9) and large (9-20) number of primary vertices. As expected from the use of tracking (as opposed to calorimeter) variables no significant dependence with pile-up is observed within statistics. Of the 16 determinations (2 working points with 8 p_{T} bins each) of performance differences between high and low number of primary vertices events, it is observed that 6 of them are positive and 10 negative, with a global mean of 0.3%. We conclude that the effect is negligible compared to other source of uncertainties.

II. b-tagging efficiency

The performance of heavy-flavor tagging in Monte Carlo events is calibrated to experimental data by means of the scale factors (SFs) measured by the b-tagging group. Such a measurement carries a systematic uncertainty, and in order to estimate its effect a conservative approach is followed: the SFs are varied in all the p_{T} bins simultaneously by one standard deviation both in the up and down directions. The result of this procedure for the distribution of two of the tracking variables used in our discriminant is illustrated in Fig. 16.

The effect of the b-tagging calibration uncertainty on the likelihood peformance is $< 1\%$, negligible with respect to the statistical uncertainty. This was indeed expected. The scale factors depend on the true flavor of the jet and on its p_{T}, but these are basically constant in the performance determination, which is based on single flavor (true b-) jets classified in p_{T}-bins.

III. Track reconstruction efficiency

This uncertainty arises from the limit in the understanding of the material layout of the Inner Detector.

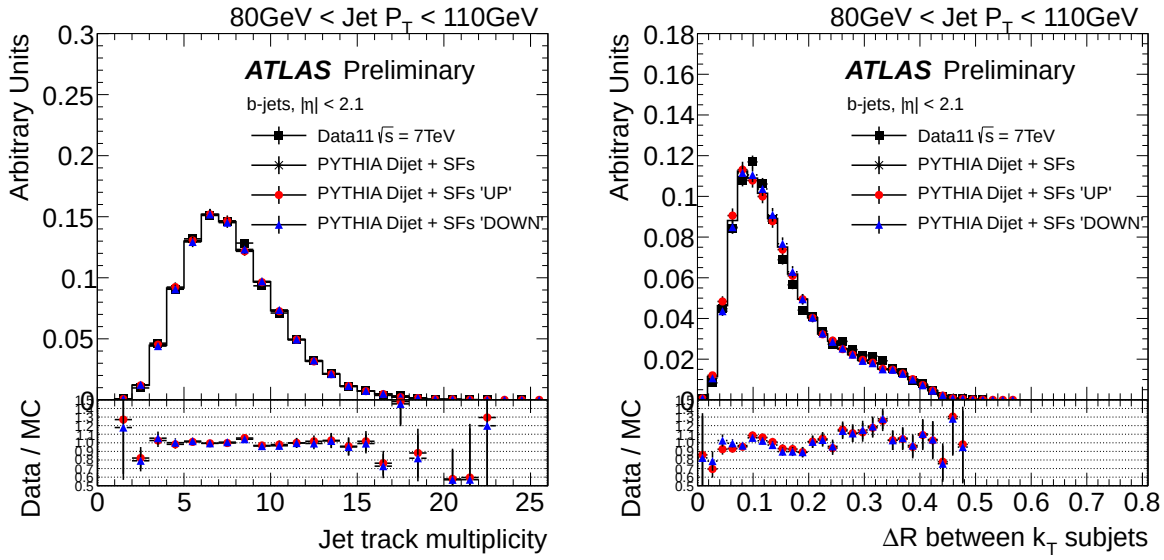

Figure 16: The effect of a variation in the *b*-tagging Scale Factors on the tracking variables distributions. Scale Factors were varied up (down) by 1-sigma to evaluate the systematic uncertainty from this source. The ratio data over MC is shown for MC PYTHIA with SFs varied up (circles) and down (triangles).

To test its impact a fraction of tracks determined from the track efficiency uncertainty was randomly removed following the method in Ref. [31].

The tracking efficiency systematics are given in bins of track η. For tracks with $p_{\mathrm{T}}^{\mathrm{track}} > 500$ MeV the uncertainties are independent of p_{T}: 2% for $|\eta^{\mathrm{track}}| < 1.3$, 3% for $1.3 < |\eta^{\mathrm{track}}| < 1.9$, 4% for $1.9 < |\eta^{\mathrm{track}}| < 2.1$, 4% for $2.1 < |\eta^{\mathrm{track}}| < 2.3$ and 7% for $2.3 < |\eta^{\mathrm{track}}| < 2.5$ [32]. All numbers are relative to the corresponding tracking efficiencies.

The tracking variables were re-calculated and the performance of the nominal likelihood was evaluated in the new sample with worse tracking efficiency. The rejection-efficiency plots, shown in Fig. 17, show a small degradation of the performance which is comparable to the statistical uncertainty. The effect is however systematically present over all 16 p_{T} bin/working points, without a clear p_{T} dependence. We have thus taken the average over p_{T}, and obtained a global systematic uncertainty of 4% both for the 50% and 60% efficiency working points.

IV. Track momentum resolution

The knowledge of the track momentum resolution is limited by the precision both in the material description of the Inner Detector and in the mapping of the magnetic field. Its uncertainty propagates to the kinematic variables used in the $g \to b\bar{b}$ tagger. In order to study this effect, track momenta are over-smeared according to the measured resolution uncertainties before computing the rejection. The actual smearing is done in $1/p_{\mathrm{T}}$, with an upper bound to the resolution uncertainty given by $\sigma(1/p_{\mathrm{T}}) = 0.02/p_{\mathrm{T}}$ [33]. The effect is found to be negligible.

V. Jet transverse momentum resolution

The jet momentum resolution was measured for 2011 data and found to be in agreement with the predictions from the PYTHIA8-based simulation [34]. The precision of this measurement, determined in p_{T} and η bins,is typically 10%. The systematic uncertainty due to the calorimeter jet p_{T} resolution was estimated by over-smearing the jet 4-momentum in the simulated data, without changing jet η or ϕ angles. The performance is found to globally decrease by 6%, without a particular p_{T} dependence.

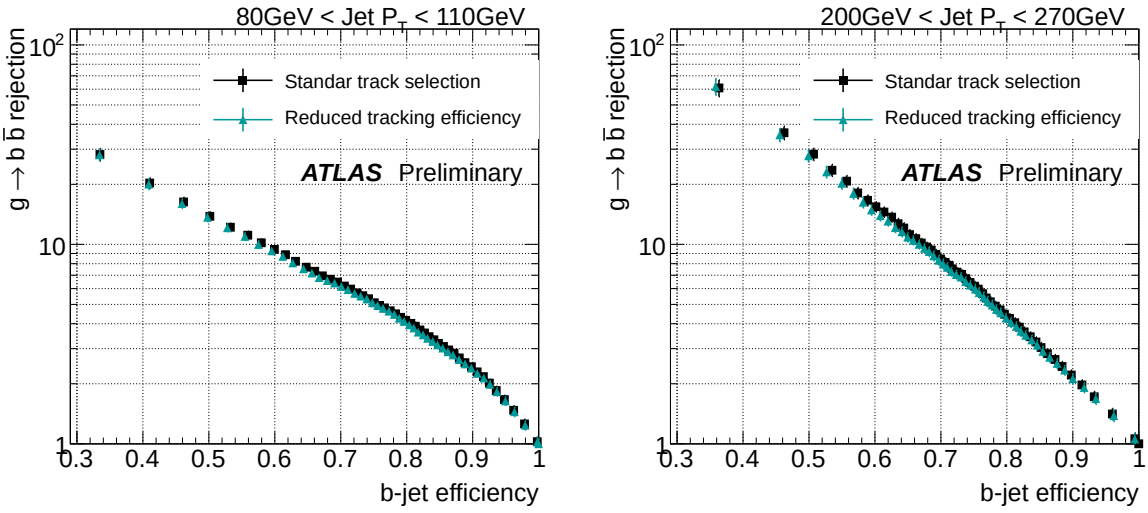

Figure 17: Rejection of $g \to b\bar{b}$ merged b-jets as a function of b-jet efficiency showing shift in likelihood performance caused by a reduction in the tracking efficiency .

The different contributions to the systematic uncertainty on the $g \to b\bar{b}$ rejection are summarized in Table 2.

Systematic source	Uncertainty
pile-up	neglible
b-tagging efficiency	neglible
track reconstruction efficiency	4%
track p_{T} resolution	neglible
jet p_{T} resolution	6%

Table 2: Systematic uncertainities in the merged b-jet rejection (common to both the 50% and the 60% efficiency working points).

10 Other Monte Carlo generators

The development, training and performance determination of the tagger has been done using Monte Carlo events generated with the PYTHIA8 event simulator, interfaced to the GEANT4 based simulation of the ATLAS detector. An immediate question is what the performance would be if studied with a different simulation. In this section we investigate this question for the Perugia tune of PYTHIA8 and the HERWIG++ event generators.

Fig. 18 shows a comparison of the likelihood rejection, at the 50% efficiency working point, between nominal PYTHIA and the alternative simulations as a function of the jet p_{T} . The larger errors are due to the reduced statistics available, which are even lower for the Perugia case than for HERWIG.

The performance in HERWIG shows a systematic trend, with agreement at low p_{T} and increasingly poor performances compared to PYTHIA as p_{T} grows. For the Perugia tune, on the other hand, there is no definite behavior, with the performance fluctuating above or below the nominal simulation for different p_{T} bins consistently with the statistical uncertainties.

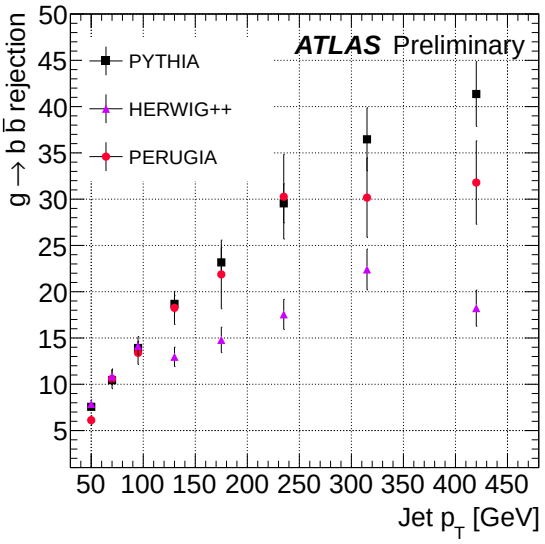

Figure 18: Rejection of $g \to b\bar{b}$ merged b-jets as a function of jet p_{T} for different Monte Carlo generators, at the 50% efficiency working point.

The reason for the systematic difference observed between the performances of PYTHIA and HERWIG can be traced to the extent with which jets are accurately modelled. Fig. 19 compares the measured jet track multiplicity distributions in b-tagged jets and the prediction from both simulations, for low and high p_{T} jets. It is observed that indeed HERWIG++ does not correctly reproduce the data, particularly at high p_{T}. The level of agreement is found to be better for track-jet width and the ΔR between the axes of the two k_t subjets in the jet, the two other variables used for discrimination.

11 Conclusions

A multivariate discriminant to identify isolated b-tagged jets containing two B-hadrons is presented. These jets are expected to arise when a gluon splits into a close-by $b\bar{b}$-pair. The method exploits the kinematic differences between "merged" $b\bar{b}$-jets and "single" b-jets, combining track-based jet shape and jet substructure variables in a likelihood classifier.

The tagger training and performance results are based on simulated events. Several variables were investigated and those showing the best discrimination power were selected for the multivariate analysis. The Monte Carlo distributions of the explored variables were validated using experimental data corresponding to an integrated luminosity of 4.7 fb^{-1} recorded by the ATLAS experiment during 2011. The agreement between data and simulation is excellent.

The peformance of the tagger in Monte Carlo events was studied in bins of the calorimeter jet p_{T}, achieving a rejection of merged jets of over 95% (90%) for a 50% single b-jet efficiency for jets with $p_{\mathrm{T}} > 150$ GeV ($p_{\mathrm{T}} > 60$ GeV).

This tool provides a handle to investigate QCD $b\bar{b}$ production and to reduce backgrounds in physics channels involving b-quarks in the final state. Future improvements comprise the study of further discrimant variables, the extension to non-isolated jets using the concept of ghost-particle matching and active area of a jet [35] for track-to-jet association and labeling, the calibration of the tagger with data, and its application to measure the fraction of gluon-splitting jets in QCD b-jet production.

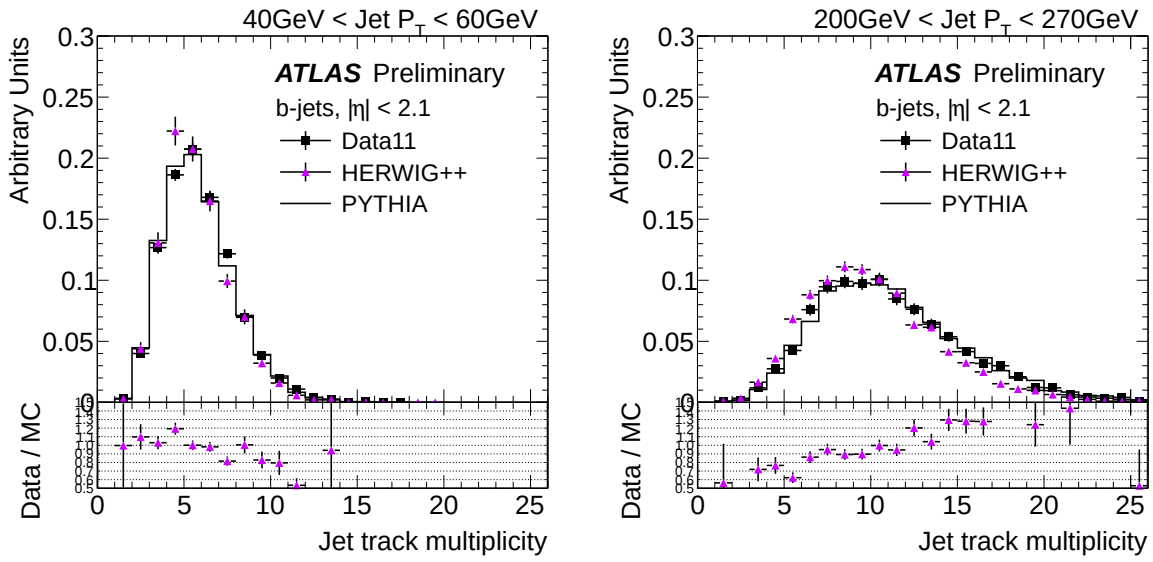

Figure 19: Distribution of the jet track multiplicity in 2 different jet p_{T} bins, for experimental data collected during 2011 (solid black points) and HERWIG++ events (solid violet triangules). The ratio data over HERWIG++ simulation is shown at the bottom of the plot. PYTHIA distribution is also shown for reference.

References

[1] ATLAS Collaboration, *Performance of Impact Parameter-Based b-tagging Algorithms with the ATLAS Detector using Proton-Proton Collisions at $\sqrt{s}$ = 7 TeV*, ATLAS-CONF-2010-091 (2010) .

[2] ATLAS Collaboration, *Performance of the ATLAS Secondary Vertex b-tagging Algorithm in 7 TeV Collision Data*, ATLAS-CONF-2010-042 (2010) .

[3] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102 (2011) .

[4] CDF Collaboration, *Measurements of Bottom Anti-Bottom Azimuthal Production Correlations in Proton-Antiproton Collisions at s**(1/2)= 1.8 TeV*, Phys.Rev.D **71** (2005) 38, arXiv:0412006 [hep-ex].

[5] S. Frixione and M. Mangano, *Heavy quark jets in hadronic collisions*, Nucl.Phys. **B483** (1997) 321–338, arXiv:9605270 [hep-ph].

[6] A. Banfi, G. Salam, and G. Zanderighi, *Accurate QCD predictions for heavy-quark jets at the Tevatron and LHC*, JHEP **0707** (2007) 026, arXiv:0704.2999 [hep-ph].

[7] A. Banfi, G. Salam, and G. Zanderighi, *Infrared safe definition of jet flavour*, Eur.Phys.J.C **47** (2006) 022, arXiv:0601139 [hep-ph].

[8] J. M. Campbell, R. Ellis, F. Maltoni, and S. Willenbrock, *Production of a W boson and two jets with one b^- quark tag*, Phys.Rev. **D75** (2007) 054015, arXiv:0611348 [hep-ph].

[9] ATLAS Collaboration, *Search for supersymmetry in pp collisions at $\sqrt{s} = 7$TeV in final states with missing transverse momentum, b-jets and no leptons with the ATLAS detector*, ATLAS-CONF-2011-098 (2011) .

[10] A. Abdesselam et al., *Boosted objects: a probe of beyond the standard model physics*, The European Physical Journal C - Particles and Fields **71** (2011) 1–19.

[11] A. Altheimer et al., *Jet Substructure at the Tevatron and LHC: New results, new tools, new benchmarks*, arXiv:1201.0008 [hep-ph].

[12] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, Phys.Rev.Lett. **100** (2008) 242001, arXiv:0802.2470 [hep-ph].

[13] T. Sjostrand, S. Mrenna, and P. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:07103820 [hep-ph].

[14] R. Corke and T. Sjöstrand, *Improved parton showers at large transverse momenta*, European Physical Journal C **69** (2010) 1, arXiv:1003.2384 [hep-ph].

[15] T. Sjöstrand and P. Z. Skands, *Transverse-momentum-ordered showers and interleaved multiple interactions*, European Physical Journal C **39** (2005) 129, 0408302 [hep-ph].

[16] B. Andersson et al., *Parton fragmentation and string dynamics*, Phys. Rep. **97** (1983) 31.

[17] *ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11*, Tech. Rep. ATL-PHYS-PUB-2011-009, CERN, Geneva, Jul, 2011.

[18] S. Agostinelli et al., *Geant4 A simulation toolkit*, Nucl. Inst. Meth. Section A **506** (2003) no. 3, 250 – 303.

[19] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur.Phys.J.C **70** (2010) 051, arXiv:1005.4568.

[20] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **0101** (2001) 010, arXiv:hep-ph/0011363 [hep-ph].

[21] P. Z. Skands, *The Perugia Tunes*, arXiv:0905.3418 [hep-ph].

[22] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[23] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, .

[24] ATLAS Collaboration, *Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data*, ATLAS-CONF-2012-020 (2012) .

[25] ATLAS Collaboration, *Jet energy scale and its systematic uncertainty for jets produced in proton-proton collisions at $\sqrt{s}$ = 7 TeV and measured with the ATLAS detector*, ATLAS-CONF-2010-056 (2010) .

[26] ATLAS Collaboration, G. Aad et al., *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, arXiv:0901.0512 [hep-ex].

[27] ATLAS Collaboration, *Calibrating the b-Tag Efficiency and Mistag Rate in 35 pb^{-1} of Data with the ATLAS Detector*, ATLAS-CONF-2011-089 (2011) .

[28] S. Catani, Y. Dokshitzer, H. Seymour, and B. Webber, *Longitudinally invariant K(t) clustering algorithms for hadron hadron collisions*, Nucl. Phys. **B406** (1993) 187.

[29] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, JHEP **1103:015** (2011) 026.

[30] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.

[31] R. Alon et al., *Backup Note for Measurement of Jet Mass and Substructure in QCD with the ATLAS Experiment*, ATL-COM-PHYS-2011-401 (2011) .

[32] ATLAS Collaboration, G. Aad et al., *Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC*, New J.Phys. **13** (2011) 053033, arXiv:1012.5104 [hep-ex].

[33] ATLAS Collaboration, *Estimating Track Momentum Resolution in Minimum Bias Events using Simulation and K_s in $\sqrt{s}$ = 900 GeV collision data*, ATLAS-CONF-2010-009 (2010) .

[34] G. Romeo, A. Schwartzman, R. Piegaia, T. Carli, and R. Teuscher, *Jet Energy Resolution from In-situ Techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy $\sqrt{s}$ = 7 TeV*, ATL-COM-PHYS-2011-240 (2011) .

[35] G. S. M. Cacciari and G. Soyez, *The Catchment Area of Jets*, JHEP **0804** (2008) 42.