



ATLAS PUB Note
ATL-PHYS-PUB-2022-035
20th July 2022



Reweighting heavy-flavor production fractions to reduce flavor modelling uncertainties for ATLAS

The ATLAS Collaboration

The ability to identify jets originating from b - and c -hadrons (flavor tagging) is one of the key experimental techniques that enables a wide range of searches and measurements by the ATLAS experiment. Generally, the tagging efficiencies in simulated Monte Carlo samples differ from the data and need to be corrected with dedicated calibration procedures. However, the tagging efficiencies also vary between the Monte Carlo shower simulations such as PYTHIA, SHERPA, and HERWIG. These differences have a direct impact on the modeling uncertainties associated with flavor tagging so it is important to reduce them if possible. In this note we introduce a method of reweighting the heavy-flavor production fractions to a common world average, which eliminates one of the largest causes of differences in tagging efficiencies between the Monte Carlo shower simulations. The heavy-flavor production fractions are the rates at which b - and c -quarks hadronize into different species and the flavor tagging efficiencies in ATLAS have been found to depend on the hadron species inside a jet. The note presents heavy-flavor production fractions in the ATLAS Monte Carlo samples, the reweighting procedure and the impact of the reweighting on flavor tagging efficiencies. Moreover, the experimental uncertainties in the heavy-flavor production fractions can also be estimated with the same reweighting procedure which gives rise to a common way of estimating these systematic uncertainties in ATLAS.

1 Introduction

The ability to identify jets originating from b - and c -hadrons (b - and c -tagging) is one of the key experimental techniques that enables a wide range of searches and measurements in collider experiments. Jets are classified as b - or c -jets based on the presence of weakly decaying b or c -hadrons within the jet. The tagging algorithms exploit several distinct properties of b - and c -hadrons that offer good discriminating power against light jets (jets originating from u , d , or s quarks or gluons). These properties are the long lifetime, larger masses, hard fragmentation function and kinematic properties of their decay products. A typical signature of b - and c -hadrons are secondary vertices (SVs) where the charged decay products of a b - or c -hadron can be traced back to a single point away from the primary vertex due to the long lifetime of b - and c -hadrons. In case of b -hadron decays there can also be tertiary vertices due to the b -hadron decays into c -hadrons. Furthermore, an important feature of the tagging algorithms is that they offer discriminating power between b - and c -jets.

Tagging efficiencies in Monte Carlo (MC) simulation samples can differ from the data and need to be corrected with dedicated calibration procedures [1–3]. However, tagging efficiencies also vary between the MC shower simulations such as PYTHIA [4], SHERPA [5] and HERWIG [6, 7]. These differences have a direct impact on the modeling uncertainties associated with flavor tagging so it is important to reduce them if possible. The main causes of differences in MC samples are: the longitudinal momentum distribution of hadrons inside a reconstructed jet (fragmentation functions), the rates at which b - and c -quarks hadronize into different species of weakly decaying heavy-flavor hadrons (production fractions), and the branching fractions of heavy-flavor hadron decays. Conventionally these differences are accounted for with the ‘MC-MC’ scale factors [8], which are parameterized as a function of jet flavor (c , b , or light) and jet transverse momentum. However, in some cases the MC-MC scale factors are relatively large (e.g. up to 1.5) so fixing the underlying differences between the MC generators is preferential if possible.

The heavy-flavor production fractions play a particularly important role in MC modelling because the properties of the different b - and c -hadrons such as lifetime and the number of charged particles in the decay vary significantly between the species. The majority of the heavy-flavor production is dominated by the mesons: B^\pm , B^0 , B_s^0 for b -hadrons and D^\pm , D^0 , D_s^\pm for c -hadrons. The fraction of b - and c -quarks hadronizing into baryons is about 10%, dominated by the Λ_c^\pm and Λ_b^0 production [9]. Production fractions are determined by non-perturbative QCD processes and can not be calculated from first principles due to the needed prohibitive computational complexity and the general lack of understanding of the exact physics processes governing them. Instead, each MC shower software is tuned to match the data from different collider experiments. However, differences in heavy-flavor production fractions may arise between the MC shower generators because they are either tuned to different data or the tuning procedure itself differs. This note addresses the differences in heavy-flavor production fractions by introducing a new analysis tool, named as the `HFProductionFractionTool`, that reweights the production fractions in MC to the common world-average values from auxiliary measurements of these observables. Re-simulating the entire MC sample with the purpose of changing the heavy-flavor production fractions would imply a large computational cost [10] which can be avoided by rescaling the event weights with this tool.

The datasets from which the world-average HF production fractions are derived are described in Section 3. These values have no kinematic dependence and assume that the production fractions are process independent. Recent measurements from ALICE and LHCb experiments at the LHC indicate that both b - and c -hadron production fractions in hadron collisions are energy dependent [11–13]. These effects appear to be significant only at low transverse momenta (< 20 GeV). The reweighting procedure introduced in this note harmonizes the production fractions in simulated samples, reducing the MC dependence of the

predicted flavor tagging performance. However, the relevant quantities (e.g. flavor tagging efficiency) still need to be calibrated to match the data.

The note is organized as follows. The production fractions in ATLAS MC samples are given in Section 2. The comparison of the production fractions in ATLAS MC to the common world-average values is presented in Section 3. Section 4 explains the reweighting procedure and highlights the design of the new analysis tool. Finally, Section 5 shows the validation of the reweighting procedure by showing its impact on the b - and c -jet tagging efficiencies in the simulated MC samples.

2 Simulated MC Samples

Several MC samples are used to evaluate the differences in the tagging efficiencies between the MC shower software and to evaluate the impact of the production-fraction reweighting procedure. Three different MC shower generator software are tested: PYTHIA [4], SHERPA [5], and HERWIG [6, 7]. In all cases, $t\bar{t}$ events with one or more W bosons decaying leptonically are used. A detailed description of the MC samples is given below and summarized in Table 1.

The validation in Section 5 is performed with POWHEG+PYTHIA 8, POWHEG+HERWIG 7.13, SHERPA 2.2.1, and SHERPA 2.2.10 samples to highlight the most important differences and conclusions. Additional MC samples for which the production fractions were also calculated are listed in this section. In the future, the ATLAS experiment will use MC samples generated with SHERPA 2.2.11 and SHERPA 2.2.12. The corresponding heavy-flavor modeling is expected to be very similar to that in Sherpa 2.2.10.

Table 1: The generators used for the simulation of the $t\bar{t}$ samples. The acronyms ME, PS and UE stand for matrix element, parton shower and underlying event, respectively.

ME generator	ME PDF	PS generator	PS PDF	UE model tune	heavy-flavor decay
POWHEG BOX v2	NNPDF3.0NLO	PYTHIA 8.230	NNPDF2.3LO	A14	EVTGEN 1.6.0
MADGRAPH5_AMC@NLO 2.6.0	NNPDF3.0NLO	PYTHIA 8.230	NNPDF2.3LO	A14	EVTGEN 1.6.0
POWHEG BOX v2	NNPDF3.0NLO	HERWIG 7.0.4	MMHT2014LO	H7UE	EVTGEN 1.6.0
POWHEG BOX v2	NNPDF3.0NLO	HERWIG 7.1.3	MMHT2014LO	HERWIG 7.1 default	EVTGEN 1.6.0
POWHEG BOX v2	NNPDF3.0NLO	HERWIG 7.2.1	MMHT2014LO	HERWIG 7.1 default	EVTGEN 1.6.0
SHERPA 2.2.1	NNPDF3.0NNLO	SHERPA 2.2.1	NNPDF3.0NNLO	SHERPA 2.2.1 default	SHERPA 2.2.1
SHERPA 2.2.10	NNPDF3.0NNLO	SHERPA 2.2.10	NNPDF3.0NNLO	SHERPA 2.2.10 default	SHERPA 2.2.10

POWHEG+PYTHIA 8: modelled using the POWHEG BOX v2 [14–17] generator at next-to-leading order (NLO) with the NNPDF3.0NLO [18] PDF set and the h_{damp} parameter¹ set to $1.5 m_{\text{top}}$ [19]. The events were interfaced to PYTHIA 8.230 [4] to model the parton shower, hadronisation, and underlying event, with parameters set according to the A14 tune [20] and using the NNPDF2.3LO set of PDFs [21]. The decays of bottom and charm hadrons were performed by EVTGEN 1.6.0 [22].

MADGRAPH5_AMC@NLO+PYTHIA 8: modelled with MADGRAPH5_AMC@NLO 2.6.0 [23] and interfaced with PYTHIA 8.230. The MADGRAPH5_AMC@NLO calculation used the NNPDF3.0NLO set of PDFs and PYTHIA used the A14 set of tuned parameters [20] and the NNPDF2.3LO set of PDFs. The decays of bottom and charm hadrons were simulated using the EVTGEN 1.6.0 program.

¹ The h_{damp} parameter is a resummation damping factor and one of the parameters that controls the matching of POWHEG matrix elements to the parton shower and thus effectively regulates the high- p_T radiation against which the $t\bar{t}$ system recoils.

POWHEG+HERWIG: three POWHEG+HERWIG configurations with different HERWIG versions are studied. In all cases, hard scatter matrix elements modelled with POWHEG BOX v2 generator and interfaced either with HERWIG 7.0.4, HERWIG 7.1.3 or HERWIG 7.2.1. The HERWIG 7.0.4 configuration uses the H7UE set of tunable parameters [7] while HERWIG 7.1.3 and HERWIG 7.2.1 use the HERWIG 7.1 default set of parameters. The PDF used in the hard scatter modelling is NNPDF3.0NLO and MMHT2014LO [24] in the shower simulation. POWHEG BOX provided matrix elements at NLO in the strong coupling constant α_s and an h_{damp} parameter value of $1.5 m_{\text{top}}$. The decays of bottom and charm hadrons were simulated using the EVTGEN 1.6.0 program.

SHERPA: two SHERPA configurations are studied: SHERPA 2.2.1 and SHERPA 2.2.10. Both configurations are simulated with the SHERPA MC generator using NLO-accurate matrix elements for up to one additional parton and LO-accurate matrix elements for up to four additional partons. The matrix elements are calculated with the COMIX [25] and OPENLOOPS [26–28] libraries. They were matched with the SHERPA parton shower [29] using the MEPS@NLO prescription [30–33] and the set of tuned parameters developed by the SHERPA authors to match the NNPDF3.0NNLO set of PDFs. The b - and c -quarks are treated as massless at matrix-element level and massive in the parton shower.

Heavy-flavor hadron production fractions in the SHERPA 2.2.1 setup are known to disagree with the world-average values as shown both in this note and in Ref. [34]. This is reflected in the poor modelling of b -jet fragmentation properties [35, 36]. A dedicated tuning was performed by the ATLAS Collaboration and implemented in the SHERPA event generator [10]. This update is already included in SHERPA 2.2.10, which is reflected in the better modeling of the heavy-flavor production fractions as seen in Section 5. Furthermore, the b -fragmentation function was altered to address an overabundance of jets with small b -momentum fraction and the heavy flavor decay tables were updated to the world-average values [9].

3 Heavy-flavor Production Fractions and Comparison to the World average

The heavy-flavor production fractions in the simulated samples are calculated by counting the number of final state weakly decaying hadrons originating from b - and c -quarks. For example, the D^\pm production fraction gives the fraction of c -quarks that hadronize into a D^\pm meson. By construction the production fractions of b - and c -quarks add up to 100%. Charm hadrons originating from b -hadron decays are excluded and all baryons are summed together into a baryon production fraction ($f(b\text{-baryon})$ and $f(c\text{-baryon})$). In case of multiple copies of the same hadron appearing in the decay chain (e.g. due to QED radiation), only the final copy is kept. The production fractions are given in Table 2. As expected there is no difference between POWHEG+PYTHIA and MADGRAPH5_AMC@NLO+PYTHIA because they use the same shower generator software. There are, however, differences between the different HERWIG and SHERPA versions.

Furthermore, the production fractions in ATLAS MC samples are compared to the world-average values. The b -hadron production fractions are taken from Particle Data Group (PDG) [9]. These are derived from hadronic Z boson decays using the data from the LEP experiments, summarized in Ref. [37]. The c -hadron production fractions are taken from Ref. [38]. The average c -hadrons production fractions are derived from a combined analysis of charm production measured in different regimes and environments: in e^+e^- collisions at B -factories, in $e^\pm p$ collisions in photoproduction, deep inelastic scattering, and in pp collisions. No kinematic or process dependence is assumed in these values. Systematic uncertainties are derived in forms of eigenvector variations from the uncertainties in the associated measurements and by using the correlation matrices published in Refs. [9] and [38]. Each eigenvector variation provides a set of alternative heavy-flavor hadron production fractions such that the sum of the individual b -hadron and

Table 2: Heavy-flavor hadron production fractions in the selected simulated samples. The statistical uncertainties in the production fractions are negligible ($<0.1\%$).

MC sample	$f(B^0)$	$f(B^\pm)$	$f(B_s^0)$	$f(b\text{-baryon})$	$f(D^\pm)$	$f(D^0)$	$f(D_s^\pm)$	$f(c\text{-baryon})$
POWHEG+PYTHIA 8	0.429	0.429	0.094	0.047	0.293	0.563	0.095	0.048
AMC@NLO +PYTHIA8	0.429	0.429	0.095	0.047	0.294	0.563	0.095	0.048
POWHEG+HERWIG 7.0.4	0.388	0.388	0.077	0.146	0.264	0.593	0.082	0.061
POWHEG+HERWIG 7.1.3	0.370	0.370	0.060	0.199	0.257	0.578	0.084	0.082
POWHEG+HERWIG 7.2.1	0.407	0.408	0.038	0.147	0.270	0.602	0.057	0.071
SHERPA 2.2.1	0.272	0.273	0.089	0.365	0.143	0.380	0.113	0.365
SHERPA 2.2.10	0.398	0.399	0.099	0.105	0.237	0.626	0.076	0.061

c -hadron fractions always remains 100%. Due to this constraint, the number of eigenvector variations is $N - 1$, where N is the number of independent heavy-flavor hadron production fractions. Therefore, there are three eigenvector variations for c -hadrons and two variations for b -hadrons. The B^0 and B^\pm production fractions are assumed to be the same because the electromagnetic decays of B^{*+} and B^{*0} and strong decays of B^{**+} and B^{**0} do not affect the relative amount of B^\pm and B^0 , unlike the excited D meson decays affect the relative rates of D^\pm and D^0 in the charm sector. The world-average productions fractions along with the eigenvector variations are given in Table 3. The deviations from the nominal values are about 2 to 3% for mesons and up to 10% for baryons.

Table 3: The world-average heavy-flavor hadron production fractions. The data is taken from PDG [9] for b -hadrons and from Ref. [38] for c -hadrons. Eigenvector variations are derived from the uncertainties and the correlation matrices given in the literature as described in the text.

MC sample	$f(B^0)$	$f(B^\pm)$	$f(B_s^0)$	$f(b\text{-baryon})$	$f(D^\pm)$	$f(D^0)$	$f(D_s^\pm)$	$f(c\text{-baryon})$
Nominal	0.408	0.408	0.100	0.084	0.240	0.609	0.080	0.071
b -hadron eig1	0.410	0.410	0.093	0.088	0.240	0.609	0.080	0.071
b -hadron eig2	0.415	0.415	0.097	0.074	0.240	0.609	0.080	0.071
c -hadron eig1	0.408	0.408	0.100	0.084	0.241	0.609	0.077	0.074
c -hadron eig2	0.408	0.408	0.100	0.084	0.244	0.611	0.078	0.067
c -hadron eig3	0.408	0.408	0.100	0.084	0.246	0.601	0.081	0.072

The comparison of the heavy-flavor production fractions in ATLAS MC samples to the world-averages is presented in Figure 1. SHERPA 2.2.1 has the worst agreement due to the value of the Heavy Baryon Enhancement parameter as discussed in Ref. [34]. It caused the fraction of baryons to be disproportionately high and the fractions of mesons to be too low. SHERPA 2.2.10 shows a better agreement with the world averages, where most of the production fractions are within the experimental systematic uncertainties. HERWIG MC samples generally agree well with the world averages except for B_s^0 and b -baryons where they can deviate from the averages up to a factor of two. The level of agreement in PYTHIA is similar to the newer SHERPA samples. Production fraction weights are determined from the ratio of the world average fractions over the predictions. In SHERPA 2.2.1 the weights have a significant impact on the effective statistical power of the simulated sample, where the relative uncertainty in the sum of weights is expected to increase by about a factor of 1.5 to 2.0 in $t\bar{t}$ events. The relative loss of statistics in other generators is up to 10% in HERWIG, 5% in PYTHIA, and less than 1% in SHERPA 2.2.10.

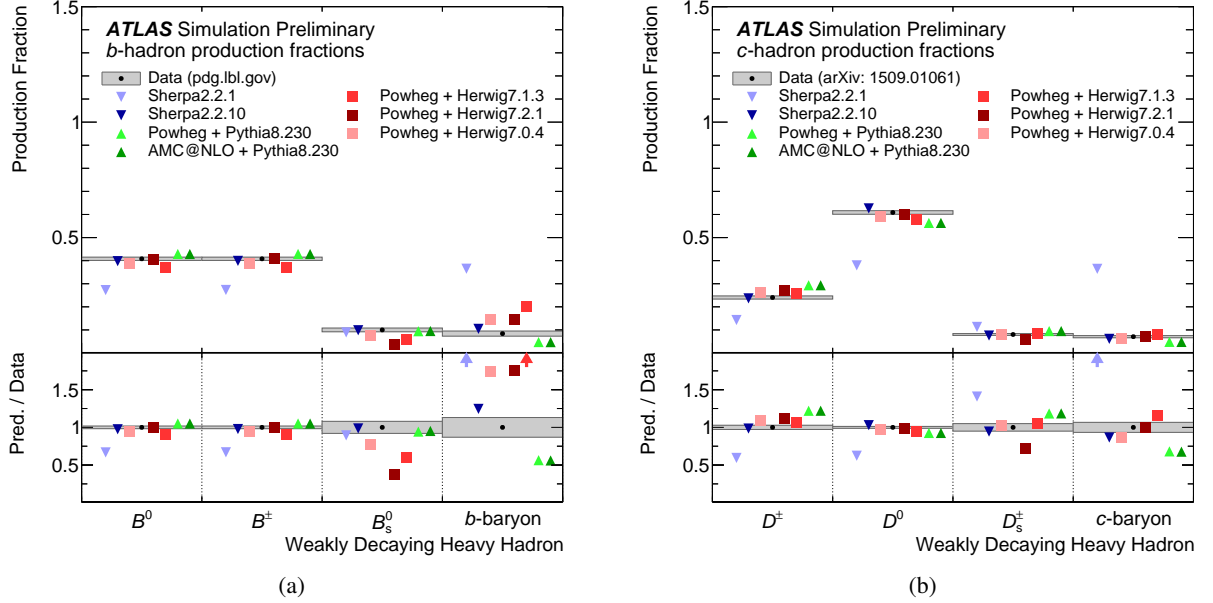


Figure 1: Heavy-flavor hadron production fractions in the selected simulated samples: (a) b -hadrons and (b) c -hadrons. The data is taken from PDG [9] for b -hadrons and from Ref. [38] for c -hadrons. The grey error bands are the uncertainties from the experimental measurements.

4 The heavy-flavor Production-fraction Reweighting Tool Design

The heavy-flavor production-fraction reweighting tool is designed to reweight the heavy-flavor production fractions in MC samples separately for each simulated event. The inputs to the tool are the heavy-flavor production fractions in ATLAS MC samples (Table 2), f^{MC} , and the world-average production fractions with the corresponding uncertainties (Table 3), f^{Data} . In addition, the tool requires access to the MC ‘truth record’ (i.e. the information about the particles simulated in the hard scatter interaction and later in the parton shower) and identifies the weakly decaying final state b - and c -hadrons in each simulated event. Fiducial cuts of $|\eta(\text{HF})| < 2.5$ and $p_{\text{T}}(\text{HF}) > 5$ GeV are applied to avoid reweighting the event weights based on heavy-flavor hadrons in the forward region outside the Inner Detector acceptance where jet flavor tagging is not possible. Furthermore, the 5 GeV cut is applied to reduce the statistical fluctuations that would otherwise be caused by low energy hadrons that are unlikely to be reconstructed in jets due to the lower energy threshold of 20 GeV for jets. Moreover, due to the low p_{T} , they are less likely to produce a significant displacement in the secondary vertex and be successfully tagged by the flavor tagging algorithms. c -hadrons originating from b -hadron decays are discarded and are not considered in the reweighting procedure. The reason for discarding them is the rate at which they are produced is determined purely from the b -hadron branching ratios and not the production fractions. Once a selection of heavy-flavor hadrons is performed, an event weight w is derived according to Eq. 1:

$$w = \prod_{\text{HF} \in V_{\text{fid}}} \frac{f^{\text{Data}}(\text{HF})}{f^{\text{MC}}(\text{HF})}, \quad (1)$$

where the event weight is multiplied with the $f^{\text{Data}}/f^{\text{MC}}$ ratio for each truth heavy-flavor hadron in the

fiducial selection, $V_{\text{fid.}}$, passing the fiducial cuts. In addition to the nominal event weight in Eq. 1 the tool provides five event weight variations corresponding to the eigenvectors in Table 3. The variations should be summed up in quadrature to derive the final uncertainty band for each observable sensitive to the heavy-flavor production fractions in MC.

The tool can either be used directly by specialized analyses that are sensitive to heavy-flavor production fractions (e.g. measurement of the associated $W + c$ production) or as part of the central heavy-flavor tagging calibration where the performance in MC is compared to that measured in data. Furthermore, the systematic uncertainties associated with the heavy-flavor production fractions provided by the tool can be used in the calibration and propagated to the scale factors used to correct the tagging efficiencies in MC. The MC-MC scale factors need to be derived with the production-fraction weights applied on both sets of simulated events. The same production fraction weights also need to be consistently applied in each analysis using the flavor tagging calibration derived in this way. The production fraction weights can be re-calculated at any stage in the analysis as long as the full truth record containing all heavy-flavor hadrons within the fiducial region and their decay productions are available.

5 Validation

The heavy-flavor production fraction reweighting tool is validated by comparing the tagging efficiencies in three $t\bar{t}$ samples generated with different MC shower software and by evaluating the impact of the heavy-flavor production-fraction weight on the agreement between the MC samples. POWHEG+PYTHIA 8, POWHEG+HERWIG 7.13, SHERPA 2.2.1, and SHERPA 2.2.10 simulated samples are compared.

5.1 Object and Event Selection

Jets are reconstructed using the ATLAS particle-flow [39] jet reconstruction algorithm. The algorithm extrapolates the charged tracks reconstructed in the Inner Detector to the calorimeter regions and removes the corresponding energy from calorimeter energy in order to remove the overlaps. The energy after subtraction is used as the input to the anti- k_t [40] jet-reconstruction algorithm with a distance parameter $R = 0.4$. The jets are required to have $p_T > 20$ GeV and $|\eta| < 2.5$. Jets from pile-up interactions are suppressed using the Jet Vertex Tagger (JVT) [41]. The jet energy scale (JES) calibration restores the jet energy to that of jets reconstructed at the particle level, as describe in Ref. [42].

Reconstructed jets are matched to truth hadrons based on geometric proximity (with $\Delta R(\text{jet}, \text{HF}) < 0.3^2$) and each reconstructed jet is categorized either as a b -jet, c -jet, or light-jet based on the truth hadron content within the jet. If the jet contains a b -hadron, it is categorized as a b -jet, as a c -jet if it contains a c -hadron and no b -hadrons, and as a light jet if no b - or c -hadrons are present. Only hadrons passing the fiducial requirement of $p_T(\text{HF}) > 5$ GeV and $|\eta(\text{HF})| < 2.5$ are considered (same as the fiducial cuts in Section 4). The tagging efficiency is then calculated as the fraction of jets of a certain type (e.g. b -jet) that are also tagged as a b -jet by the DL1r tagging algorithm. In order to remove the effect of multiple hadrons

² ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z -axis along the beam pipe. The x -axis points from the IP to the centre of the LHC ring, and the y -axis points upwards. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z -axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

within a jet (e.g. the poorly understood $g \rightarrow cc/bb$ process), only jets containing exactly one truth weakly decaying heavy-flavor hadron are considered in the validation.

Jets with $|\eta| < 2.5$ and $p_T > 20$ GeV containing b -hadrons are identified by a deep neural network tagger, DL1r [2, 43], that uses displaced tracks, secondary vertices and decay topologies. The 70% efficiency working point (WP) is used as benchmark. It has an efficiency of identifying b -jets in a $t\bar{t}$ sample of 70% and the measured rejection factor of c -jets (light-jets) is about 12 (580).

5.2 Results

As a first step of the validation of the production-fraction reweighting tool the heavy-flavor production fractions are calculated in event samples with at least one reconstructed jet. These do not necessarily match the production fractions given in Table 1 because of the requirement of a reconstructed jet which could have different efficiencies depending on the heavy-flavor species in which the b - or c -quarks hadronized. The event weights derived with the heavy-flavor production-fraction reweighting tool, described in Section 4, are applied and Figure 2 shows the production fractions before and after the reweighting. The production fractions before the reweighting show large differences between the MC shower generators. The largest outlier is the SHERPA 2.2.1 sample with the too large baryon fraction. After the weight from the tool is applied, the large disagreement among the samples is eliminated, showing that the reweighting procedure works as intended. For D_s^\pm , c -baryons, and b -baryons in the HERWIG sample, small differences up to 15% with respect to the other generators exist after the reweighting. This remaining discrepancy originates from different production fractions between the inclusive hadrons and hadrons with the jet selection. Given the small fractions of these hadrons, the impact of this discrepancy is much smaller than the effect of the reweighting.

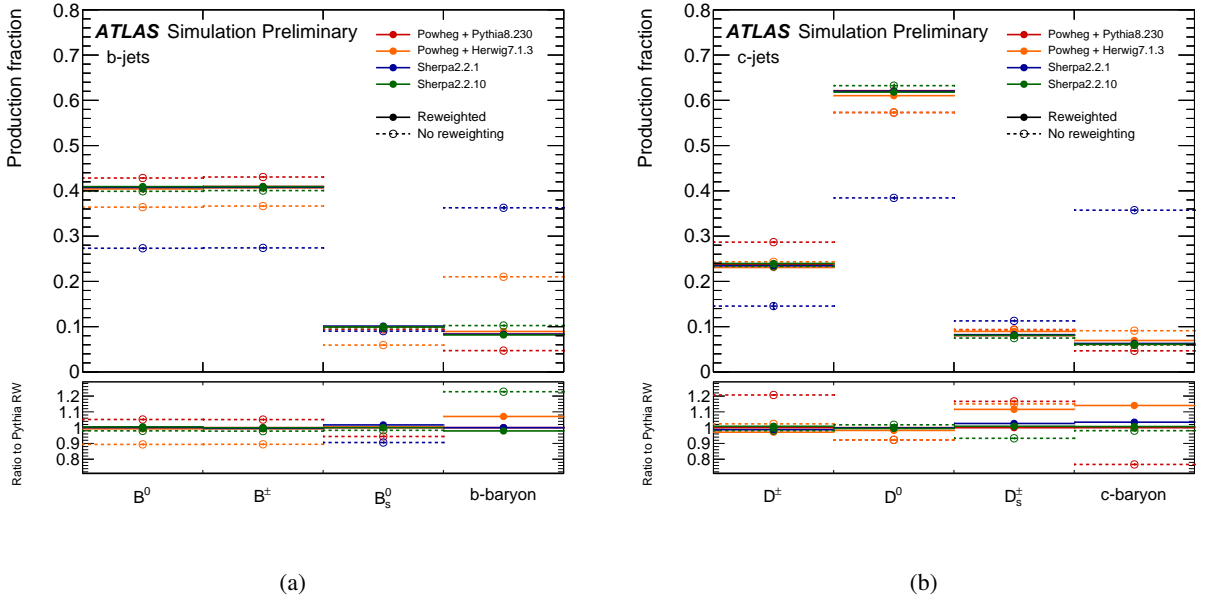


Figure 2: The heavy-flavor production fractions of heavy-flavor hadrons inside the reconstructed jets in the $t\bar{t}$ MC samples. The dashed lines represent the production fractions in the out-of-the-box MC and the solid lines show the production fractions after the heavy-flavor production-fraction reweighting is applied.

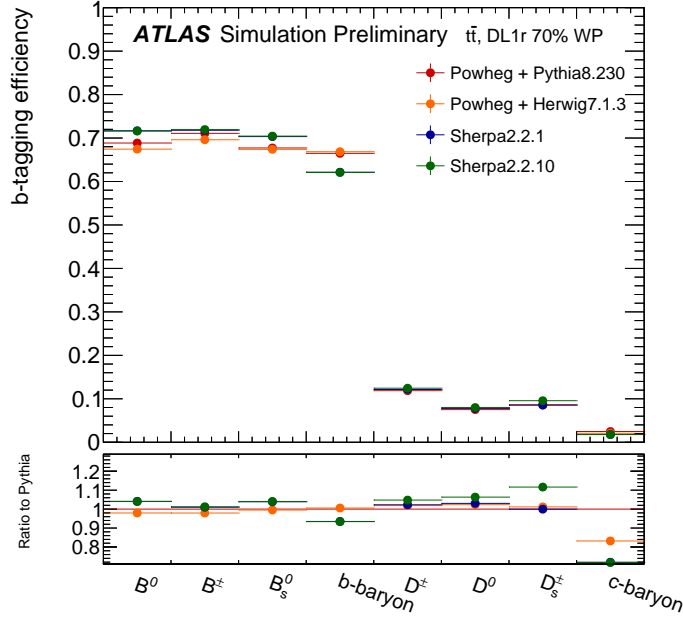


Figure 3: The b -tagging efficiency of b - and c -jets depending on the species of the truth hadron within the jet. The 70% b -tagging working point of the DL1r tagger is used.

The b -tagging efficiency of b -jets and the mistag rate of c -jets is shown in Figure 3 separately for each hadron species. By construction the tagging efficiencies per hadron species are independent of the production fractions. The tagging efficiency of b -jets is centered around 70%, corresponding to the 70% b -tagging working point. The efficiency is the highest for jets containing a B^\pm meson and the lowest for jets with a b -baryon. The c -jet mistag rate is about 10%, also corresponding to the expected rejection factor of the b -tagging working point. The difference in the efficiencies depending on the hadron species is larger for c -jets likely due to larger differences in lifetimes and decay topologies of c -hadrons in comparison to b -hadrons. For example, the relative difference between the D^\pm and D^0 efficiency is more than 30%, while it is almost the same for B^0 and B^\pm . Due to these different efficiencies it is important to correctly model the jet flavor composition in order to accurately describe the tagging efficiency of inclusive jets in MC. Moreover, the individual tagging efficiencies vary up to 5% between the MC shower generators, caused by two leading effects: different branching ratios and different fragmentation functions.

Lastly, the effect of the production fraction reweighting on the tagging efficiencies of b - and c -jets is studied. Figure 4 shows the b -tagging efficiency as a function of jet p_T for b - and c -jets, irrespective of the hadron species. The b -jet tagging efficiencies were already within 2% between the generators before the reweighting and the level of agreement stays mostly the same after the weights are applied. Tagging efficiency in HERWIG is about 2% lower than in PYTHIA and in SHERPA it is about 2% higher. There were larger differences in c -jet mistag rates, especially between SHERPA 2.2.1 and the rest of the MC shower generators. After the weights are applied SHERPA 2.2.1 matches with other samples within about 2%. This shows that the heavy-flavor production fractions played the largest role in the initial discrepancy between the samples for c -jets. However, there is little impact on the tagging efficiency of other samples because the production fractions were already quite close to the world average values and the differences in the tagging

efficiencies between the individual hadrons are not large enough cause mismodelling in the inclusive jet tagging efficiencies. SHERPA 2.2.10 still shows a discrepancy of up to 10% at high p_T . These remaining differences were found to be caused by the different c -hadron branching ratios in SHERPA 2.2.10 and PYTHIA samples; on average, the multiplicity of charged final state decay products in c -hadron decays was found to be larger in SHERPA 2.2.10, causing the c -jets mistag rates to be higher. Furthermore, small differences in flavor tagging efficiencies occur also due to the different fragmentation functions.

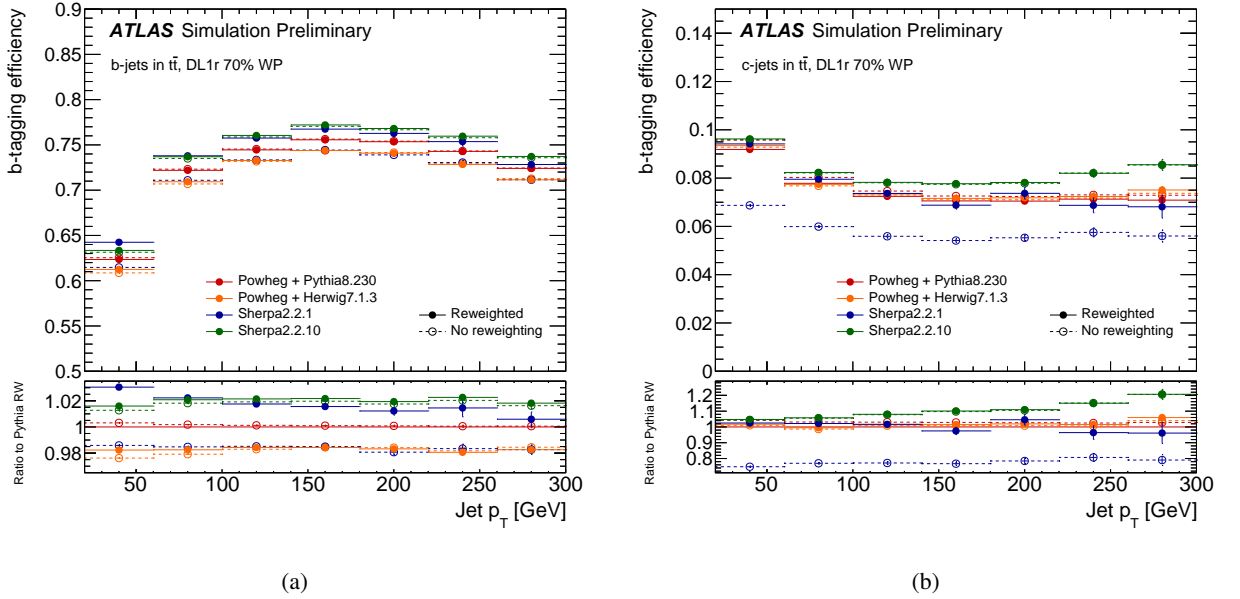


Figure 4: The b -tagging efficiencies of b - and c -jets as a function of jet p_T . The dashed lines represent the production fractions in the out-of-the-box MC and the solid lines show the production fractions after the heavy-flavor production fraction reweighting is applied.

In addition to reweighting the heavy-flavor production fractions to a common average, the tool also provides a set of eigenvector variations corresponding to the systematic uncertainties associated with the auxiliary production-fraction measurements. The effect on the tagging efficiencies is evaluated by calculating the efficiency for each eigenvector variation in Table 3. The total uncertainty band around the nominal value is then obtained by summing the differences with respect to the nominal value in quadrature and symmetrizing the sum. This is shown in Figure 5 for b - and c -jets. The resulting uncertainty on the tagging efficiency is less than 1%, and it is smaller for b -jets than c -jets. This is a direct consequence of individual c -hadrons having a larger spread in the tagging efficiency compared to b -hadrons, as shown in Figure 3. The systematic uncertainty from heavy-flavor production fractions on the tagging efficiency is therefore smaller than the remaining differences between the generators after the reweighting. This uncertainty is much smaller than the initial differences in production fractions between the generators, which shows that reweighting them is a good first step towards reducing the currently over-estimated hadronization uncertainties. However, differences between the generators of up to 2% in b -jet tagging efficiencies and up to 10% in c -jet mistag rates remain after the reweighting. They could be further reduced by harmonizing the heavy-flavor hadron decay tables and by reducing the differences in heavy-flavor fragmentation functions between the MC samples.

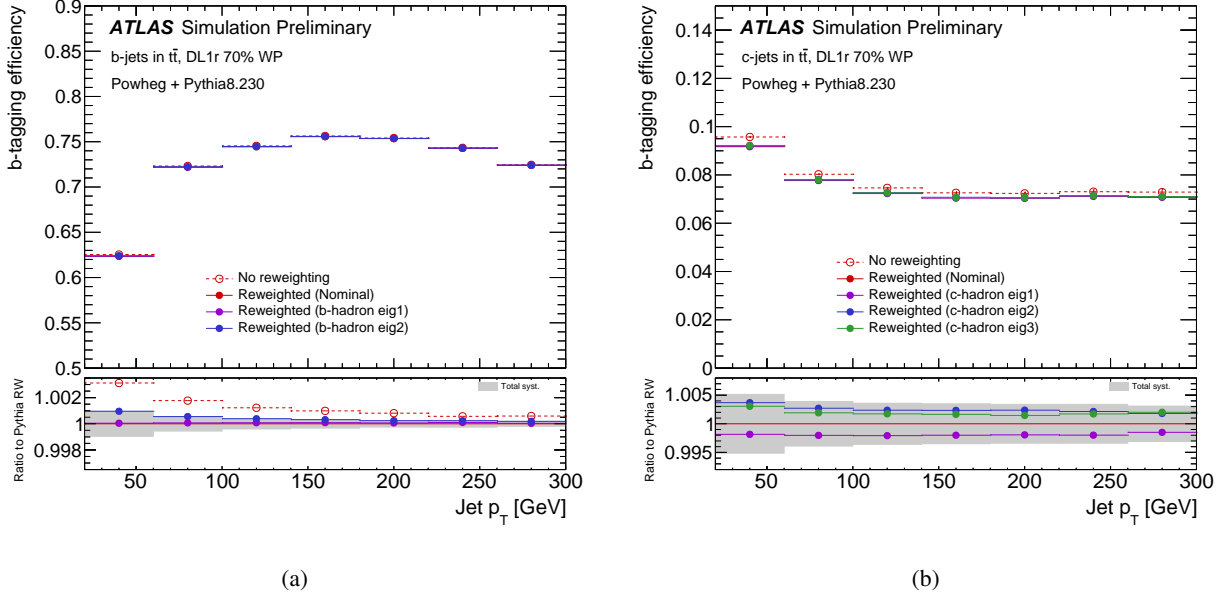


Figure 5: The effect of the production-fraction eigenvector variations on the b - and c -jet flavor tagging efficiency. The gray band shows the total uncertainty which is the symmetrized quadratic sum of the individual variations. Evaluated using the POWHEG+PYTHIA 8 MC sample.

6 Conclusions

The heavy-flavor production-fraction reweighting tool is presented. It is designed to reweight the heavy-flavor production fractions in MC samples to common world-average values. The differences in production fractions stem from differences in the tuning of each MC shower generator software and may have a large effect on quantities such as the heavy-flavor tagging efficiencies. The reweighting procedure eliminates the differences in the tagging efficiencies caused by production fractions and therefore improves the MC modelling related to heavy-flavor tagging. Before the reweighting, the relative differences in c -jet mistag rates were up to 40% between SHERPA 2.2.1 and other generators. After the reweighting, SHERPA 2.2.1 tagging efficiencies and c -jet mistag rates agrees within 2% with PYTHIA and HERWIG. Production fraction weights, however, do not have a large impact on tagging efficiencies in PYTHIA, HERWIG, and SHERPA 2.2.10 where the HF production fractions were already relatively close to the world averages before reweighting. Remaining differences of up to 2% are seen in b -jet tagging efficiencies and up to 10% in c -jet mistag rates. These are likely caused by different branching ratios and fragmentation functions between the generators. In addition, the tool estimates the impact of experimental systematic uncertainties from the auxiliary measurements of the heavy-flavor production fractions by providing five eigenvector variations. The systematic uncertainty is below 1% both for b - and c -jet tagging efficiencies. The tool can either be used directly by specialized analyses that are sensitive to heavy-flavor production fractions (e.g. measurement of the associated $W + c$ production) or as part of the central heavy-flavor tagging calibration where the MC modelling is evaluated with respect to the data.

References

- [1] ATLAS Collaboration, *Measurement of the c -jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS detector*, *Eur. Phys. J. C* **82** (2021) 95, arXiv: [2109.10627](https://arxiv.org/abs/2109.10627) [[hep-ex](#)] (cit. on p. 2).
- [2] ATLAS Collaboration, *ATLAS b -jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 970, arXiv: [1907.05120](https://arxiv.org/abs/1907.05120) [[hep-ex](#)] (cit. on pp. 2, 8).
- [3] ATLAS Collaboration, *Calibration of light-flavour b -jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2018-006, 2018, URL: <https://cds.cern.ch/record/2314418> (cit. on p. 2).
- [4] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159, arXiv: [1410.3012](https://arxiv.org/abs/1410.3012) [[hep-ph](#)] (cit. on pp. 2, 3).
- [5] E. Bothmann et al., *Event generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) 034, arXiv: [1905.09127](https://arxiv.org/abs/1905.09127) [[hep-ph](#)] (cit. on pp. 2, 3).
- [6] M. Bähr et al., *Herwig++ physics and manual*, *Eur. Phys. J. C* **58** (2008) 639, arXiv: [0803.0883](https://arxiv.org/abs/0803.0883) [[hep-ph](#)] (cit. on pp. 2, 3).
- [7] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J. C* **76** (2016) 196, arXiv: [1512.01178](https://arxiv.org/abs/1512.01178) [[hep-ph](#)] (cit. on pp. 2–4).
- [8] ATLAS Collaboration, *Monte Carlo to Monte Carlo scale factors for flavour tagging efficiency calibration*, ATL-PHYS-PUB-2020-009, 2020, URL: <https://cds.cern.ch/record/2718610> (cit. on p. 2).
- [9] P. D. Group and P. A. e. a. Zyla, *Review of Particle Physics*, *Progress of Theoretical and Experimental Physics* **2020** (2020), 083C01, ISSN: 2050-3911, eprint: <https://academic.oup.com/ptep/article-pdf/2020/8/083C01/33653179/ptaa104.pdf>, URL: <https://doi.org/10.1093/ptep/ptaa104> (cit. on pp. 2, 4–6).
- [10] ATLAS Collaboration, *Modelling and computational improvements to the simulation of single vector-boson plus jet processes for the ATLAS experiment*, (2021), arXiv: [2112.09588](https://arxiv.org/abs/2112.09588) [[hep-ex](#)] (cit. on pp. 2, 4).
- [11] *Charm-quark fragmentation fractions and production cross section at midrapidity in pp collisions at the LHC*, *Phys. Rev. D* **105** (1 2022) L011103, URL: <https://link.aps.org/doi/10.1103/PhysRevD.105.L011103> (cit. on p. 2).
- [12] Λ_c^+ production in pp and in p -Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, *Phys. Rev. C* **104** (5 2021) 054905, URL: <https://link.aps.org/doi/10.1103/PhysRevC.104.054905> (cit. on p. 2).
- [13] *Measurement of b hadron fractions in 13 TeV pp collisions*, *Phys. Rev. D* **100** (3 2019) 031102, URL: <https://link.aps.org/doi/10.1103/PhysRevD.100.031102> (cit. on p. 2).
- [14] S. Frixione, G. Ridolfi and P. Nason, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, *JHEP* **09** (2007) 126, arXiv: [0707.3088](https://arxiv.org/abs/0707.3088) [[hep-ph](#)] (cit. on p. 3).
- [15] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, arXiv: [hep-ph/0409146](https://arxiv.org/abs/hep-ph/0409146) (cit. on p. 3).

- [16] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070, arXiv: [0709.2092 \[hep-ph\]](#) (cit. on p. 3).
- [17] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, arXiv: [1002.2581 \[hep-ph\]](#) (cit. on p. 3).
- [18] R. D. Ball et al., *Parton distributions for the LHC run II*, *JHEP* **04** (2015) 040, arXiv: [1410.8849 \[hep-ph\]](#) (cit. on p. 3).
- [19] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, ATL-PHYS-PUB-2016-020, 2016, URL: <https://cds.cern.ch/record/2216168> (cit. on p. 3).
- [20] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: <https://cds.cern.ch/record/1966419> (cit. on p. 3).
- [21] R. D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244, arXiv: [1207.1303 \[hep-ph\]](#) (cit. on p. 3).
- [22] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152 (cit. on p. 3).
- [23] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, arXiv: [1405.0301 \[hep-ph\]](#) (cit. on p. 3).
- [24] L. Harland-Lang, A. Martin, P. Motylinski and R. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, *Eur. Phys. J. C* **75** (2015) 204, arXiv: [1412.3989 \[hep-ph\]](#) (cit. on p. 4).
- [25] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, *JHEP* **12** (2008) 039, arXiv: [0808.3674 \[hep-ph\]](#) (cit. on p. 4).
- [26] F. Buccioni et al., *OpenLoops 2*, *Eur. Phys. J. C* **79** (2019) 866, arXiv: [1907.13071 \[hep-ph\]](#) (cit. on p. 4).
- [27] F. Cascioli, P. Maierhöfer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601, arXiv: [1111.5206 \[hep-ph\]](#) (cit. on p. 4).
- [28] A. Denner, S. Dittmaier and L. Hofer, *COLLIER: A fortran-based complex one-loop library in extended regularizations*, *Comput. Phys. Commun.* **212** (2017) 220, arXiv: [1604.06792 \[hep-ph\]](#) (cit. on p. 4).
- [29] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani–Seymour dipole factorisation*, *JHEP* **03** (2008) 038, arXiv: [0709.1027 \[hep-ph\]](#) (cit. on p. 4).
- [30] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, *JHEP* **09** (2012) 049, arXiv: [1111.1220 \[hep-ph\]](#) (cit. on p. 4).
- [31] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *QCD matrix elements + parton showers. The NLO case*, *JHEP* **04** (2013) 027, arXiv: [1207.5030 \[hep-ph\]](#) (cit. on p. 4).
- [32] S. Catani, F. Krauss, B. R. Webber and R. Kuhn, *QCD Matrix Elements + Parton Showers*, *JHEP* **11** (2001) 063, arXiv: [hep-ph/0109231](#) (cit. on p. 4).

- [33] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, **JHEP** **05** (2009) 053, arXiv: [0903.1219 \[hep-ph\]](#) (cit. on p. 4).
- [34] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling with Sherpa and MG5_aMC@NLO*, ATL-PHYS-PUB-2017-007, 2017, URL: <https://cds.cern.ch/record/2261938> (cit. on pp. 4, 5).
- [35] ATLAS Collaboration, *Measurements of jet observables sensitive to b-quark fragmentation in $t\bar{t}$ events at the LHC with the ATLAS detector*, (2022), arXiv: [2202.13901 \[hep-ex\]](#) (cit. on p. 4).
- [36] ATLAS Collaboration, *Measurement of b-quark fragmentation properties in jets using the decay $B^\pm \rightarrow J/\psi K^\pm$ in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, **JHEP** **12** (2021) 131, arXiv: [2108.11650 \[hep-ex\]](#) (cit. on p. 4).
- [37] *Averages of b-hadron, c-hadron, and τ -lepton properties as of 2018*, **Eur. Phys. J. C** **81** (2021) 226, arXiv: [1909.12524 \[hep-ex\]](#) (cit. on p. 4).
- [38] M. Lisovyi, A. Verbytskyi and O. Zenaiev, *Combined analysis of charm-quark fragmentation-fraction measurements*, **Eur. Phys. J. C** **76** (2016) 397, arXiv: [1509.01061 \[hep-ex\]](#) (cit. on pp. 4–6).
- [39] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, **Eur. Phys. J. C** **77** (2017) 466, arXiv: [1703.10485 \[hep-ex\]](#) (cit. on p. 7).
- [40] M. Cacciari, G. P. Salam and G. Soyez, *The anti- k_t jet clustering algorithm*, **JHEP** **04** (2008) 063, arXiv: [0802.1189 \[hep-ph\]](#) (cit. on p. 7).
- [41] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATL-CONF-2014-018, 2014, URL: <https://cds.cern.ch/record/1700870> (cit. on p. 7).
- [42] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, **Eur. Phys. J. C** **81** (2020) 689, arXiv: [2007.02645 \[hep-ex\]](#) (cit. on p. 7).
- [43] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013, 2017, URL: <https://cds.cern.ch/record/2273281> (cit. on p. 8).

Appendix

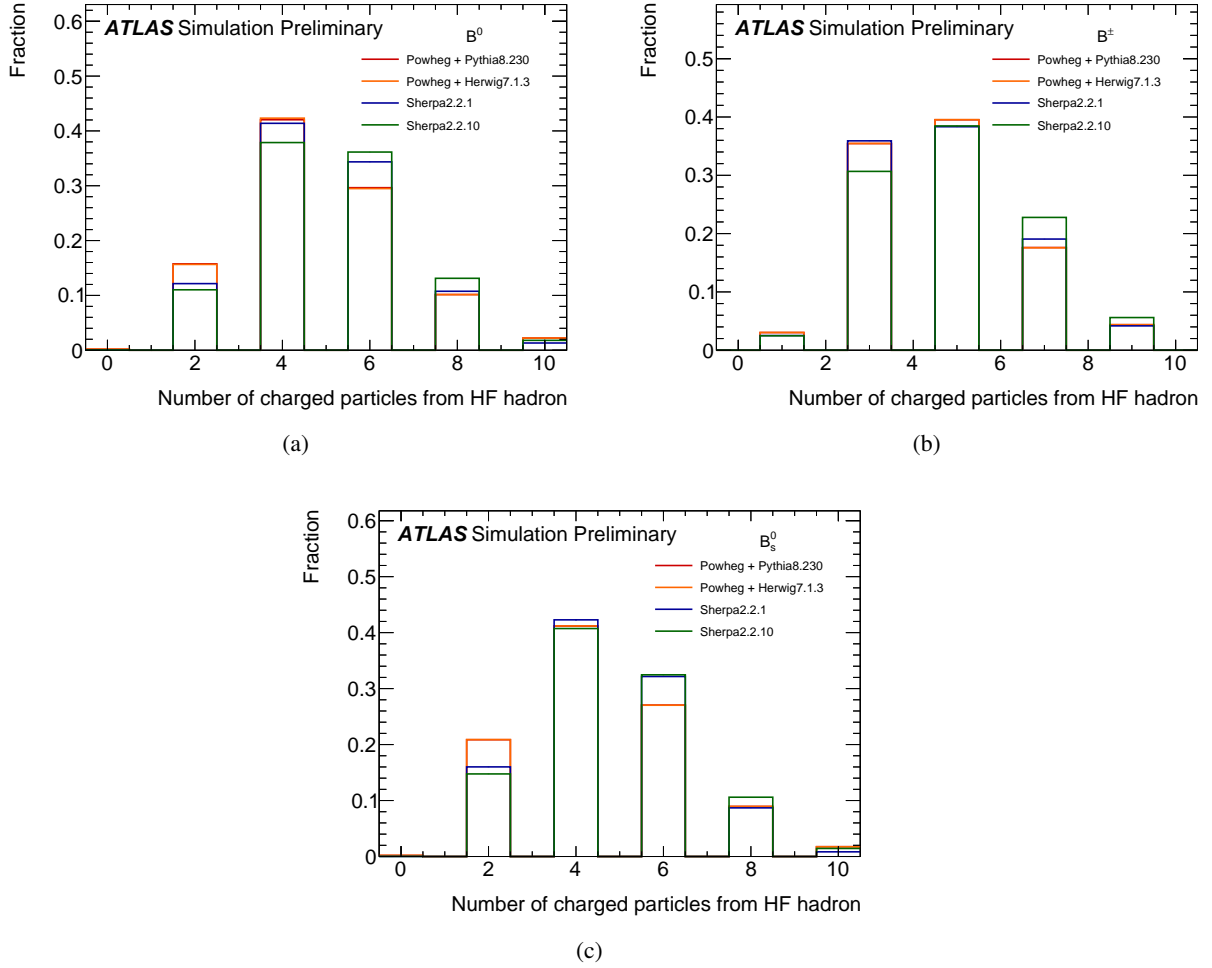
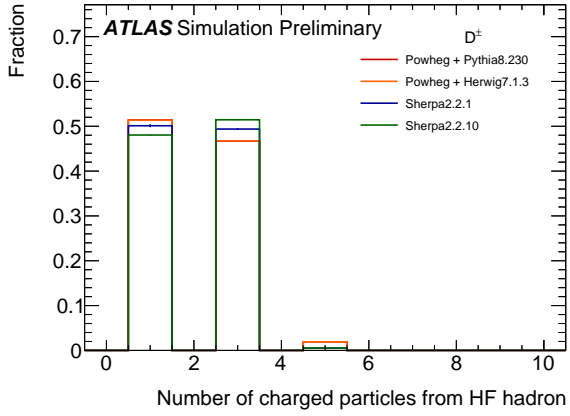
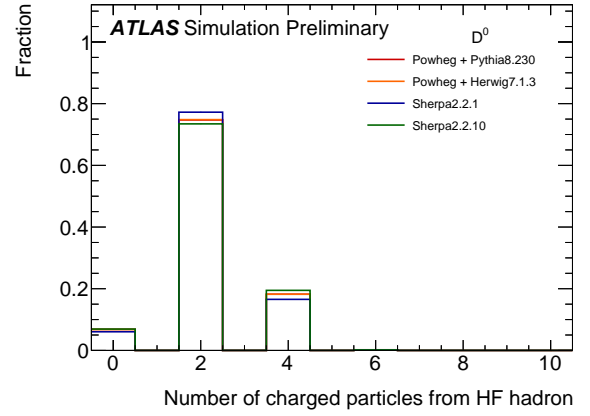


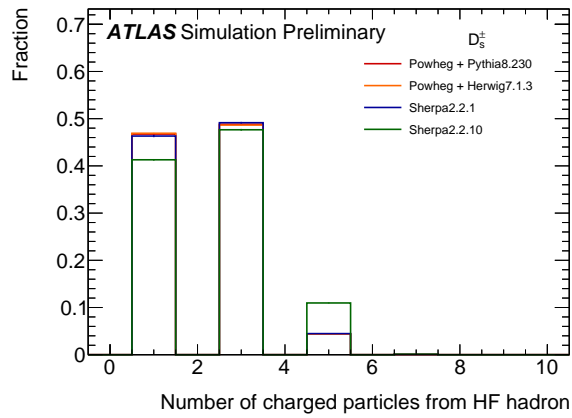
Figure 6: The distribution of multiplicity of charged particles in the b -meson weak decays for multiple MC shower generators: (a) B^0 mesons, (b) B^\pm mesons, and (c) B_s^0 mesons. Evaluated using the POWHEG+PYTHIA 8 MC sample.



(a)



(b)



(c)

Figure 7: The distribution of multiplicity of charged particles in the c -meson weak decays for multiple MC shower generators: (a) D^\pm mesons, (b) D^0 mesons, and (c) D_s^\pm mesons. Evaluated using the POWHEG+PYTHIA 8 MC sample.