

Multidifferential study of identified (π^\pm, K^\pm, p^\pm) and nonidentified charged hadron distributions in Z-tagged jets in pp collisions at $\sqrt{s} = 13$ TeV

Sookhyun Lee, Christine A. Aidala

University of Michigan, Ann Arbor, United States

Abstract

The distributions of charged hadrons within jets recoiling against a Z boson are measured in 13 TeV proton-proton collision data from 2016. The charged hadron distributions are studied longitudinally and transverse to the jet axis for jets with transverse momentum $p_T > 20$ GeV/ c and in the pseudorapidity range $2.5 < \eta < 4$. Identified charged pion, kaon, and proton distributions in jets are measured for the first time in proton-proton collisions, including the joint distributions in two kinematic variables simultaneously. In addition, non-identified charged hadron results in 1D are compared to previous LHCb measurements performed at 8 TeV [1] and extended to 2D joint distributions for the higher-statistics 13 TeV data set.

Contents

1	Introduction	1
2	Event Samples	4
2.1	Data	4
2.2	Monte Carlo Simulation	5
3	Event Selection	6
3.1	Cuts	6
3.2	Kinematic Distributions	7
4	Jet Reconstruction Performance and Efficiencies	11
4.1	Jet Directional Resolution	11
4.2	Jet p_T Resolution	11
4.3	Jet Reconstruction Efficiency	12
5	Muon Reconstruction Efficiencies	13
5.1	Trigger efficiencies	14
5.2	Muon Identification efficiencies	14
5.3	Muon track reconstruction efficiencies	16
6	Hadron reconstruction efficiency and purity within jets	17
7	Particle Identification Corrections	20
8	Unfolding	26
9	Systematic Uncertainties	30
9.1	Normalization	30
9.2	Jets	31
9.2.1	Jet energy scale	31
9.2.2	Jet energy resolution	33
9.3	Unfolding	34
9.4	Hadrons	35
9.4.1	Track Selection Purity	35
9.4.2	Track reconstruction efficiency	37
9.4.3	Particle identification	37
9.4.4	Magnet polarity	38
9.5	Summary of systematic uncertainties	38
10	Results	42
10.1	Nonidentified charged hadrons (π^\pm , K^\pm and p^\pm combined)	42
10.2	Identified charged hadrons - π^\pm , K^\pm and p^\pm	51

1 Introduction

This note documents measurements of identified charged hadron distributions within Z -tagged jets at $\sqrt{s} = 13$ TeV, following up on the initial measurements of this kind at LHCb performed with nonidentified charged hadrons at $\sqrt{s} = 8$ TeV [1]. The transverse motion of partons inside the nucleon has gained enormous attention over the past few decades since the strikingly large asymmetries at large $x_F = 2p_L/\sqrt{s}$ observed in the transverse single-spin asymmetry (TSSA) measurements for particle production in hadronic collisions involving transversely polarized proton beams (Fig. 1). The experimental observations contradicted predictions of asymmetries of order $\sim \alpha_s \frac{m_q}{Q}$ based on naive perturbative QCD calculations [2], where Q is a hard scale and m_q is the mass of a quark.

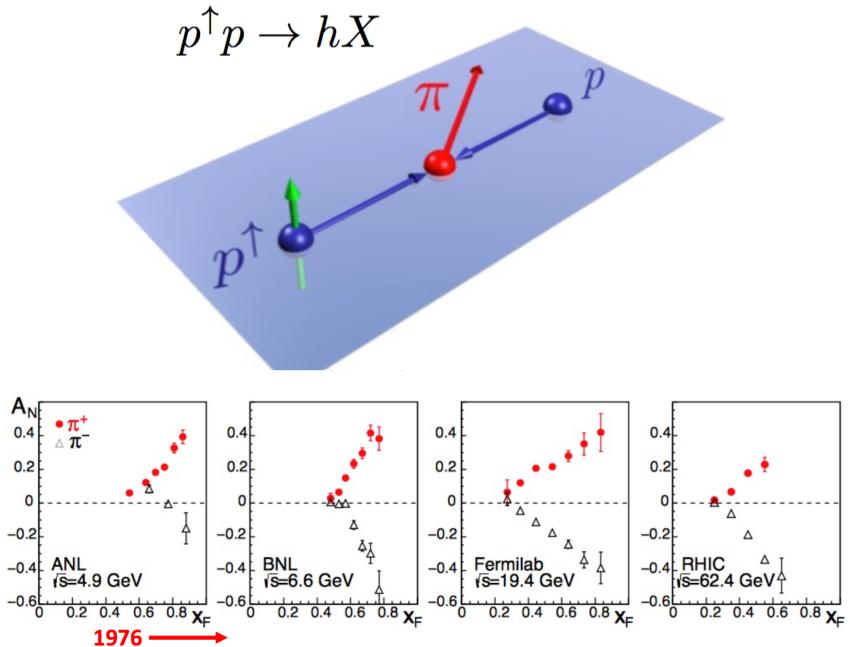


Figure 1: Transverse single-spin asymmetry (TSSA) A_N measurements for charged pion production in hadronic collisions involving a transversely polarized proton beam [3]. The A_N is defined as $\frac{d\sigma^\uparrow - d\sigma^\downarrow}{d\sigma^\uparrow + d\sigma^\downarrow}$, where $d\sigma^{\uparrow(\downarrow)}$ is the differential cross section of hadron production when one of the two incoming protons has spin up (down), *i.e.* transversely polarized with respect to its momentum.

One of the two mechanisms that can explain the large asymmetries takes into account the correlation between partonic transverse motion inside the nucleon, encoded in the transverse momentum dependent parton distribution functions (TMD PDFs), and the spin of the nucleon. The counterpart of this phenomenon for the fragmenting partons is also being widely studied in terms of TMD fragmentation functions (TMD FFs) that renders description possible of three-dimensional motion of the final state hadrons with respect to the hard-scattered partons and their interactions with the spin of the hadrons. In $p + p$ collisions, looking at hadron distributions within jets can directly probe the TMD FFs, avoiding convolutions with the TMD PDFs.

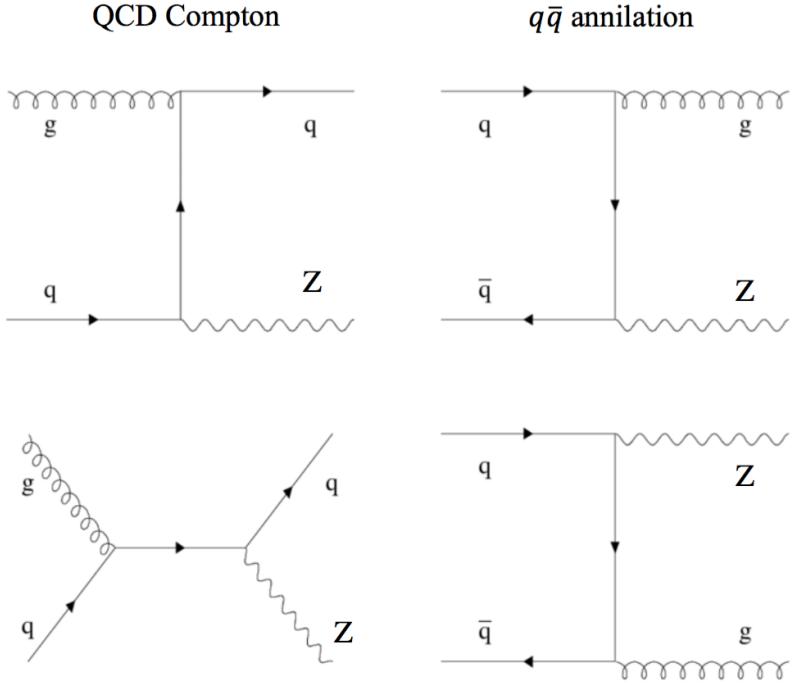


Figure 2: Leading order partonic processes for Z +jet production.

20 The main goal of this analysis is to make measurements of observables that can be used
 21 to constrain unpolarized quark transverse-momentum-dependent (TMD) fragmentation
 22 functions (FF) that is related to quark TMD Fragmenting Jet Function (FJF), well defined
 23 within the soft collinear effective theory (SCET) framework [4]. Generally, an advantage of
 24 measuring jet substructure observables is that it can access more differential information on
 25 the physical objects of interest. The first measurements using jet substructure techniques
 26 in order to access TMD FFs were performed by the STAR collaboration [5]. Their
 27 measurements are sensitive to the Collins (polarized TMD) fragmentation functions.
 28 Recent Belle results on transverse momentum dependent cross section measurements of
 29 charged hadrons [6] intended for accessing unpolarized TMD FFs are also closely related
 30 to our measurements.

31 The unpolarized TMD FFs depend on z , j_T and theory scale μ , *i.e.* $\hat{D}_i^h(z, j_T, \mu)$,
 32 therefore our main observables naturally involve z and j_T , experimentally defined in Eq. 1
 33 and also found in Ref. [1].

34

$$z = \frac{\vec{p}_h \cdot \vec{p}_{jet}}{\vec{p}_{jet}^2}, \quad j_T = \frac{|\vec{p}_h \times \vec{p}_{jet}|}{\vec{p}_{jet}} \quad (1)$$

35

- 36 • z : the longitudinal momentum fraction of jet momentum carried by a hadron.
- 37 • j_T : the transverse component of hadron momentum with respect to the jet axis.

38

39 Leading order partonic processes are shown in Fig. 2. The dominant hard partonic
 40 process for Z +jet production in the LHCb acceptance ($2 < \eta < 5$) at 13 TeV is $qg \rightarrow Zq$,

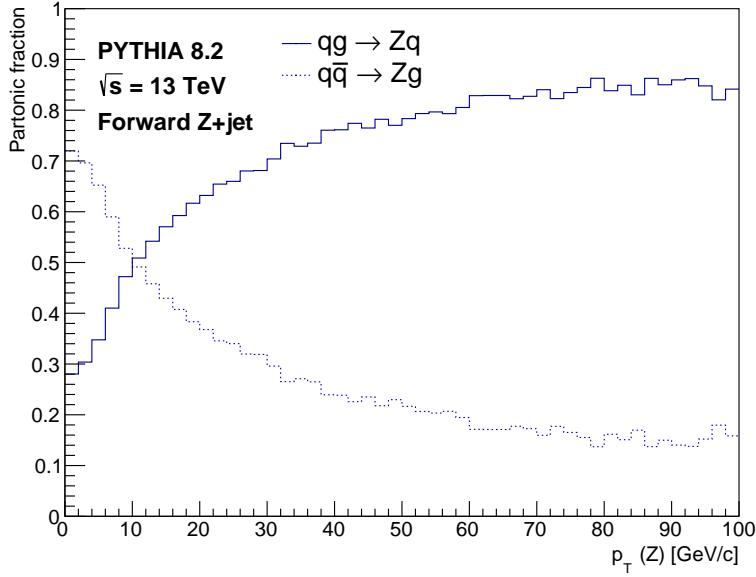


Figure 3: Composition of partonic processes for Z +jet events at LHCb.

similarly to 8 TeV, as can be seen in Fig. 3. Z -tagged jets at LHCb are therefore predominantly initiated by quarks and enhance sensitivity to the quark TMD FFs. Identified charged hadron distributions provide additional information on the valence versus sea quark's role initiating the parton shower that leads to the creation of jets. Measurements that are sensitive to electromagnetic charge such as positive-to-negative charge ratio or jet charge [7] can offer a more systematic way of decomposing jets into different parton flavors.

One caveat to interpreting these measurements is that TMD factorization breaking is predicted for jet production in $p+p$ [8], in which case non-perturbative TMD distribution and fragmentation functions can no longer be identified as universal objects in different collision processes.

The results of jet fragmentation function measurements we will be presenting in this note are listed in Eq. 2, where \mathcal{PS} indicates “phase space” (see [4] for similar notation). By dividing the differential cross section by the one integrated over the observable of interest, experimental corrections related to the jet reconstruction, not associated with the hadrons inside jets, *e.g.* muons efficiencies and jet reconstruction efficiencies, cancel out.

$$F(z_h) = \frac{d\sigma}{d\mathcal{PS}dz_h} / \frac{d\sigma}{d\mathcal{PS}}, \quad f(z_h, j_T) = \frac{d\sigma}{d\mathcal{PS}dz_h}, \quad K(j_T) = \frac{d\sigma}{d\mathcal{PS}dj_T} / \frac{d\sigma}{d\mathcal{PS}} \quad (2)$$

where $h = \{\pi^\pm, K^\pm, p^\pm, \text{all three combined}\}$.

Experimentally, these quantities can be expressed as shown in Eq. 3, where N_{Z+jet} represents the number of Z -tagged jets in each jet p_T bin and N_h the number of hadrons in each hadronization variable bin. For an apples-to-apples comparison with the 8 TeV results, we will present 1-dimensional distributions for non-identified charged hadrons as well.

$$F = \frac{1}{N_{Z+jet}} \frac{N_h(z_h)}{\Delta z_h}, \quad f = \frac{1}{N_{Z+jet}} \frac{N_h(z_h, j_T)}{\Delta z_h \Delta j_T}, \quad K = \frac{1}{N_{Z+jet}} \frac{N_h(j_T)}{\Delta j_T} \quad (3)$$

64 Various data corrections that will be discussed in this note are summarized in Eq. 4,
65 with $\Delta\mathcal{O}$ representing the bin width for variables of interest.

$$\frac{1}{N_{Z+jet}} \frac{N_h(\mathcal{O})}{\Delta\mathcal{O}} = \frac{\epsilon_{Z\mu^+\mu^-}\epsilon_{jet}}{U_{jet}N_{Z+jet}^{reco}} U_{jet-track} \frac{c_{PID}N_{track}^{reco}(\mathcal{O})}{c_{hadron}\Delta\mathcal{O}}, \quad (4)$$

66 where ϵ_μ is the muon reconstruction efficiency, ϵ_{jet} is the jet reconstruction efficiency,
67 U_{jet} is the unfolding factor for jet p_T , $U_{jet-track}$ is the unfolding factor for hadronization
68 variables simultaneously unfolded with jet variables, c_{PID} is the PID correction, c_{hadron} is
69 the hadron reconstruction efficiency and purity correction. Finally, the new elements in
70 this analysis as compared to the published 8 TeV results are summarized in Table 1.
71

Table 1: Summary of differences between published results [1] and this analysis.

	old (published)	new (this analysis)
\sqrt{s}	8 TeV	13 TeV
j_T and z distributions	1D	1D and 2D
PID	non-identified h^\pm	non-identified and identified π^\pm, K^\pm, p^\pm

72 2 Event Samples

73 2.1 Data

74 The analysis presented here is based on the $p + p$ data collected in 2016 at $\sqrt{s} =$
75 13 TeV during the Run 2 data taking period. This data selection corresponds to an
76 integrated luminosity of 1.64 fb^{-1} . Systematic uncertainties dominate the uncertainties on
77 the measurements, *e.g.* jet-energy-scale (JES)/jet-energy-resolution (JER) uncertainties
78 and the uncertainties on the particle ID correction factors at low ($< 10 \text{ GeV}/c$) and
79 high momentum ($> 100 \text{ GeV}/c$), therefore only 2016 data is used. The event selection
80 and analysis of jet and muon reconstruction closely follow the established $Z + \text{jet}$ cross
81 section analyses [9, 10]. The analysis was performed using data that were produced
82 with `Reco16` and `Stripping28r2` and passed `Z02MuMuLine` selection criteria. At least
83 one of the decay muons from Z was required to be Triggered-On-Signal (TOS) all
84 the way through the trigger, on the following lines: `L0MuonEW`, `Hlt1SingleMuonHighPT`
85 and `Hlt2EWSingleMuonVHighPt`. Jet substructure measurements (quarkonia in jets [11],
86 hadrons in jets, *etc.*) performed with the LHCb detectors in the low pileup ($\nu \sim 1.6$ for
87 Run 2, even less than $\nu \sim 2.5$ for Run 1 period) environment have a great advantage over
88 other experiments at LHC. As has been done in these previous analyses, requiring the

presence of only one primary vertex (nPV) removes issues with high fake jet rate. Jets from additional primary vertices that accidentally meet the back-to-back configuration requirement can make up for the fake jets in this analysis. Requiring nPV = 1 retains $\sim 40\%$ in statistics as seen in Fig. 4.

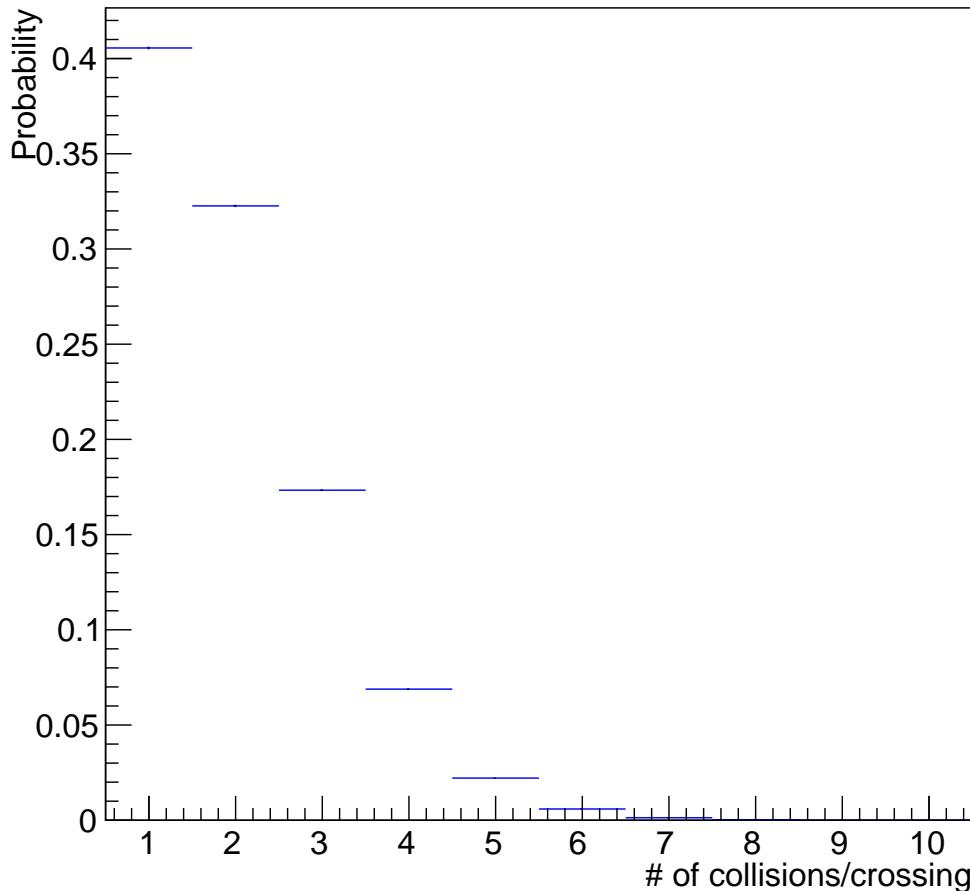


Figure 4: Probability distribution of number of collisions per bunch crossing for Run II 2016 data at LHCb.

2.2 Monte Carlo Simulation

Monte Carlo simulation was used to estimate systematic uncertainties on various data corrections, and to study the detector response to jets and charged tracks for unfolding. The MC used for this analysis was MC16, standard MC for data taken in 2016 Run 2, specifically the one reconstructed with `Reco16` and processed with `Sim09j` version and `Stripping28r2`, the same Reconstruction and Stripping selection that was applied to data. The generation code corresponds to 42112022 which requires at least one decay muon from a Z in the LHCb acceptance. This MC is based on PYTHIA8.2 [12] and EVTGEN [13], and includes a simulation of the LHCb detector [14], as well as the interactions of the

102 detector with the particles in the event using GEANT4 [15]. Events are simulated using
103 the CT09MCS [16] parametrisation of PDFs.

104 3 Event Selection

105 3.1 Cuts

106 Event selection is based on fiducial cuts, event topology and kinematic or quality
107 cuts applied to jets, muons and charged hadrons. The decay muons from Z bosons were
108 required to be in the nominal LHCb fiducial area, while the center of jets had to be inside
109 the (fiducial boundary - jet radius) to ensure all of the constituent particles are within
110 the fiducial area.

111 Loose track quality cuts such as momentum resolution (relative uncertainty on the
112 momentum) less than 10% and the probability of track fit greater than 0.1% were required
113 for each of the muon pairs to veto fake tracks resulting from accidental matching of hits
114 in the detector. Identification of decay muons was done by passing `isMuon` criteria and
115 the invariant mass of the track pairs within the Z boson mass window.

116 The common particle object `StdJet` provided by the Jet Reconstruction working
117 group [17] was used for this analysis. Jets were reconstructed using the `anti-kt`
118 algorithm with a jet radius of 0.5. We used `StdJets` as `HltStdJet` was not available until
119 after we had evaluated performance using `StdJet` and the analysis progressed. The
120 results of evaluation are documented in Sec.4 in detail and showed that `StdJet` performed
121 reasonably well, better than Run 1 on some criteria including fake jet rate partly thanks
122 to the unique primary vertex requirement per event and lower pileup environment. A
123 minimum p_T cut of 15 GeV/c was applied to reconstructed jets. The p_T range of this
124 analysis starts from 20 GeV/c. The p_T threshold was determined taking the following
125 factors into consideration. (1) The jet reconstruction efficiency is high enough, greater
126 than 80% for a given threshold. (2) In order to correct for smearing effects, *i.e.* unfold
127 the p_T distribution, one needs to make measurements in a slightly extended p_T range
128 such that a large fraction of smearing can be accounted for. A threshold of 15 GeV/c met
129 these requirements. We applied a tighter eta cut in order to avoid degraded performance
130 for ‘tracks in a jet’ due to edge effects that cannot be corrected for in calibrations. Studies
131 from previous Run1 non-identified charged hadron distribution analysis found that the
132 usual (2.2, 4.2) had to be tightened and we applied the same cuts.

133 In order to enhance back-to-back topology for Z and jet for a given event, a minimum
134 azimuthal distance between the two had to be met. Also, decay muons were required to
135 be well separated from the jet. The cuts discussed are summarized below.

136
137 Z bosons:

- 138
139 • $p_T(\mu) > 20 \text{ GeV}/c,$
- 140
141 • $2.0 < \eta(\mu) < 4.5$
- 142
143 • $\frac{\sigma_p}{p} < 0.1,$

- Track Prob(χ^2) > 0.1 %,
- `isMuon`,
- $60 < M_{\mu^+\mu^-} < 120 \text{ GeV}/c^2$

Jets:

- `StdJets`,
- $p_T(\text{jet}) > 15 \text{ GeV}/c$,
- $2.5 < \eta(\text{jet}) < 4.0$,

Event topology:

- $\Delta\phi(\text{jet}-\mu) > 0.4$
- $\Delta\phi(\text{Z-jet}) > \frac{7}{8}\pi$

Tracks :

- Track $\chi^2/\text{ndf} < 3$
- $4 < p < 1000 \text{ GeV}/c$
- $p_T > 0.25 \text{ GeV}/c$
- $\text{ProbNNGhost} < 0.5$

3.2 Kinematic Distributions

In this section, basic kinematic distributions of the Z bosons and jets fulfilling all the requirements described previously are compared between the data and MC. Generally, good agreement is seen between simulation and data, indicating that underlying physics processes and detector effects in the MC simulation are compatible with the corresponding ones in data. The jet transverse momentum and pseudorapidity distributions are found in Fig. 5 and Fig. 6. The Z boson transverse momentum and rapidity distributions are shown in Fig. 7 and Fig. 8. These distributions are from events that pass the back-to-back configuration requirement for the Z and jet, corresponding to approximately 14K events for both the data and MC. Notably, as can be seen in Fig. 9, the MC simulation describes the azimuthal distance between the Z and jet in data; for the 7 TeV data there had been an overestimation of back-to-back Z +jet events [9]. Figure 10 shows the rapidity difference between the Z and jet in MC compared to data. In these two figures, there are ten times the statistics that were in the event sample requiring the back-to-back configuration for the Z and jet.

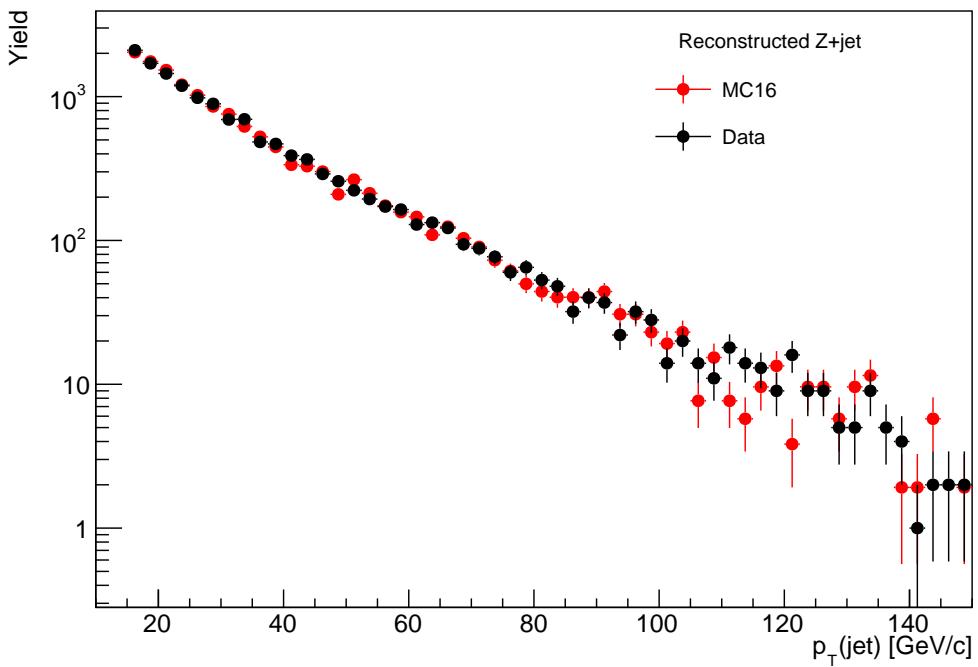


Figure 5: The jet transverse momentum distribution in data (black) and simulation (red) shown in log scale.

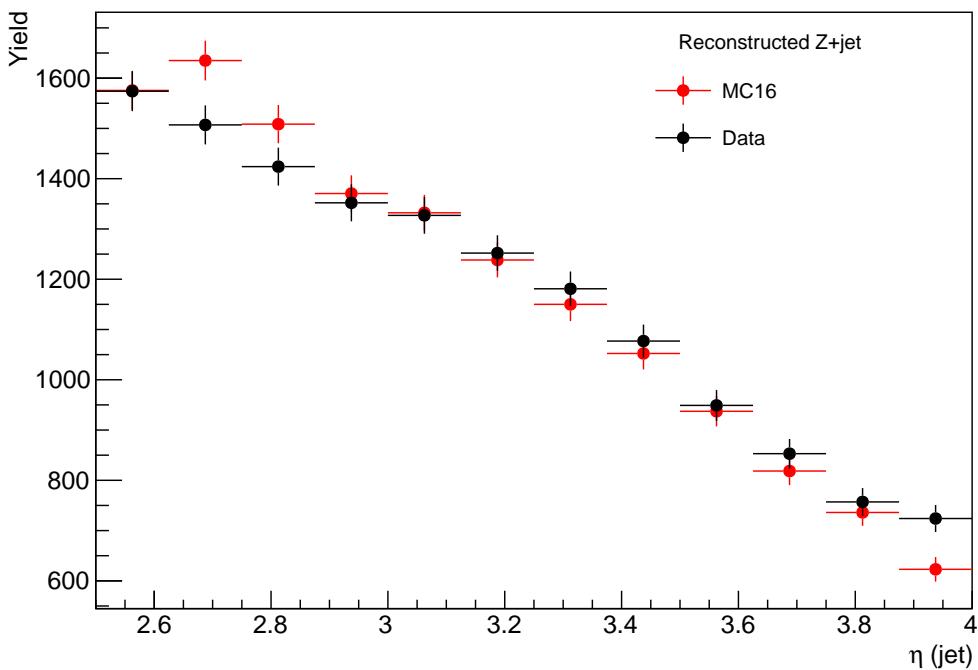


Figure 6: The jet pseudorapidity distribution in data (black) and simulation (red).

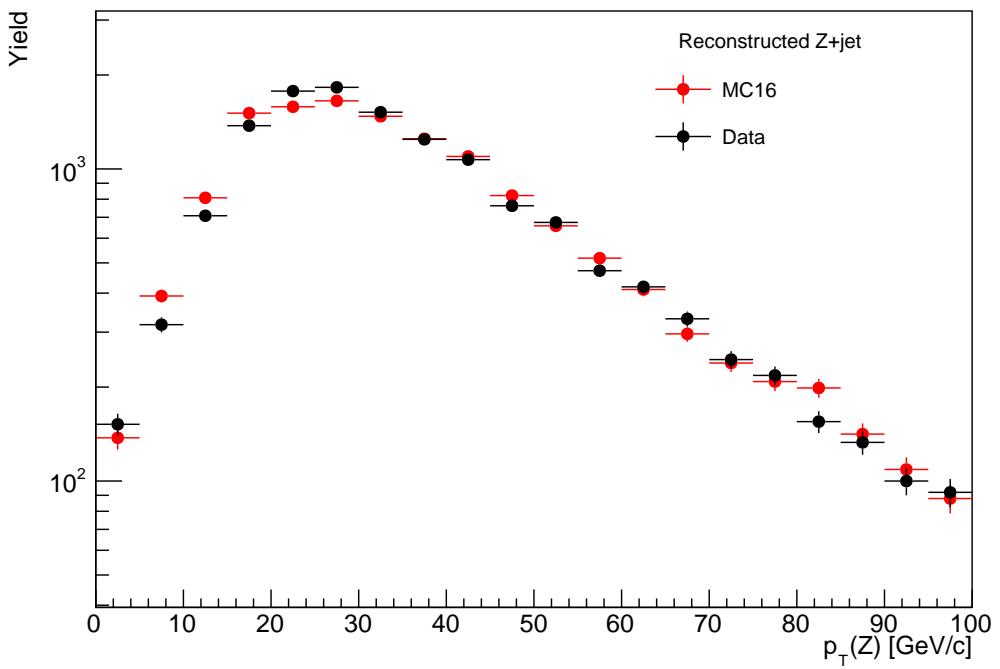


Figure 7: The Z boson transverse momentum distribution in data (black) and simulation (red) shown in log scale.

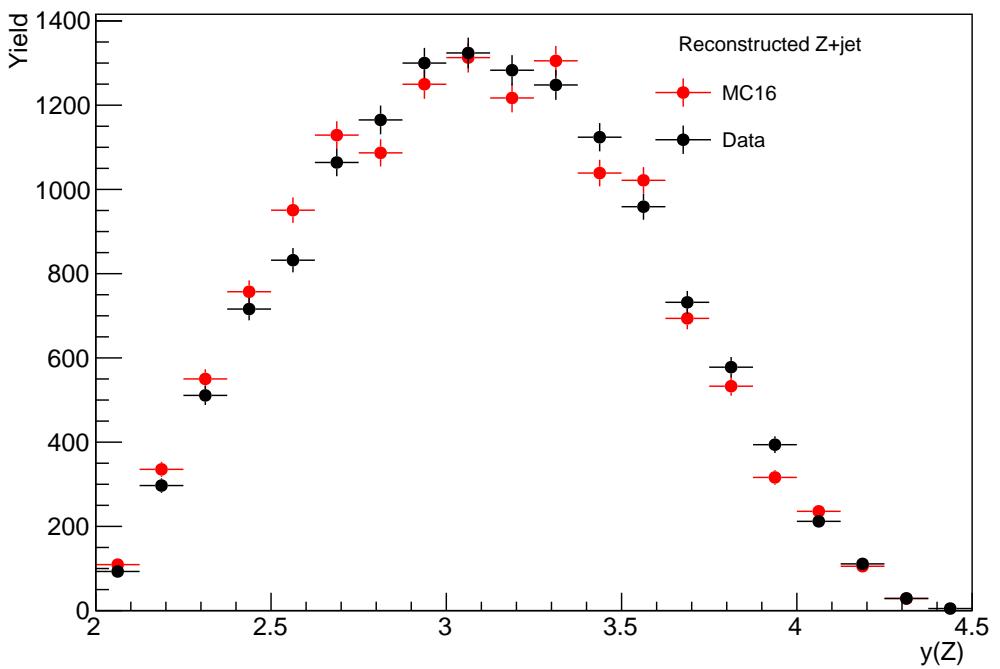
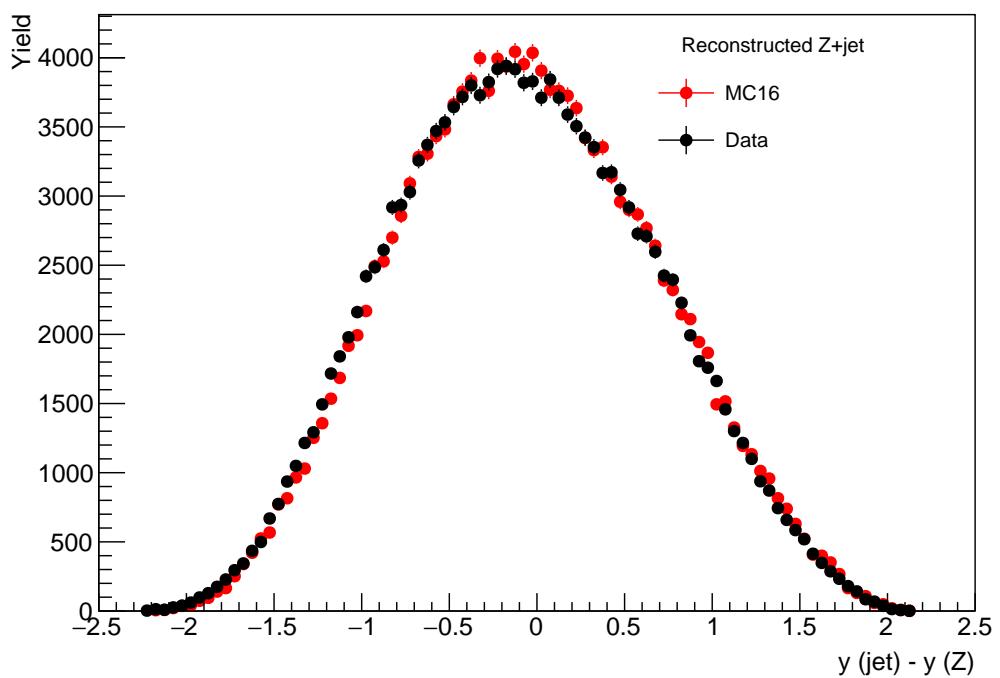
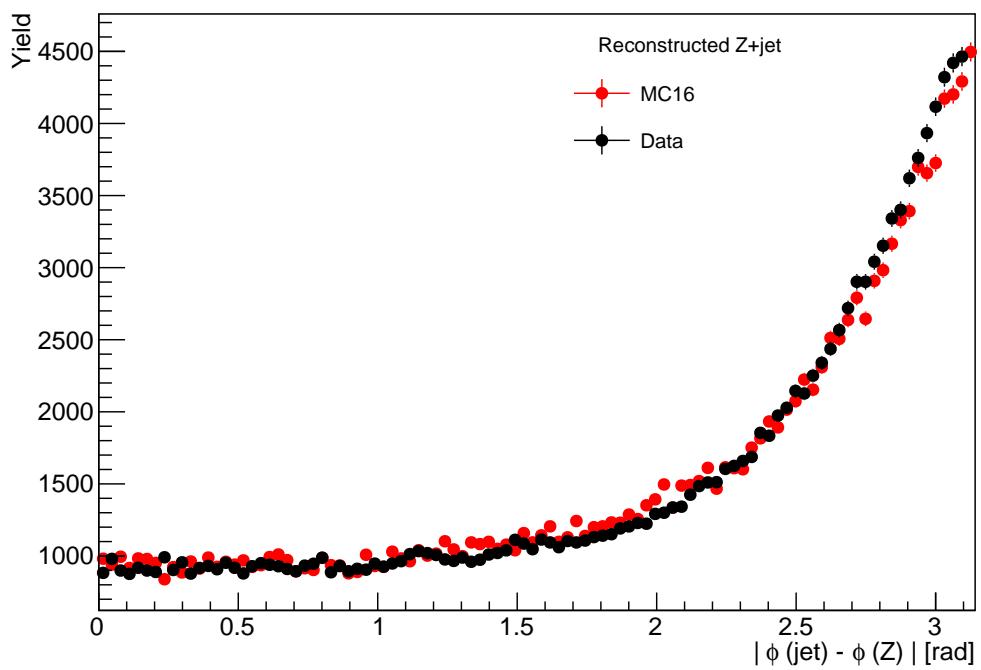


Figure 8: The Z boson rapidity distribution in data (black) and simulation (red).



190 4 Jet Reconstruction Performance and Efficiencies

191 Jets are reconstructed at LHCb using the anti- kt clustering algorithm [18] with $R = 0.5$.
 192 The inputs to the clustering algorithm are the four-momentum vectors of charged particle
 193 as well as neutral particle objects selected by the particle flow algorithm [19]. Candidate
 194 jets also have to pass JetID cuts that are imposed to reduce fake jet rates. The jet energy
 195 correction [20] provided by the [Tools for Jet Reconstruction Working Group](#) apply
 196 a correction factor k that is binned in the clustered jet p_T , jet η , jet ϕ , the fraction of
 197 the jet energy carried by charged particles and the number of primary vertices in the
 198 event. The correction factor k is determined in such a way that $p_T(\text{reco}) = p_T(\text{truth})$ in
 199 simulation. This section will present some of the results of performance evaluation.

200 4.1 Jet Directional Resolution

201 A subset of $Z + \text{jet}$ events where jet $p_T > 10 \text{ GeV}/c$ and $2.0 < \eta(\text{jet}) < 4.5$, $\Delta\phi(\text{jet}-\mu) >$
 202 0.4 and $\Delta\phi(\text{jet}-Z) > \frac{7}{8}\pi$ is used for jet performance evaluation. The truth jets were
 203 matched to the hardest p_T reconstructed jets within a radius of 0.5 in $\eta - \phi$ space. The
 204 directional resolution is defined as the distance in $\eta - \phi$ space between truth jet and
 205 reconstructed jet that are matched. Figure 11 shows that 93.3% of reconstructed jets fall
 206 within ΔR of 0.1 .

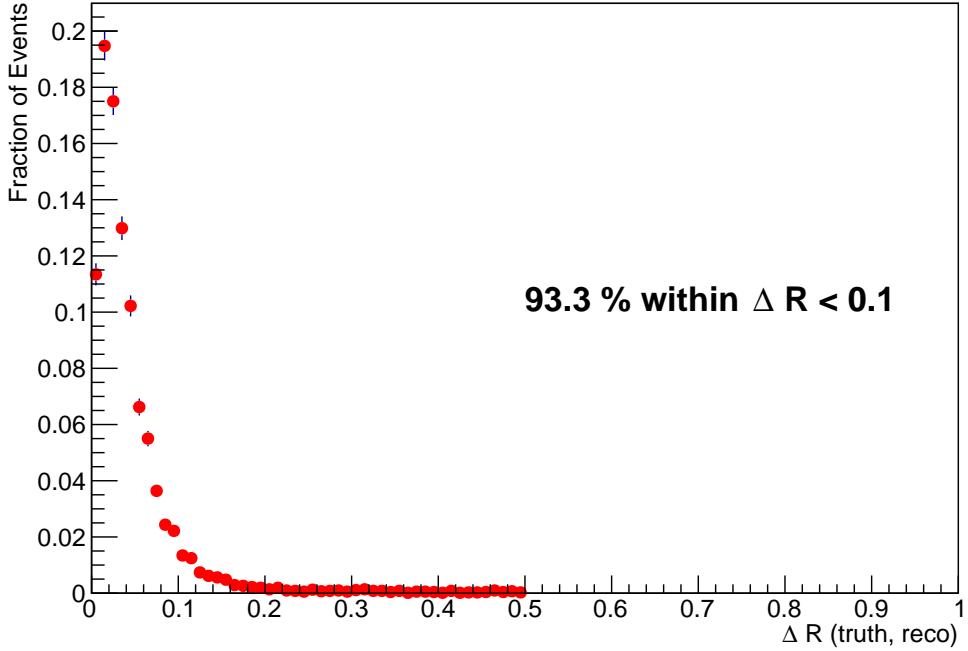


Figure 11: The direction resolution of reconstructed jets in simulation.

207 4.2 Jet p_T Resolution

208 The transverse momentum resolution of reconstructed jets is estimated by looking at
 209 the distribution of $p_T(\text{reco})/p_T(\text{truth})$ as shown in Fig.12. The peak is clearly seen at 1 ,

indicating the jet reconstruction performs well. The resolution, when defined as the half width at half maximum, is overall $\sim 15\%$ while there is slight dependence on jet p_T , η , and the number of primary vertices in the event.

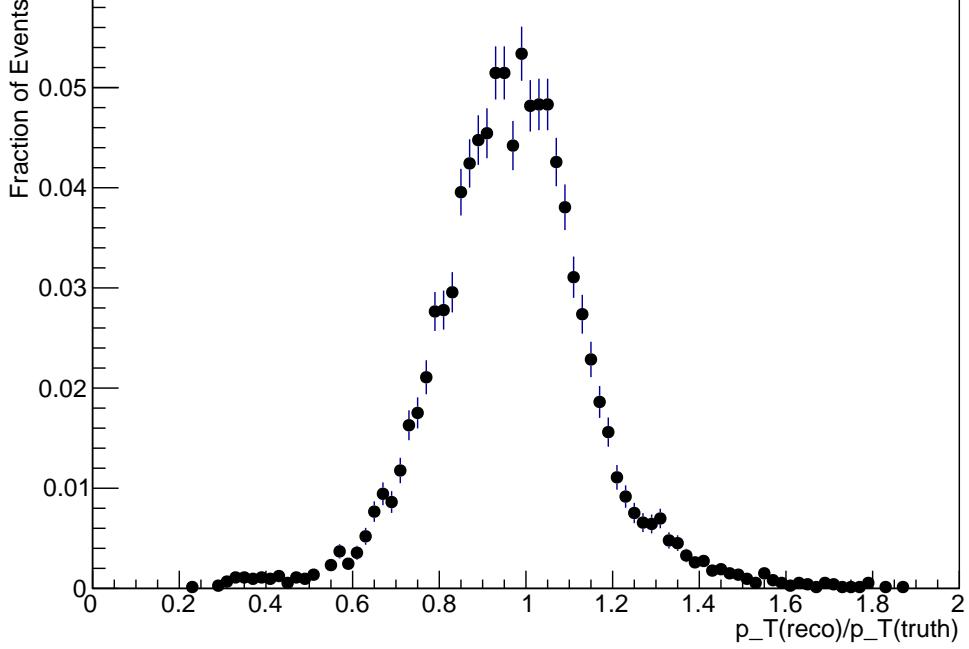


Figure 12: The jet transverse momentum resolution in simulation.

4.3 Jet Reconstruction Efficiency

While the unfolding corrections accounting for jet p_T migrations due to finite resolution effects and imperfect reconstruction of jets will be discussed in a dedicated section 9.3, we determine in this section the jet reconstruction efficiencies in selecting the events we consider in the unfolding procedure. The selected events passed the standard cuts laid out for this analysis in Sec. 3.1, including the minimum p_T threshold of 15 GeV/c and $2.5 < \eta < 4.0$. The efficiency is defined as the probability that given a true jet in the said acceptance, we reconstruct a jet that passes JetID cuts in addition to the standard cuts. The reconstruction efficiencies for jets of truth $p_T > 20$ GeV/c are found as a function of truth jet p_T and shown in Fig. 13.

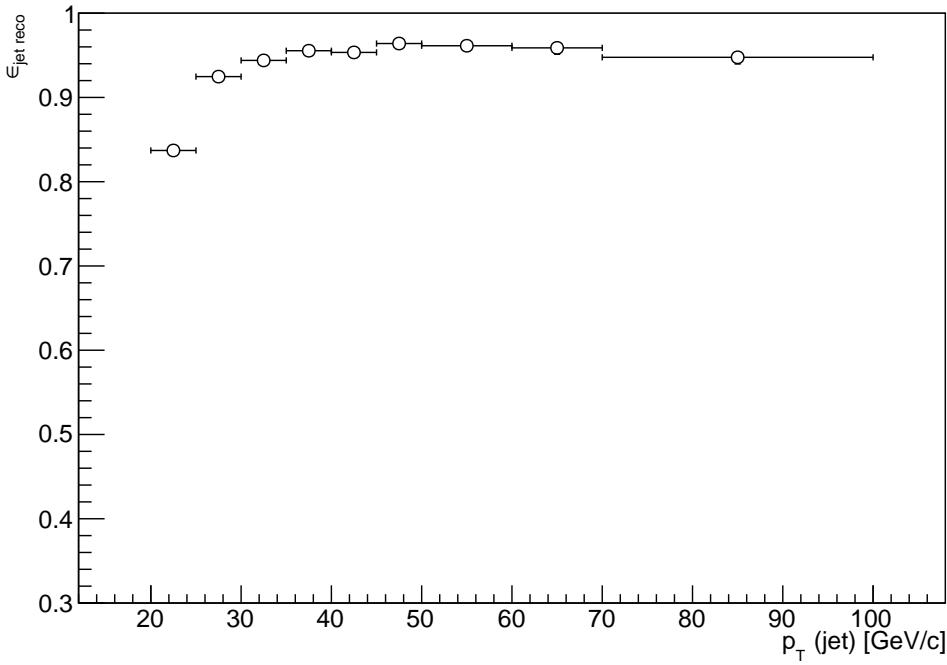


Figure 13: The jet reconstruction efficiency as a function of $p_T(jet)$, from simulation.

223 5 Muon Reconstruction Efficiencies

224 In this analysis, muon efficiencies play a role in the normalization, *i.e.* the number of
 225 $Z+jet$ events, as well as the distributions of charged hadrons. One may consider these
 226 efficiencies canceled out between the jet distribution and charged hadron distribution
 227 because the muon efficiencies are determined independently from the jet efficiencies.
 228 However, even within a given jet p_T bin, there may exist dependence on the p_T of Z in
 229 the differential cross section with respect to hadronization variables due to correlations
 230 between the jet p_T and Z p_T . Therefore, these corrections were done separately for the
 231 distributions of jets and charged hadrons.

232
 233 Methods for the determination of reconstruction efficiencies relevant to high p_T muons
 234 are well established and documented for $W^\pm \rightarrow \mu^\pm \nu$ and $Z \rightarrow \mu^+ \mu^-$ decay measurements.
 235 These studies showed that there are three components contributing to total single muon
 236 reconstruction efficiency, namely, muon trigger efficiency ϵ_μ^{trig} , muon identification efficiency
 237 ϵ_μ^{ID} and muon track reconstruction efficiency ϵ_μ^{track} . The total single muon reconstruction
 238 efficiency is then expressed as the product of these individual components as shown in
 239 Eq. 5. All of these efficiency components depend on the kinematics of the decay muons.
 240 Note that muon candidates in this section are required to be $2.0 < \eta < 4.5$ and $p_T >$
 241 20 GeV/ c .

$$\epsilon_{Z\mu\mu} = \epsilon_{Z\mu\mu}^{trig} \cdot \epsilon_{Z\mu\mu}^{ID} \cdot \epsilon_{Z\mu\mu}^{track} \quad (5)$$

242 In order to measure reconstruction efficiencies, Z to $\mu\mu$ events in data and MC were
 243 used. The data sample sizes for muon trigger, muon identification and muon tracking

²⁴⁴ efficiencies are 25K, 82K and 205K events, respectively. A large set of inclusive Z to $\mu\mu$
²⁴⁵ MC samples analyzed in Ref. [21] was used for comparisons with data. The statistical
²⁴⁶ uncertainties on these measurements made using the MC samples are less than 0.1% in
²⁴⁷ all bins.

²⁴⁸ 5.1 Trigger efficiencies

²⁴⁹ For the Z boson measurement either muon can be responsible for the event passing
²⁵⁰ the trigger stage, so the overall trigger efficiency for the Z boson candidate is determined
²⁵¹ as unity minus the probability of neither decay muon being detected,

$$\epsilon_{Z\mu^+\mu^-}^{trig} = 1 - (1 - \epsilon_{\mu^+}^{trig})(1 - \epsilon_{\mu^-}^{trig}), \quad (6)$$

²⁵² where the single muon trigger efficiency can depend on the kinematics of each decay muon.
²⁵³ The overall Z boson efficiency is then the product of this combined trigger efficiency with
²⁵⁴ the tracking and ID efficiencies of each decay muon.

Table 2: The tag-and-probe conditions for the determination of ϵ_μ^{trig}

Tag	Probe	Event
Long track isMuon	Long track isMuon	$60 < M_{\mu\mu} < 120 \text{ GeV}/c^2$ $ \Delta\phi_{\mu\mu} > 0.1 \text{ rad}$
Triggered-on-signal (TOS) $\sigma_p/\ p\ < 0.1$ Track Prob(χ^2) > 0.1 %	$\sigma_p/\ p\ < 0.1$ Track Prob(χ^2) > 0.1 %	

²⁵⁵ The tag-and-probe selections used in the tag-and-probe method are tabulated in
²⁵⁶ Table 2. The single muon trigger efficiencies applied to this analysis were determined
²⁵⁷ using a data with the tag-and-probe method as a function of kinematic variables. The
²⁵⁸ results are compared with the ones in simulation in Fig. 14. The jumps and discontinuities
²⁵⁹ seen in the efficiencies are caused by detector geometry effects, e.g. gaps and support
²⁶⁰ structure. The largest inefficiency for the trigger is known to come from the L0 hardware
²⁶¹ level trigger. The new L0 trigger (L0MuonEW) implemented in Run 2, which avoided
²⁶² the SPD multiplicity cuts of L0Muon, required a tighter pT cut. As the pT resolution
²⁶³ of L0 is quite poor and hardware based, the L0 trigger efficiency is not expected to be
²⁶⁴ well modelled in the MC, therefore a large difference seen in trigger efficiencies. However,
²⁶⁵ this does not affect our analysis as we determine the trigger efficiencies using data-driven
²⁶⁶ method. For the determination of systematic uncertainties attributed to the tag-and-probe
²⁶⁷ method, we refer to the analysis note on the high p_T muon efficiencies in Run 2 [21].

²⁶⁸ 5.2 Muon Identification efficiencies

²⁶⁹ The muon identification efficiency ϵ_μ^{ID} is defined as the efficiency that a decay muon
²⁷⁰ passes **isMuon** criteria. The efficiencies for Z can be obtained by multiplying the efficiencies
²⁷¹ of the two opposite-charged muons. In this study, an inclusive $Z \rightarrow \mu^+\mu^-$ event sample
²⁷² was used to determine the single muon identification efficiency on a track-by-track basis

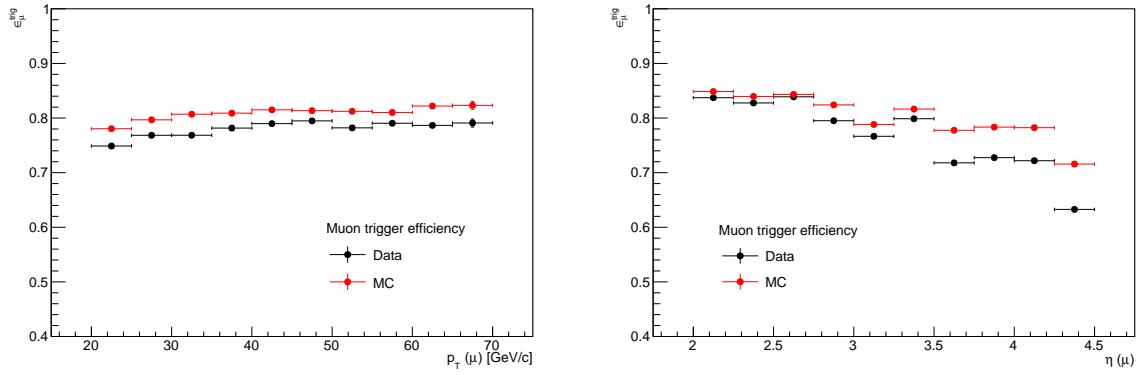


Figure 14: The single muon trigger efficiencies in data (black) and simulation (red).

with the tag-and-probe method. The tag must be an identified and triggered muon that passes track quality cuts imposed for this analysis, and the probe is a long track with the same track quality requirements. The tag-and-probe selections are summarized in Table 3. The probe selection had to be tighter in order to reduce background contamination coming from heavy flavor decay muons. The pseudorapidity and p_T dependent single muon identification efficiencies determined with these selections are shown in Fig. 15. A dip in the muon identification efficiency at the largest η bin comes from the effects of tracks passing the edge of the RICH detector.

Table 3: The tag-and-probe selections for the determination of ϵ_μ^{ID}

Tag	Probe	Event
Long track isMuon $\text{ProbNN}\mu > 0.68$ TOS $\sigma_p/\ p\ < 0.1$ Track Prob(χ^2) > 0.1 % IP $\chi^2 < 1$.	Long track $\sigma_p/\ p\ < 0.1$ Track Prob(χ^2) > 0.1 %	$80 < M_{\mu\mu} < 100 \text{ GeV}/c^2$ $ \Delta\phi_{\mu\mu} > 0.1 \text{ rad}$ PV $\chi^2/\text{d.o.f} < 3$

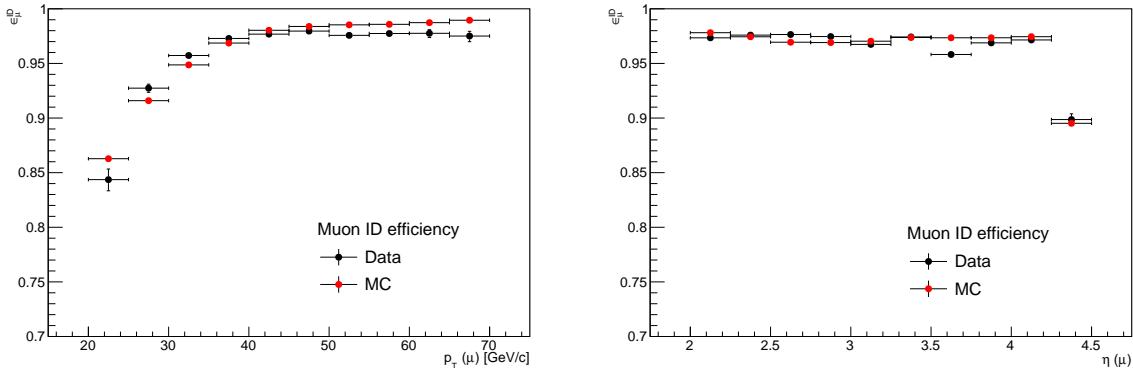


Figure 15: The single muon identification efficiencies in data (black) and simulation (red).

281 5.3 Muon track reconstruction efficiencies

282 For the determination of muon track reconstruction efficiency, a tag-and-probe method
 283 called the muonTT method [22] was chosen. In this method, a track reconstructed
 284 by combining hits in the muon stations and the Trigger Tracker (TT), known as a
 285 muonTT track, is selected as a probe. A tag muon track must fulfill the TOS and `isMuon`
 286 requirements. A muonTT track is matched to a long track if they share greater than 40% of
 287 muon hits and 60% of TT hits, and the invariant mass of the tag and the probe track has to
 288 be larger than $40 \text{ GeV}/c^2$ [23]. More detailed studies on systematic uncertainties attributed
 289 to this method are available in Ref. [21]. As stated at the beginning of this section, only
 290 relative p_T dependence affects this analysis and the absolute scale of efficiencies has no
 291 effects. The selection criteria for the tag and probe tracks are summarized in Table 4.
 292 The results are shown in Fig. 16. The final efficiency $\epsilon_{Z\mu\mu}$ is obtained by multiplying the
 293 efficiencies of the two opposite-charged muons.

Table 4: The tag and probe selections for the determination of ϵ_μ^{track}

Tag	Probe	Event
Long track	muonTT track	$80 < M_{\mu\mu} < 100 \text{ GeV}/c^2$
<code>isMuon</code>		$ \Delta\phi_{\mu\mu} > 0.1 \text{ rad}$
TOS		$\text{PV } \chi^2/\text{d.o.f} < 3$
$\sigma_p/\ p\ < 0.1$		
Track Prob(χ^2) > 0.1 %		

294 Finally, the Z reconstruction efficiency from taking into account all three efficiencies
 295 is shown as a function of Z pt in Fig. 17. The Z efficiencies at a given Z pt are a range
 296 of values as they are determined by pairs of muons with different p_T .

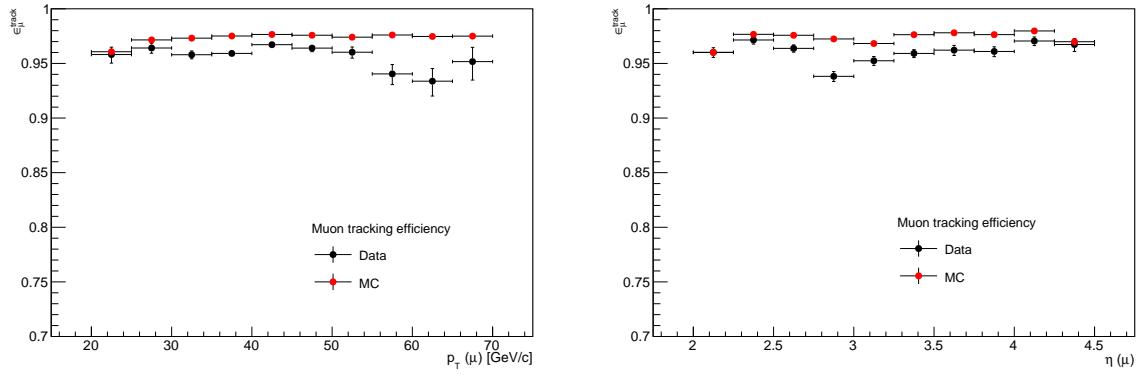


Figure 16: The single muon track reconstruction efficiencies in data (black) and simulation (red).

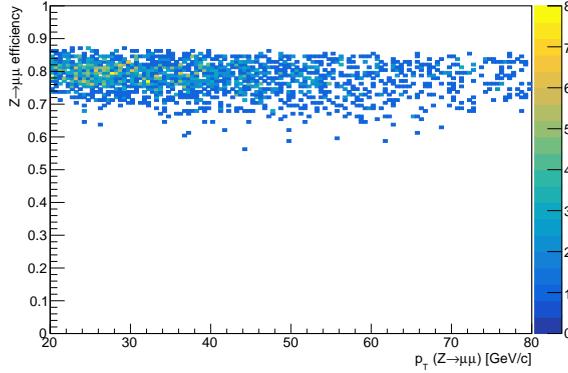


Figure 17: The Z reconstruction efficiency.

297 6 Hadron reconstruction efficiency and purity within 298 jets

299 Hadron-level corrections break down into a few components including (a) the
300 in-reco-jet efficiency, (b) the track reconstruction efficiency, (c) the in-jet-radius efficiency
301 and (d) the track matching efficiency. Additionally (e) the hadron-in-jet-purity is also
302 considered. Dedicated discussions on particle identification corrections follow in the next
303 section. The corrections are calculated as in Eq. 7,
304

$$c_{\text{hadron}} = \frac{\epsilon_{\text{h-efficiency}}}{C_{\text{h-in-jet-purity}}}, \quad (7)$$

305 where $\epsilon_{\text{h-efficiency}} = \epsilon_{\text{in-reco-jet}} \cdot \epsilon_{\text{track-reco}} \cdot \epsilon_{\text{in-jet-radius}} \cdot \epsilon_{\text{track-matching}}$.

306
307
308 The starting point or denominator of $\epsilon_{\text{h-efficiency}}$ requires a hadron from a particle
309 level/truth jet (not necessarily need to be reconstructed). The sub-factors progressively
310 consider whether the truth particle is inside a reconstructed & matched jet ($\epsilon_{\text{in-reco-jet}}$),

311 the track is successfully reconstructed ($\epsilon_{track-reco}$) and inside the jet radius ($\epsilon_{in-jet-radius}$),
 312 and has a reconstructed track that is matched to the truth particle ($\epsilon_{track-matching}$). Lastly,
 313 additional tracks gained in a given jet by fake tracks or mis-association are corrected for
 314 in $c_{h-in-jet-purity}$, which is defined as the ratio of 'reconstructed tracks that are matched
 315 to truth particles that belong to the truth jet' to 'any reconstructed tracks identified as
 316 charged hadrons that belong to the reconstructed jet'.

317 A jet containing a proton with a moderately high momentum is less likely to be
 318 successfully reconstructed than one containing a pion, as shown in Fig. 18. One of the
 319 reasons for this is understood by considering track multiplicities per jet between different
 320 particle species Fig. 19, mainly accounting for the difference between pions vs. Kaons.
 321 Another cause for the distinctive difference between Kaons vs. protons is the differences in
 322 track reconstruction efficiency. These differences are accounted for in $\epsilon_{in-reco-jet}$, defined as
 323 the probability of a hadron track belonging to a successfully reconstructed jet at the truth
 324 level. Whether tracks were reconstructed with a momentum within reasonable resolution
 325 is not taken into consideration in this definition.

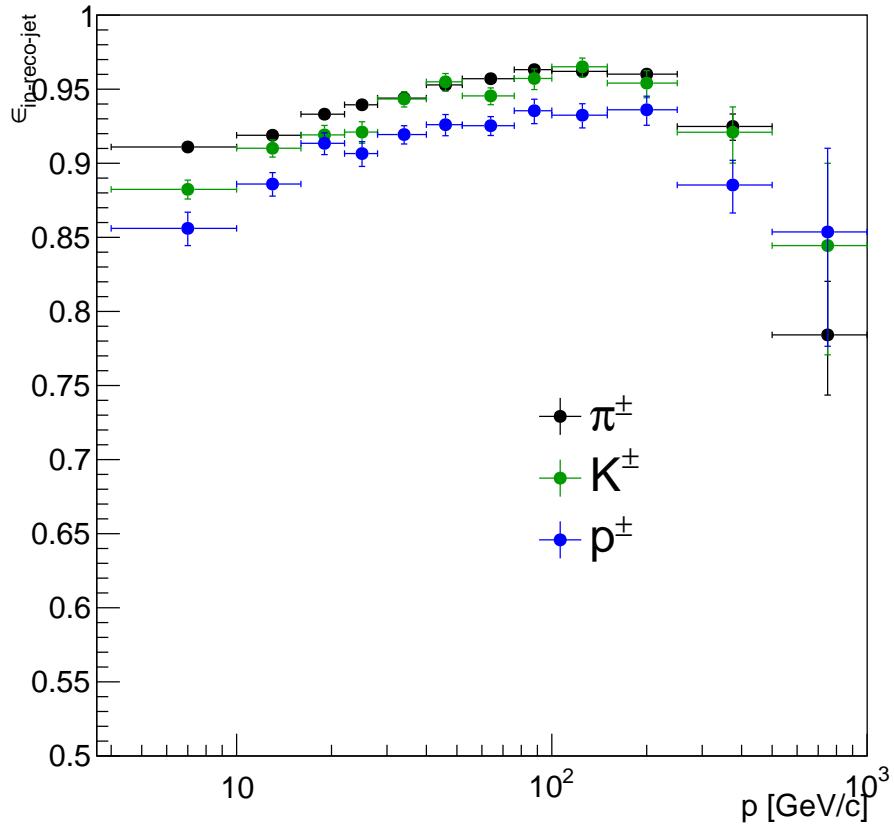


Figure 18: $\epsilon_{in-reco-jet}$.

326 The efficiency of a track being reconstructed and passing track cuts is accounted for
 327 in the track reconstruction efficiency $\epsilon_{track-reco}$. A $\sim 20\%$ nuclear interaction length in
 328 detector materials at LHCb substantially impacts charged hadron tracking efficiencies,
 329 and there is an additional factor of a decay probability, notably at 8% for K^\pm . The track

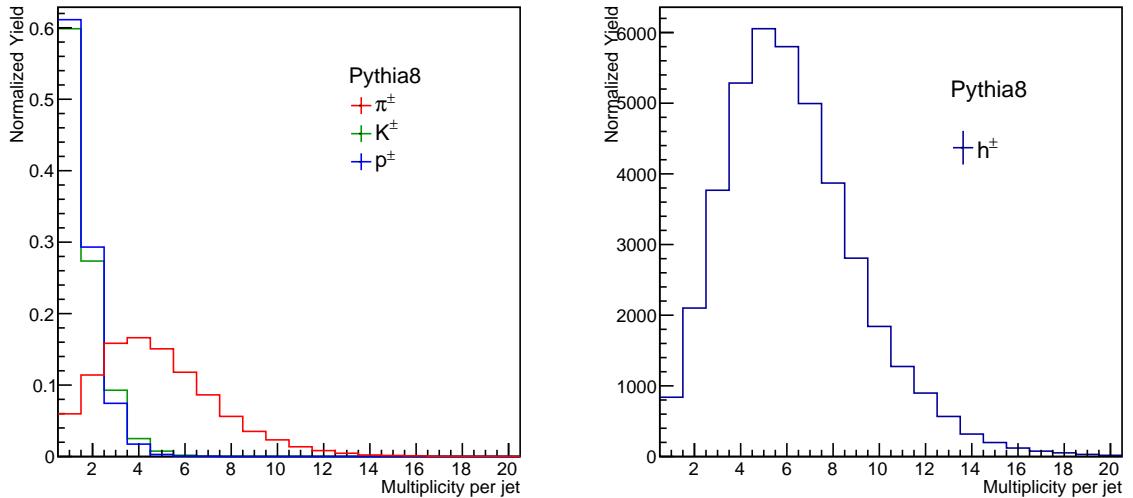


Figure 19: Identified charged hadron multiplicity per jet.

reconstruction efficiencies averaged over the entire momentum range for π^\pm , K^\pm and p^\pm are 79%, 77% and 63%, respectively. Charge dependence seen in protons vs. antiprotons is due to different interaction probabilities in the detector material.

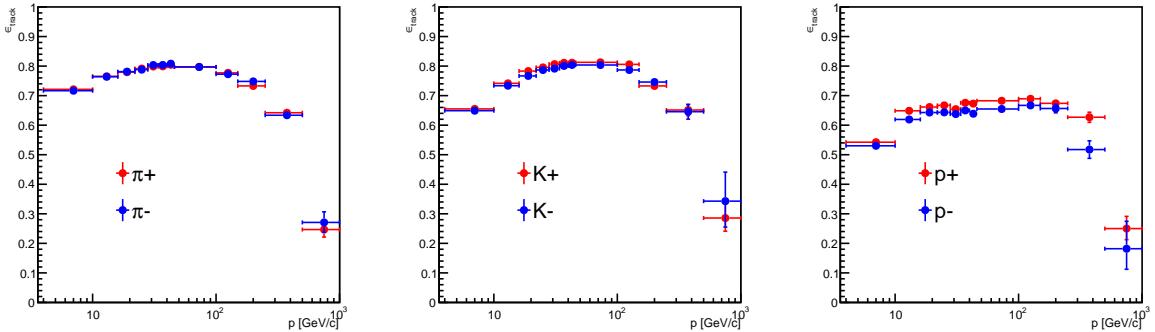


Figure 20: The hadron track reconstruction efficiency $\epsilon_{\text{track-reco}}$ as a function of momentum for the three different hadron species.

The in-jet-radius efficiency $\epsilon_{\text{in-jet-radius}}$, defined as the probability of a track that belongs to a reconstructed jet at the truth level actually being within the jet radius from the reconstructed jet axis, is largely determined by the directional resolution of jets. Thanks to the excellent directional resolution of reconstructed jets, this efficiency is fairly high, above 99.8% throughout the kinematic phase space in momentum and pseudorapidity. The track-matching efficiency $\epsilon_{\text{track-matching}}$ is generally higher than 98%. Combined hadron efficiencies ϵ_{hadron} are obtained for each charge and particle species separately and shown in Fig. 21.

Lastly, the hadron-in-jet-purity $c_{\text{h-in-jet-purity}}$ is shown in Fig. 22. The hadron purity is higher than 95% in most regions in $p\text{-}\eta$ space populated with abundant charged particles.

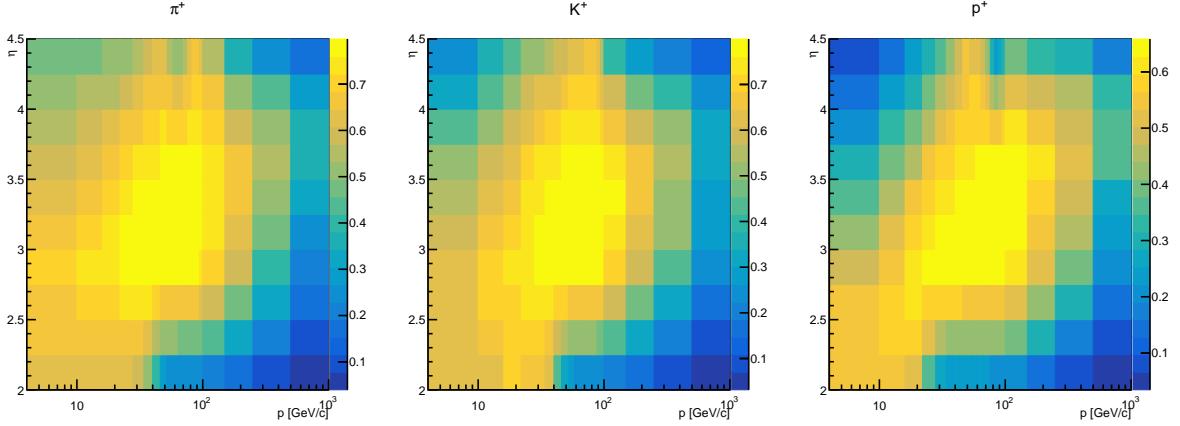


Figure 21: The total hadron efficiency $\epsilon_{h\text{-efficiency}}$ as a function of momentum and pseudorapidity for three different hadron species.

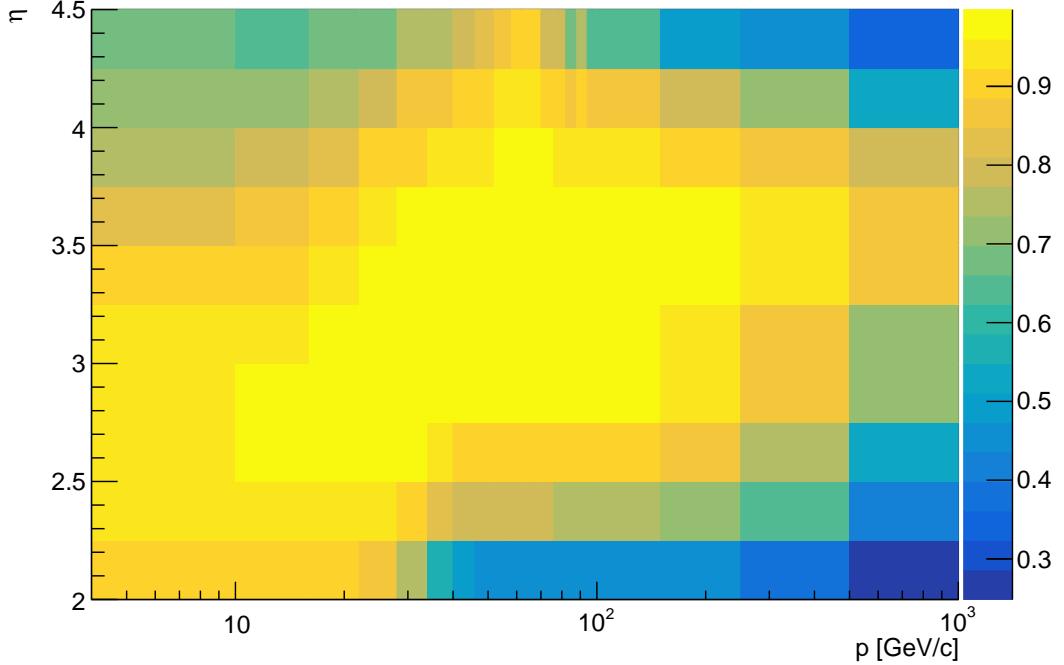


Figure 22: The purity of tracks inside jets $c_{h\text{-in-jet-purity}}$ that accounts for tracks that do not belong to the truth jet matched to the reconstructed jet.

343 7 Particle Identification Corrections

344 Identification of charged hadrons is done utilizing signatures in various detector
 345 subsystems such as the RICH, CALO and Muon chambers. Two separate RICH detectors
 346 play a crucial role in $\pi^\pm/K^\pm/p^\pm$ identification, with RICH-1 (C_4F_{10} radiator) covering
 347 low p (2-60 GeV/c) and RICH-2 (CF_4) covering the high p region (15-100 GeV/c). Note
 348 that kaons and protons gradually run out of statistics above a momentum of 100 GeV/c in
 349 this analysis, leaving π^\pm as the predominant species in the momentum range between 100

350 and 1000 GeV/ c as can be seen in Fig. 23. In addition to EM and Hadronic calorimeters,
 351 the SPD and Preshower subdetectors make up the CALO part of the PID detectors.

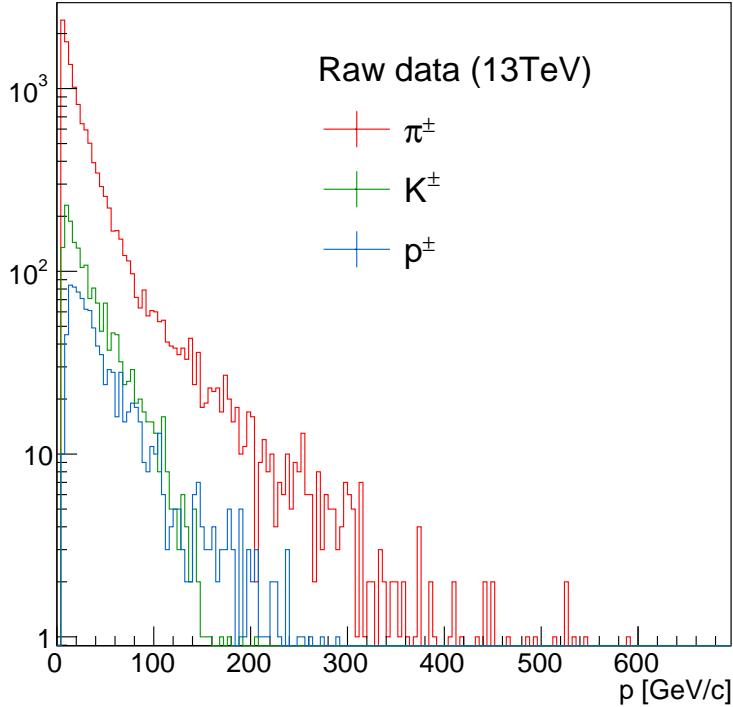


Figure 23: Momentum distributions for π^\pm , K^\pm and p^\pm .

352 Global PID variables are constructed from the PID information obtained from the
 353 MUON, RICH and CALO systems. The simplest approach adopted initially was to take
 354 the difference in combined log-likelihood (**DLL**) relative to the pion hypothesis, where
 355 the combined log-likelihood is produced by simply multiplying the likelihoods given by
 356 individual systems ($L = L_{RICH} \cdot L_{CALO} \cdot L_{MUON}$). A more sophisticated approach developed
 357 using multivariate techniques builds a new type of global PID variables called **probNN**
 358 utilizing additionally tracking information. The latter method provides probability
 359 distributions for kaons, pions, protons, muons and electrons. In this analysis, both
 360 types of variables were used to optimize identification efficiencies of $\pi^\pm/K^\pm/p^\pm$. Particle
 361 identification efficiencies were obtained using the PIDCalib package [24]. For Run 2
 362 analyses, dedicated calibration samples [25] had been obtained to be used for both offline
 363 and online analysis. Brunel variables with the MC15TuneV1 with the default calibration
 364 sample available in Turbo16 were used in this analysis and the cuts used for charged
 365 hadron identification are summarized in Table 5. All charged tracks were required to
 366 be identified as one of the particles $e/\pi/K/p/\mu$ in default PID given in Davinci. For
 367 the identification of charged pions, inequality cuts have been used in order to remove
 368 particles with high ProbNNe (high probNNmu) while retaining high pion statistics. The
 369 2D distributions of probNNpi-probNNe and probNNpi-probNNmu are shown in Fig. 24
 370 that visualize populated regions in the 2D probNN space.

371 The PIDCalib package provides efficiencies as a function of momentum, eta and track

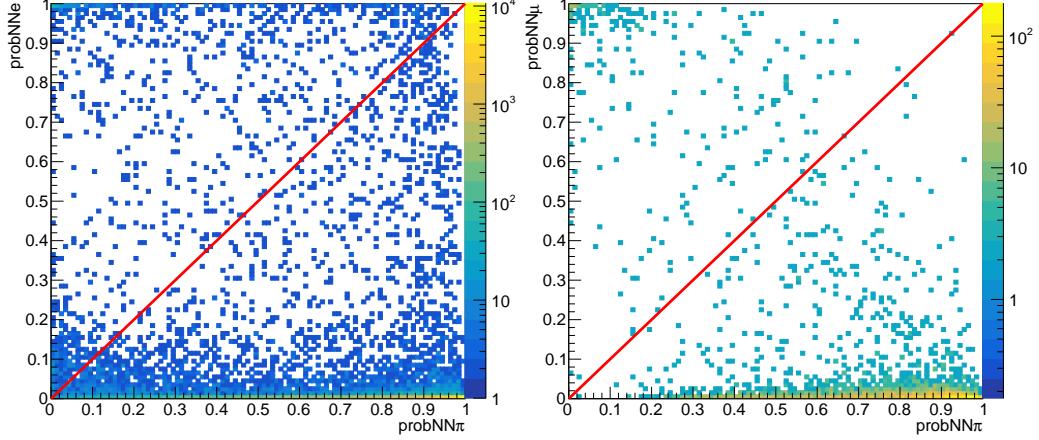


Figure 24: The 2D distributions of (left) probNNpi-probNNne and (right) probNNpi-probNNmu. The red diagonal line represent the cuts applied.

Table 5: Summary of PID cuts

Particle	Cuts
π^\pm	ProbNNpi > ProbNNne & ProbNNpi > ProbNNmu & !K & !p [†]
K^\pm	DLLK > 1 & DLLK \geq DLLp & ProbNNK > ProbNNp & ProbNNK > ProbNNpi
p^\pm	DLLp > 2 & DLLp \geq DLLK & ProbNNp > ProbNNK & ProbNNp > ProbNNpi

[†] !K and !p indicate not being identified as K and p in custom definition.

372 multiplicity such that efficiencies can be re-weighted according to the distributions in data
 373 relevant for an analysis of interest. The weights for track multiplicity per event applied to
 374 data are shown in Fig. 25. The re-weighted efficiencies as a function of track momentum
 375 obtained from PIDCalb can be seen in Fig. 26.

376 Mis-identification of non-hadronic particles such as electrons or muons as any of
 377 the three charged hadron species is negligible, 0.9%, 0.01% 0.04% for pions, kaons and
 378 protons, respectively. The rate of long-lived hyperons (Ξ , Ω and Σ) being misidentified
 379 as $\pi^\pm/K_\pm/p^\pm$ is less than 0.01% below 100 GeV/c and increases with p upto 1.0% at
 380 500 GeV/c. The loss of charged hadrons due to decay in flight, etc. is accounted for in
 381 tracking efficiency corrections. The rate of pions being mis-identified as electrons was
 382 determined using PIDCalib to be less than 1%. A matrix of particle identification efficiency
 383 is constructed using efficiencies obtained from PIDCalib for a given track momentum bin.
 384 The momentum distributions of reconstructed π^\pm , K^\pm and p^\pm particles that pass all the
 385 track cuts were unfolded by solving an equation shown in Eq. 8.

$$\vec{x}_{reco} = A \vec{x}_{true} \quad (8)$$

$$\text{where } A = \begin{Bmatrix} \epsilon_{\pi \rightarrow \pi} & \epsilon_{K \rightarrow \pi} & \epsilon_{p \rightarrow \pi} \\ \epsilon_{\pi \rightarrow K} & \epsilon_{K \rightarrow K} & \epsilon_{p \rightarrow K} \\ \epsilon_{\pi \rightarrow p} & \epsilon_{K \rightarrow p} & \epsilon_{p \rightarrow p} \end{Bmatrix}, \vec{x}_{true} = \begin{Bmatrix} x_\pi^{true} \\ x_K^{true} \\ x_p^{true} \end{Bmatrix}, \text{ and } \vec{x}_{reco} = \begin{Bmatrix} x_\pi^{reco} \\ x_K^{reco} \\ x_p^{reco} \end{Bmatrix}.$$

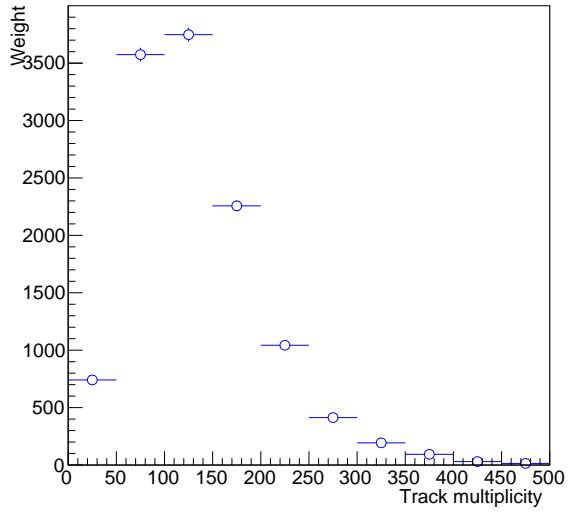


Figure 25: The track multiplicity per event with a jet containing π^\pm , K^\pm or p^\pm .

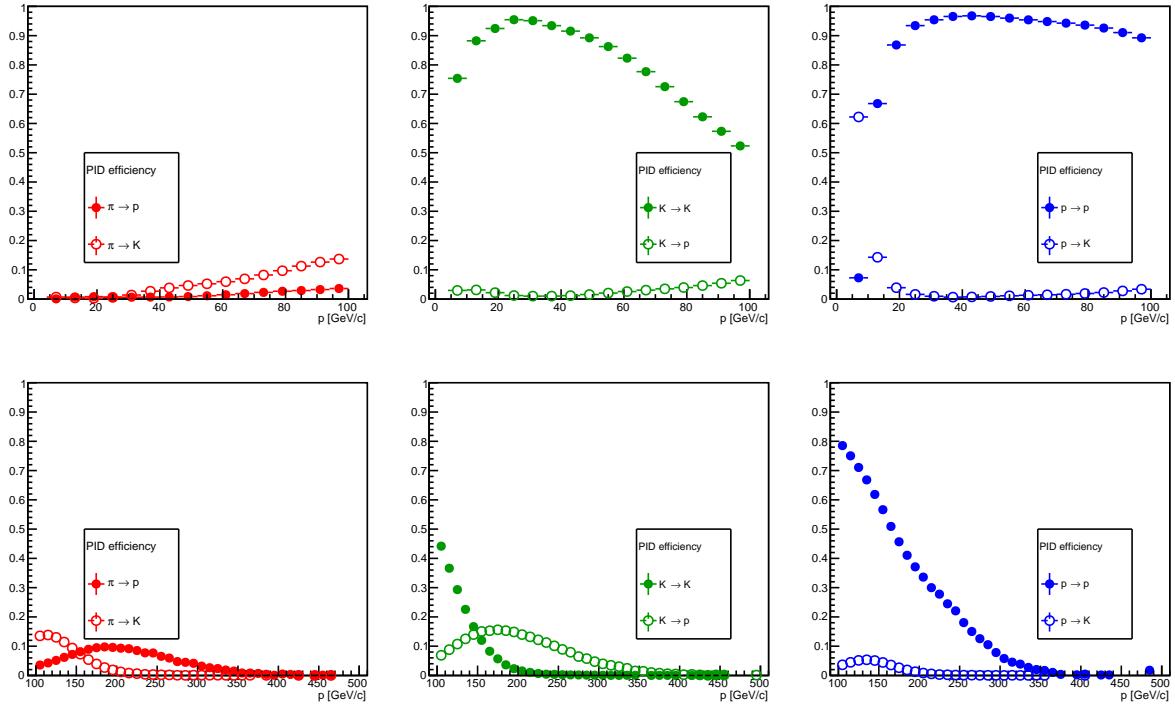


Figure 26: Particle identification efficiencies for π^\pm , K^\pm and p^\pm . (Top) $p < 100 \text{ GeV}/c$, (Bottom) $p > 100 \text{ GeV}/c$

386

387

388 The unfolding method introduced above has been validated by performing a closure
 389 test of which results are shown in Fig. 27. In addition to the corrections attributed to the
 390 unfolding of momentum distributions, the PID dependent track reconstruction efficiencies

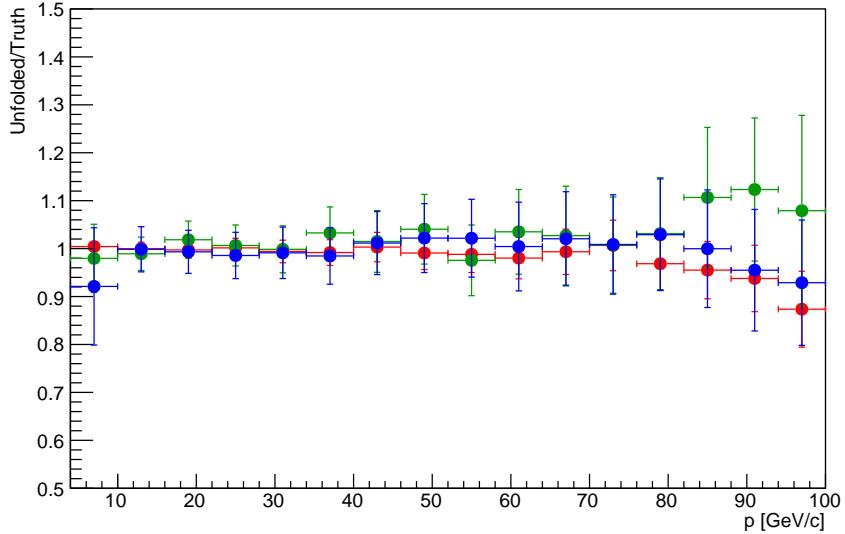
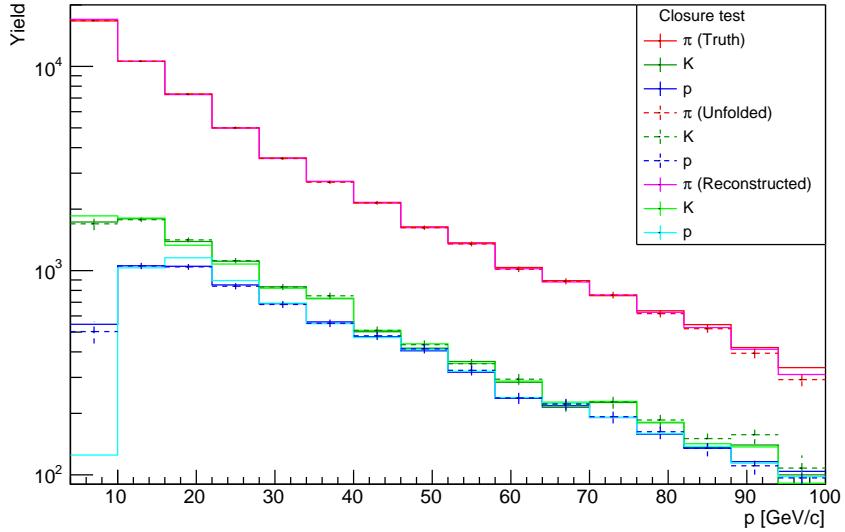


Figure 27: Closure test for unfolding of π^\pm , K^\pm and p^\pm momentum distributions.

391 obtained based on the truth PID of a particle, but assigned based on the reconstructed
 392 PID of a particle, have to be corrected. The correct efficiencies are obtained by multiplying
 393 the PID matrix by the efficiencies for truth PID as shown in Eq. 9. Consequently, the
 394 total PID correction factor C_{PID} is expressed in terms of the two components as in Eq. 10.

$$\vec{\epsilon}_{\text{reco}} = A \vec{\epsilon}_{\text{true}}. \quad (9)$$

$$C_{\text{PID}}^j(i) = C_{\text{PID-unfold}}^j(i) \cdot C_{\text{PID-efficiency}}^j(i) \quad (10)$$

395 where $C_{\text{PID-unfold}}^j(i) = \frac{(\vec{x}_{\text{true}})^j(i)}{(\vec{x}_{\text{reco}})^j(i)}$ and $C_{\text{PID-efficiency}}^j(i) = \frac{(\vec{\epsilon}_{\text{true}})^j(i)}{(\vec{\epsilon}_{\text{reco}})^j(i)}$

396 The index i runs over the momentum bin as PID unfolding is performed momentum-bin
 397 by momentum-bin. The division of the two vectors is done element-wise, *i.e.* the index j
 398 runs over particle species, pions, Kaons or protons.

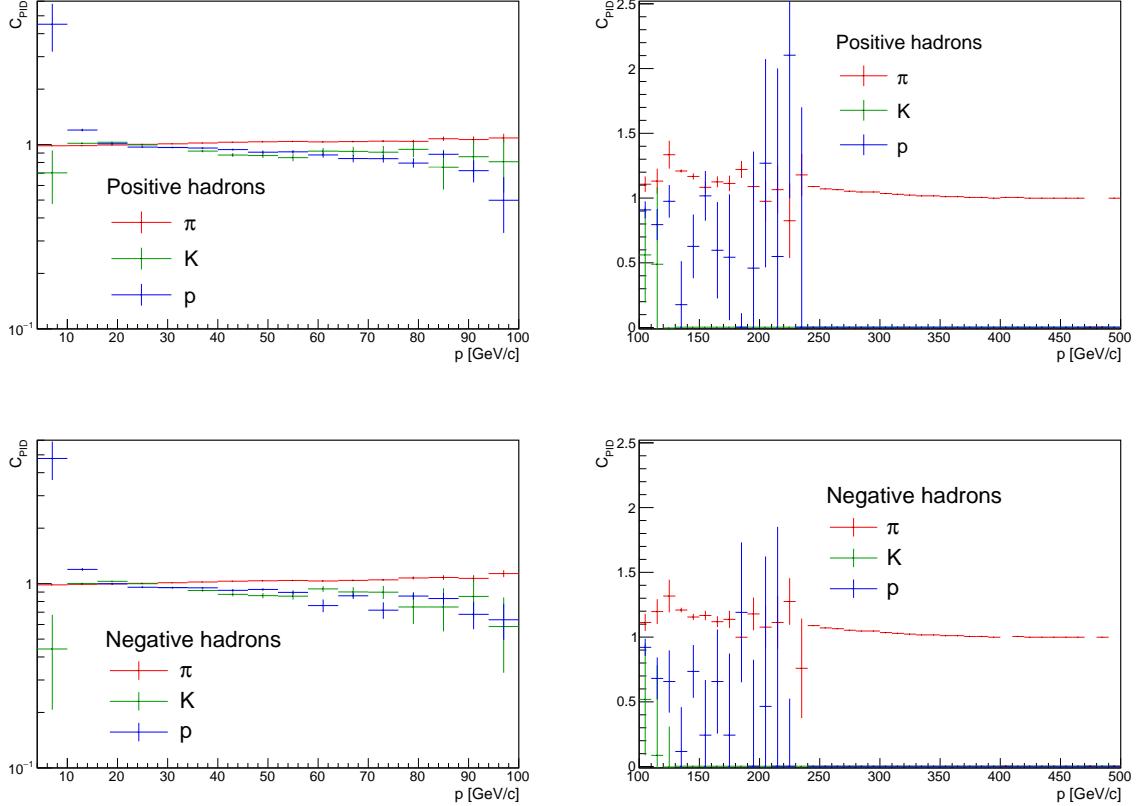


Figure 28: Total correction factors for π^\pm , K^\pm and p^\pm as a function of momentum. Positive hadrons are shown on the top and negative hadrons on the bottom. (Left) $p < 100$ GeV/ c , (Right) $p > 100$ GeV/ c

399 The total correction factors for each particle species and charge are shown in Fig. 28.
 400 Uncertainties displayed on data points are dominated by statistical uncertainties on the
 401 particle yields. They were estimated using a bootstrap method, where a new particle
 402 identification matrix equation has been reconstructed repeatedly for a large set of different
 403 yields generated from the sampling distributions.

404 8 Unfolding

405 Detector resolution effects that cause smearing in hadron distributions for a given
 406 jet p_T bin and jet p_T spectrum were corrected by unfolding procedures. Corrections for
 407 the latter, denoted as U_{jet} , are purely at the jet reconstruction level and agnostic about
 408 hadrons inside, whereas the former, denoted as $U_{\text{jet-hadron}}$, is more involved because it is
 409 affected by an interplay between the resolution effects on hadron reconstruction and jet
 410 reconstruction. While the resolution of charged hadron momentum is better than one
 411 percent up to $\sim 1000 \text{ GeV}/c$, the hadronization variables that are constructed in terms
 412 of jet momentum are susceptible to the relatively larger resolution of reconstructed jet
 413 momentum. The two components combined make up the total unfolding corrections
 414 U_{unfold} as in Eq. 11

415

$$U_{\text{unfold}} = U_{\text{jet}} \cdot U_{\text{jet-hadron}} \quad (11)$$

416

417 Two different methods were tried for this analysis: a binned iterative Bayesian unfolding
 418 approach and an unbinned Bayesian-based machine learning (ML) unfolding approach
 419 using external package in Ref. [26]. For final results, the latter method is adopted for
 420 reasons that will become obvious.

421

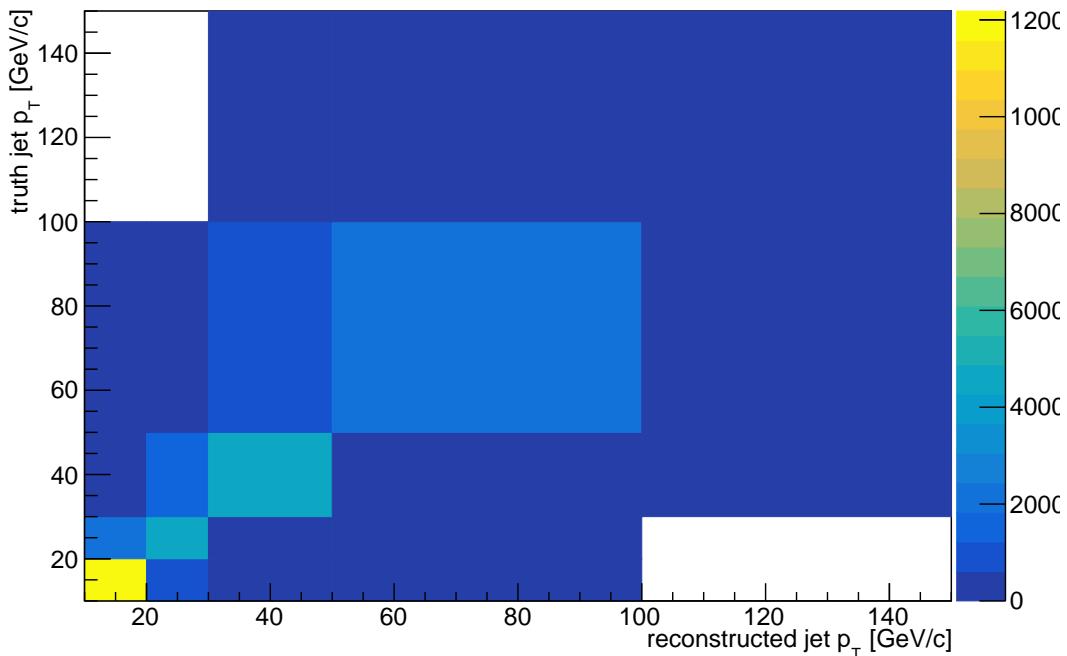


Figure 29: Response matrix for jet p_T .

422 In the binned approach, a response matrix is created that maps out the migration
 423 of true kinematic variables to reconstructed variables, whereas layers of deep neural
 424 networks (DNN) are trained to learn the migration pattern in the unbinned ML approach.

425 An example of the response matrix for the jet p_T distribution is shown in Fig. 29. This
 426 was used to perform the unfolding of the jet p_T distribution with an iterative Bayesian
 427 method [27] provided in the RooUnfold package [28]. In common, the range of jet p_T is
 428 chosen wider than the actual minimum p_T requirement in order to take into account the
 429 loss and gain of counts caused by smearing effects at edge bins. Best results are achieved
 430 by covering the entire p_T range of reconstructed jets that correspond to true jets with p_T
 431 above the minimum cut off.

432

433 The jet yields that serve as normalization were unfolded using the binned Bayesian
 434 method as well as the DNN method to compare. Common pseudodata were used
 435 for constructing a response matrix (binned Bayesian) and training networks (DNN).
 436 The hadronization variables normalized by the jet yield were unfolded only using the
 437 DNN method due to limited statistics for the multidimensional measurements. The
 438 reconstruction efficiencies (muon efficiencies and jet reconstruction efficiencies for jet yield
 439 normalization, and hadron reconstruction efficiency \times purity for charged hadrons) and
 440 PID corrections were applied prior to performing the unfolding procedure.

441

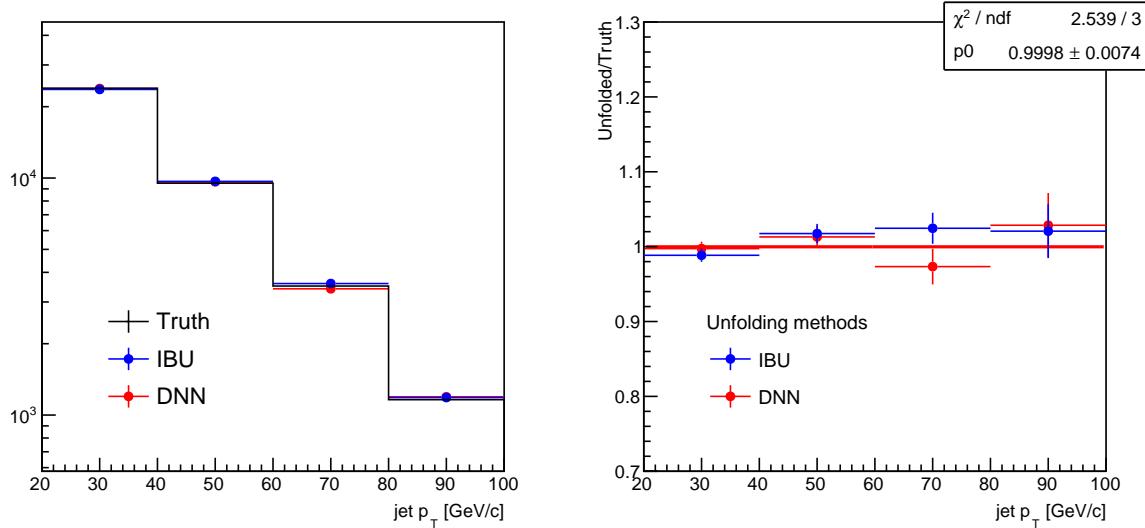


Figure 30: Comparisons of (left) unfolded jet p_T spectra between the iterative binned-Bayesian method (IBU) and the DNN method and (right) their ratios to the truth. The result of fitting a constant to the DNN-unfolded data is shown.

442 Figure 30 compares unfolded jet p_T distributions obtained from the two different
 443 methods. For multidimensional unfolding, jet p_T , jet η , hadron z and hadron j_T variables
 444 are simultaneously unfolded.

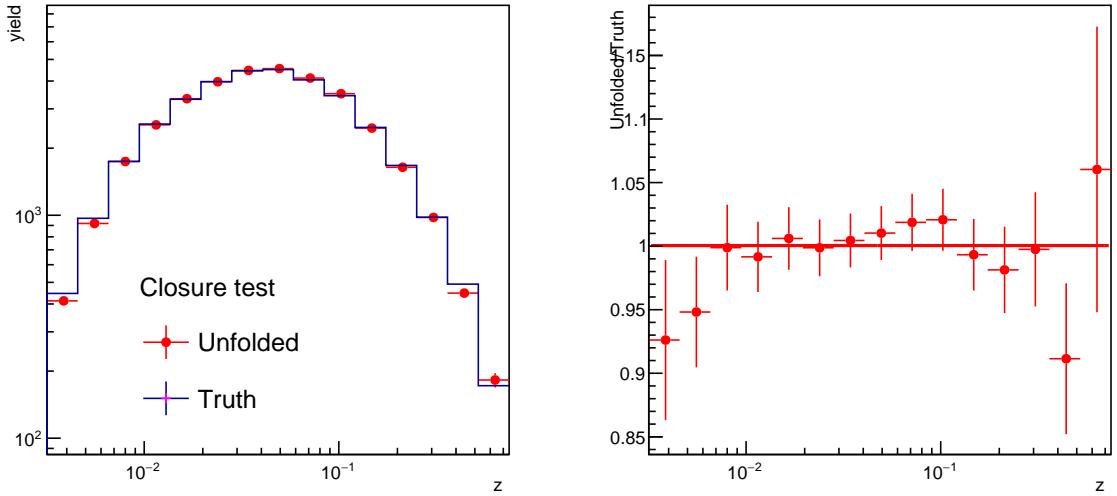


Figure 31: Unfolded and truth z distributions (left) and their ratio (right).

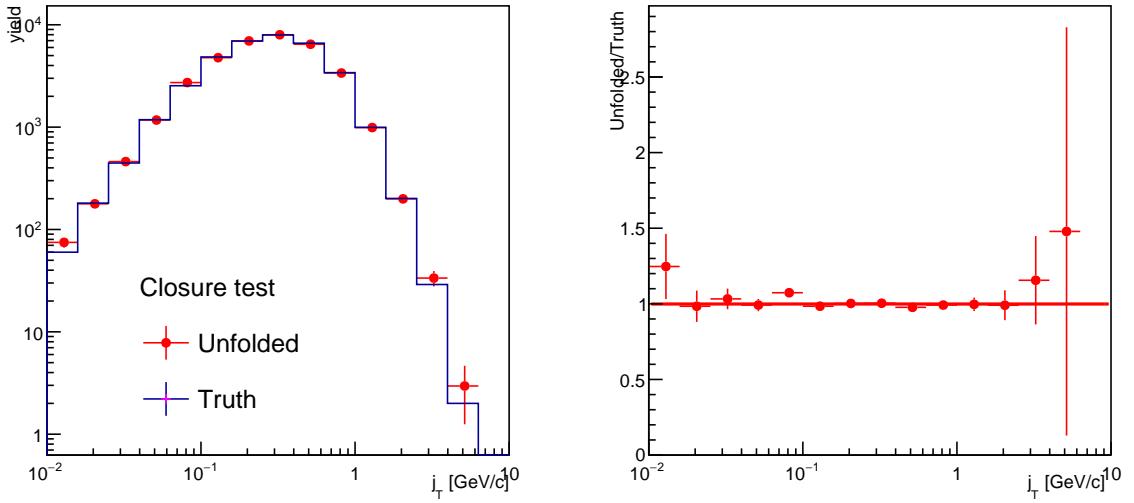


Figure 32: Unfolded and truth j_T distributions (left) and their ratio (right).

445 Figures 31 and 32 show results of closure tests for the two hadronization variables z
 446 and j_T projected to one dimension. Two independent MC data sets with opposite magnet
 447 polarity were used, one for training and the other for testing. When fitted to a constant,
 448 the ratio of unfolded to truth distributions is consistent with unity within 0.1%.

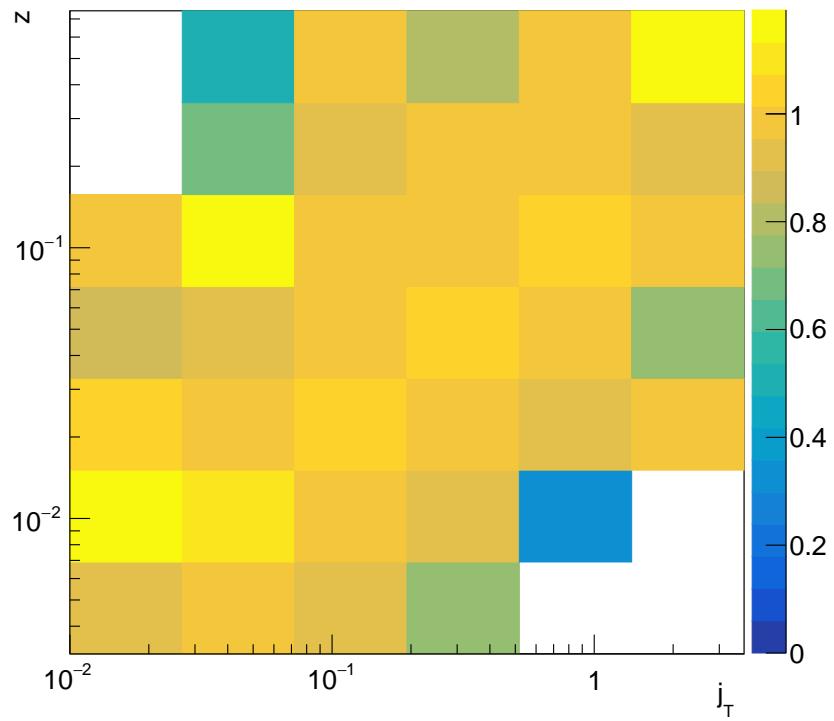


Figure 33: Unfolded to truth ratio of z - j_T distribution. j_T is shown in GeV/c .

449 Closure test results for the z - j_T distribution are shown in Fig. 33. While statistical
 450 fluctuations in less populated bins are present, the weighted average of the unfolded-to-
 451 truth ratio over all bins is consistent with unity within 0.5%.

452

453

454 9 Systematic Uncertainties

455 9.1 Normalization

456 Systematic uncertainties on the normalization, i.e. the number of jets associated with
457 a Z boson, come from those on the jet reconstruction efficiency studied in Sec. 4.3, muon
458 reconstruction efficiency documented in Sec. 5 and Ref. [29] and the purity of $Z +$ jet
459 sample.

460
461 Most of the uncertainties on the normalization are correlated between the denom-
462 inator and the numerator, and therefore are cancelled out. For instance, an overall
463 p_T -independent increase of jet reconstruction efficiencies or muon efficiencies by a certain
464 amount will result in a multiplicative correction factor for both the jet and hadron
465 distributions. Remaining sources of uncertainties that are not cancelled out include
466 background contributions from fake jets and physics background in $Z +$ jet selection and
467 any disagreement between data and models in MC.

468
469 The fake jet rate is determined in simulation by counting the number of reconstructed
470 jets that pass the selection criterion but do not have an associated truth jet matched to
471 them. Requiring only one primary vertex in an event results in very low fake jet rate of
472 0.2%. Since the selection criteria used to identify Z bosons are the same as Ref. [30],
473 the purity of the Z can be taken to be the conservative limit of the previous analysis as
474 98.6% (99.1 from Table 2 and additional 0.5 from Fig. 16). Therefore, the total purity
475 of the $Z +$ jet selection is $= 0.986 \times (1 - 0.002) = 0.984$. The total impurity of 1.6% is
476 assigned as a systematic uncertainty.

477
478 The reconstructed jets used in this analysis were required to pass the jet identification
479 cuts that aim to reduce fake jet rates. The requirements are 1) no one particle
480 contains more than 80% of the jet energy (MTF) 2) the jet contains at least one
481 track with $p_T > 1.2$ GeV (MPT) 3) charged particles make up at least 10% of the
482 jet p_T (CPF) and 4) each jet contains at least two particles associated to the same
483 PV as the boson (NTRK). The systematic uncertainties on the jet efficiencies are
484 dominated by these identification cuts. Similarly to previous analysis [31] [1], these
485 cuts were tightened to 60%, 2.4 GeV, 20% and 3 tracks for the MTF, MPT CPF
486 and NTRK, respectively. The relative difference in the fraction of accepted events
487 between the data and MC is used as the level of agreement. This multiplied by the
488 efficiency of the jet ID cut was assigned as a systematic uncertainty. Depending on
489 the jet p_T and the jet ID variables, the largest uncertainty varies from 1.1 to 1.8%.
490 Conservatively, 1.8% was assigned to account for the worst p_T region. The reconstruction
491 inefficiency before imposing the jet ID cut that is determined in simulation is canceled
492 out between the hadron distributions (numerator) and the jet distributions (denomina-
493 tor), therefore is not considered as a source of systematic uncertainty in this analysis.
494 The distributions of the four different variables in data and simulation are shown in Fig. 34.

495
496 Lastly, the systematic uncertainty on the muon track reconstruction efficiency was not
497 added to the total systematic uncertainties as the muon efficiencies are canceled between
498 the numerator and denominator of the observables. A total normalization uncertainty is

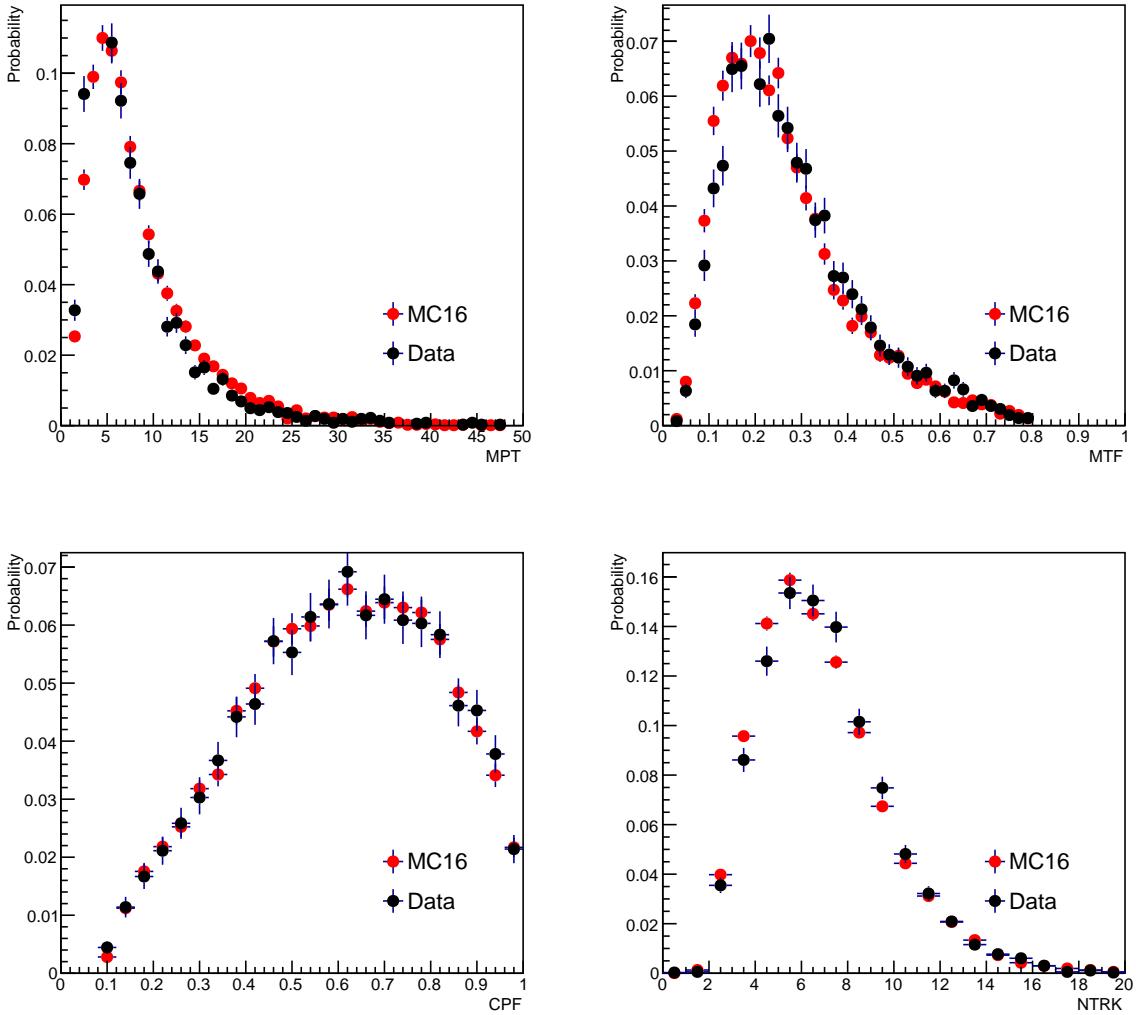


Figure 34: The four Jet-ID variables in data (black) and reconstructed MC (red). In general, a good agreement is seen in all four distributions, indicating the detector response to jets in simulation is reasonably consistent with the one in data.

499 obtained by summing the three components in quadrature and found to be 2.4%.

500 9.2 Jets

501 9.2.1 Jet energy scale

502 At LHCb, two different approaches have been taken in previous analyses that
 503 determine how well MC models the detector response to jets, the first of which uses
 504 the p_T balance between the Z boson and recoiled jet [32] and the other analyzes
 505 charged/neutral/ghost particles contributing to jets [31]. The two methods were shown to
 506 reach consistent results, and we take the first approach in this analysis.

507
 508 In events where there is little activity other than the jet and Z boson produced
 509 back-to-back in ϕ , the p_T of the Z boson can be used as a proxy for the p_T of the jet. This

property is being widely used for calibrating the jet energy scale at LHC experiments. To enhance the back-to-back event topology, additional requirements were placed with respect to the standard Z +jet event selection; the azimuthal distance between the leading jet and the Z boson has to be greater than $\frac{7\pi}{8}$, and additionally, the p_T of the sub-leading jet has to be less than 25% of the p_T of the hardest jet.

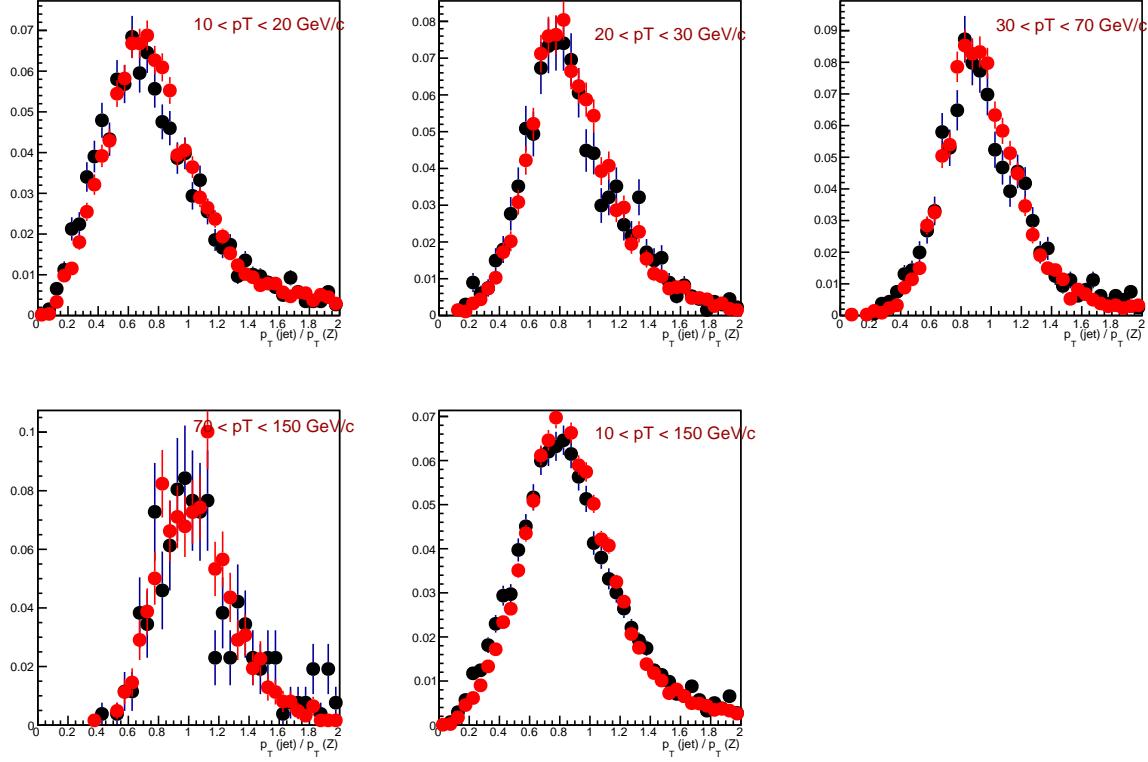


Figure 35: The Z -jet transverse momentum balance in data (black) and reconstructed MC (red).

The systematic uncertainty on the jet energy scale is determined by looking at the level of agreement between the MC and data in the p_T balance of the jet against the Z boson p_T . Fig. 35 compares the $\frac{p_T(\text{jet})}{p_T(Z)}$ distribution between the data and MC at detector level for different regions of the jet p_T . A good agreement is shown, indicating the detector response to jets is well modeled in simulation. Fig. 36 additionally shows the ratio of $\frac{p_T(\text{jet})}{p_T(Z)}$ distributions in data to MC. The region where the p_T imbalance is below 0.4 is considered to be contaminated with fake jets, therefore is disregarded [32]. The data to MC ratio of the mean $\frac{p_T(\text{jet})}{p_T(Z)}$ is consistent with unity within 2%. The scale β defined as the ratio of the new scale to the original scale was varied until the χ^2 is minimized. At β values that minimize the χ^2 , the difference of the means in data and MC, computed within a truncated range between 0.5 and 1.6, was seen to be minimized. The best scale β_{best} found for different jet p_T regions of 10-20 GeV, 20-30 GeV, 30-70 GeV and 70-150 GeV was 1.00 ± 0.01 , 0.98 ± 0.01 , 0.98 ± 0.01 , 1.01 ± 0.02 , respectively. The uncertainties on the β_{best} were determined by taking the difference between β_{best} and β , where the latter corresponds to the 1σ variation of the mean p_T balance obtained from the fit in Fig. 36. The results are consistent with the previous studies documented in Sec.10.2 in Ref. [32]. A systematic uncertainty of 3% is therefore assigned on the jet energy scale in order to account for the

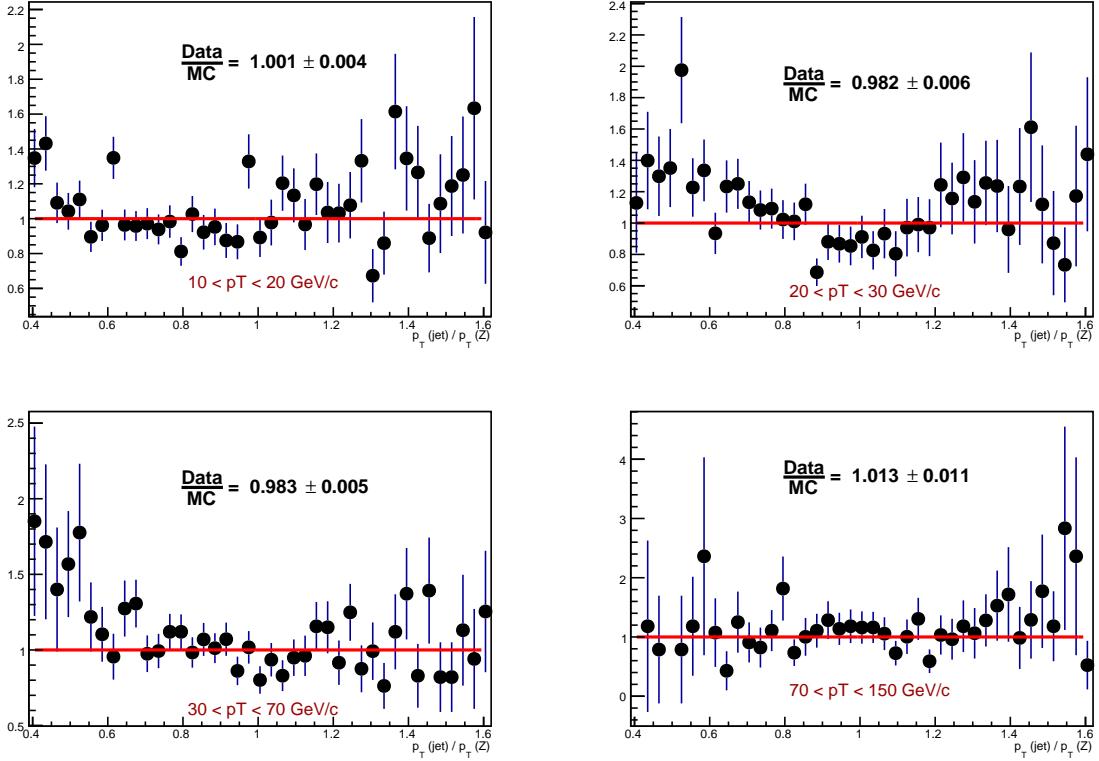


Figure 36: The ratio of $\frac{p_T(\text{jet})}{p_T(Z)}$ distributions in data to MC for different jet p_T bins. The mean of the p_T imbalance in data is consistent with the one in MC within 2%.

spread and uncertainty in these values of β_{best} . To evaluate the effect of the systematic uncertainty on the hadronization variables, the momentum scale of reconstructed jets has been varied by 3% in simulation and the unfolding and re-analysis of data have been repeated using the new detector response. The estimated uncertainties from this source are generally less than 10% with the exception of a few bins at the edge (as large as 50%) with the rising pattern depending on the statistics of bins for z , j_T and r while overall uncertainties grow with jet p_T , as shown in Figs. 37 and 38. Other effects such as flavor dependence of the jet reconstruction performance are not considered because they are canceled out.

9.2.2 Jet energy resolution

The uncertainties attributed to the jet energy resolution are found by smearing the detector response to jets in their four momenta and redoing the entire chain of the analysis. This smearing is performed in simulation by taking a random number from a Gaussian distribution, with a mean of 1.0 and a width of 0.1. A different random number is used for each jet in a given event. The 10% resolution was determined based on the level of agreement in the width of the $p_T(\text{jet})/p_T(Z)$ distributions between the data and simulation, similarly to Ref. [32]. Smearing the jet p_T by taking a random number from a Gaussian distribution of unit mean and width of 10% corresponds to the limit within which a consistency between the data and simulation is observed. Systematic

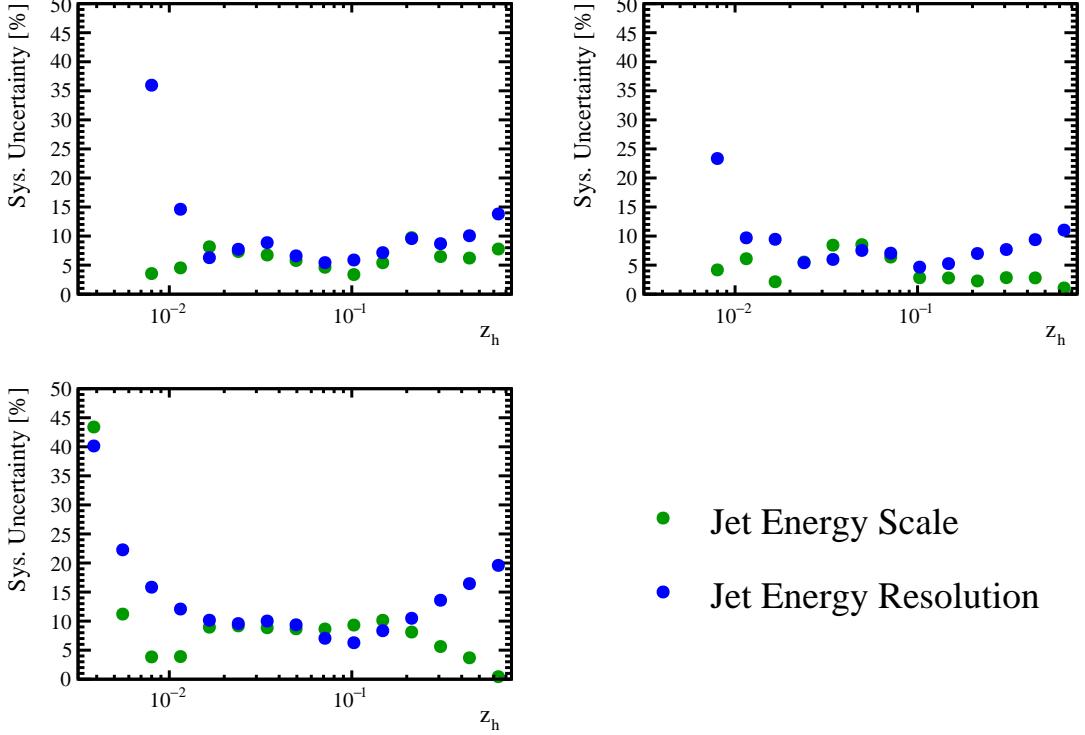


Figure 37: The jet energy scale and jet energy resolution systematic uncertainties on the hadron z_h distribution for different jet p_T bins.

uncertainties on each point are determined by computing the standard deviation of the unfolded distributions using smeared responses with respect to the nominally unfolded distributions. The smearing procedure of jet energy is repeated 10 times independently. The resulting uncertainties on hadronization variables are well within 10%, similarly to the effects of jet energy scale with the largest uncertainties in edge bins at the 50% level, as shown in Figs. 37 and 38.

9.3 Unfolding

Closure tests were performed in Sec. 8 to evaluate the level to which we can trust the unfolded distributions. Any non-closure can result from incorrectly simulating detector responses to the jet production in MC, *e.g.* resolution effects, non-uniformity, mismatching to fake jets, etc. The systematic uncertainties are determined by taking a standard deviation of weighted absolute differences in the ratio between the unfolded and truth distributions.

The DNN unfolding method adopted in this analysis is much less subject to the effects of insufficient statistics in data than other conventional binned methods. This is particularly beneficial for multidimensional analyses in which there can be regions where statistics run out as a result of multidimensional binning. This either leads to disproportionately large uncertainties on the unfolded distributions or simply renders unfolding not feasible. The fluctuation in the unfolded-to-truth distributions is reduced from as high as 200% in low statistics bins seen in a previous analysis [1] to less than 10% as seen in Fig. 31 - Fig. 32 for jet p_T between 20 and 100 GeV/ c . Similar results

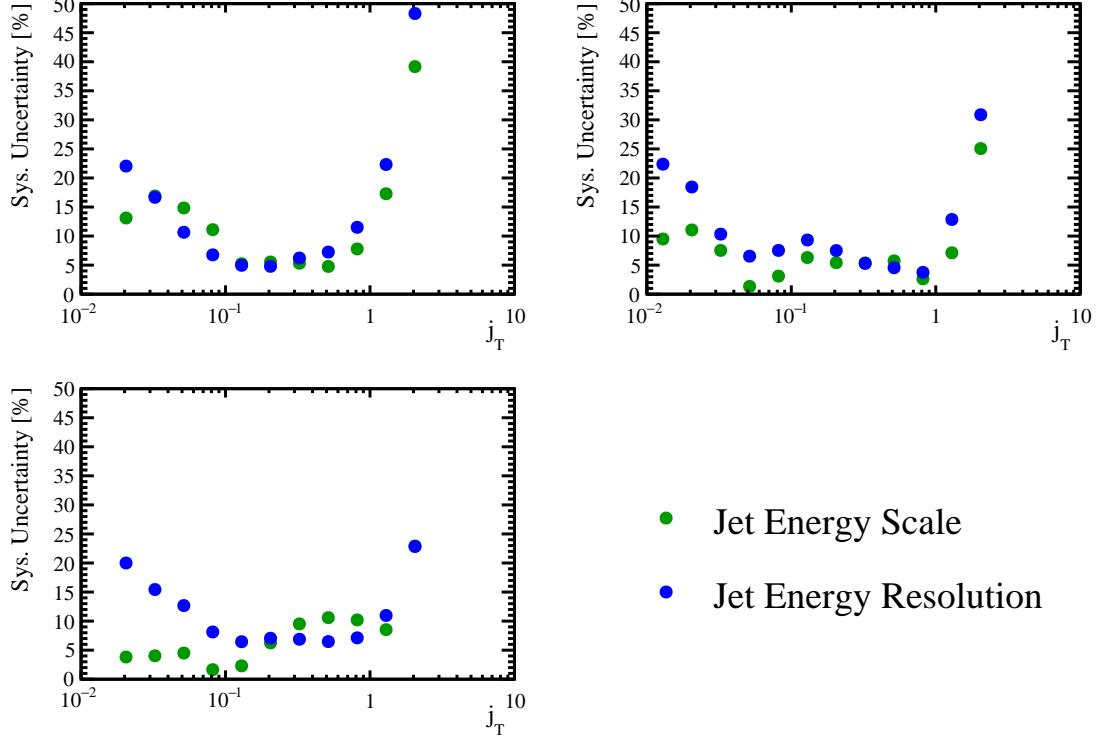


Figure 38: The jet energy scale and jet energy resolution systematic uncertainties on the hadron j_T distribution for different jet p_T bins.

were obtained for each jet p_T bins used for this analysis.

The uncertainties on the unfolded number of jets and the non-normalized hadronization variables are added in quadrature to give a total uncertainty of $\sqrt{1.1^2 + 0.8^2} = 1.4\%$. A slightly larger uncertainty coming from the number of jets is expected due to larger migration and momentum resolution of the jets than the charged tracks.

9.4 Hadrons

There are three sources of systematic uncertainties around charged hadrons - the first two are related to track reconstruction, *e.g.* efficiency and purity defined in Eq. 7 in Sec. 6 and the last to particle identification studied in Sec. 7. These corrections were determined as a function of particle momentum.

9.4.1 Track Selection Purity

Charged tracks in the particle flow algorithm are nominally required to have their probability of being a fake track, `ProbNNghost`, less than 0.5. The fake tracks result from wrong combinations of well reconstructed track segments in the upstream and downstream regions of the magnet due to the long lever arm in between. The mismatched fake tracks remain the most abundant category after fake track rejection according to studies performed in Ref. [33]. The uncertainty on the purity of the track selection caused by possible discrepancies in the `ProbNNghost` distributions between the data and MC is estimated by imposing a tighter ghost track rejection cut of `ProbNNghost` <

591 0.3. Any difference in the distributions between the nominal and tightened cut would
 592 indicate systematic effects are present. The momentum and η dependent changes in the
 593 cut efficiency and purity are also taken into account when comparing the nominal and
 594 tight ProbNNghost requirements. The measure of discrepancy is determined by taking a
 595 combined ratio as follows.

$$ratio = \frac{N_{track}^{reco}}{c_{hadron}}(\text{ProbNNghost} < 0.3) / \frac{N_{track}^{reco}}{c_{hadron}}(\text{ProbNNghost} < 0.5) \quad (12)$$

596 This quantifier is ideally unity in the case of perfect agreement between the data
 597 and MC. Fig. 39 shows this ratio translated into z_h and j_T distributions. The ratios are
 598 consistent within 2σ . Therefore no systematic effects are present.

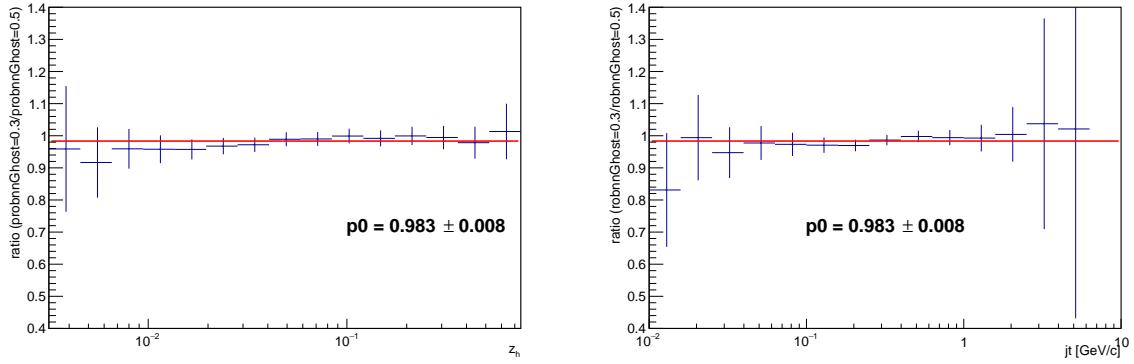


Figure 39: The combined ratio defined in Eq. 12 translated into z_h and j_T distributions with jet p_T integrated between 20 and 100 GeV/ c .

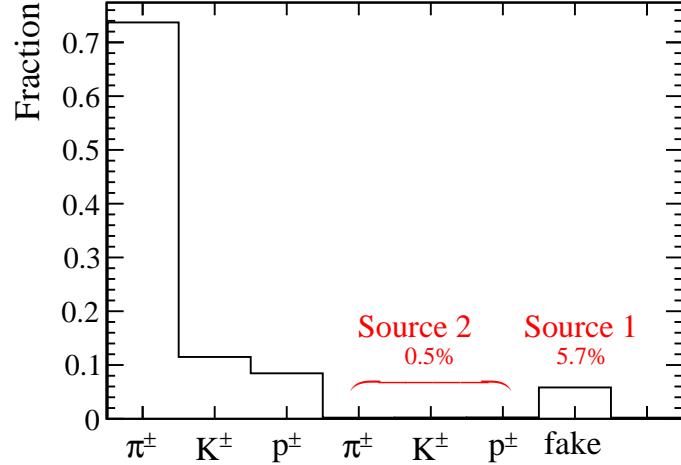


Figure 40: Composition of hadron candidates that consist of signal tracks, background tracks (Source 1 and Source 2). Source 2 includes charged hadrons not associated with a jet of interest (0.5%) and Source 1 is fake tracks (5.7%).

599 The track selection purity is the dominant component in $c_{h-in-jet-purity}$ while the
 600 effects of hadrons not associated with a jet of interests are less than 0.5% as can be seen

601 in Fig. 40. The ghost tracks are made of mismatched track segments and it does not
602 depend on the particle species. For this reason, no systematic uncertainties are assigned
603 for applying a particle species independent correction factor c_{hadron} in both identified and
604 non-identified hadron distributions.

605 9.4.2 Track reconstruction efficiency

606 There are two main sources of systematic uncertainties attributed to track reconstruc-
607 tion efficiency - the binning effect and the uncertainties on the material budget. The
608 track reconstruction efficiency is determined bin-by-bin in MC and then the resulting
609 efficiency is smoothed to give a more continuous distribution as a function of p and η .
610 This alleviates the non-physical discontinuity created between adjacent bins in regions of
611 insufficient statistics. To estimate a systematic uncertainty on the track reconstruction
612 efficiency, the analysis is performed again with the raw efficiency before the smoothing
613 procedure. The ratio of the distributions with the nominal smoothed track efficiency to
614 the distributions determined with the raw efficiency is taken as a systematic uncertainty.
615 Figures 41 and 42 show the uncertainties determined from the ratios of these distributions
616 as a function of the hadronization variables. The figures show that the change is small
617 at higher z_h and j_T , similar to the track selection purity case. For these observables,
618 the uncertainty can reach up to 15% in kinematic regions where the track efficiency was
619 measured to be changing rather drastically bin-to-bin, *e.g.* at the lower ($< 20 \text{ GeV}/c$)
620 and higher ($> 100 \text{ GeV}/c$) momentum bins. The effects are more pronounced at low
621 momentum than higher momentum due to larger statistics there. The kinematic region
622 that is most affected is the lowest j_T bins for the highest jet p_T bin.

623 In addition to the binning effects, a PID dependent uncertainty attributed to the
624 uncertainties on the material budget in detectors is also considered. They are estimated
625 by the inefficiency caused by hadronic interactions, 15% (pions), 12.7% (Kaons) and 33%
626 (protons) multiplied by 10% (uncertainty in material budget) [34]. The results are 1.50,
627 1.27 and 3.3% for pions, kaons and protons, respectively.

628 9.4.3 Particle identification

629 Systematic uncertainties on the particle identification corrections determined by using
630 the bootstrap method in Sec. 7 were propagated to the hadronization variables for identified
631 charged hadrons. The analysis was repeated multiple times for each of the cases where
632 the corrections factors varied by +1 and -1 times σ . As expected from uncertainties as a
633 function of track momentum, pions are least affected by PID corrections followed by kaons
634 and protons. Generally, larger uncertainties are seen towards the lower (higher) limit of z_h
635 due to an excess of low momentum (high momentum) particles. Large uncertainties are
636 attributed to the turn-on of PID efficiency curves for Kaons and protons at low momentum
637 less than $8 \text{ GeV}/c$, whereas they are limited by statistics at high momentum greater than
638 $100 \text{ GeV}/c$. This trend is not as strong in the j_T distributions as in z_h . Figures 43 and 44
639 show the systematic uncertainties on the hadron z_h and j_T distributions due to the PID
640 corrections.

641 9.4.4 Magnet polarity

642 As a crosscheck, the total hadron correction factors for magnet polarity Up and Down
643 were compared as a function of particle momentum. In addition, the charged particle
644 momentum spectra normalized by the number of jets ($15 \text{ GeV}/c < p_T < 100 \text{ GeV}/c$) were
645 compared between the two magnet polarities. No dependence on the magnet polarity was
646 found and results are shown in Fig. 45.

647 9.5 Summary of systematic uncertainties

648 The systematic uncertainties from various sources are classified into three categories
649 and are summarized in Tables 6-8. The first type of systematic uncertainty is normalization
650 uncertainty. The second type is point-to-point uncertainties where the uncertainty on
651 each data point moves up or down independently. The uncertainties of the third type are
652 correlated between data points.

Table 6: Summary of normalization uncertainties

Source	Uncertainty in %
Jet reco. + jet ID	1.8
Purity of Z +jet sample	1.6
Sum	2.4

Table 7: Summary of point-to-point uncertainties

Source	Uncertainty in %
Track reconstruction efficiency	0.2-8.7
Particle Identification	0.8-15.0 (π^\pm), 0.1-34.6 (K^\pm), 0.5-52.5 (p^\pm)

Table 8: Summary of correlated uncertainties

Source	Uncertainty in %
JES	0.4-43.4
JER	3.8-48.3
Unfolding	1.4

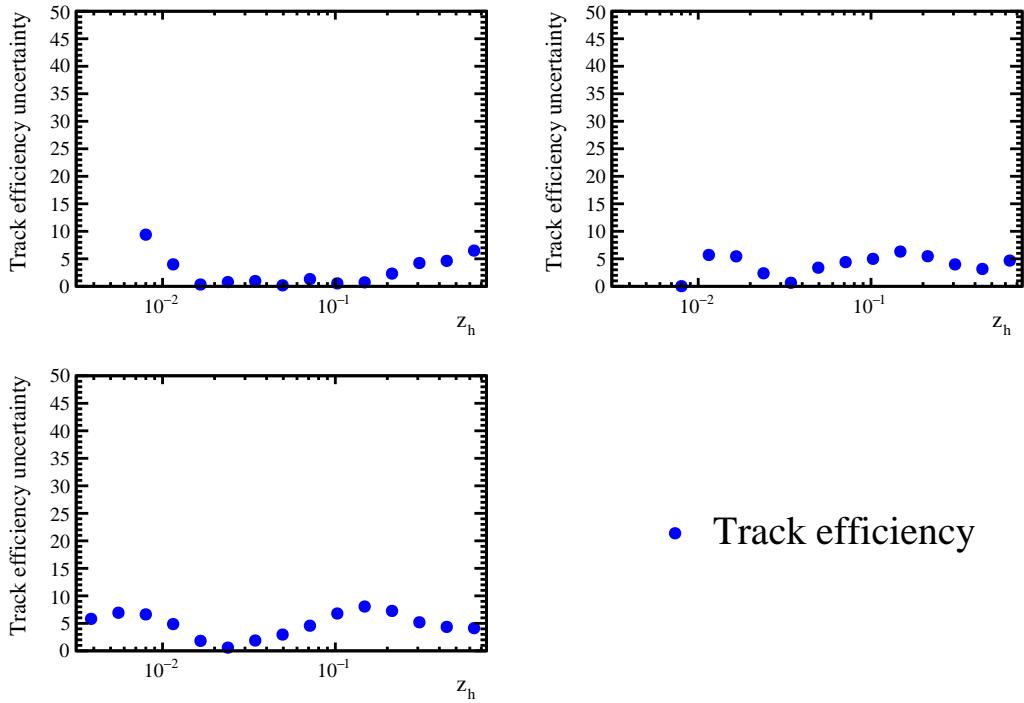


Figure 41: The hadron track reconstruction (purity and efficiency) systematic uncertainties on the hadron z_h distribution for different jet p_T bins.

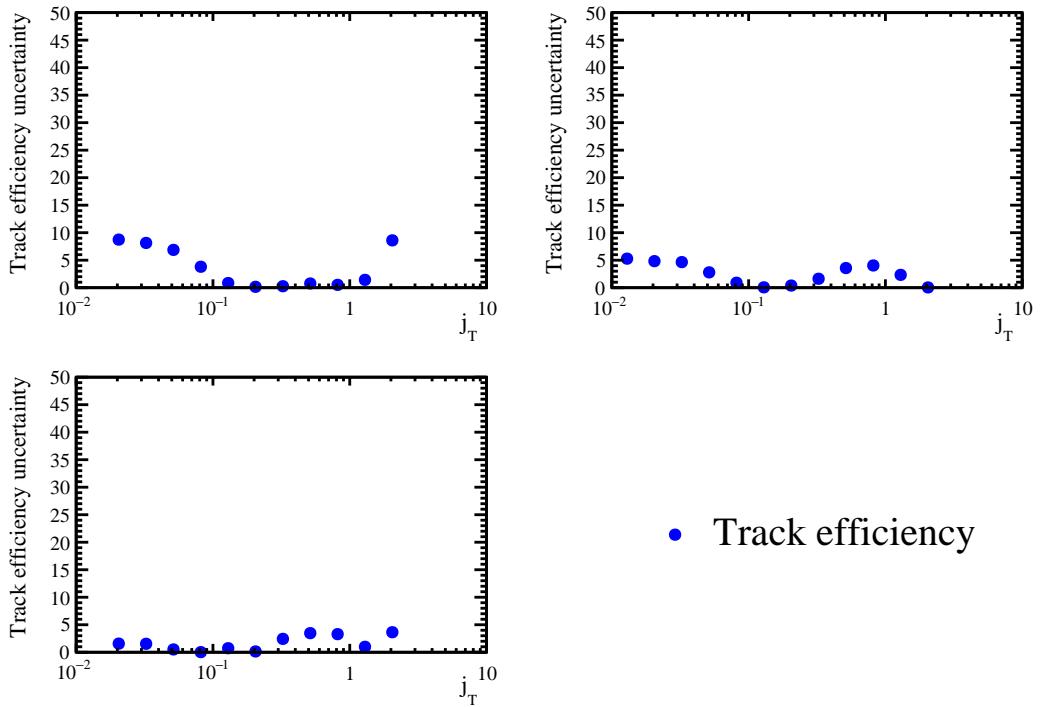


Figure 42: The hadron track reconstruction (purity and efficiency) systematic uncertainties on the hadron j_T distribution for different jet p_T bins.

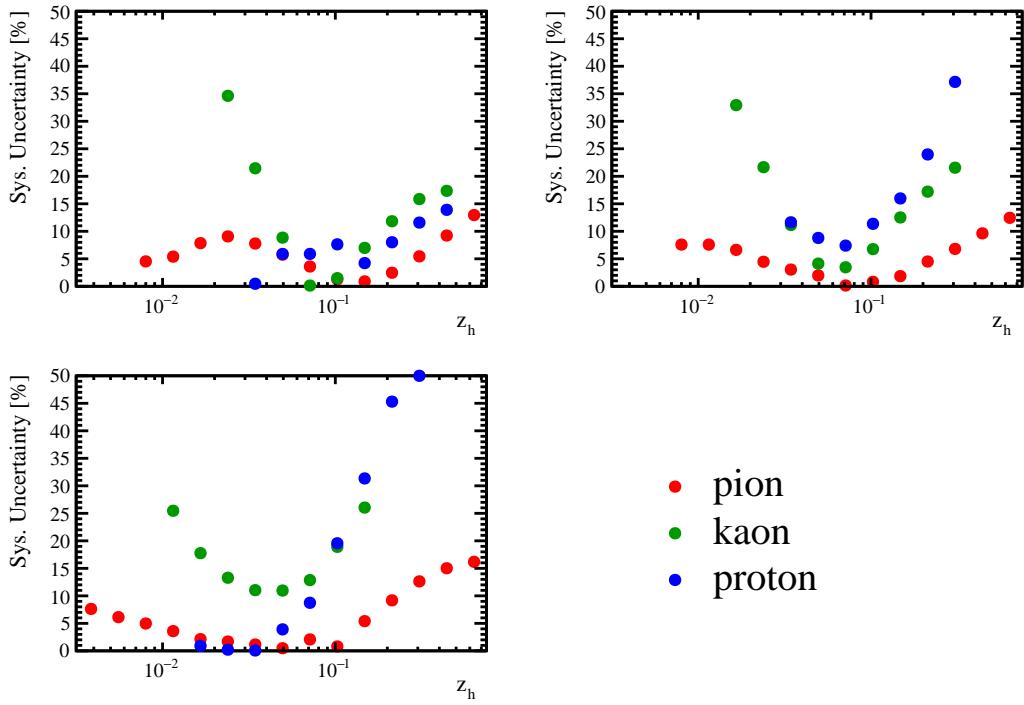


Figure 43: The PID systematic uncertainties on the hadron z_h distribution for different jet p_T bins.

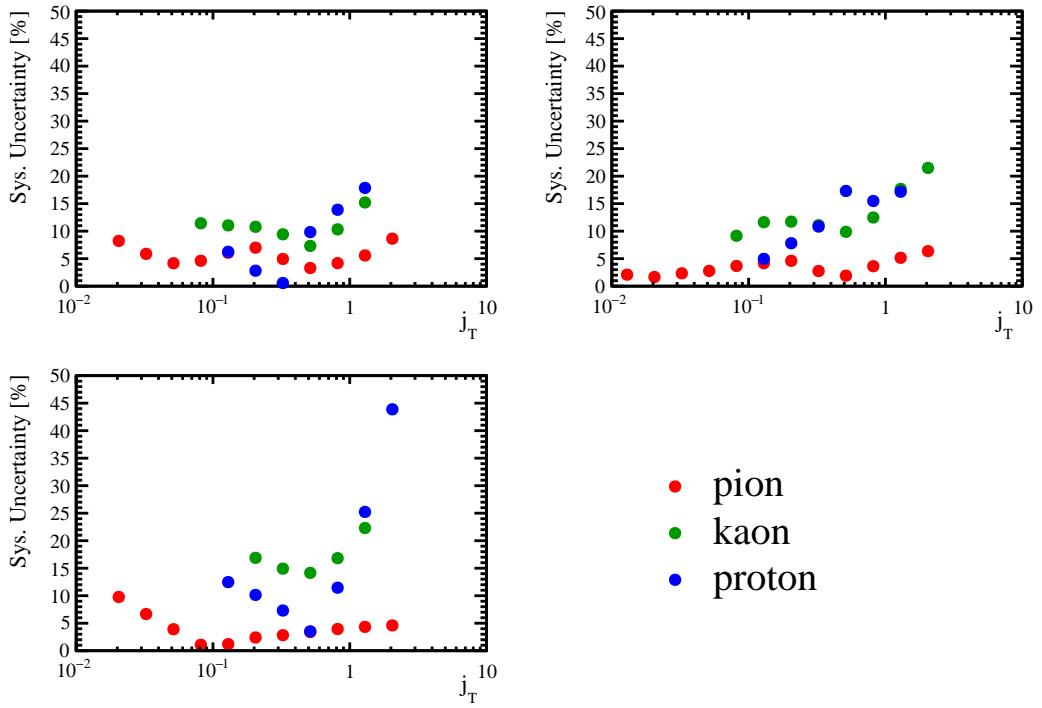


Figure 44: The PID systematic uncertainties on the hadron j_T distribution for different jet p_T bins.

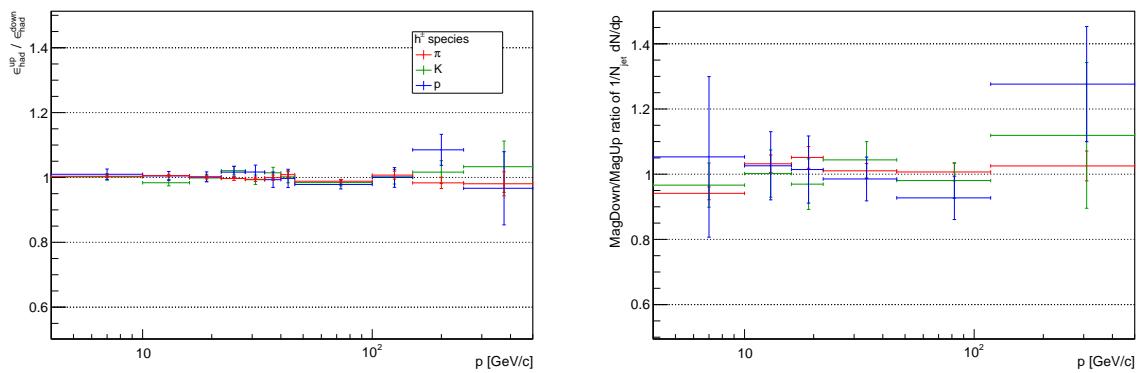


Figure 45: The relative (left) hadron correction factors and (right) normalized yields between data sets taken with the two opposite magnet polarities.

653 10 Results

654 10.1 Nonidentified charged hadrons (π^\pm , K^\pm and p^\pm combined)

655 The results of the nonidentified charged hadron z distributions are shown in Fig. 46
656 - Fig. 47. Figure 46 compares the hadronic longitudinal momentum fraction of jet
657 momentum along the jet axis in 13 TeV and 8 TeV for a given jet p_T , where an
658 excess is seen at lower z for the 13 TeV data. This can be possibly explained by an
659 increased level of soft particles generated at higher beam energy. Figure 47 shows the
660 jet p_T dependence of the z distributions in 13 TeV and 8 TeV, separately. In both
661 energies, it is clear that more energy available in higher jet p_T goes into creating low
662 momentum particles, which translates into low z particles. Pythia8 (Fig. 48 and Fig. 49)
663 generally describes the data as a function of z reasonably well in the peak region, while it
664 consistently under-predicts in the higher z region. Note that the data fully corrected for de-
665 tector effects are compared to the particle level in MC with identical kinematic cuts applied.
666

667 Figures 50-51 show the results of the hadronic transverse momentum distribution with
668 respect to the jet axis, j_T . The agreement between 13 TeV and 8 TeV is consistent with
669 universality of fragmenting jet functions. In Fig. 51, a slight widening of the j_T distribution
670 is seen as jet p_T increases, consistent with expectations from Collins-Soper-Sterman
671 evolution. The average number of charged hadrons per jet integrated over j_T , which is
672 obtained by adding up each data point multiplied by the bin width over j_T for the lowest
673 jet p_T bin, is 5.6. This is consistent with the one integrated over z within 0.2%, indicating
674 robustness of measurements. At the mid and highest jet p_T bins, the average numbers
675 of charged hadrons per jet are 7.3 and 7.7, respectively. Pythia8 (Fig. 52 and Fig. 53)
676 slightly underpredicts the distributions at low j_T in the middle and highest jet p_T bin.
677

678 For the first time, fully 2-dimensional j_T - z_h joint distributions have been measured.
679 Results are shown in Fig. 54 and Fig. 55. A broad correlation is seen between the
680 longitudinal momentum fraction and transverse momentum, which can be pictured such
681 that particles of higher thrust along the jet axis are more likely to have a larger transverse
682 momentum components whereas soft particles are likely to be collinear to the jet axis.
683 The centroids of the distributions migrate towards lower z and slightly higher j_T with jet
684 p_T . The joint distributions sliced in z (j_T) are also shown in Fig. 55 as a function of j_T
685 (z). Alternatively, the three dimensional images of the measured TMD jet fragmentations
686 into hadrons are shown in Fig. 56. The latter figures clearly show that the shift of the
687 centroids of the distributions with jet p_T results from additional soft particle productions
688 at low $\log_{10}(z_h)$ (less than -1.4) and an increased j_T for a given z_h of a particle at large
689 $\log_{10}(z_h)$ (greater than -1.4).

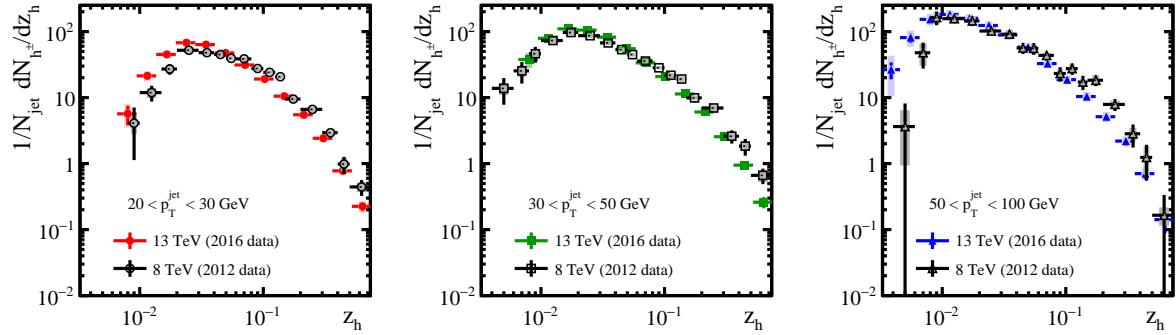


Figure 46: The distributions of z_h for different jet p_T bins. In each panel, the present 13 TeV results are shown in color, while the previously published 8 TeV results are shown in black.

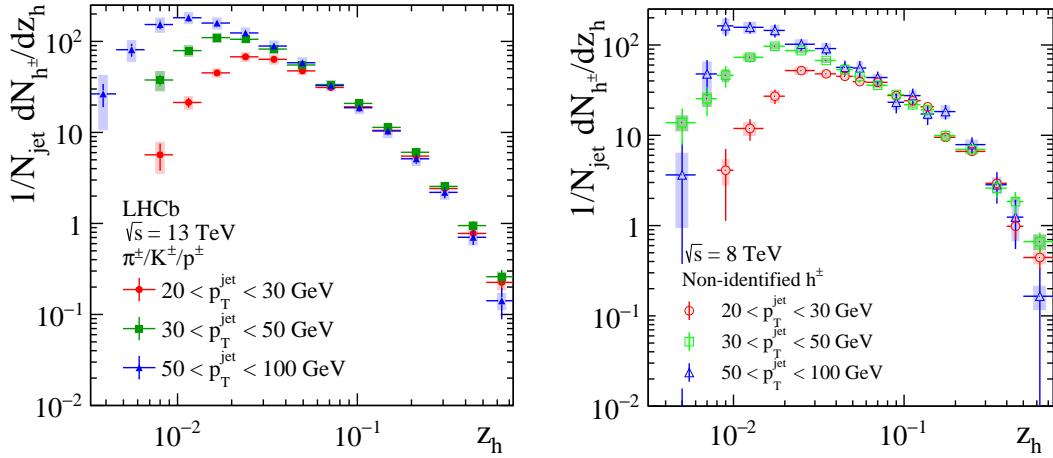


Figure 47: z_h distributions for the three different jet p_T bins shown on the same plot. 13 TeV results are shown on the left, and 8 TeV on the right.

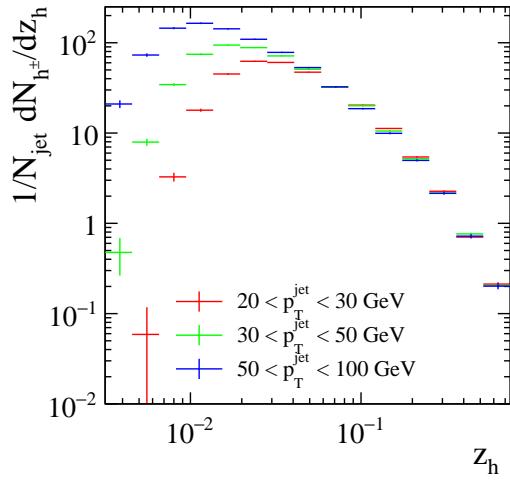


Figure 48: The distributions of z_h for different jet p_T bins from Pythia8.

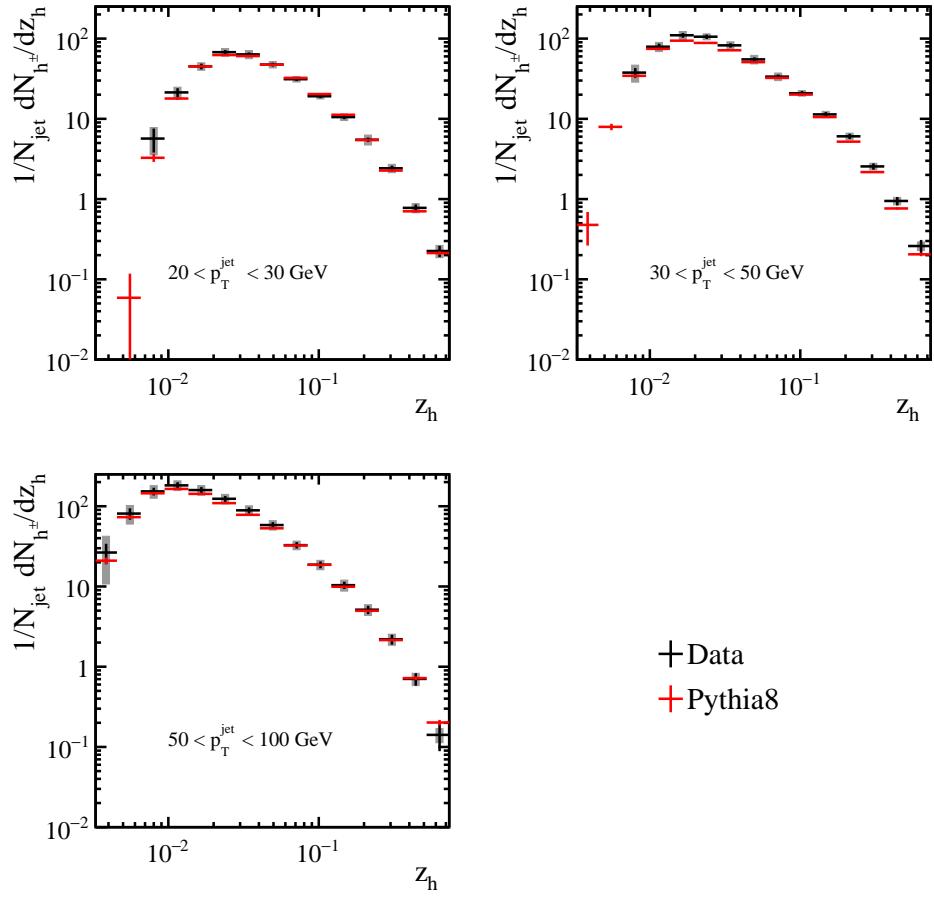


Figure 49: The distributions of z_h for different jet p_T bins, with Pythia compared to the 13 TeV data.

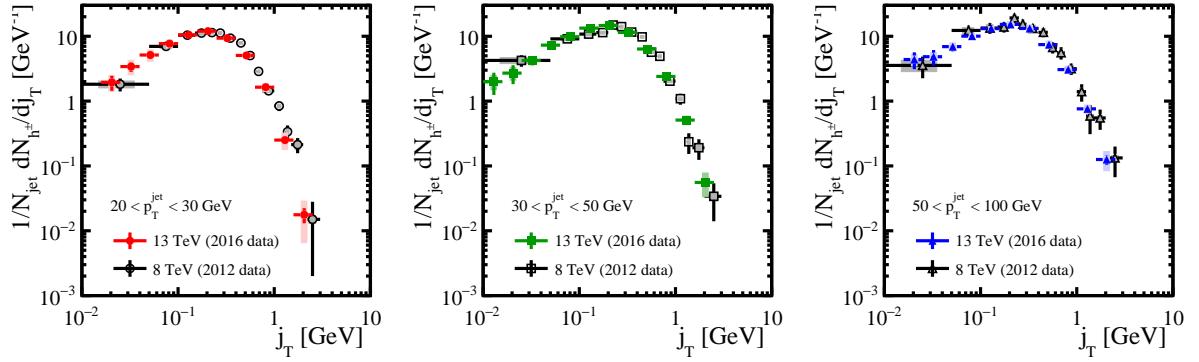


Figure 50: The distributions of j_T for different jet p_T bins. In each panel, the present 13 TeV results are shown in color, while the previously published 8 TeV results are shown in black.

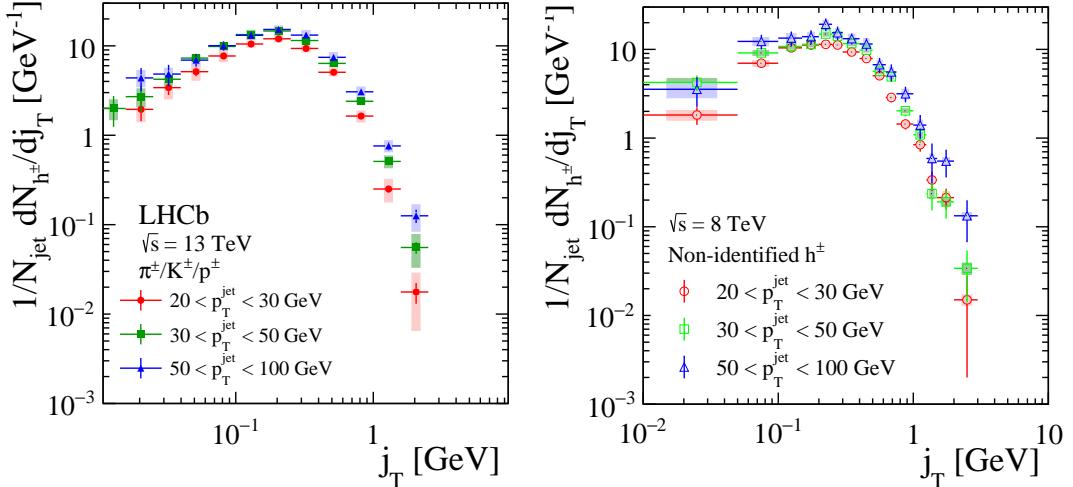


Figure 51: j_T distributions for the three different jet p_T bins shown on the same plot. 13 TeV results are shown on the left, and 8 TeV on the right.

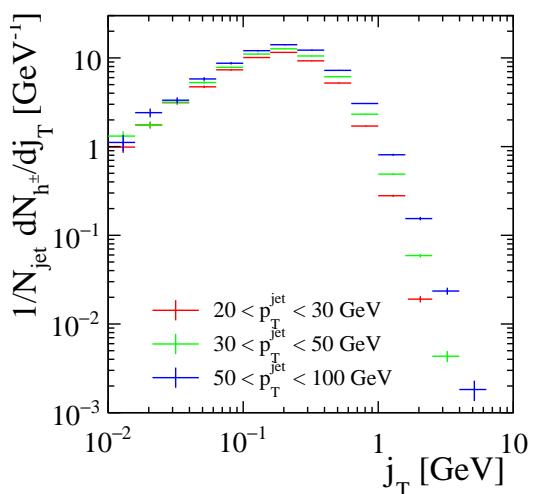


Figure 52: The distributions of j_T for different jet p_T bins from Pythia8.

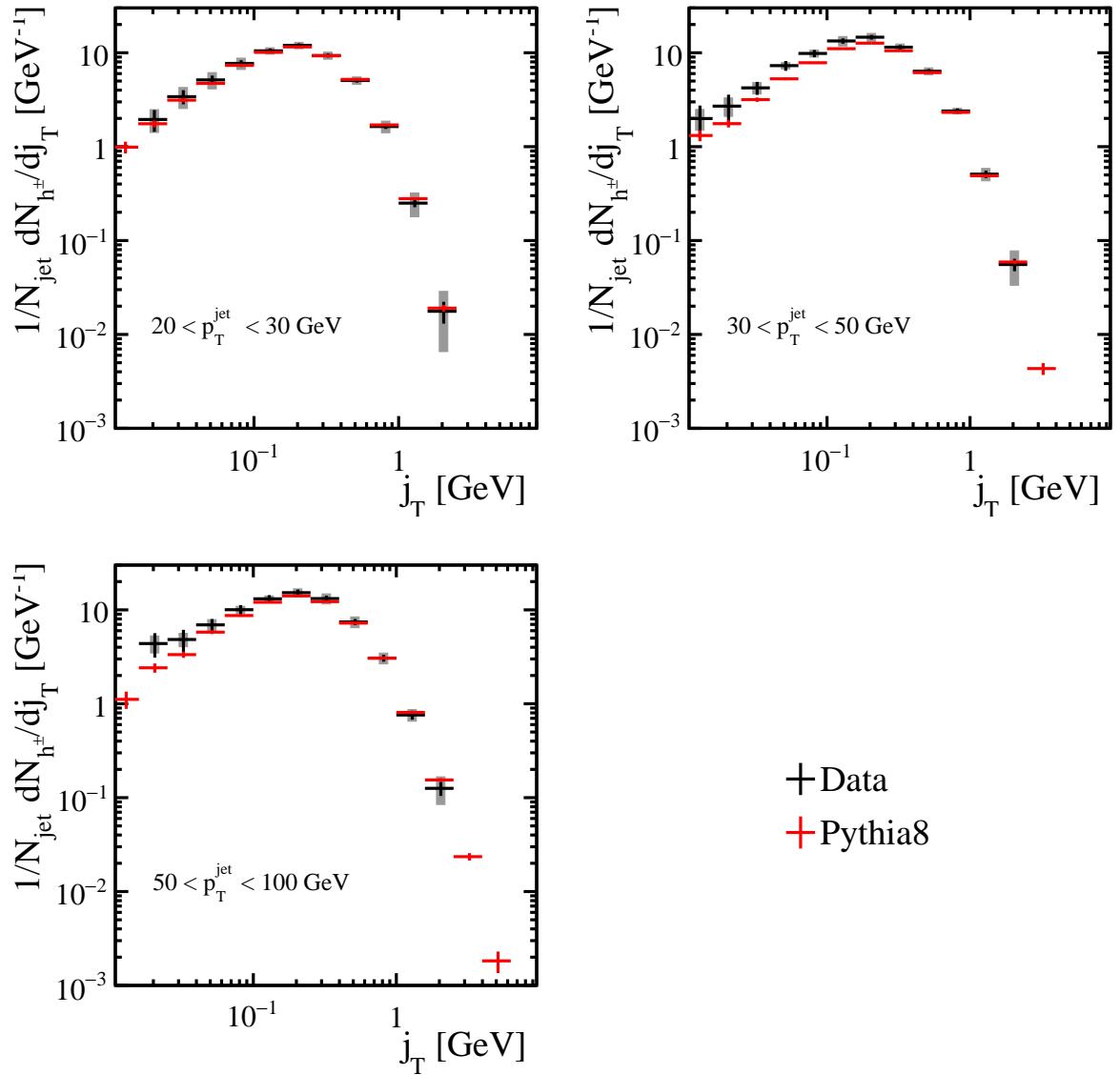


Figure 53: The distributions of j_T for different jet p_T bins, with Pythia compared to the 13 TeV data.

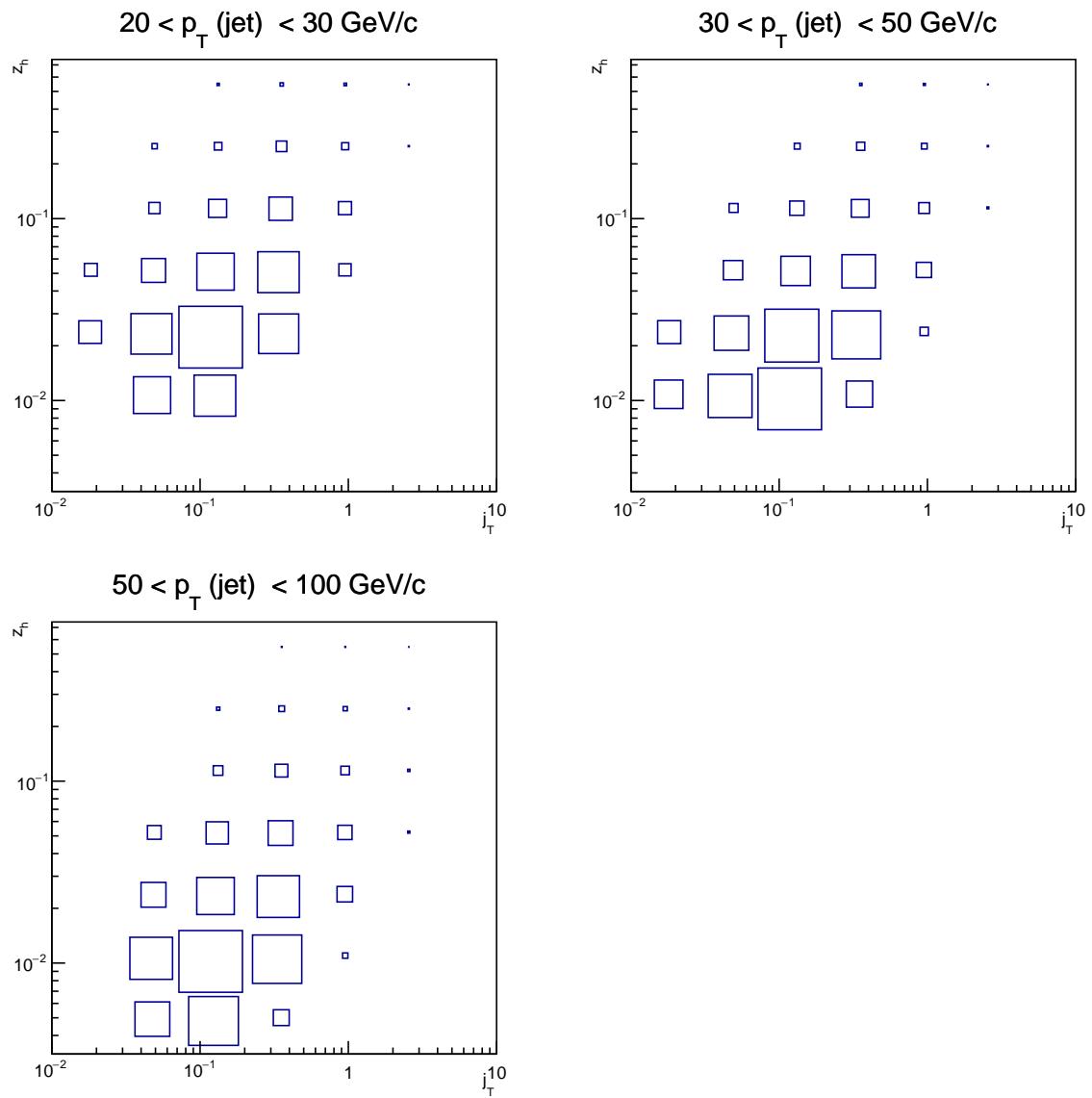


Figure 54: The z_h vs. j_T distributions for different jet p_T bins.

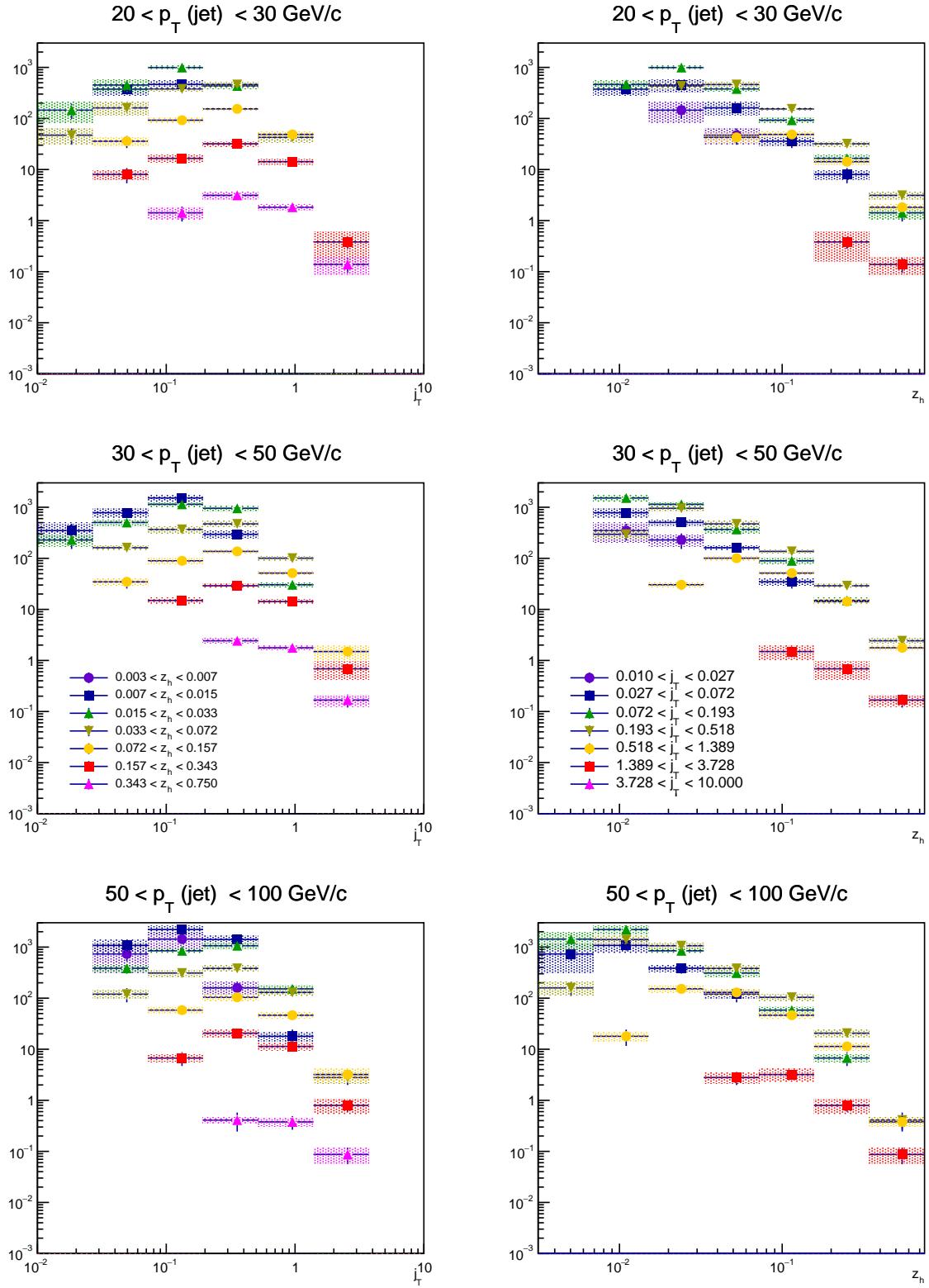


Figure 55: The j_T distributions in z_h bins (left) and z_h distributions in j_T bins (right) for different jet p_T bins.

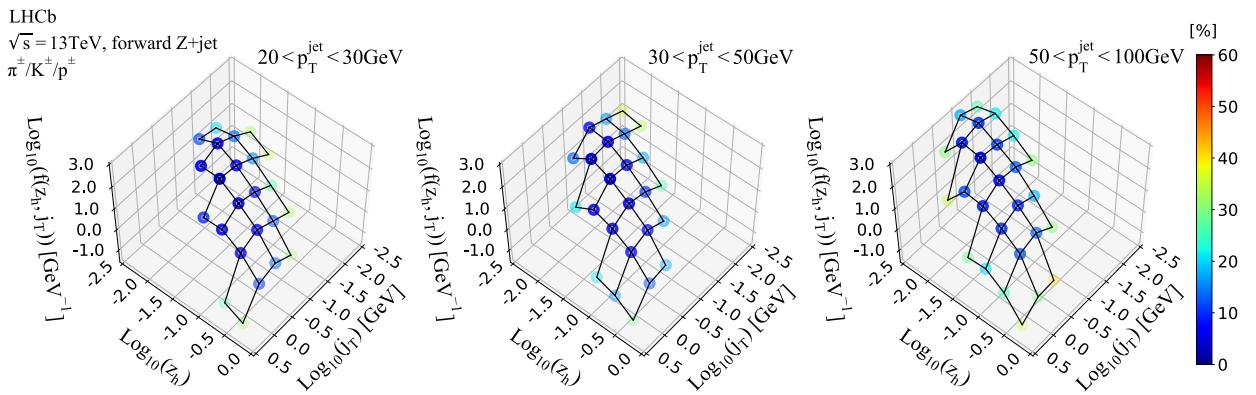


Figure 56: The z_h vs. j_T distributions for three different jet p_T bins: (left) $20-30 \text{ GeV}/c$ (middle) $30-50 \text{ GeV}/c$ (right) $50-100 \text{ GeV}/c$. The quadratic sum of statistical and systematic uncertainties are color-coded in the side bar.

690 **10.2 Identified charged hadrons - π^\pm , K^\pm and p^\pm**

691 In this analysis, the hadronization with identified charged hadrons within a jet
692 has been measured for the first time in $p + p$ collisions. The measurements provide
693 sensitivity to the different kinematics of hadronization for pions, kaons, and protons
694 in a light-quark-dominated jet sample and will constrain hadron-in-jet fragmentation
695 functions for identified particles. Figure 57 shows a compilation of 1-dimensional z and
696 j_T spectra for pions, kaons and protons, separately.

697

698 A common feature in the z distributions is that at the lower limit in z , a damping
699 of the hadron spectra is seen. The lower limit in z is determined by the minimum
700 momentum required for hadrons as well as the hadron mass, which are correlated with
701 the minimum momentum of the parton that hadronized to form a particle. In this limit,
702 hadron mass effects play a role. The same momentum cut for heavier particles translates
703 into an increased lower limit of parton momentum. This explains the increased shift of z
704 threshold seen in Fig. 58 for kaon and proton z spectra.

705

706 Above a certain z value, scaling behavior is seen in all three particles. This is
707 compatible with the predictions by purely perturbative parton splitting before color
708 coherence effects take over towards the end of the parton shower.

709

710 The shape of the j_T distributions do not drastically depend on jet p_T , only the overall
711 magnitude increases with jet p_T due to more energy available to create particles with the
712 largest increase visible at highest j_T .

713

714 Figures 60-62 compare the results of this analysis with Pythia8. Overall, a good
715 agreement is shown in the shape of the distributions. Pythia8 underpredicts pion
716 production while it overpredicts kaon and proton production.

717

718 The z vs. j_T distributions for identified hadrons shown in Fig. 63 exhibit a similar
719 trend as nonidentified hadrons. Kaons and protons populate higher z and j_T regions that
720 are more localized than for pions. Kaons and protons are more likely to take a large
721 fraction of parton momentum because they are less likely produced at the end of the QCD
722 cascade. The sliced j_T (z) distributions at fixed z (j_T) for π^\pm , K^\pm and p^\pm are shown in
723 Fig. 64 - Fig. 66. Alternatively, the three dimensional images of the measured TMD jet
724 fragmentations into pions, Kaons and protons in jets with a pT between 20 and 30 GeV/ c
725 are shown in Fig. 67.

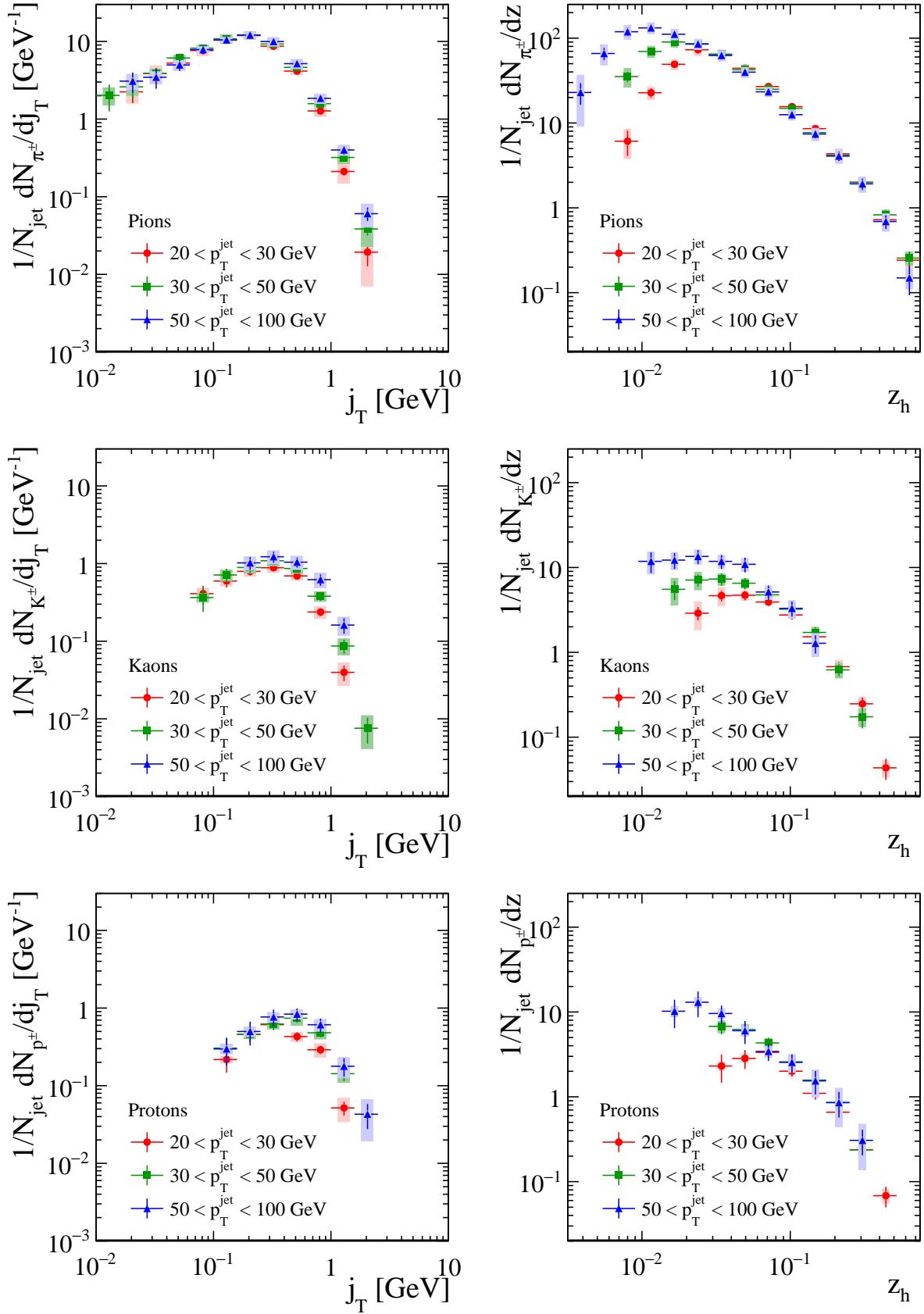


Figure 57: The j_T distributions (left) and z distributions (right) for pions (top), kaons (middle) and protons (bottom).

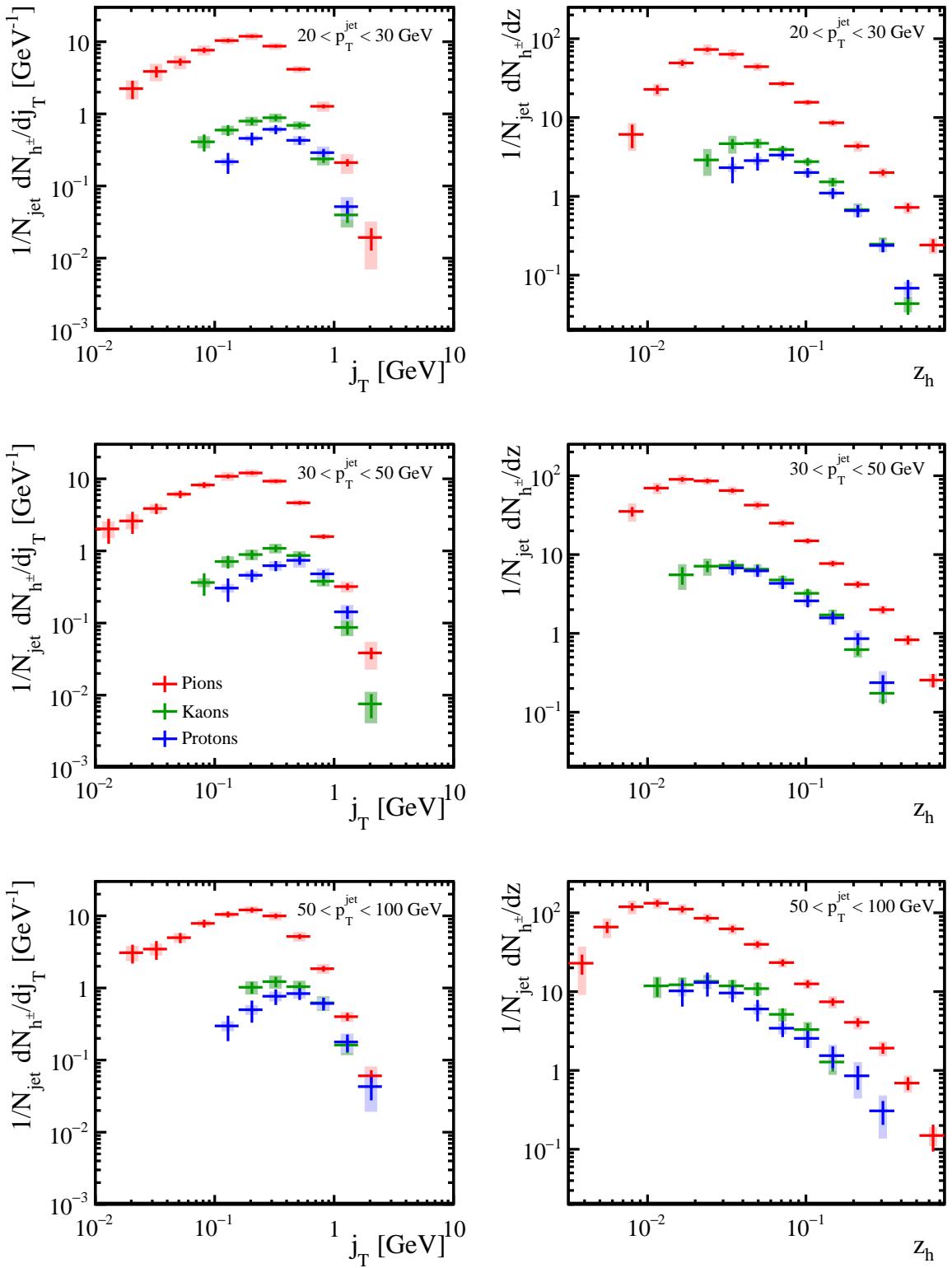


Figure 58: The comparison of the j_T (left) and z (right) distributions measured in data between different particle species.

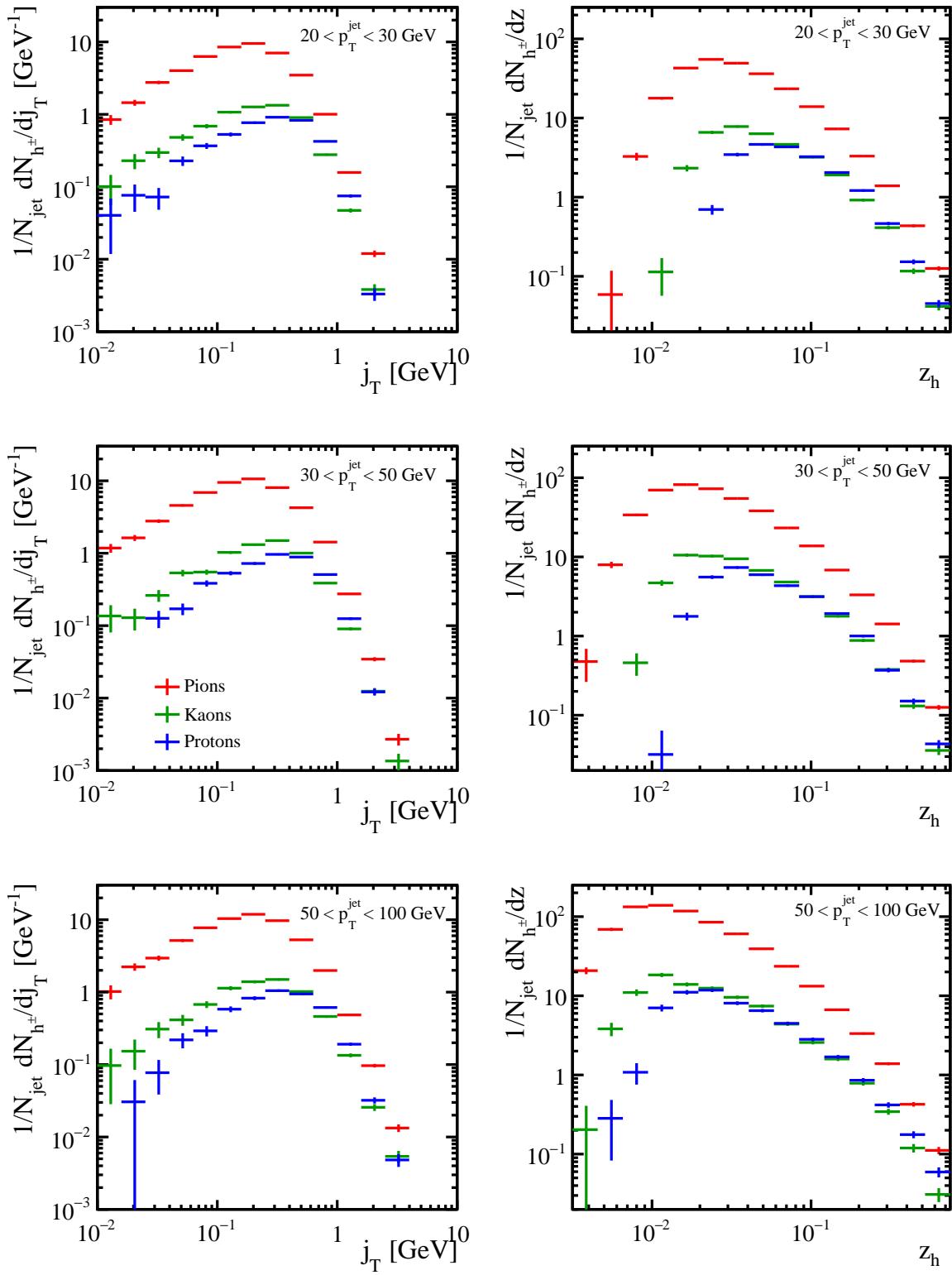


Figure 59: The comparison of the j_T (left) and z (right) distributions measured in Pythia8 between different particle species.

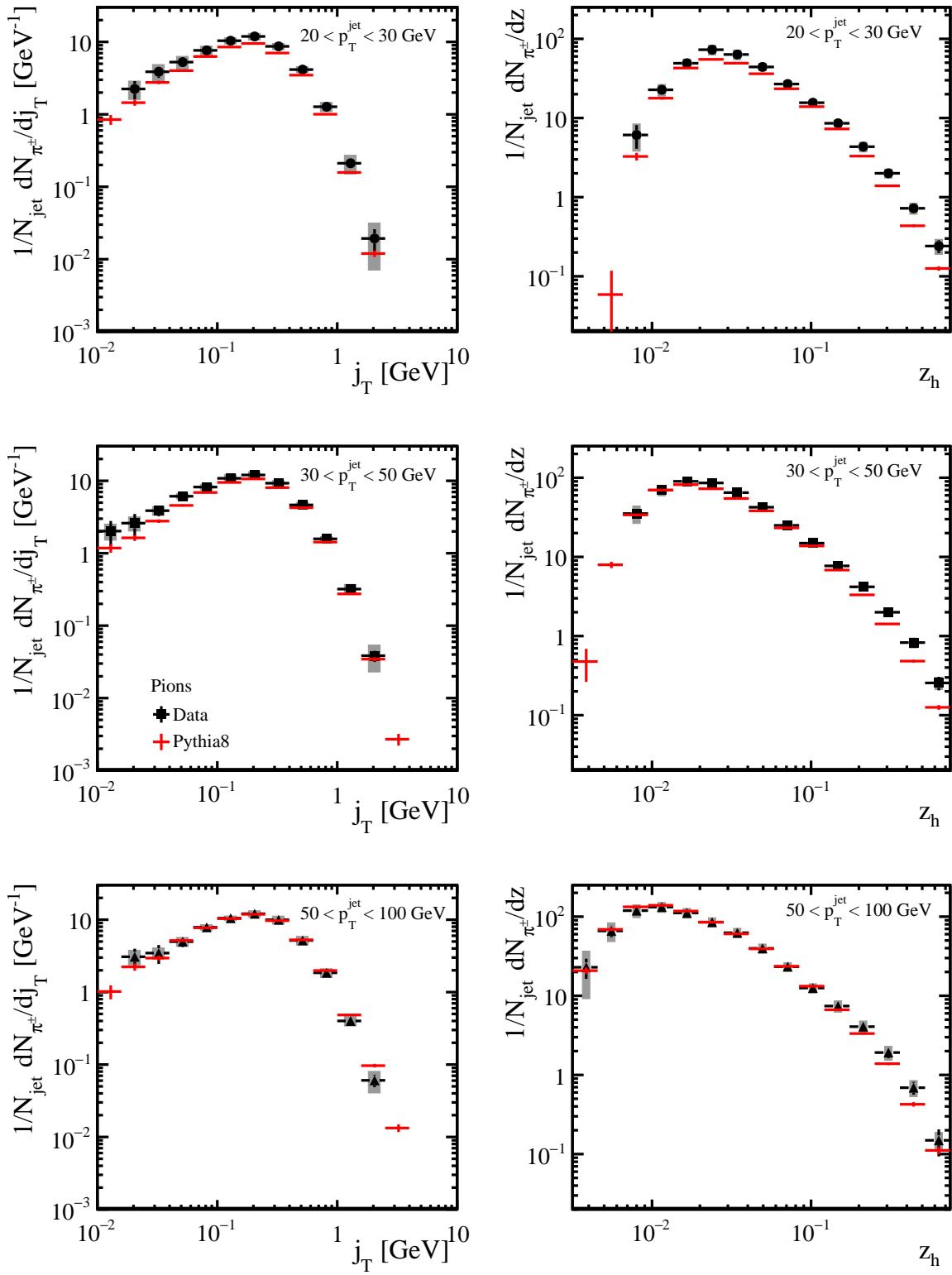


Figure 60: The comparison of the j_T (left) and z (right) distributions for pions measured in data and Pythia8.

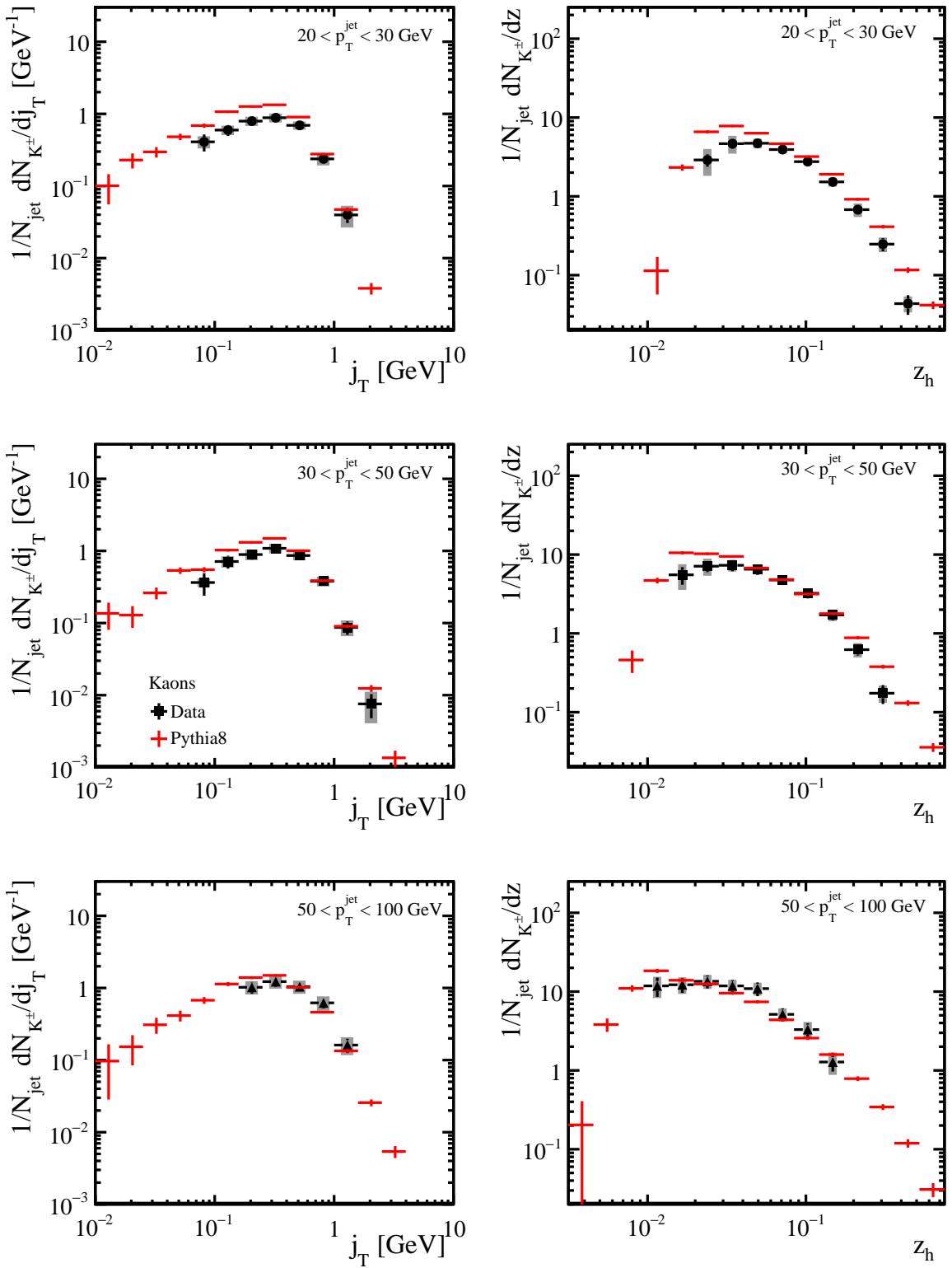


Figure 61: The comparison of the j_T (left) and z (right) distributions for kaons measured in data and Pythia8.

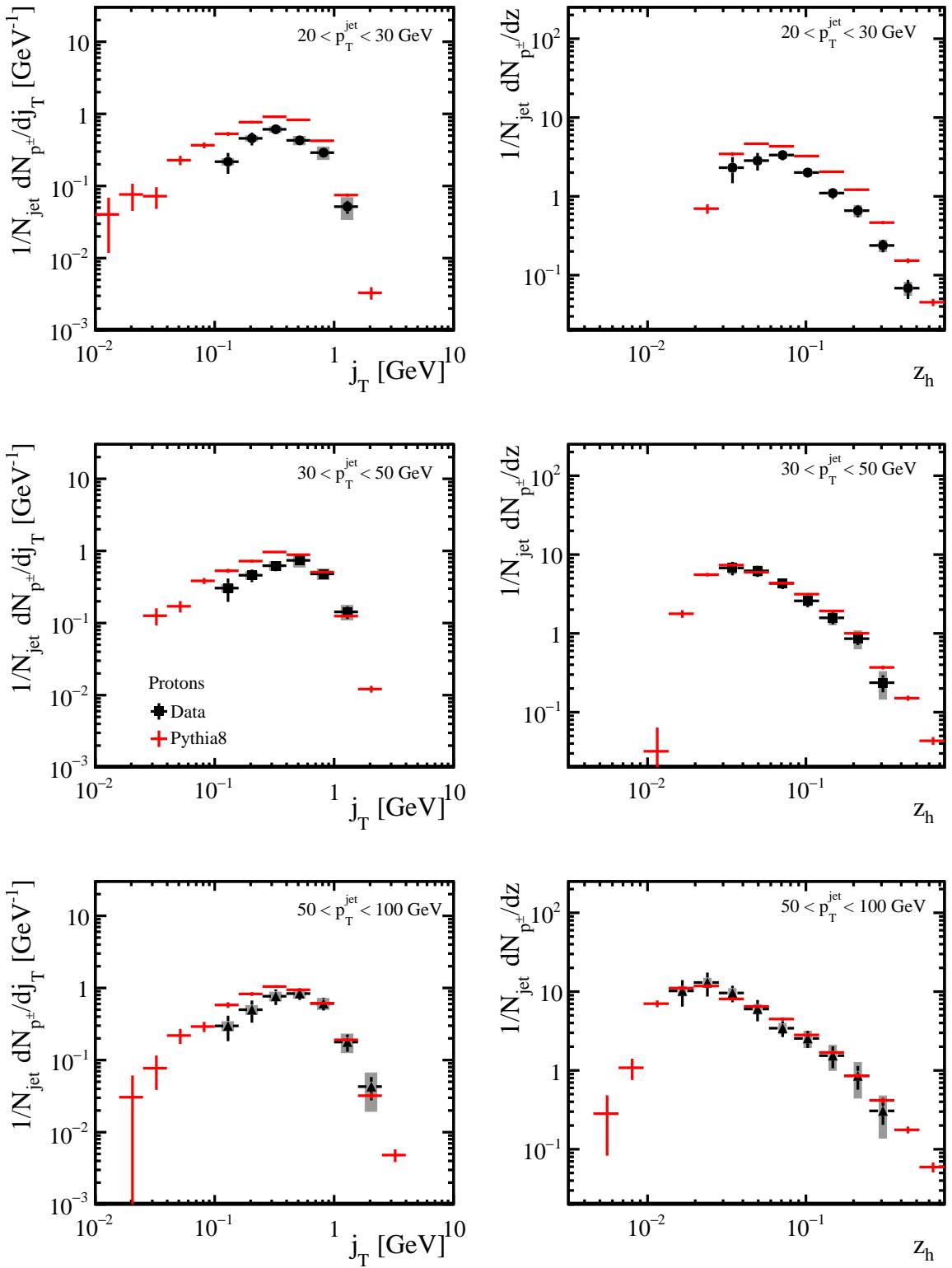


Figure 62: The comparison of the j_T (left) and z (right) distributions for protons measured in data and Pythia8.

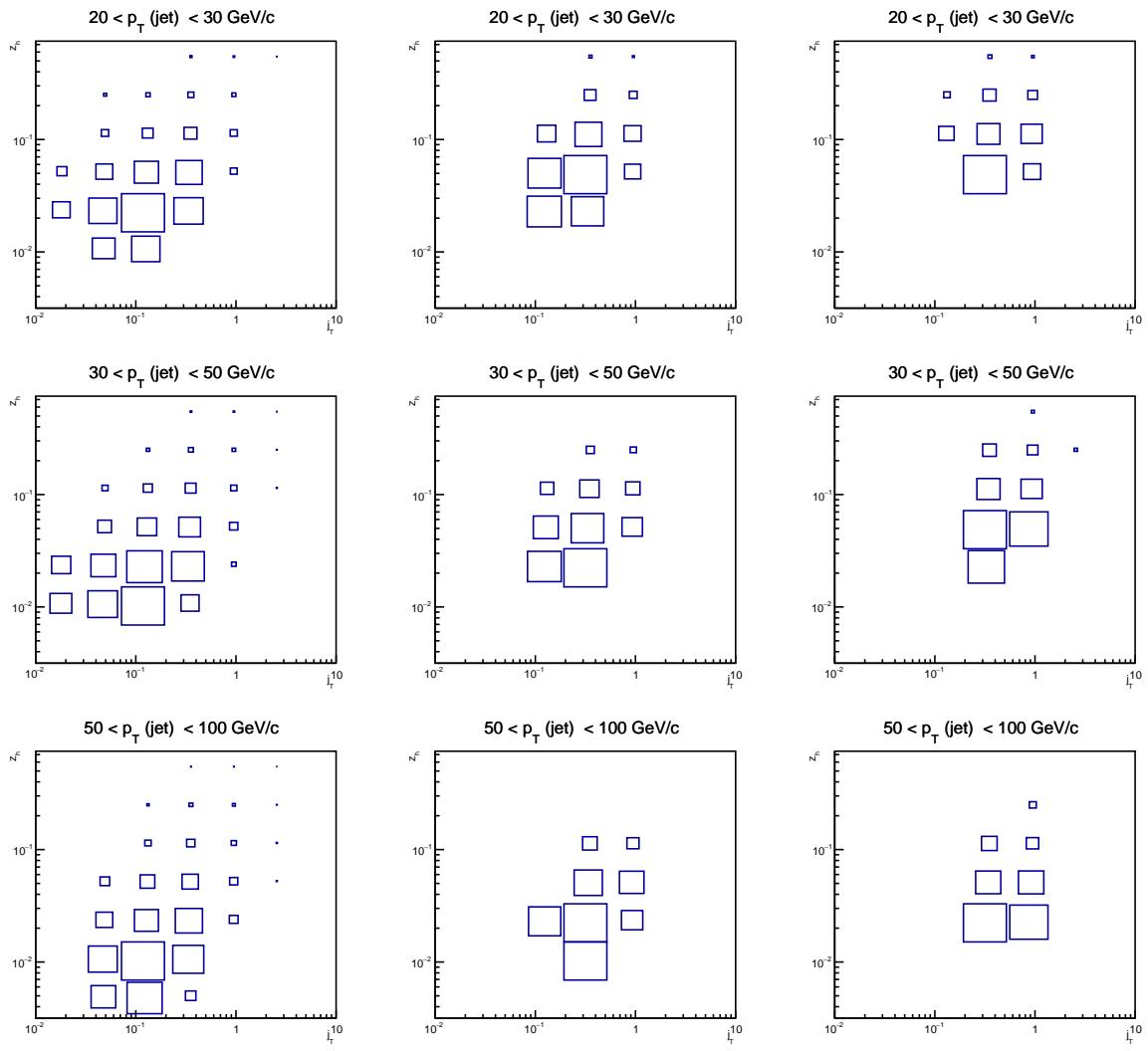


Figure 63: The z vs. j_T distributions for identified pions (left), kaons (middle) and protons (right).

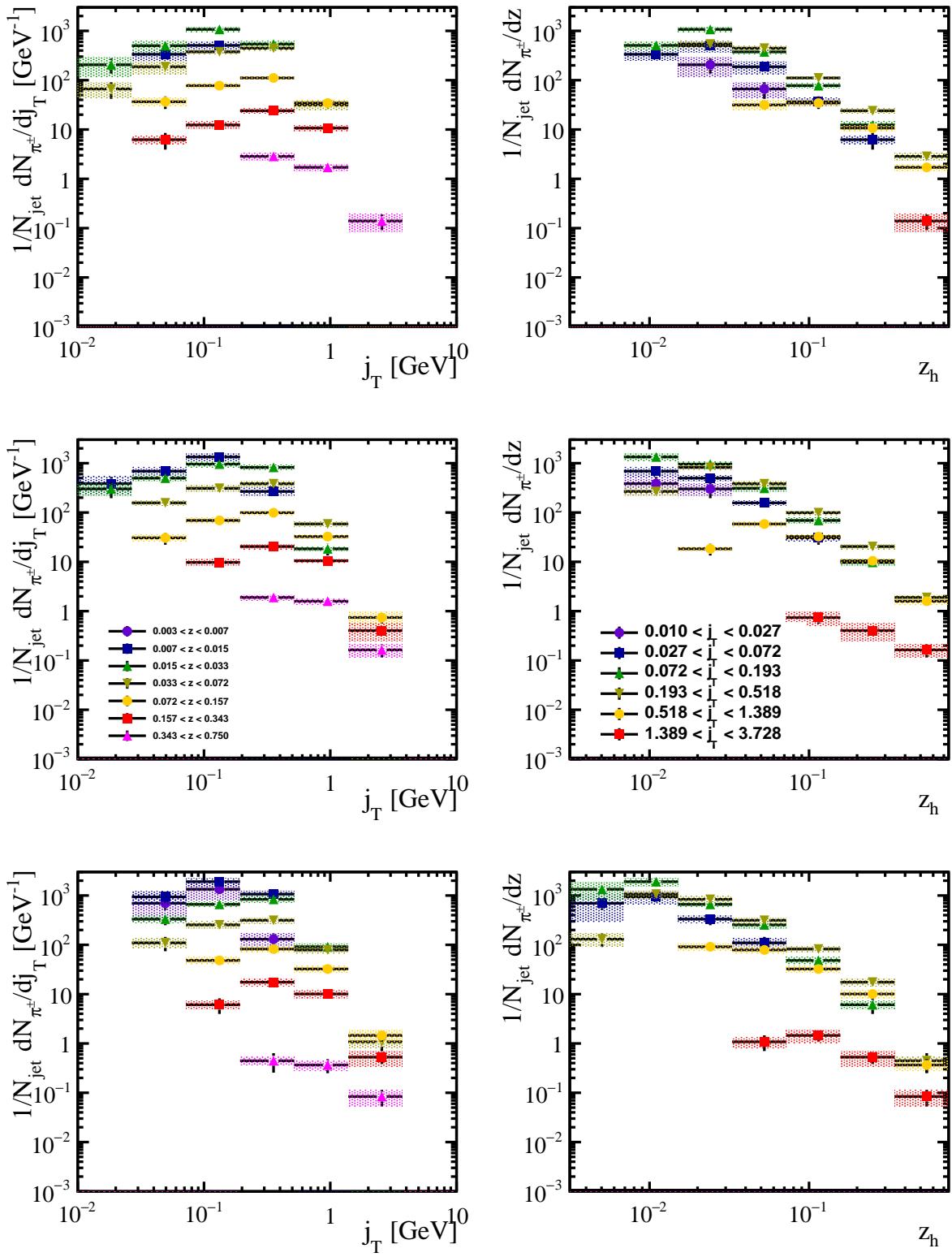


Figure 64: The j_T distributions in z_h bins (left) and z_h distributions in j_T bins (right) for different jet p_T bins for π^\pm .

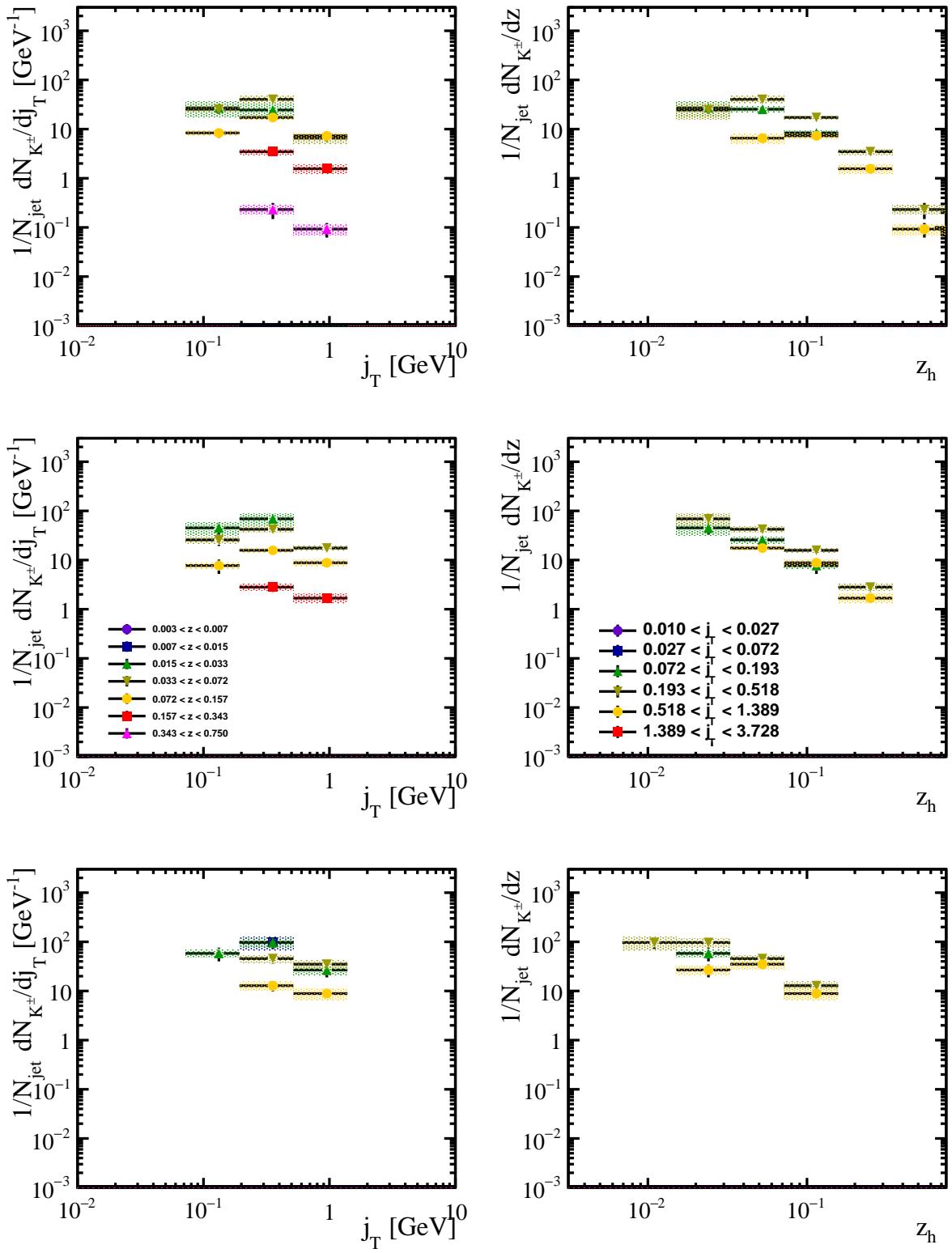


Figure 65: The j_T distributions in z_h bins (left) and z_h distributions in j_T bins (right) for different jet p_T bins for K^\pm .

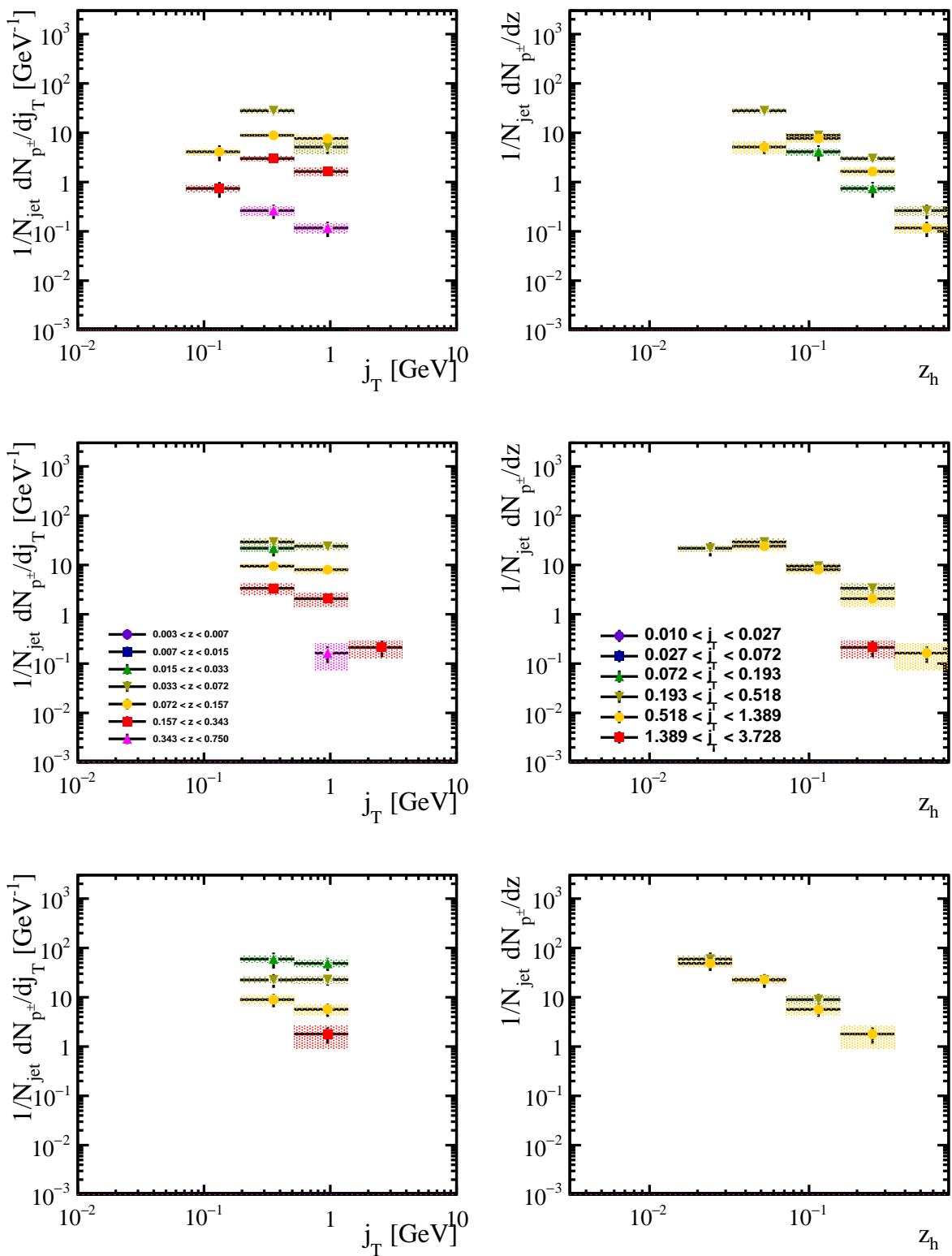


Figure 66: The j_T distributions in z_h bins (left) and z_h distributions in j_T bins (right) for different jet p_T bins for p^\pm .

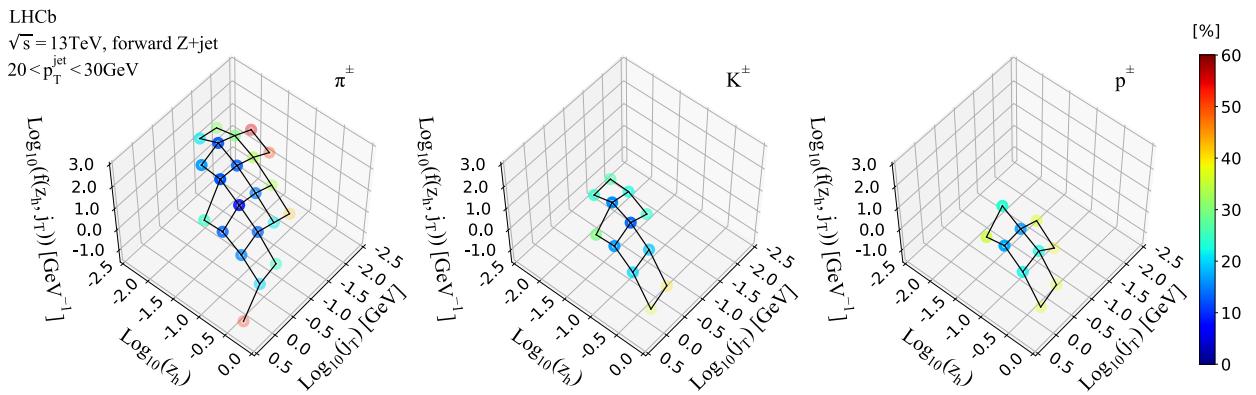


Figure 67: The z_h vs. j_T distributions for pions (left), Kaons (middle) and protons (right) inside jets with p_T between 20 and 30 GeV/c. The quadratic sum of statistical and systematic uncertainties are color-coded in the side bar.

726 References

- 727 [1] LHCb collaboration, R. Aaij *et al.*, *Measurement of charged hadron production in*
728 *Z-tagged jets in proton-proton collisions at $\sqrt{s} = 8 \text{ TeV}$* , Phys. Rev. Lett. **123** (2019)
729 232001, [arXiv:1904.08878](#).
- 730 [2] G. L. Kane, J. Pumplin, and W. Repko, *Transverse Quark Polarization in Large*
731 *p(T) Reactions, e+ e- Jets, and Leptoproduction: A Test of QCD*, Phys. Rev. Lett.
732 **41** (1978) 1689.
- 733 [3] C. A. Aidala, S. D. Bass, D. Hasch, and G. K. Mallot, *The Spin Structure of the*
734 *Nucleon*, Rev. Mod. Phys. **85** (2013) 655, [arXiv:1209.2803](#).
- 735 [4] Z.-B. Kang, K. Lee, J. Terry, and H. Xing, *Jet fragmentation functions for z-tagged*
736 *jets*, Physics Letters B **798** (2019) 134978, [arXiv:1906.07187](#).
- 737 [5] STAR collaboration, R. Aaij *et al.*, *Azimuthal transverse single-spin asymmetries*
738 *of inclusive jets and charged pions within jets from polarized-proton collisions at*
739 *$\sqrt{s} = 500 \text{ GeV}$* , Phys. Rev. D **97** (2018) 032004, [arXiv:1708.07080](#).
- 740 [6] Belle Collaboration, R. Seidl *et al.*, *Transverse momentum dependent production cross*
741 *sections of charged pions, kaons and protons produced in inclusive e^+e^- annihilation*
742 *at $\sqrt{s} = 10.58 \text{ GeV}$* , Phys. Rev. D **99** (2019) 112006, [arXiv:1902.01552](#).
- 743 [7] Z.-B. Kang, X. Liu, S. Mantry, and D. Y. Shao, *Jet Charge: A Flavor Prism*
744 *for Spin Asymmetries at the Electron-Ion Collider*, Phys. Rev. Lett. **125** 242003,
745 [arXiv:2008.00655](#).
- 746 [8] T. C. Rogers and P. J. Mulders, *No Generalized TMD-Factorization in Hadro-*
747 *Production of High Transverse Momentum Hadrons*, Phys. Rev. D **81** (2010) 094006,
748 [arXiv:1001.2977](#).
- 749 [9] LHCb collaboration, R. Aaij *et al.*, *Study of forward Z+jet production in pp collisions*
750 *at $\sqrt{s} = 7 \text{ TeV}$* , JHEP **01** (2014) 033, [arXiv:1310.8197](#).
- 751 [10] LHCb collaboration, R. Aaij *et al.*, *Measurement of forward W and Z boson production*
752 *in association with jets in proton-proton collisions at $\sqrt{s} = 8 \text{ TeV}$* , JHEP **05** (2016)
753 131, [arXiv:1605.00951](#).
- 754 [11] LHCb collaboration, R. Aaij *et al.*, *Study of J/ψ production in jets*, Phys. Rev. Lett.
755 **118** (2017) 192001, [arXiv:1701.05116](#).
- 756 [12] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.2*, Comput.
757 Phys. Commun. **191** (2015) 159, [arXiv:1410.3012](#).
- 758 [13] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth.
759 **A462** (2001) 152.
- 760 [14] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and*
761 *experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

- [15] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- [16] H.-L. Lai *et al.*, *Parton distributions for event generators*, Journal of High Energy Physics **04** (2010) 035.
- [17] *Jet Reconstruction Working Group*, <https://twiki.cern.ch/twiki/bin/view/LHCbPhysics/JetReco>.
- [18] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet user manual*, The European Physical Journal C **72** (2012) , arXiv:hep-ph/1111.6097.
- [19] W. Barter *et al.*, *Jet reconstruction and performances at LHCb*, 2012. LHCb-INT-2012-015.
- [20] M. Santana Rangel, R. Ticse Torres, V. Coco, and C. Potterat, *Jet energy studies for Run I simulation*, 2013. LHCb-INT-2013-036.
- [21] M. Xu, H. Yin, and S. Farry, *The measurements of high- p_T muon reconstruction efficiency for the LHCb Run-2 data-sets*, 2020. LHCb-INT-2020-029.
- [22] J. Van Tilburg, *Track simulation and reconstruction in LHCb*, PhD thesis, NIKHEF, Amsterdam, 2005.
- [23] S. Farry and N. Chiapolini, *A measurement of high- p_T muon reconstruction efficiencies in 2011 and 2012 data*, 2014. LHCb-INT-2014-030.
- [24] L. Anderlini *et al.*, *The PIDCalib package* LHCb-PUB-2016-021. CERN-LHCb-PUB-2016-021, 2016.
- [25] L. Anderlini, V. V. Gligorov, O. Lupton, and B. Sciascia, *Calibration samples for particle identification at LHCb in Run 2* LHCb-PUB-2016-005, 2016.
- [26] A. Andreassen *et al.*, *OmniFold: A Method to Simultaneously Unfold All Observables*, Phys. Rev. Lett. **124** (2020) 182001, arXiv:1911.09107.
- [27] G. D'Agostini, *A multidimensional unfolding method based on bayes' theorem*, Nucl. Instrum. Meth. **A** (1995) 487.
- [28] T. Adye, *Unfolding algorithms and tests using roounfold*, arXiv:1105.1160 (2011).
- [29] C. Aidala *et al.*, *Measurement of the angular coefficients of $\mu^+\mu^-$ pairs in the Z boson mass region at LHCb*, 2020. LHCb-ANA-2020-075.
- [30] W. Barter *et al.*, *Measurement of forward Z boson production in proton-protons collisions at $\sqrt{s} = 13$ TeV*, 2016. LHCb-ANA-2016-026.
- [31] W. Barter, S. Bifani, and S. Farry, *Measurement of forward Z and W production in association with jets in proton-protons collisions at $\sqrt{s} = 8$ TeV*, 2015. LHCb-ANA-2015-067.
- [32] W. Barter, A. Bursche, and D. Ward, *Study of forward Z + jet production in pp collisions at $\sqrt{s} = 7$ TeV*, 2013. LHCb-ANA-2013-056.

- 797 [33] M. De Cian, S. Farry, P. Seyfert, and S. Stahl, *Fast neural-net based fake track*
798 *rejection in the LHCb reconstruction*, 2017. LHCb-PUB-2017-011.
- 799 [34] *PID and CO Working Group*, <https://twiki.cern.ch/twiki/bin/view/LHCbPhysics/ChargedPID>.