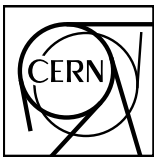
EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH



CERN-EP-2022-242
08 November 2022

Measurement of the angle between jet axes in pp collisions at $\sqrt{s} = 5.02$ TeV

ALICE Collaboration

Abstract

This article reports measurements of the angle between differently defined jet axes in pp collisions at $\sqrt{s} = 5.02$ TeV carried out by the ALICE Collaboration. Charged particles at midrapidity are clustered into jets with resolution parameters $R = 0.2$ and 0.4. The jet axis, before and after Soft Drop grooming, is compared to the jet axis from the Winner-Takes-All (WTA) recombination scheme. The angle between these axes, ΔR_{axis}, probes a wide phase space of the jet formation and evolution, ranging from the initial high-momentum-transfer scattering to the hadronization process. The ΔR_{axis} observable is presented for $20 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 100$ GeV/c, and compared to predictions from the PYTHIA 8 and Herwig 7 event generators. The distributions can also be calculated analytically with a leading hadronization correction related to the non-perturbative component of the Collins–Soper–Sterman (CSS) evolution kernel. Comparisons to analytical predictions at next-to-leading-logarithmic accuracy with leading hadronization correction implemented from experimental extractions of the CSS kernel in Drell–Yan measurements are presented. The analytical predictions describe the measured data within 20% in the perturbative regime, with surprising agreement in the non-perturbative regime as well. These results are compatible with the universality of the CSS kernel in the context of jet substructure.

1 Introduction

Jets play a fundamental role in the study of quantum chromodynamics (QCD). Jets form when partons (quarks and gluons) scattered in high-momentum-transfer (hard) processes fragment into lower-energy (softer) partons. The fragmentation creates a shower of partons, until the average energy per particle falls below the scale at which color-neutral hadrons emerge. Jet substructure, which studies the radiation patterns inside jets, is a prolific field in both experiment [1] and theory [2]. The large difference between the energy scale of the hard-scattered parton and the measured hadrons leaves a significant phase space for jet formation and evolution [3]. Therefore, a multitude of jet-substructure measurements probing different regions of this phase space is needed to characterize the internal substructure of jets and advance our understanding of QCD. Numerous analyses have been carried out by the ALICE [4–12], ATLAS [13–21], CMS [22–29], and LHCb [30, 31] Collaborations at the LHC, as well as at RHIC [32–34].

In this article, a novel jet-substructure observable proposed in Ref. 35 corresponding to the angle between two definitions of the axis of a jet:

$$\Delta R_{\text{axis}} \equiv \sqrt{(y_{\text{axis}\,1} - y_{\text{axis}\,2})^2 + (\varphi_{\text{axis}\,1} - \varphi_{\text{axis}\,2})^2} \tag{1}$$

is studied. Here, a given axis corresponds to a set of coordinates in the rapidity (y) and azimuth (φ) plane. This is illustrated in Fig. 1. The "Standard" axis is determined by clustering the jet constituents with the anti-k_{T} algorithm [36] and the E recombination scheme. Alternatively, the "Groomed" axis can be determined by first using a systematic procedure to remove the soft wide-angle radiation in the jet and then determining again the axis of the constituents that remain after grooming with the anti-k_{T} algorithm and the E recombination scheme. Given the removal of soft wide-angle radiation, this quantity is less sensitive to non-perturbative effects. The grooming procedure used in this analysis is reviewed below.

A third way to define the jet axis corresponds to reclustering the jet (initially clustered with the anti-k_{T} algorithm and E recombination scheme) with the Cambridge–Aachen (C/A) algorithm [37, 38] ensuring that the resulting jet includes all constituents from the original jet. The C/A algorithm clusters particles exclusively based on their spatial separation (in the y–φ plane) without taking into account their energies/momenta. Thus, particles closest in distance are clustered first, which results in an angular-ordered clustering sequence. Subsequently, the constituents are recombined with the Winner-Takes-All (WTA) transverse-momentum recombination scheme [39]. This consists of going through the clustering history and combining the pair of prongs in each $2 \to 1$ merging by assigning to the merged branch the direction of the harder of the two prongs and transverse momentum (p_{T}) corresponding to the sum of

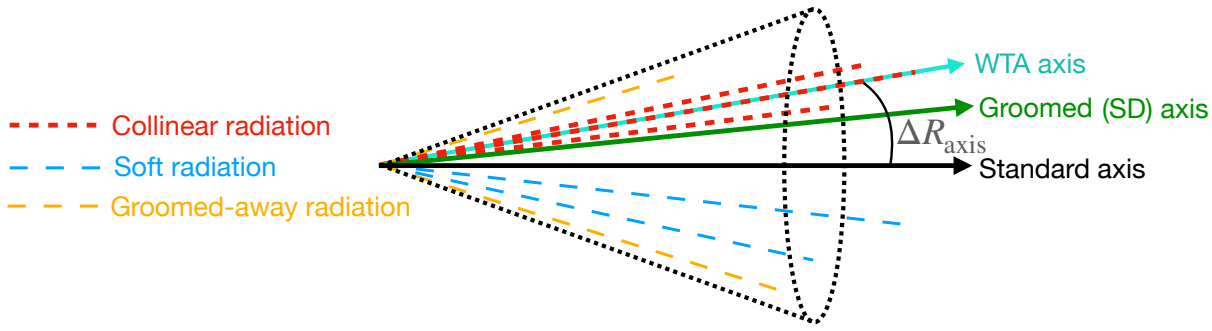


Figure 1: Representation of different jet axes. The colored dashed lines represent particles comprising the jet in the initial sample constructed with the anti-k_{T} algorithm and E recombination scheme. These define the "Standard" jet axis. The grooming procedure removes soft wide-angle radiation (represented by orange lines), and the resulting axis ("Groomed (SD)") is defined by the remaining particles (red and blue lines). Finally, the "WTA" axis, which is determined with all the particles in the initial (ungroomed) jet, tends to be aligned with the most-energetic particle in the jet. The ΔR_{axis} observable is determined from the angle between any pair of these axes.

the two transverse momenta. The WTA scheme is infrared and collinear (IRC) safe and the resulting axis is insensitive to soft radiation at leading power of ΔR_{axis}, which makes it amenable to perturbative calculations [35].

The ΔR_{axis} observable is IRC safe and can be analytically calculated [35]. The sensitivity of ΔR_{axis} to non-perturbative effects can be controlled by changing the pair of axes that are used to determine the observable and also by varying the grooming. For instance, the difference between the Standard and Groomed axes specifically probes the influence of the groomed soft wide-angle radiation on the jet direction. Consequently, this observable can be used both to test the performance of analytic predictions and to constrain non-perturbative models used in event generators.

The ΔR_{axis} is sensitive to Transverse-Momentum-Dependent (TMD) physics [35]. The leading hadronization correction for ΔR_{axis} is related to the non-perturbative component of the Collins–Soper–Sterman (CSS) evolution kernel or rapidity anomalous dimension. Here, leading means logarithmically enhanced in the calculation of ΔR_{axis}, in contrast to other nonperturbative components that are not logarithmically enhanced and thus expected to be small. The CSS kernel is a process-independent non-perturbative function constrained from fits to measured data that governs the non-perturbative evolution in rapidity and encodes information about soft-gluon exchanges between partons in vacuum [40]. As such, extractions from diverse physics processes such as Drell–Yan and semi-inclusive deep-inelastic scattering [41–44] should adequately describe the hadronization corrections for the jet substructure observable presented in this study. This measurement can be included in future global fits to further constrain the CSS kernel. Furthermore, it has been proposed that the CSS kernel can be determined from lattice QCD [45, 46], and such results can be benchmarked with our measurements.

The measurements presented here can also serve as a reference for measurements in heavy-ion collisions. The internal structure of jets fragmenting in the strongly interacting, deconfined state of matter formed in heavy-ion collisions is modified relative to jet fragmentation in vacuum [47, 48]. Comparing to pp collisions allows us to study the medium properties.

In this analysis, the Soft Drop grooming procedure [49] is used to determine the Groomed axis (referred to as "SD" axis from here on). The jet is first reclustered with the C/A algorithm, with the same resolution parameter used in the original clustering of the anti-k_{T}-based sample. This tree is then recursively declustered starting from the largest-distance splitting and at each $1 \to 2$ splitting, the condition

$$\frac{\min(p_{\text{T},1}, p_{\text{T},2})}{p_{\text{T},1} + p_{\text{T},2}} > z_{\text{cut}} \left(\frac{\Delta R_{1,2}}{R} \right)^{\beta} \tag{2}$$

is checked. Here, $p_{\text{T},1}$ and $p_{\text{T},2}$ are the transverse momenta of each prong, and $\Delta R_{1,2}$ is the angular distance between them, calculated as in Eq. 1. The jet resolution parameter is denoted by R. The free parameters z_{cut} and β determine how asymmetric and how wide the splitting can be, respectively. If the condition is not satisfied at a given splitting, the softer branch is removed from the tree and the procedure continues through the harder prong. When a splitting that satisfies the condition is found, the Soft Drop procedure stops and the remaining branch defines the groomed jet. If there isn't a single splitting that satisfies the Soft Drop condition, then the grooming procedure fails. These jets are labeled "untagged". In this analysis, a scan in Soft Drop parameters is carried out, both fixing $z_{\text{cut}} = 0.1$ and varying β from 0 to 3 in increments of 1 and fixing $\beta = 1$ and varying z_{cut} from 0.1 to 0.3 in increments of 0.1.

This article presents the first measurement of the ΔR_{axis} observable, which is carried out in pp collisions at a center-of-mass energy $\sqrt{s} = 5.02$ TeV with the ALICE detectors using charged-particle jets. The document is structured as follows. Section 2 describes the experimental setup used for the measurement and details of the datasets. Section 3 outlines the steps taken in the data analysis. Section 4 describes variations implemented in the analysis to determine systematic uncertainties. Section 5 presents our results and a discussion of the findings. Section 6 summarizes our findings and introduces the conclusions

of this work.

2 Experimental setup and data sets

This analysis uses proton–proton (pp) collisions at $\sqrt{s} = 5.02$ TeV, recorded in 2017 with the ALICE detector at the CERN LHC [50]. A detailed description of the ALICE detector and its performance can be found in Refs. 51, 52. The event sample consists of minimum-bias events triggered by a coincidence of hits in the two V0 scintillator detectors [53], which cover an azimuthal acceptance of $0 < \varphi < 2\pi$ and pseudorapidity $2.8 < \eta < 5.1$ (V0A) and $-3.7 < \eta < -1.7$ (V0C). The primary event vertex is required to be within ± 10 cm of the nominal interaction point. The data sample contains 870 million events, corresponding to an integrated luminosity of 18.0(4) nb^{-1} [54].

Tracks were reconstructed using the Inner Tracking System (ITS) [55] and Time Projection Chamber (TPC) [56]. The track sample consists of two track categories. In the first category, tracks are required to include at least one hit in the silicon pixel detector (SPD) of the ITS, cross at least 70 (out of 159) TPC readout pad rows, and have at least 80% of the geometrically findable space points in the TPC, among other quality criteria [57]. In the second category, tracks that do not contain any SPD hits but otherwise satisfy the track-selection criteria are refit with a constraint to the primary vertex. The inclusion of this second category produces a track sample approximately uniform in azimuth, while preserving a similar p_{T} resolution to tracks with SPD hits. Tracks with $p_{\mathrm{T,track}} > 0.15$ GeV/c, $|\eta| < 0.9$, and $0 < \varphi < 2\pi$ are assigned the charged-pion mass and included in the analysis.

The ALICE detector tracking efficiency in pp collisions grows from $\approx 67\%$ at $p_{\mathrm{T,track}} = 0.15$ GeV/c to $\approx 84\%$ at $p_{\mathrm{T,track}} = 1$ GeV/c, and remains above $\approx 75\%$ for higher $p_{\mathrm{T,track}}$ [51]. The track momentum resolution increases from $\approx 1\%$ at $p_{\mathrm{T,track}} = 1$ GeV/c to $\approx 4\%$ at $p_{\mathrm{T,track}} = 4$ GeV/c.

3 Analysis method

3.1 Jet reconstruction

Charged-particle jets were reconstructed from tracks with $0.15 < p_{\mathrm{T,track}} < 100$ GeV/c, using the anti-k_{T} algorithm [36], and the E recombination scheme, for resolution parameters $R = 0.2$ and 0.4 using the FastJet package [58]. Only jets within the fiducial acceptance of the TPC ($|\eta_{\mathrm{jet}}| < 0.9 - R$) were analyzed, and the resulting distributions are reported in four $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$ intervals with edges at $[20, 40, 60, 80, 100]$ GeV/c. The different axes used to construct the ΔR_{axis} observable are determined from the jets in this sample. The underlying-event was not subtracted.

The jet reconstruction performance was estimated using PYTHIA 8 [59] with Monash 2013 tune [60] events ("truth" level) propagated through a GEANT 3 [61] model of the ALICE detector ("detector" level). In this sample, final-state particles are defined as those with a mean proper lifetime $c\tau > 1$ cm [62]. Particles (tracks) at truth (detector) level were individually clustered into jets, and the jets were matched between the two populations by requiring that the distance between their Standard axes in the y–φ plane satisfied $\Delta R < 0.6R$ and that the match was unique. The matched jets were used to construct the jet energy scale (JES $\equiv (p_{\mathrm{T,det}}^{\mathrm{ch\,jet}} - p_{\mathrm{T,truth}}^{\mathrm{ch\,jet}})/p_{\mathrm{T,truth}}^{\mathrm{ch\,jet}}$), jet energy resolution (JER $\equiv \sigma(p_{\mathrm{T,det}}^{\mathrm{ch\,jet}})/p_{\mathrm{T,truth}}^{\mathrm{ch\,jet}}$), and jet reconstruction efficiency ($\varepsilon_{\mathrm{reco}}$), where $p_{\mathrm{T,truth}}^{\mathrm{ch\,jet}}$ and $p_{\mathrm{T,det}}^{\mathrm{ch\,jet}}$ refer to the transverse momentum of the truth- and detector-level jets, respectively, and $\sigma(p_{\mathrm{T,det}}^{\mathrm{ch\,jet}})$ corresponds to the standard deviation of the $p_{\mathrm{T,det}}^{\mathrm{ch\,jet}}$ spectrum for a given value of $p_{\mathrm{T,truth}}^{\mathrm{ch\,jet}}$. Table 1 shows values characterizing the jet-reconstruction performance. In the case of the JES, this distribution is peaked at 0, but has an asymmetric tail that shifts the mean value, Δ_{JES}, due to tracking inefficiencies. The unfolding procedure described in the next section corrects for these JES and JER effects.

Table 1: Approximate values for the figures of merit characterizing the jet-reconstruction performance in this analysis. Δ_{JES} is the mean value of the JES spectrum. See text for details.

R	$p_{\mathrm{T}}^{\mathrm{ch\,jet}}$ (GeV/c)	Δ_{JES} (%)	JER (%)	$\varepsilon_{\mathrm{reco}}$ (%)
0.4	20	−13	21	97
	40	−16	21	100
	60	−19	22	100
	80	−21	23	100
	100	−25	24	100
0.2	20	−12	21	94
	40	−16	23	100
	60	−20	24	100
	80	−23	25	100
	100	−27	26	100

3.2 Corrections

To obtain distributions free of detector effects (truth level), the measured spectra (detector level) were unfolded using an iterative procedure based on Bayes' theorem [63, 64], implemented in the RooUnfold package [65]. This unfolding procedure accounts for effects such as track p_{T} resolution, tracking inefficiencies, and particle interactions in the detector volume. The input to this procedure consisted of a 4-dimensional response matrix (RM) that maps the correspondence between detector and truth levels for $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$ and ΔR_{axis}. The RM was created using the simulated jets described in the previous section. The $p_{\mathrm{T,det}}^{\mathrm{ch\,jet}}$ axes were constructed in the range $[10, 130]$ GeV/c to capture bin-migration effects. In the ΔR_{axis} cases in which one of the two axes is the SD axis, the untagged jets were included in the unfolding procedure. The unfolding iterations were fixed at the number that minimizes the quadratic sum of the statistical and systematic uncertainties. In some cases, a few more iterations were carried out until a convergence within 5% between subsequent iterations was achieved.

The unfolding procedure was validated by performing a series of closure tests. A refolding test was performed to check that the detector-level spectrum can be recovered by reversing the unfolding procedure. In this test, the unfolded result was multiplied by the response matrix and the resulting spectrum was compared to the detector-level spectrum. Additionally, a statistical-closure test was performed to confirm that the unfolding procedure is robust against statistical fluctuations in the data. In this test, the simulated detector-level spectrum was smeared by an amount equal to the statistical uncertainty of the measured data. Subsequently, the smeared spectrum was unfolded, and the agreement between the resulting distribution and the truth-level simulated distribution was assessed. Finally, a shape-closure test was performed to account for the fact that the true detector-level distribution may be different than that from the generator, and evaluate the robustness against this shape. In this test, the shapes of the simulated detector- and truth-level spectra were scaled by the prior-scaling function described in the systematic-uncertainty section. Subsequently, the scaled detector-level spectrum was unfolded, and the resulting spectrum was compared to the scaled truth-level distribution. In all these tests, closure within the statistical uncertainties was achieved.

4 Systematic uncertainties

Three sources of systematic uncertainties are considered in this analysis: those arising from the unfolding procedure, uncertainties in the tracking efficiency, and the choice of event generator used to produce the response matrix.

The unfolding uncertainty is estimated by performing four variations of the unfolding procedure:

1. The analysis is repeated with a number of iterations varied by ± 2 around the nominal value and the average difference with respect to the nominal spectrum is taken as an uncertainty.
2. The analysis is repeated with a prior distribution multiplied by $(p_{\mathrm{T}}^{\mathrm{ch\,jet}})^{\pm 0.5} \times (\pm A\Delta R_{\mathrm{axis}} \mp B)$, where the parameters A and B are selected for each ΔR_{axis} variation such that the endpoints of the spectrum are scaled by $\approx \pm 20\%$. The maximum difference between these two variations and the nominal result is taken as an uncertainty.
3. The analysis is repeated with the transverse-momentum range $p_{\mathrm{T,det}}^{\mathrm{ch\,jet}} \in [5, 135]$ GeV/c and the difference with respect to the nominal analysis ($p_{\mathrm{T,det}}^{\mathrm{ch\,jet}} \in [10, 130]$ GeV/c) is taken as an uncertainty.
4. The analysis is repeated with a different ΔR_{axis} binning at detector level (slightly finer or coarser granularity) and the difference with respect to the main result is taken as an uncertainty.

Given that these four variations probe the same underlying source of uncertainty, the total unfolding uncertainty is defined as the standard deviation $\sigma_{\mathrm{unfolding}} \equiv \sqrt{\sum_{i=1}^{4} \sigma_i^2/4}$, where σ_i corresponds to the uncertainty from the individual variations.

The uncertainty on the ALICE tracking efficiency is 3–4%, determined by varying the track selection parameters and possible imperfections in the description of the ITS–TPC matching efficiency in the simulation. Consequently, the analysis was repeated with a RM populated with jets clustered over a track sample with 4% of all tracks randomly discarded, taking the more conservative value. Differences with respect to the nominal analysis were assigned as the systematic uncertainty due to the tracking efficiency.

To assess the model dependence of the analysis on the generated spectra used to compute the RM (based on PYTHIA 8 with Monash 2013 tune events propagated through GEANT 3), the analysis was repeated using RMs based on Herwig 7 (default tune) and PYTHIA 8 with Monash 2013 tune constructed using a fast simulation, and the difference between the two unfolded results was assigned as a systematic uncertainty.

The total systematic uncertainty is taken as the sum in quadrature of the contributions due to the unfolding, tracking efficiency, and choice of event generator. Table 2 summarizes the range of the systematic uncertainties in different ΔR_{axis} intervals for jets of $R = 0.2$ and 0.4 in the transverse momentum range $40 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 60$ GeV/c. Most often, the dominant systematic uncertainty originates from the uncertainty on the tracking efficiency.

5 Results and discussion

The ΔR_{axis} distributions are reported as normalized differential cross sections:

$$\frac{1}{\sigma_{\mathrm{jets}}(p_{\mathrm{T}}^{\mathrm{ch\,jet}})} \frac{\mathrm{d}\sigma}{\mathrm{d}\Delta R_{\mathrm{axis}}} \left(p_{\mathrm{T}}^{\mathrm{ch\,jet}}\right) \equiv \frac{1}{N_{\mathrm{jets}}(p_{\mathrm{T}}^{\mathrm{ch\,jet}})} \frac{\mathrm{d}N}{\mathrm{d}\Delta R_{\mathrm{axis}}} \left(p_{\mathrm{T}}^{\mathrm{ch\,jet}}\right), \tag{3}$$

for jets of $R = 0.2$ and 0.4 in 20-GeV/c-wide $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$ intervals in the 20–100 GeV/c range. Here, $N_{\mathrm{jets}}(p_{\mathrm{T}}^{\mathrm{ch\,jet}})$ is the number of inclusive jets in a given event sample in the transverse momentum interval centered at $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$. In the groomed cases, the normalization factor N_{jets} is obtained by including the number of untagged jets (i.e. without any splitting in the jet that satisfies the SD condition, so that the grooming process fails) in the unfolding procedure. Therefore, the number of jets appearing in the ΔR_{axis} distributions differs from N_{jets} by the number of jets that are untagged.

Table 2: Range of the estimated value of relative systematic uncertainties in ΔR_{axis} intervals in the transverse momentum range $40 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 60$ GeV/c. The unfolding, tracking efficiency, and generator systematic uncertainties can be found in the columns labeled Unf., Trk. Eff., and Gen., respectively. In the case of the groomed observables, the grooming parameters are specified as $(z_{\mathrm{cut}}, \beta)$. The displayed uncertainties correspond to the lowest and highest values for a given setting.

	$R = 0.2$				$R = 0.4$			
ΔR_{axis}	Unf.	Trk. Eff.	Gen.	Total	Unf.	Trk. Eff.	Gen.	Total
WTA–Standard	0–3%	0–8%	0–5%	1–10%	1–4%	0–6%	0–4%	1–6%
WTA–SD (0.1, 0)	1–4%	0–4%	0–3%	2–6%	1–4%	1–4%	0–4%	1–6%
WTA–SD (0.1, 1)	0–3%	0–5%	0–3%	1–5%	1–4%	0–4%	0–4%	2–6%
WTA–SD (0.1, 2)	0–2%	0–5%	0–3%	1–5%	1–4%	1–5%	0–4%	2–6%
WTA–SD (0.1, 3)	0–2%	0–5%	0–3%	1–6%	1–3%	0–6%	0–3%	1–6%
WTA–SD (0.2, 1)	0–4%	0–5%	0–4%	2–6%	1–4%	1–4%	0–4%	2–6%
WTA–SD (0.3, 1)	1–5%	1–4%	0–4%	2–6%	0–3%	1–4%	0–5%	1–6%
Standard–SD (0.1, 0)	1–6%	0–3%	0–2%	1–7%	1–3%	0–4%	0–2%	2–4%
Standard–SD (0.1, 1)	0–4%	0–5%	0–2%	1–6%	1–4%	0–5%	1–2%	2–7%
Standard–SD (0.1, 2)	1–4%	1–5%	0–2%	1–7%	1–4%	0–5%	0–2%	1–6%
Standard–SD (0.1, 3)	1–5%	1–5%	0–3%	1–7%	1–3%	1–4%	0–4%	2–5%
Standard–SD (0.2, 1)	1–4%	0–4%	0–2%	2–6%	0–2%	0–4%	0–3%	1–5%
Standard–SD (0.3, 1)	2–4%	1–4%	0–2%	2–6%	0–3%	0–5%	0–3%	1–6%

Sections 5.1 and 5.2 present the experimental results as well as comparisons to predictions from MC generators and analytical calculations, respectively.

5.1 Comparison to MC generators

Figure 2 compares the measured Standard–SD distributions with predictions from Monte Carlo event generators, for jets of $R = 0.4$ (top) and 0.2 (bottom) in $40 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 60$ GeV/c. Results for different $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$ intervals are presented in Appendix A. The vertical error bars are statistical uncertainties, and the rectangles indicate the total systematic uncertainties. The solid (dashed) lines show results from PYTHIA 8 (Herwig 7). The bottom two panels in these figures correspond to the ratios of data to PYTHIA 8 and Herwig 7 distributions. The left panel of Fig. 2 shows the case in which z_{cut} is fixed at 0.1 and β is varied from 0 to 3. In the right panel, β is fixed at 1 and z_{cut} is varied from 0.1 to 0.3. These spectra are plotted with a logarithmic y-axis to better exhibit the entire distribution. Figure 3 shows the equivalent comparison for the case of the WTA-relative distributions.

Overall, the Standard–SD distributions are narrow and peaked at very small values. This implies that the Standard and SD axes are aligned and grooming does not significantly impact the jet direction. However, as the grooming becomes more aggressive (i.e. higher z_{cut} or smaller β), the alignment between the Standard and SD axes worsens somewhat. This trend is present for both $R = 0.2$ and R = 0.4. There is a maximum at $\Delta R_{\mathrm{axis}}^{\mathrm{Standard-SD}} = 0$ that corresponds to jets for which the first splitting after reclustering the jet with the C/A algorithm already satisfies the SD condition. As a result, the difference with respect to the Standard axis is exactly 0.

The shape of the $\Delta R_{\mathrm{axis}}^{\mathrm{Standard-SD}}$ spectra is better described by Herwig 7 than by PYTHIA 8. In the $40 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 60$ GeV/c range, bin-by-bin deviations with respect to PYTHIA 8 (Herwig 7) reach values up to $\approx 24\%$ ($\approx 7\%$). Given that this observable is particularly sensitive to soft-radiation effects, our data can be used to further constrain the hadronization models in these generators.

The distributions relative to the WTA axis are broader ($0 < \Delta R_{\mathrm{axis}} \lesssim R/2$) and peak at larger ΔR_{axis}, showing substantial deviation between the WTA and Standard/SD jet axes. Additionally, these distributions show very low sensitivity to the parameters chosen in the Soft Drop grooming procedure.

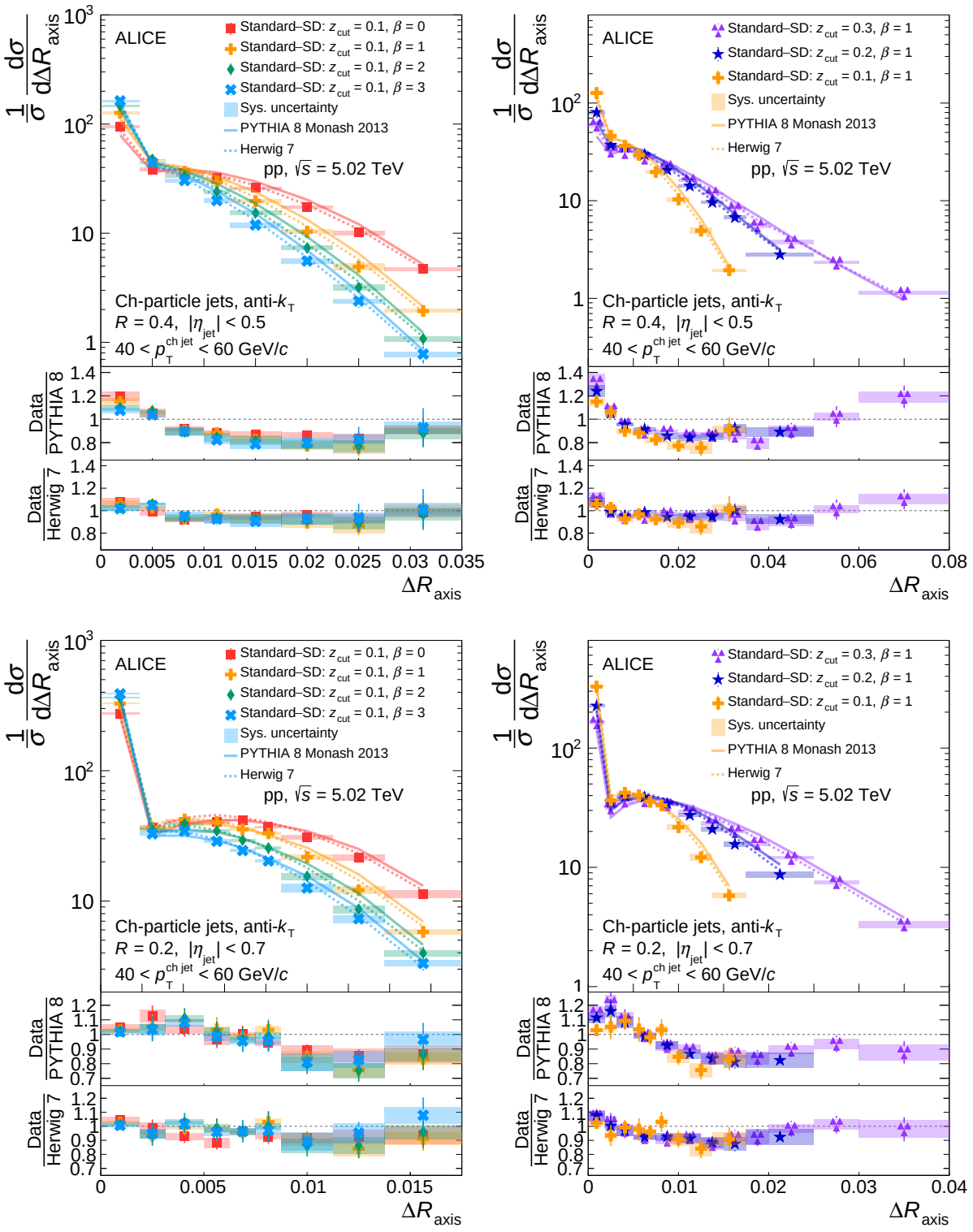


Figure 2: Comparison between Standard–SD measured distributions and Monte Carlo event generators for jets of $R = 0.4$ (top) and 0.2 (bottom) in $40 < p_T^{ch\ jet} < 60$ GeV/c. Left: distributions with $z_{cut} = 0.1$ and varying β. Right: distributions with $\beta = 1$ and varying z_{cut}.

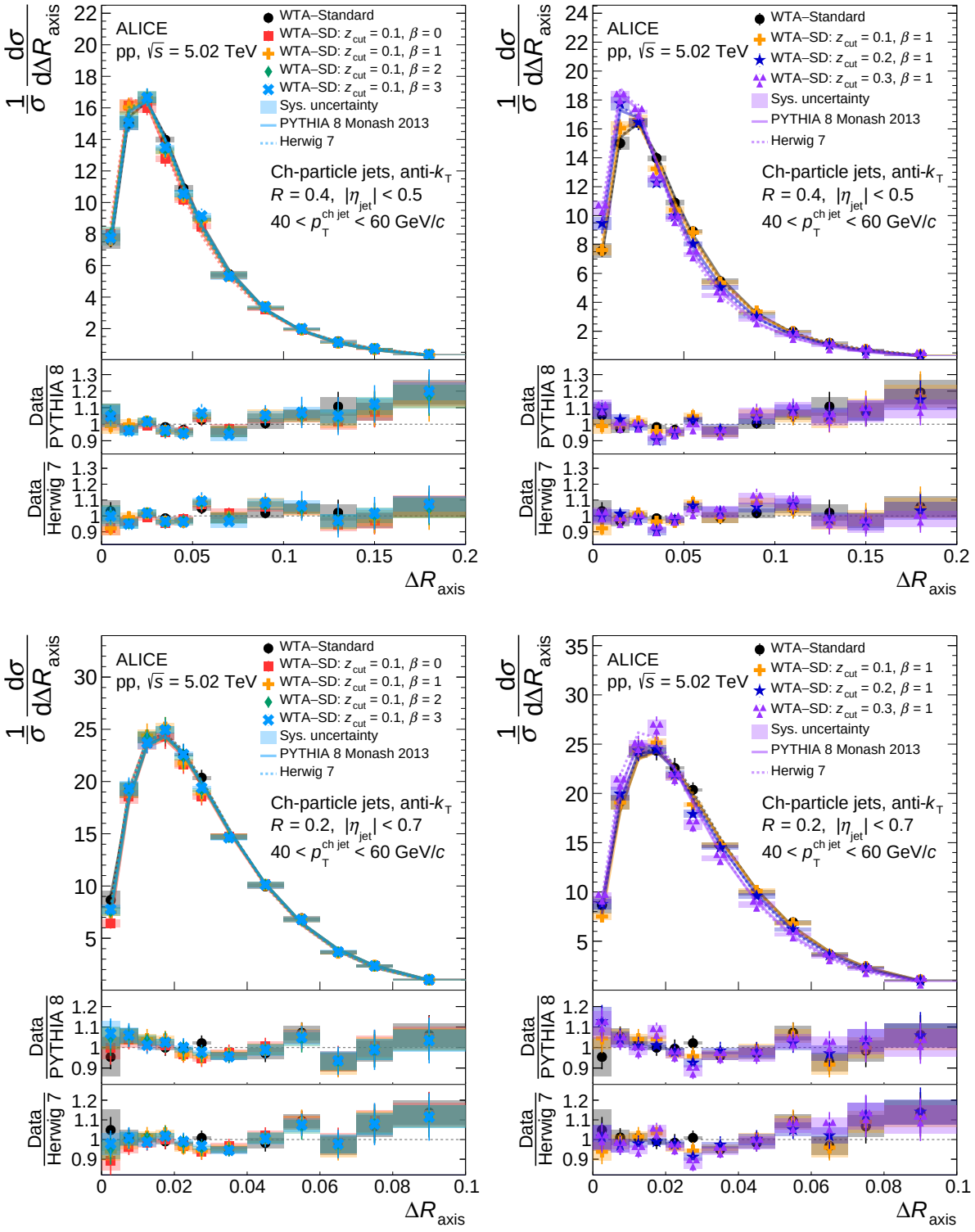


Figure 3: Comparison between WTA-relative measured distributions and Monte Carlo event generators for jets of $R = 0.4$ (top) and 0.2 (bottom) in $40 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 60$ GeV/c. Left: distributions with $z_{\mathrm{cut}} = 0.1$ and varying β. Right: distributions with $\beta = 1$ and varying z_{cut}.

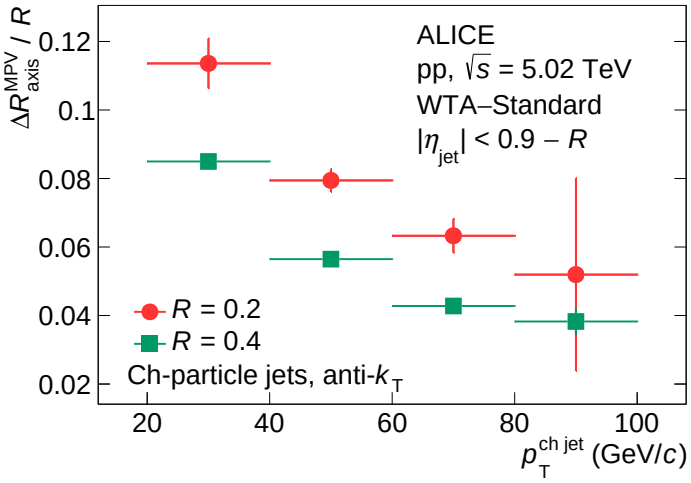


Figure 4: Distribution of the most-probable value of the WTA-Standard spectra as a function of $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ for $R = 0.2$ and $R = 0.4$.

The $\Delta R_{\mathrm{axis}}^{\mathrm{WTA-Standard}}$ and $\Delta R_{\mathrm{axis}}^{\mathrm{WTA-SD}}$ predictions from PYTHIA 8 and Herwig 7 generally agree with the data, deviating by up to $\approx 10\%$ in the $40 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 60$ GeV/c interval and up to $\approx 40\%$ in several other $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ intervals. The largest deviations are at low values of ΔR_{axis}, corresponding to the region in which non-perturbative effects are significant. It should be noted, though, that the normalization convention implies that conclusions can only be made about the overall shape and not about a discrepancy in a specific ΔR_{axis} interval.

The axis differences with respect to WTA shift to lower ΔR_{axis} values at higher $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$, implying that the WTA and Standard or SD axes are more aligned in more energetic jets. This is summarized in Fig. 4, where the most-probable value (MPV) of the $\Delta R_{\mathrm{axis}}^{\mathrm{WTA-Standard}}$ distribution normalized to the jet resolution parameter R is shown as a function of $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$. The values in this figure are determined by repeatedly fitting a Landau distribution to the $\Delta R_{\mathrm{axis}}^{\mathrm{WTA-Standard}}$ spectra changing each time the fit range. A histogram is then filled with the peak-position values extracted from each fit. From this histogram, the mean and standard deviation are extracted as the central value and uncertainty, respectively. The values decrease from $\approx 11\%$ (8.5%) at low $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ to $\approx 5\%$ (4%) at high $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ for $R = 0.2$ (0.4).

5.2 Comparison to analytical calculations

The ΔR_{axis} observable has been calculated in the Soft Collinear Effective Theory (SCET) framework [66]. In this framework, the jet-production cross section is factorized into parton distribution functions (PDF), "hard", and "jet" contributions in order to separate physics processes at different scales. The PDFs encode the probability of finding a parton with a given flavor and momentum fraction from a proton and are non-perturbative objects extracted from global fits to measured data (see, e.g. Refs. 67, 68). The SCET calculations presented here use the CT14 NLO PDF set [69]. The hard contribution encodes the short-distance physics, i.e. the hard scattering of one parton from each colliding proton, and the distribution of the resulting partons. Finally, the jet function describes the evolution of a final-state parton from the hard scattering into a collimated jet. Large logarithms are resummed to Next-to-Leading Logarithmic (NLL) accuracy for the Standard–SD ΔR_{axis}, and NLL$'$ for the WTA–Standard and WTA–SD ΔR_{axis}, including the contribution from non-global logarithms [70]. NLL$'$ refers to the inclusion in the resummation of terms that formally only enter at Next-to-Next-to-Leading Logarithmic (NNLL) accuracy, but their inclusion in the NLL calculation improves the theoretical uncertainty. Specifically, logarithms of the jet resolution parameter R, the grooming parameter z_{cut}, and the observable ΔR_{axis} are resummed. The resummed result is presented without matching to the fixed-order calculation in the high-ΔR_{axis} region.

Unlike other TMD-sensitive jet substructure observables, such as hadron-in-jet fragmentation [71, 72], ΔR_{axis} does not depend on collinear fragmentation functions. Thus, a significant source of uncertainty is removed from these calculations.

The leading hadronization correction for these observables includes terms of the form $\exp[-g_K(b_\perp; b_\perp^{\mathrm{max}})]$ [35], where $b_\perp$ is the Fourier conjugate of k_{T}, the projection of the transverse momentum of a jet axis transverse to the other axis in ΔR_{axis}. The function $g_K(b_\perp; b_\perp^{\mathrm{max}})$ is the non-perturbative component of the Collins–Soper–Sterman (CSS) evolution kernel or rapidity anomalous dimension. In the so-called b_* prescription [73, 74] it is often parametrized as

$$g_K(b_\perp; b_\perp^{\mathrm{max}}) = g_2(b_\perp^{\mathrm{max}}) b_\perp^2. \tag{4}$$

In this work, the universality of the CSS kernel is tested by verifying that the ΔR_{axis} observables are well described across several resolution parameters R, grooming settings, and $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$ intervals with the same parameters in $g_K(b_\perp; b_\perp^{\mathrm{max}})$ corresponding to $b_\perp^{\mathrm{max}} = 1.5\,(\mathrm{GeV}/c)^{-1}$ and $g_2(b_\perp^{\mathrm{max}}) = 0.18$ from a global data analysis to Drell–Yan lepton pair and Z^0-boson production [75].

The analytic predictions are provided by the Authors of Ref. 35 for the kinematics of our measurement at hadron level (i.e. after hadronization, in contrast to parton level) for full jets (i.e. including both charged and neutral hadrons), and without including background effects. To do a comparison to the measured distributions, the analytic predictions are corrected using data from Monte Carlo event generators. Final-state particles from pp events are clustered into full and charged-particle jets following the same procedure as in the data analysis (i.e. using the anti-k_{T} algorithm with the E recombination scheme for a given R), and the resulting jets are required to satisfy $|\eta_{\mathrm{jet}}| < 0.9 - R$ and $p_{\mathrm{T}}^{\mathrm{jet}} > 5$ GeV/c. The full and charged-particle jets are then geometrically matched following the matching procedure of the RM from the data analysis. This sample is used to construct a 4D response matrix that maps the $p_{\mathrm{T}}^{\mathrm{jet}}$ and ΔR_{axis} dependence from full- to charged-particle-jet levels for each observable.

The analytic predictions are provided as normalized densities $(1/\sigma_{\mathrm{jet}})(\mathrm{d}\sigma_{\mathrm{jet}}/\mathrm{d}\Delta R_{\mathrm{axis}})$ for the ΔR_{axis} observable in different $p_{\mathrm{T}}^{\mathrm{jet}}$ intervals (of 5 GeV/c width) for the nominal case, as well as for the results obtained by systematically varying the scales that appear in the calculation to account for theoretical uncertainties. The first step in the correction corresponds to multiplying the spectrum by the average value of the inclusive cross section in the considered $p_{\mathrm{T}}^{\mathrm{jet}}$ interval to obtain the distribution for $\mathrm{d}\sigma_{\mathrm{jet}}/\mathrm{d}\Delta R_{\mathrm{axis}}$. The cross section used was calculated at Next-to-Leading Order (NLO) with resummation of logarithms of the jet radius at NLL [76]. The scaled distributions are stored in 2D histograms in p_{T} and ΔR_{axis} and multiplied by the 4D response matrices described in the previous paragraph to obtain the analytic predictions for charged-particle jets. Subsequently, the resulting 2D histograms are corrected bin-by-bin to account for multi-parton interactions (MPI), and projected onto the observable axis for a range in $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$ equal to that of the measured distribution. The final theory prediction corresponds to the curve obtained from the nominal calculation. Additionally, the other results obtained from the systematic scale variations are equally corrected, and the envelope of all resulting distributions is taken as the theoretical uncertainty on the calculations.

The use of a Monte Carlo event generator in these corrections introduces a model dependence, the significance of which is explored by applying the correction procedure with two different generators: PYTHIA 8 with the Monash 2013 tune [60] and Herwig 7 with the default tune [77].

Figure 5 shows a subset of the comparison between the measured distributions and the corrected analytic predictions. Equivalent comparisons for the rest of all available predictions are presented in Appendix B. The black markers correspond to the distributions determined from measured data. These distributions are identical to those in Fig. 3 up to a normalization factor defined below. The vertical error bars correspond to the statistical uncertainties, and the rectangles correspond to the total systematic uncertainties. The colored curves correspond to the SCET-based analytic predictions corrected for charge and MPI

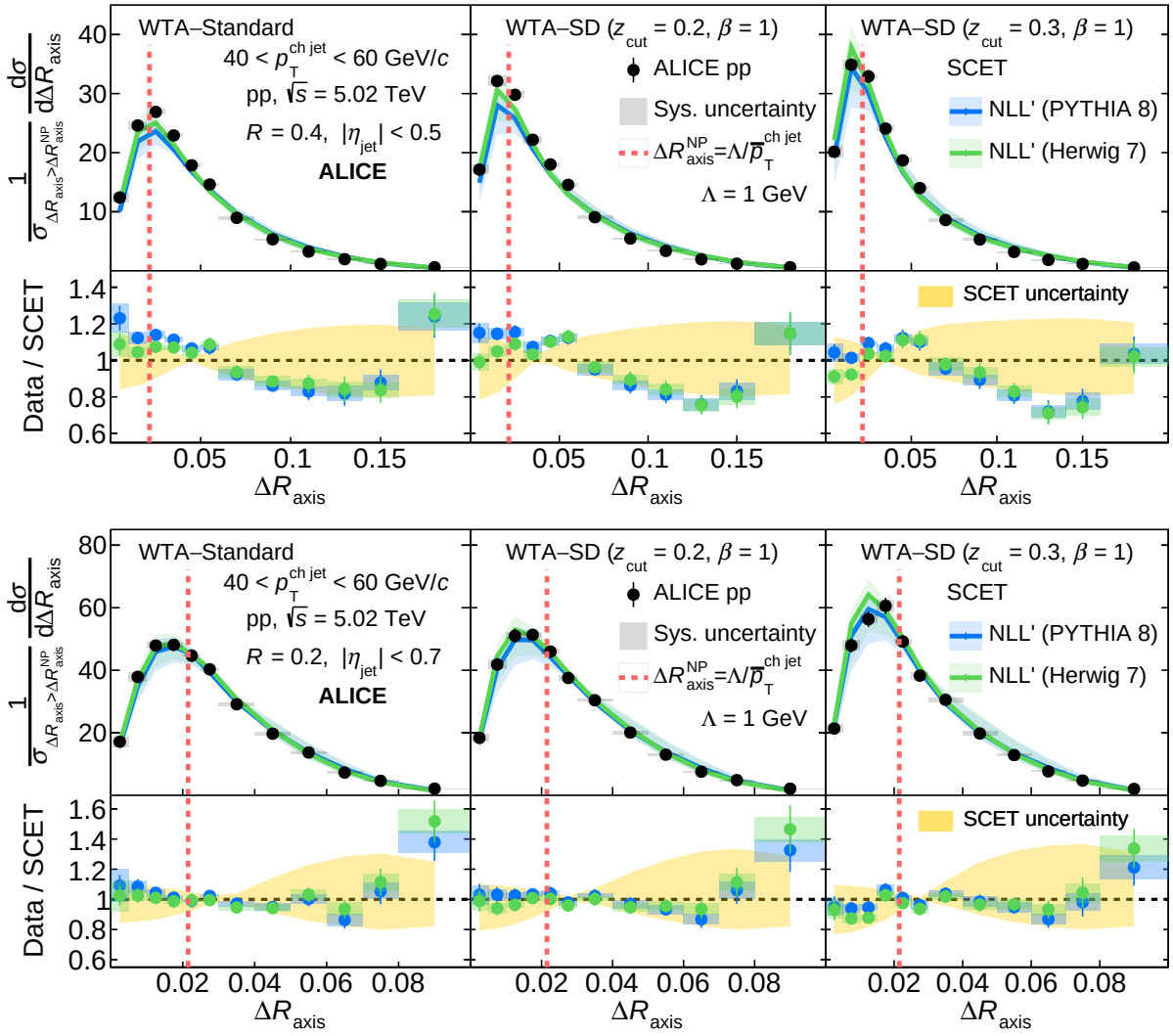


Figure 5: Comparison between measured distributions and analytic predictions for the difference between the WTA axis and the Standard (left), SD ($z_{\rm cut} = 0.2, \beta = 1$) (center), and SD ($z_{\rm cut} = 0.3, \beta = 1$) (right) axes for jets of $R = 0.4$ (top) and 0.2 (bottom) in the transverse momentum range $40 < p_{\rm T}^{\rm ch\ jet} < 60$ GeV/c. The black markers correspond to the distributions determined from measured data. The vertical error bars correspond to the statistical uncertainties, and the rectangles correspond to the total systematic uncertainties. The colored curves correspond to the SCET-based analytic predictions corrected for charge and MPI effects using two event generators (PYTHIA 8 and Herwig 7). The vertical dashed line defines the approximate boundary between the non-perturbative and perturbative regions. Both the measured and analytic predictions are normalized so that the integral $\int_{\Delta R_{\rm axis}^{\rm NP}}^{R/2} {\rm d}\Delta R_{\rm axis} ({\rm d}\sigma/{\rm d}\Delta R_{\rm axis}) = 1$. The bottom panels show the data/SCET ratios. The colored rectangles correspond to the systematic uncertainty of the measured distribution. The size of the theoretical uncertainty in the analytic predictions is shown as a yellow band.

effects using two event generators (PYTHIA 8 and Herwig 7).

Differences between the SCET predictions corrected with either Monte Carlo event generator are very small. This is due to the fact that, since the input calculations are provided at hadron level, the most significant correction is done to the $p_{\rm T}$ scale of the jet, and this correction is well modeled by both generators. Thus, the resulting distributions are not significantly model dependent.

The analytic calculations are only expected to describe the measured distributions in the perturbative region. The predictions presented here become non-perturbative approximately at $\Delta R_{\rm axis} \lesssim \Delta R_{\rm axis}^{\rm NP} =$

$\Lambda/p_{\mathrm{T}}^{\mathrm{ch\,jet}}$ [35], where Λ corresponds to the scale at which the strong coupling constant becomes non-perturbative. The red vertical line in Fig. 5 corresponds to this value calculated using $\Lambda = 1$ GeV and $p_{\mathrm{T}}^{\mathrm{ch\,jet}} = \overline{p}_{\mathrm{T}}^{\mathrm{ch\,jet}}$, the average jet-transverse-momentum value in the interval. The comparison is done by normalizing both the measured distributions and analytic predictions so that $\int_{\Delta R_{\mathrm{axis}}^{\mathrm{NP}}}^{R/2} \mathrm{d}\Delta R_{\mathrm{axis}} (\mathrm{d}\sigma/\mathrm{d}\Delta R_{\mathrm{axis}}) = 1$. The ΔR_{axis} interval that contains the value $\Delta R_{\mathrm{axis}}^{\mathrm{NP}}$ is defined to belong to the non-perturbative region and excluded from the integral. In reality, there is no sharp boundary between the perturbative and non-perturbative regimes, but a continuous transition. This vertical line is therefore an indicative value defined for the sake of the comparison.

The Standard–SD variable is also IRC safe and therefore calculable in the SCET framework. However, these calculations are computationally expensive and the results are not available at the time of this article. These distributions are particularly sensitive to soft effects, and thus also to the hadronization corrections.

5.3 Discussion of analytic calculations

The analytic predictions describe the measured distributions in the perturbative regime within uncertainties for all variations of the observable (i.e. jet-resolution parameter, grooming setting and $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$) considered in the analysis. The agreement is excellent (largest deviations observed are within 10 to 20% in the lower p_{T} bins, where the smaller systematic uncertainties allow us to draw more precise conclusions) given the low $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$ values from this measurement. Additionally, even though the distributions are normalized to their integral in the perturbative region, significant agreement is also found in the non-perturbative region, where the calculations are most sensitive to the non-perturbative correction. This agreement persists independently of jet-resolution parameter, grooming setting, and $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$. There are larger shape differences between the data and analytic predictions for larger R.

The agreement at lower ΔR_{axis} values is increasingly due to the inclusion of the non-perturbative correction from Eq. 4. The measured distributions are compatible with the universal behavior of the non-perturbative component of the CSS evolution kernel to the extent this hadronization correction is well modeled. This is the first time this has been verified in jet substructure observables. The measurements presented in this article should be included in future global fits to further increase the precision of the experimental extractions of the non-perturbative component of the CSS kernel.

The theoretical uncertainties in the analytic predictions are significant, and reducing them can lead to a higher-precision test. Specifically, as shown in the ratio plots from Fig. 5, while the measured distributions and the calculations agree within uncertainties, there are systematic shape differences between them. For instance, there is a sudden rise in the interval at the highest ΔR_{axis} which can be related to the lack of matching to a fixed-order calculation or power corrections in the observable in the SCET prediction. However, because the distributions are self-normalized, the disagreement cannot be pinned to a specific ΔR_{axis} value.

6 Conclusions

The first measurement of the angle between the Standard, Soft Drop groomed, and Winner-Take-All charged-jet axes is reported. The measurement was carried out in pp collisions at $\sqrt{s} = 5.02$ TeV with the ALICE detector for jets of $R = 0.2$ and 0.4. This analysis focused on jets below 100 GeV/c, where non-perturbative effects are more important than in highly energetic jets. Jet grooming does not substantially change the jet axis direction. The WTA and Standard jet axes differ substantially, though, showing that the Standard jet axis does not generally point in the direction of the highest-momentum jet fragment. However, the WTA and Standard or SD axes become more aligned in more energetic jets. The motivation to use the WTA axis arises from its insensitivity to soft radiation; it is amenable to perturbative calculations and is expected to be less modified in the medium created in ultrarelativistic nucleus–nucleus

collisions.

The distributions from PYTHIA 8 and Herwig 7 show overall good agreement with the data; both groomed and ungroomed jet axis differences with WTA are generally described within the uncertainties. The Standard–SD jet axis difference is better described by Herwig 7 than by PYTHIA 8, suggesting that Herwig better reproduces the soft splittings which are groomed away by Soft Drop. There is also good agreement between the analytic predictions and the measured spectra, even in the region of ΔR_{axis} where non-perturbative physics is important. Within the precision of the measurement and analytic calculations, our result is compatible with the universality of the non-perturbative component of the Collins–Soper–Sterman (CSS) TMD evolution kernel. This study constitutes the first experimental verification of such compatibility in the context of jet-substructure measurements.

These measurements shed light on the interplay between perturbative and non-perturbative effects. The majority of jets produced in pp collisions at this center-of-mass energy are initiated by gluons. It will be illuminating to compare these distributions to those from quark-initiated jets. This can be done by measuring ΔR_{axis} in, for example, heavy-flavor jets or photon–jet coincidences.

The angle between different jet axes may also reflect medium modification of jet angular substructure in heavy-ion collisions. The results presented here will provide a valuable baseline for Pb–Pb studies.

References

[1] R. Kogler *et al.*, "Jet Substructure at the Large Hadron Collider: Experimental Review", *Rev. Mod. Phys.* **91** no. 4, (2019) 045003, arXiv:1803.06991 [hep-ex].

[2] A. J. Larkoski, I. Moult, and B. Nachman, "Jet Substructure at the Large Hadron Collider: A Review of Recent Advances in Theory and Machine Learning", *Phys. Rept.* **841** (2020) 1–63, arXiv:1709.04464 [hep-ph].

[3] S. Marzani, G. Soyez, and M. Spannowsky, *Looking inside jets: an introduction to jet substructure and boosted-object phenomenology*, vol. 958. Springer, (2019). arXiv:1901.10342 [hep-ph].

[4] **ALICE** Collaboration, B. B. Abelev *et al.*, "Charged jet cross sections and properties in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. D* **91** no. 11, (2015) 112012, arXiv:1411.4969 [nucl-ex].

[5] **ALICE** Collaboration, S. Acharya *et al.*, "Jet fragmentation transverse momentum measurements from di-hadron correlations in $\sqrt{s}$ = 7 TeV pp and $\sqrt{s_{\rm NN}}$ = 5.02 TeV p–Pb collisions", *JHEP* **03** (2019) 169, arXiv:1811.09742 [nucl-ex].

[6] **ALICE** Collaboration, S. Acharya *et al.*, "Charged jet cross section and fragmentation in proton-proton collisions at $\sqrt{s}$ = 7 TeV", *Phys. Rev. D* **99** no. 1, (2019) 012016, arXiv:1809.03232 [nucl-ex].

[7] **ALICE** Collaboration, S. Acharya *et al.*, "Measurement of the production of charm jets tagged with D^0 mesons in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **08** (2019) 133, arXiv:1905.02510 [nucl-ex].

[8] **ALICE** Collaboration, S. Acharya *et al.*, "Exploration of jet substructure using iterative declustering in pp and Pb–Pb collisions at LHC energies", *Phys. Lett. B* **802** (2020) 135227, arXiv:1905.02512 [nucl-ex].

[9] **ALICE** Collaboration, S. Acharya *et al.*, "Measurements of the groomed and ungroomed jet angularities in pp collisions at $\sqrt{s}$ = 5.02 TeV", *JHEP* **05** (2022) 061, arXiv:2107.11303 [nucl-ex].

[10] **A Large Ion Collider Experiment, ALICE** Collaboration, S. Acharya *et al.*, "Measurement of the groomed jet radius and momentum splitting fraction in pp and Pb−Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV", *Phys. Rev. Lett.* **128** no. 10, (2022) 102001, arXiv:2107.12984 [nucl-ex].

[11] **ALICE** Collaboration, S. Acharya *et al.*, "First measurements of N-subjettiness in central Pb-Pb collisions at $\sqrt{s_{\rm NN}}$ = 2.76 TeV", *JHEP* **10** (2021) 003, arXiv:2105.04936 [nucl-ex].

[12] **ALICE** Collaboration, S. Acharya *et al.*, "Direct observation of the dead-cone effect in quantum chromodynamics", *Nature* **605** no. 7910, (2022) 440–446, arXiv:2106.05713 [nucl-ex]. [Erratum: Nature 607, E22 (2022)].

[13] **ATLAS** Collaboration, G. Aad *et al.*, "Jet mass and substructure of inclusive jets in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS experiment", *JHEP* **05** (2012) 128, arXiv:1203.4606 [hep-ex].

[14] **ATLAS** Collaboration, G. Aad *et al.*, "ATLAS Measurements of the Properties of Jets for Boosted Particle Searches", *Phys. Rev. D* **86** (2012) 072006, arXiv:1206.5369 [hep-ex].

[15] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of jet shapes in top-quark pair events at $\sqrt{s}$ = 7 TeV using the ATLAS detector", *Eur. Phys. J. C* **73** no. 12, (2013) 2676, arXiv:1307.5749 [hep-ex].

[16] **ATLAS** Collaboration, M. Aaboud *et al.*, "Measurement of the Soft-Drop Jet Mass in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector", *Phys. Rev. Lett.* **121** no. 9, (2018) 092001, arXiv:1711.08341 [hep-ex].

[17] **ATLAS** Collaboration, M. Aaboud *et al.*, "Measurement of jet-substructure observables in top quark, W boson and light jet production in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *JHEP* **08** (2019) 033, arXiv:1903.02942 [hep-ex].

[18] **ATLAS** Collaboration, G. Aad *et al.*, "Properties of jet fragmentation using charged particles measured with the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. D* **100** no. 5, (2019) 052011, arXiv:1906.09254 [hep-ex].

[19] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of soft-drop jet observables in pp collisions with the ATLAS detector at $\sqrt{s}$ =13 TeV", *Phys. Rev. D* **101** no. 5, (2020) 052007, arXiv:1912.09837 [hep-ex].

[20] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of the Lund Jet Plane Using Charged Particles in 13 TeV Proton–Proton Collisions with the ATLAS Detector", *Phys. Rev. Lett.* **124** no. 22, (2020) 222002, arXiv:2004.03540 [hep-ex].

[21] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of the jet mass in high transverse momentum $Z(\rightarrow \mathrm{b\bar{b}})\gamma$ production at $\sqrt{s} = 13$ TeV using the ATLAS detector", *Phys. Lett. B* **812** (Jan, 2021) 135991, arXiv:1907.07093 [hep-ex].

[22] **CMS** Collaboration, S. Chatrchyan *et al.*, "Shape, Transverse Size, and Charged Hadron Multiplicity of Jets in pp Collisions at 7 TeV", *JHEP* **06** (2012) 160, arXiv:1204.3170 [hep-ex].

[23] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of jet fragmentation into charged particles in pp and PbPb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV", *JHEP* **10** (2012) 087, arXiv:1205.5872 [nucl-ex].

[24] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of Jet Fragmentation in PbPb and pp Collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV", *Phys. Rev. C* **90** no. 2, (2014) 024908, arXiv:1406.0932 [nucl-ex].

[25] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Measurements of jet charge with dijet events in pp collisions at $\sqrt{s} = 8$ TeV", *JHEP* **10** (2017) 131, arXiv:1706.05868 [hep-ex].

[26] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Measurement of the Splitting Function in pp and Pb-Pb Collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *Phys. Rev. Lett.* **120** no. 14, (2018) 142302, arXiv:1708.09429 [nucl-ex].

[27] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Measurements of the differential jet cross section as a function of the jet mass in dijet events from proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **11** (2018) 113, arXiv:1807.05974 [hep-ex].

[28] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Jet Shapes of Isolated Photon-Tagged Jets in Pb-Pb and pp Collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV", *Phys. Rev. Lett.* **122** no. 15, (2019) 152001, arXiv:1809.08602 [hep-ex].

[29] **CMS** Collaboration, A. Tumasyan *et al.*, "Study of quark and gluon jet substructure in Z+jet and dijet events from pp collisions", *JHEP* **01** (2022) 188, arXiv:2109.03340 [hep-ex].

[30] **LHCb** Collaboration, R. Aaij *et al.*, "Study of J/ψ Production in Jets", *Phys. Rev. Lett.* **118** no. 19, (2017) 192001, arXiv:1701.05116 [hep-ex].

[31] **LHCb** Collaboration, R. Aaij *et al.*, "Measurement of charged hadron production in Z-tagged jets in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev. Lett.* **123** no. 23, (2019) 232001, arXiv:1904.08878 [hep-ex].

[32] **STAR** Collaboration, J. Adam *et al.*, "Measurement of groomed jet substructure observables in p+p collisions at $\sqrt{s}$ =200 GeV with STAR", *Phys. Lett. B* **811** (2020) 135846, arXiv:2003.02114 [hep-ex].

[33] **STAR** Collaboration, M. Abdallah *et al.*, "Invariant Jet Mass Measurements in pp Collisions at $\sqrt{s} = 200$ GeV at RHIC", *Phys. Rev. D* **104** no. 5, (2021) 052007, arXiv:2103.13286 [hep-ex].

[34] **STAR** Collaboration, M. S. Abdallah *et al.*, "Differential measurements of jet substructure and partonic energy loss in Au+Au collisions at $\sqrt{S_{NN}}$ =200 GeV", *Phys. Rev. C* **105** no. 4, (2022) 044906, arXiv:2109.09793 [nucl-ex].

[35] P. Cal, D. Neill, F. Ringer, and W. J. Waalewijn, "Calculating the angle between jet axes", *JHEP* **04** (2020) 211, arXiv:1911.06840 [hep-ph].

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

[37] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, "Better jet clustering algorithms", *JHEP* **08** (1997) 001, arXiv:hep-ph/9707323.

[38] M. Wobisch and T. Wengler, "Hadronization corrections to jet cross-sections in deep inelastic scattering", in *Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting)*, pp. 270–279. 4, 1998. arXiv:hep-ph/9907280.

[39] D. Bertolini, T. Chan, and J. Thaler, "Jet Observables Without Jet Algorithms", *JHEP* **04** (2014) 013, arXiv:1310.7584 [hep-ph].

[40] A. A. Vladimirov, "Self-contained definition of the Collins–Soper kernel", *Phys. Rev. Lett.* **125** no. 19, (2020) 192002, arXiv:2003.02288 [hep-ph].

[41] F. Landry, R. Brock, P. M. Nadolsky, and C. P. Yuan, "Tevatron Run-1 Z boson data and Collins–Soper–Sterman resummation formalism", *Phys. Rev. D* **67** (2003) 073016, arXiv:hep-ph/0212159.

[42] P. Sun, J. Isaacson, C. P. Yuan, and F. Yuan, "Nonperturbative functions for SIDIS and Drell–Yan processes", *Int. J. Mod. Phys. A* **33** no. 11, (2018) 1841006, arXiv:1406.3073 [hep-ph].

[43] A. Bacchetta, F. Delcarro, C. Pisano, M. Radici, and A. Signori, "Extraction of partonic transverse momentum distributions from semi-inclusive deep-inelastic scattering, Drell–Yan and Z-boson production", *JHEP* **06** (2017) 081, arXiv:1703.10157 [hep-ph]. [Erratum: JHEP 06, 051 (2019)].

[44] V. Bertone, I. Scimemi, and A. Vladimirov, "Extraction of unpolarized quark transverse momentum dependent parton distributions from Drell–Yan/Z-boson production", *JHEP* **06** (2019) 028, arXiv:1902.08474 [hep-ph].

[45] M. A. Ebert, I. W. Stewart, and Y. Zhao, "Determining the Nonperturbative Collins–Soper Kernel From Lattice QCD", *Phys. Rev. D* **99** no. 3, (2019) 034505, arXiv:1811.00026 [hep-ph].

[46] P. Shanahan, M. Wagman, and Y. Zhao, "Collins–Soper kernel for TMD evolution from lattice QCD", *Phys. Rev. D* **102** no. 1, (2020) 014511, arXiv:2003.06063 [hep-lat].

[47] B. V. Jacak and B. Müller, "The exploration of hot nuclear matter", *Science* **337** (Jul, 2012) 310–314.

[48] W. Busza, K. Rajagopal, and W. van der Schee, "Heavy Ion Collisions: The Big Picture, and the Big Questions", *Ann. Rev. Nucl. Part. Sci.* **68** (2018) 339–376, arXiv:1802.04801 [hep-ph].

[49] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft Drop", *JHEP* **05** (2014) 146, arXiv:1402.2657 [hep-ph].

[50] L. Evans (ed.) and P. Bryant (ed.), "LHC Machine", *JINST* **3** (2008) S08001.

[51] **ALICE** Collaboration, B. B. Abelev *et al.*, "Performance of the ALICE Experiment at the CERN LHC", *Int. J. Mod. Phys. A* **29** (2014) 1430044, arXiv:1402.4476 [nucl-ex].

[52] **ALICE** Collaboration, K. Aamodt *et al.*, "The ALICE experiment at the CERN LHC", *JINST* **3** (2008) S08002.

[53] **ALICE** Collaboration, E. Abbas *et al.*, "Performance of the ALICE VZERO system", *JINST* **8** (2013) P10016, arXiv:1306.3130 [nucl-ex].

[54] **ALICE** Collaboration, S. Acharya *et al.*, "ALICE 2017 luminosity determination for pp collisions at $\sqrt{s} = 5$ TeV", ALICE-PUBLIC-2018-014. https://cds.cern.ch/record/2648933.

[55] **ALICE** Collaboration, K. Aamodt *et al.*, "Alignment of the ALICE Inner Tracking System with cosmic-ray tracks", *JINST* **5** (2010) P03003, arXiv:1001.0502 [physics.ins-det].

[56] J. Alme *et al.*, "The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events", *Nucl. Instrum. Meth. A* **622** (2010) 316–367, arXiv:1001.1950 [physics.ins-det].

[57] **ALICE** Collaboration, S. Acharya *et al.*, "Measurement of charged jet cross section in pp collisions at $\sqrt{s} = 5.02$ TeV", *Phys. Rev. D* **100** no. 9, (2019) 092004, arXiv:1905.02536 [nucl-ex].

[58] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J. C* **72** (2012) 1896, arXiv:1111.6097 [hep-ph].

[59] T. Sjöstrand *et al.*, "An introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159–177, arXiv:1410.3012 [hep-ph].

[60] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 Tune", *Eur. Phys. J. C* **74** no. 8, (2014) 3024, arXiv:1404.5630 [hep-ph].

[61] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini, *GEANT 3: user's guide Geant 3.10, Geant 3.11; rev. version*. CERN, Geneva, Sep, 1987. https://cds.cern.ch/record/1119728.

[62] **ALICE** Collaboration, "The ALICE definition of primary particles",. https://cds.cern.ch/record/2270008.

[63] G. D'Agostini, "A multidimensional unfolding method based on Bayes' theorem", *Nucl. Instrum. Meth. A* **362** no. 2, (1994) 487 – 498.

[64] G. D'Agostini, "Improved iterative Bayesian unfolding", in *Alliance Workshop on Unfolding and Data Correction*. 10, 2010. arXiv:1010.0632 [physics.data-an].

[65] https://hepunx.rl.ac.uk/~adye/software/unfold/RooUnfold.html.

[66] C. W. Bauer, D. Pirjol, and I. W. Stewart, "Soft collinear factorization in effective field theory", *Phys. Rev. D* **65** (2002) 054022, arXiv:hep-ph/0109045.

[67] H.-L. Lai *et al.*, "New parton distributions for collider physics", *Phys. Rev. D* **82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[68] J. Gao *et al.*, "CT10 next-to-next-to-leading order global analysis of QCD", *Phys. Rev. D* **89** no. 3, (2014) 033009, arXiv:1302.6246 [hep-ph].

[69] S. Dulat *et al.*, "New parton distribution functions from a global analysis of quantum chromodynamics", *Phys. Rev. D* **93** no. 3, (2016) 033006, arXiv:1506.07443 [hep-ph].

[70] M. Dasgupta and G. P. Salam, "Resummation of nonglobal QCD observables", *Phys. Lett. B* **512** (2001) 323–330, arXiv:hep-ph/0104277.

[71] R. Bain, Y. Makris, and T. Mehen, "Transverse Momentum Dependent Fragmenting Jet Functions with Applications to Quarkonium Production", *JHEP* **11** (2016) 144, arXiv:1610.06508 [hep-ph].

[72] Z.-B. Kang, X. Liu, F. Ringer, and H. Xing, "The transverse momentum distribution of hadrons within jets", *JHEP* **11** (2017) 068, arXiv:1705.08443 [hep-ph].

[73] J. C. Collins, D. E. Soper, and G. F. Sterman, "Transverse Momentum Distribution in Drell-Yan Pair and W and Z Boson Production", *Nucl. Phys. B* **250** (1985) 199–224.

[74] Y. Makris, D. Neill, and V. Vaidya, "Probing Transverse-Momentum Dependent Evolution With Groomed Jets", *JHEP* **07** (2018) 167, arXiv:1712.07653 [hep-ph].

[75] A. V. Konychev and P. M. Nadolsky, "Universality of the Collins–Soper–Sterman nonperturbative function in gauge boson production", *Phys. Lett. B* **633** (2006) 710–714, arXiv:hep-ph/0506225.

[76] Z.-B. Kang, F. Ringer, and I. Vitev, "The semi-inclusive jet function in SCET and small radius resummation for inclusive jet production", *JHEP* **10** (2016) 125, arXiv:1606.06732 [hep-ph].

[77] S. Gieseke, C. Rohr, and A. Siodmok, "Colour reconnections in Herwig++", *Eur. Phys. J. C* **72** (Nov, 2012) 2225, arXiv:1206.0041 [hep-ph].

A Comparison to Monte Carlo event generators

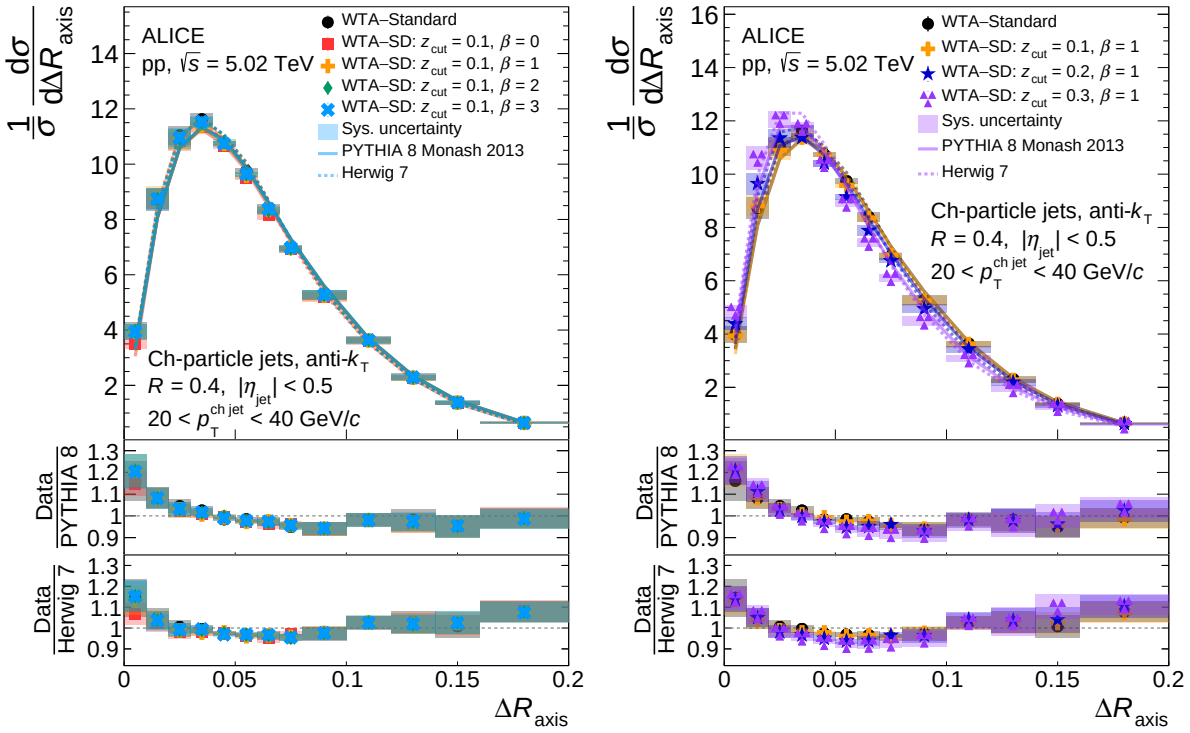


Figure A.1: Comparison between WTA-relative measured distributions and Monte Carlo event generators for jets of $R = 0.4$ in $20 < p_{\mathrm{T}}^{\mathrm{jet}} < 40$ GeV/c. Left: distributions with $z_{\mathrm{cut}} = 0.1$ and varying β. Right: distributions with $\beta = 1$ and varying z_{cut}.

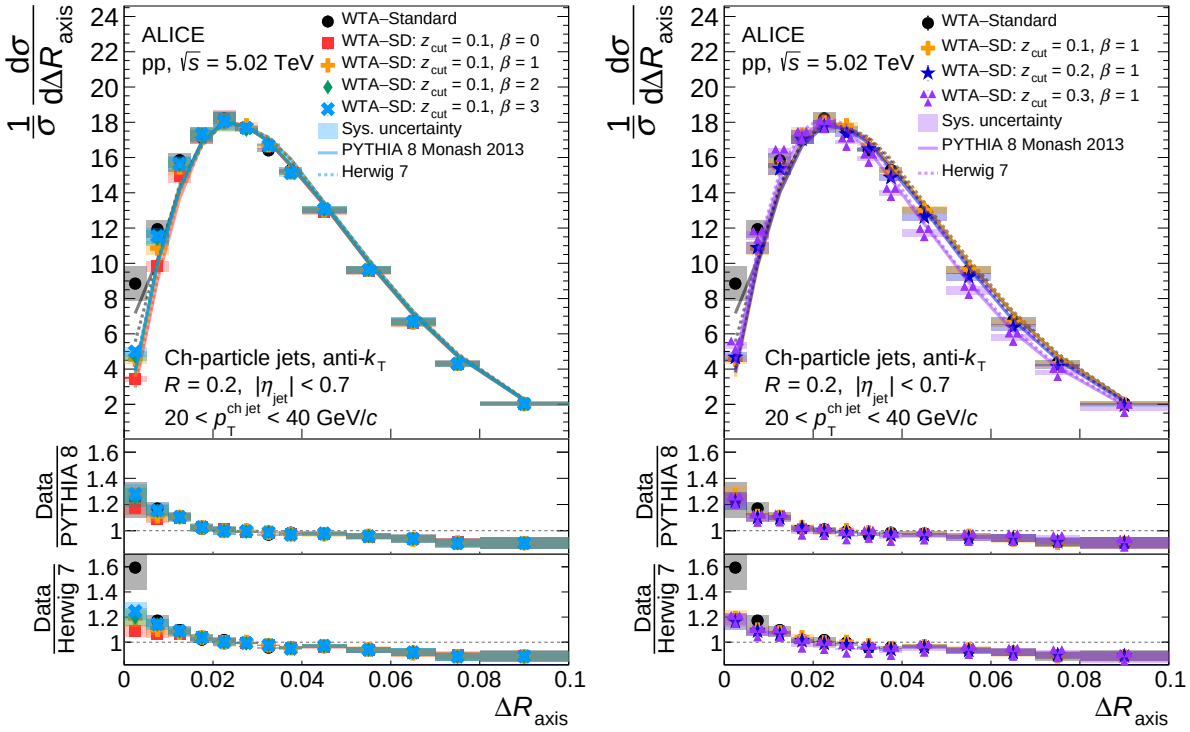


Figure A.2: Same as Figure A.1 for $R = 0.2$.

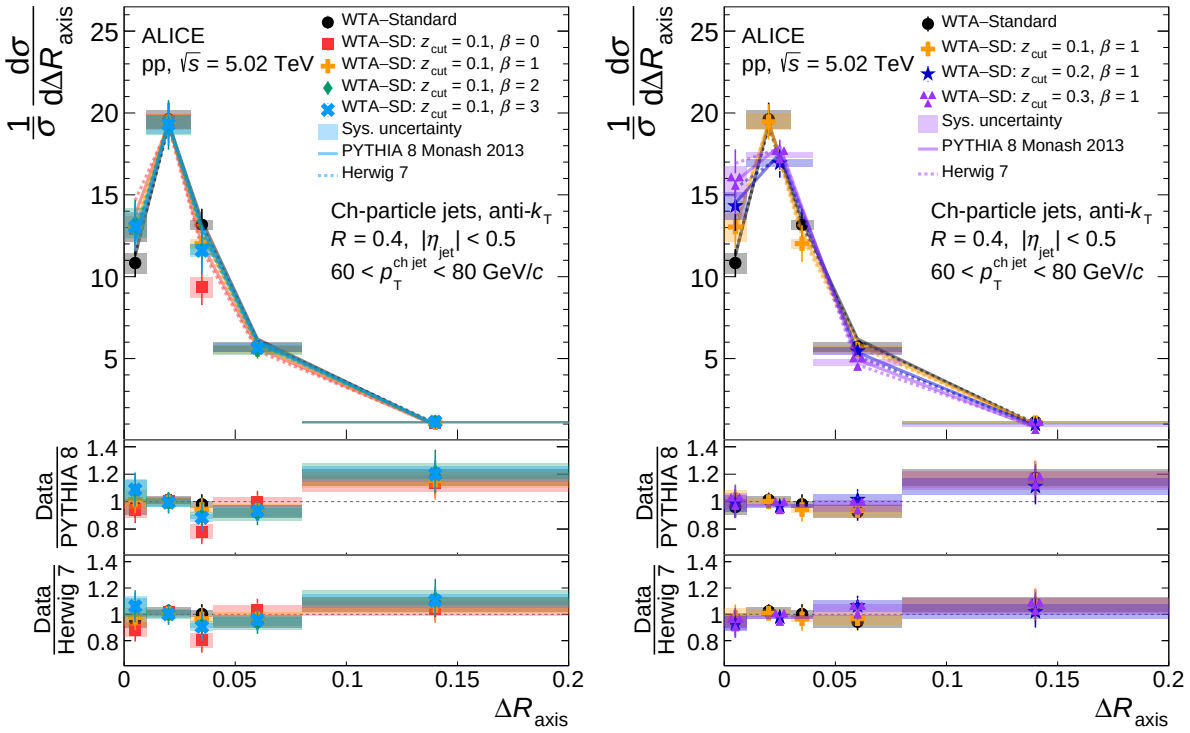


Figure A.3: Same as Figure A.1 for $60 < p_{\mathrm{T}}^{\mathrm{jet}} < 80$ GeV/c.

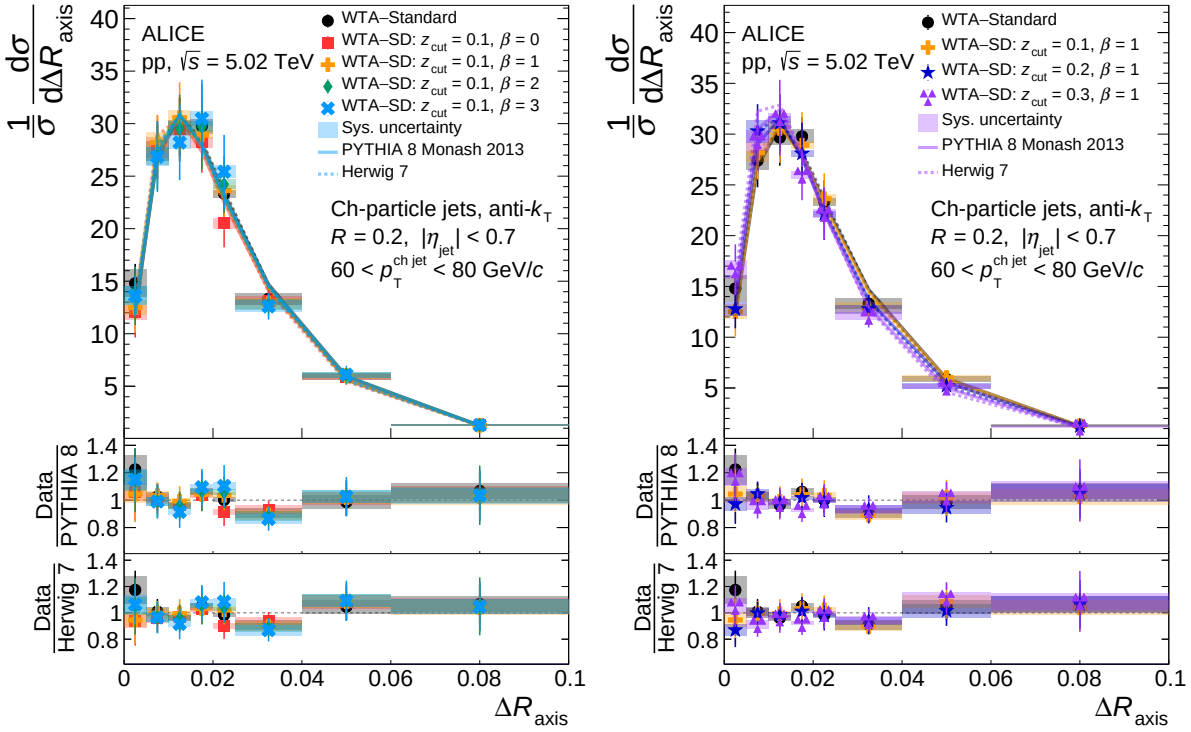


Figure A.4: Same as Figure A.1 for $R = 0.2$ for $60 < p_{\mathrm{T}}^{\mathrm{jet}} < 80$ GeV/c.

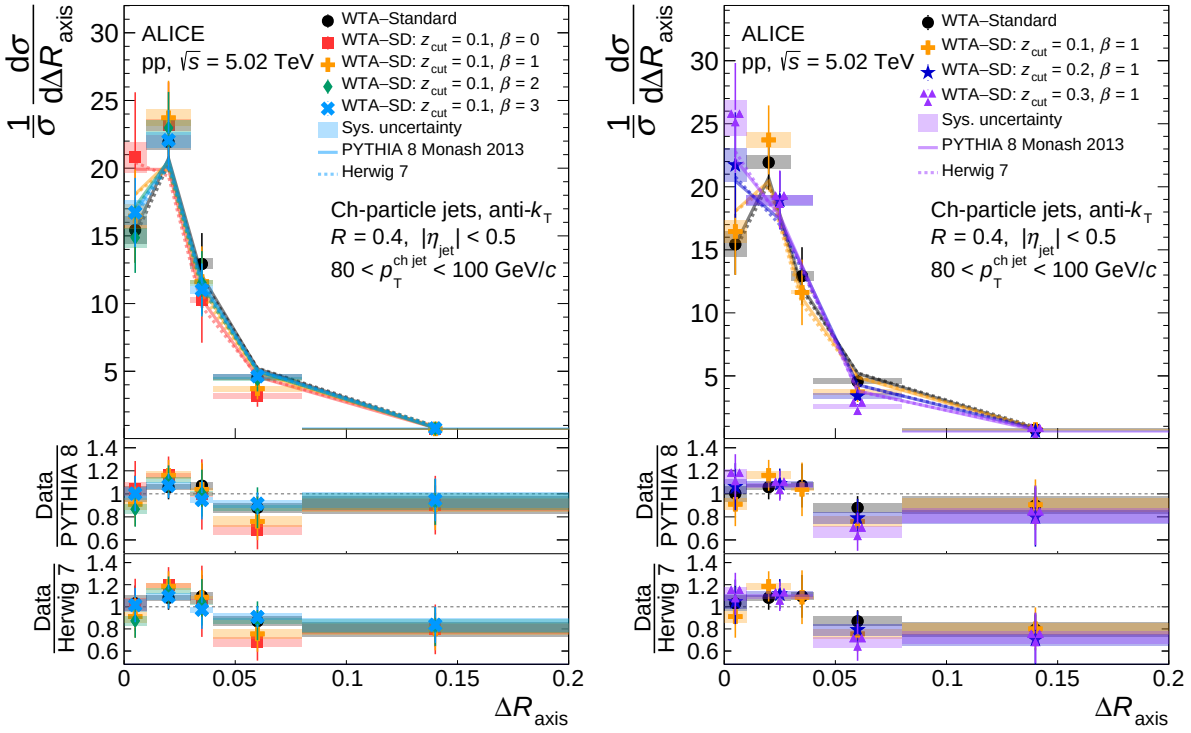


Figure A.5: Same as Figure A.1 for $80 < p_{\mathrm{T}}^{\mathrm{jet}} < 100$ GeV/c.

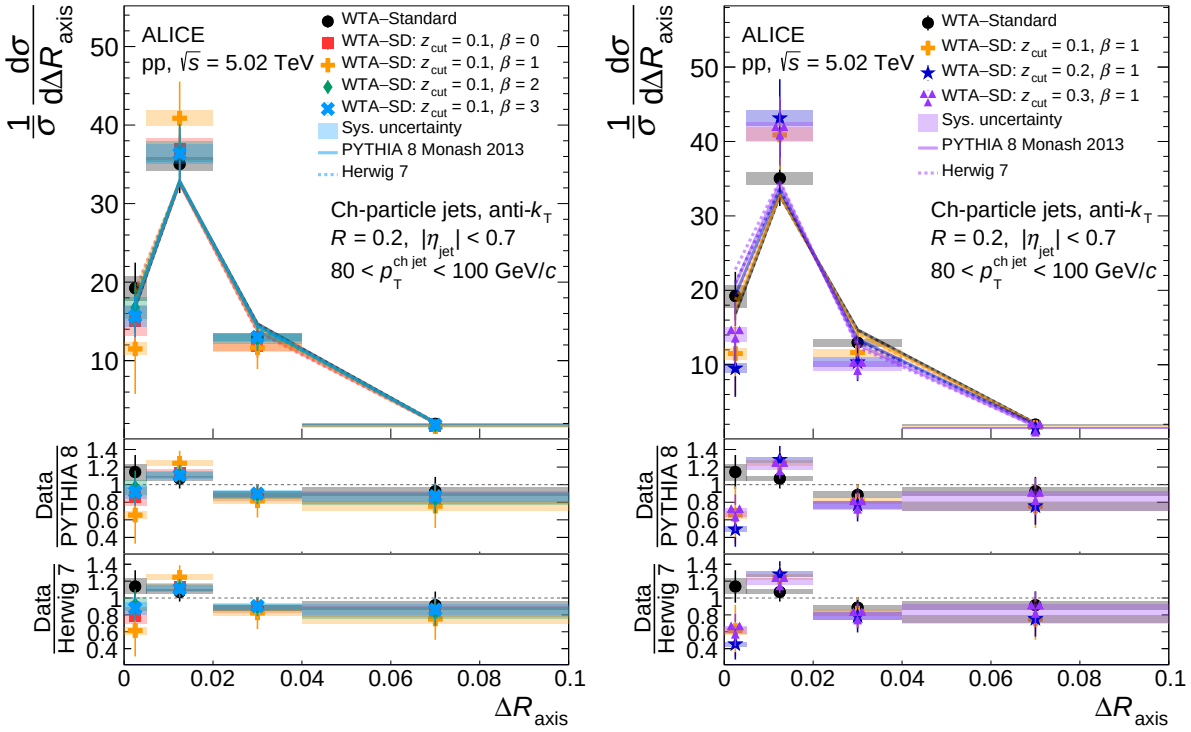


Figure A.6: Same as Figure A.1 for $R = 0.2$ for $80 < p_{\mathrm{T}}^{\mathrm{jet}} < 100$ GeV/c.

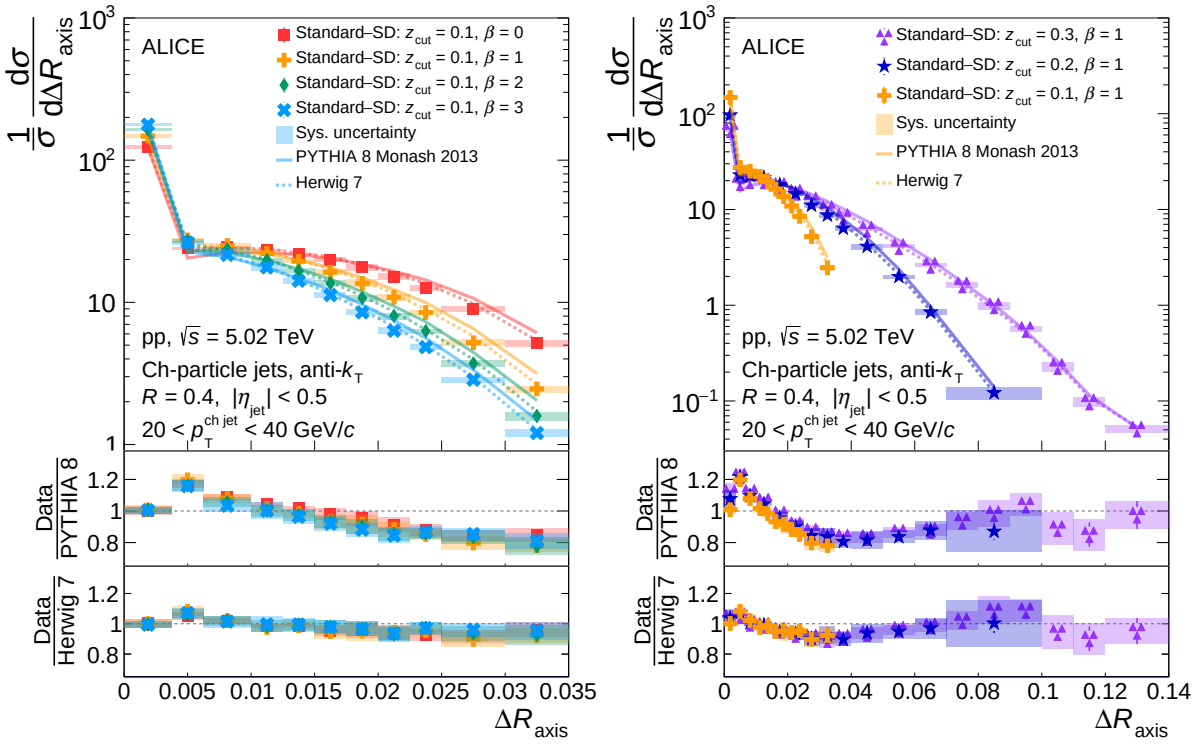


Figure A.7: Comparison between Standard−SD measured distributions and Monte Carlo event generators for jets of $R = 0.4$ in $20 < p_{\mathrm{T}}^{\mathrm{jet}} < 40$ GeV/c. Left: distributions with $z_{\mathrm{cut}} = 0.1$ and varying β. Right: distributions with $\beta = 1$ and varying z_{cut}.

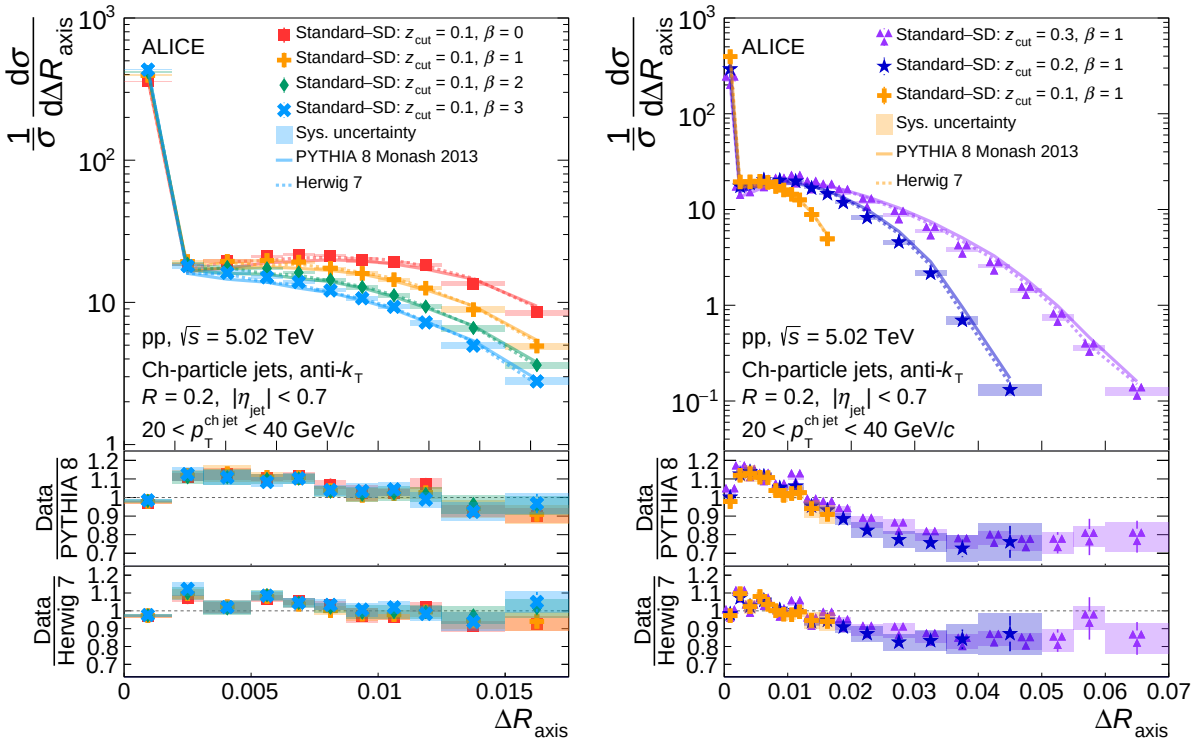


Figure A.8: Same as Figure A.7 for $R = 0.2$.

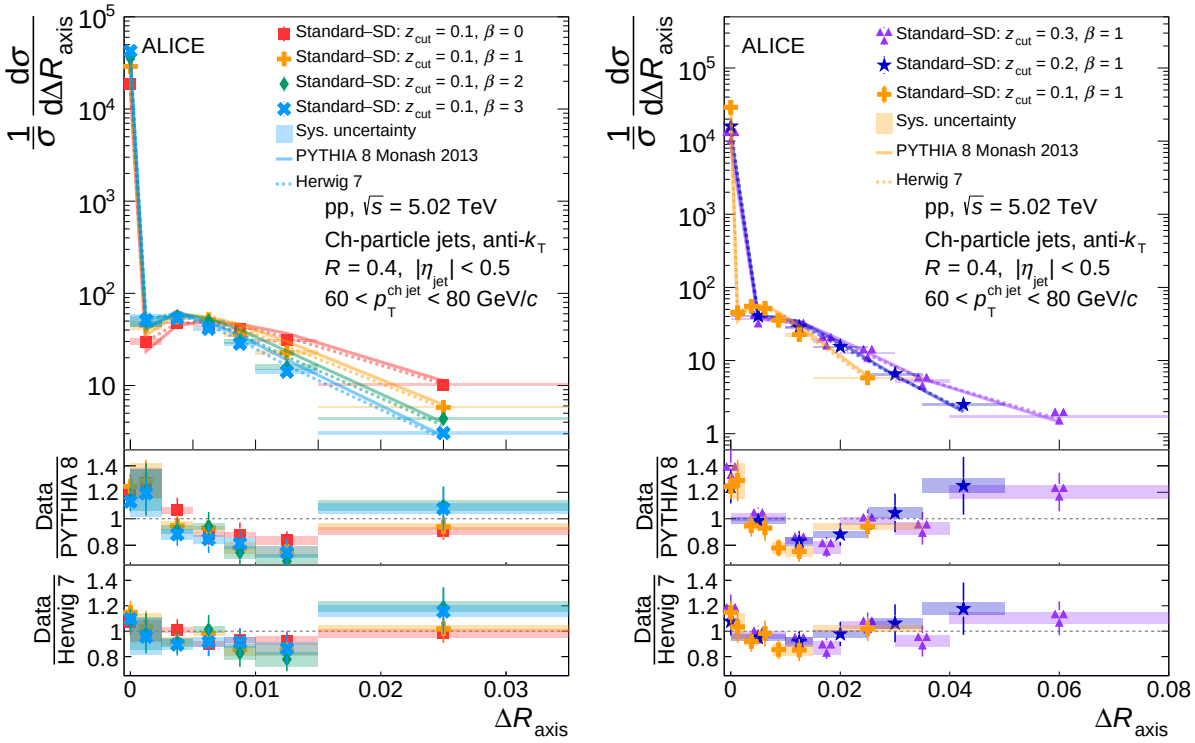


Figure A.9: Same as Figure A.9 for $60 < p_T^{jet} < 80$ GeV/c.

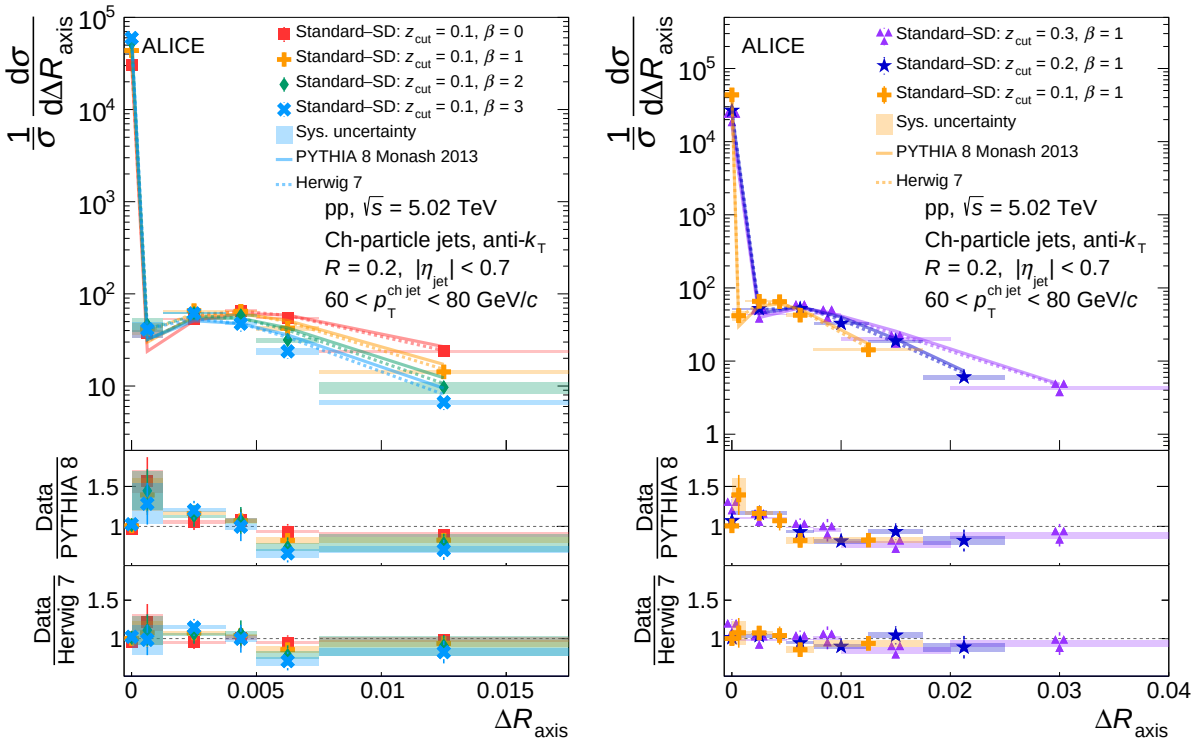


Figure A.10: Same as Figure A.9 for $R = 0.2$ for $60 < p_{\mathrm{T}}^{\mathrm{jet}} < 80$ GeV/c.

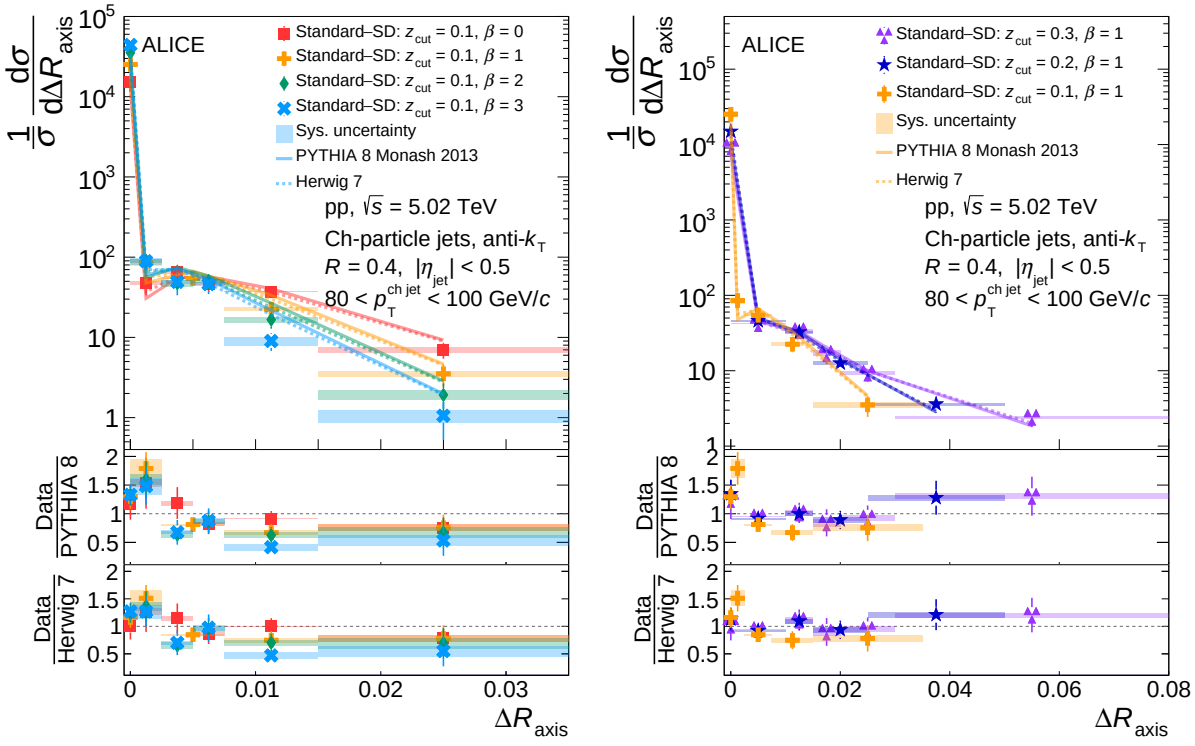


Figure A.11: Same as Figure A.9 for $80 < p_{\mathrm{T}}^{\mathrm{jet}} < 100$ GeV/c.

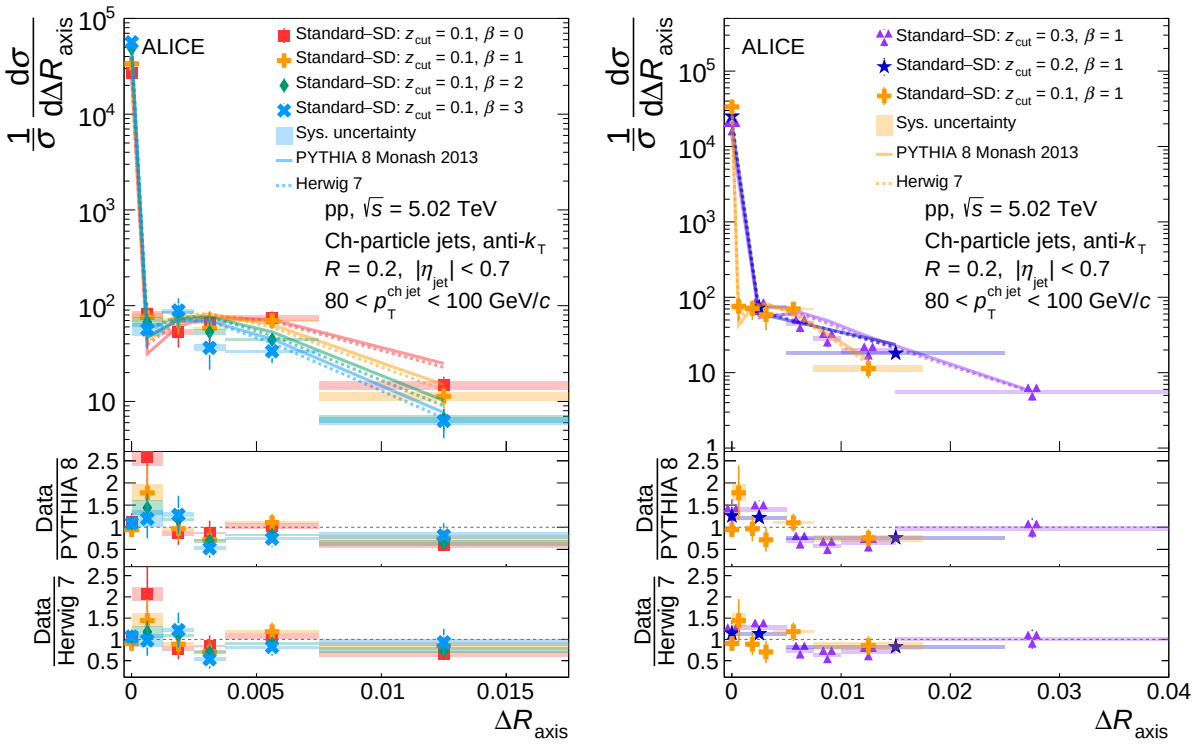


Figure A.12: Same as Figure A.9 for $R = 0.2$ for $80 < p_{\mathrm{T}}^{\mathrm{jet}} < 100$ GeV/c.

B Comparison to analytic predictions

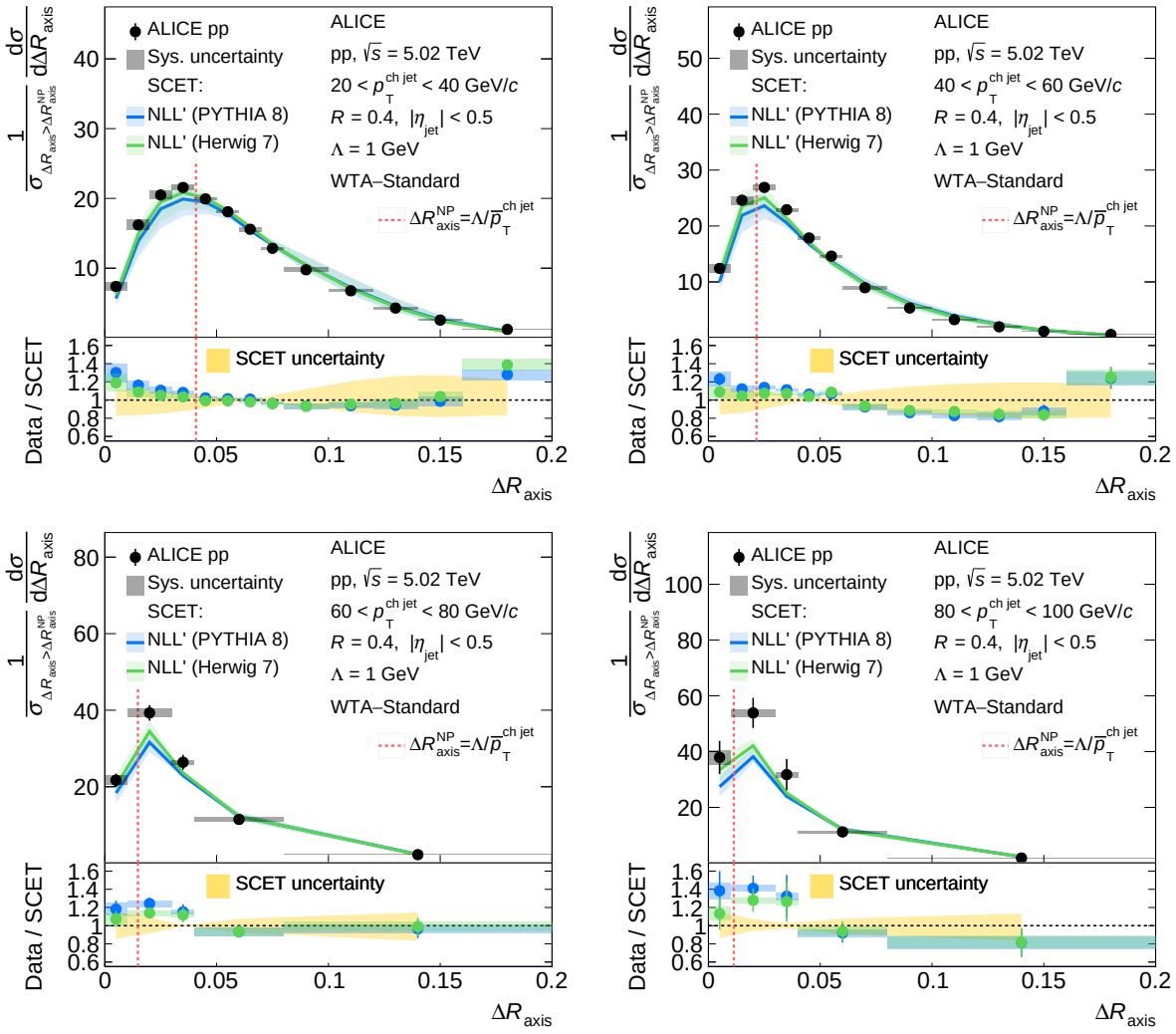


Figure B.1: Comparison between measured distribution and analytic prediction for the Standard - WTA case with $R = 0.4$. Each panel corresponds to a different $p_{\mathrm{T}}^{\mathrm{jet}}$ bin. The measured distribution is shown in black. The colored distributions correspond to the SCET prediction corrected with different MC event generators.

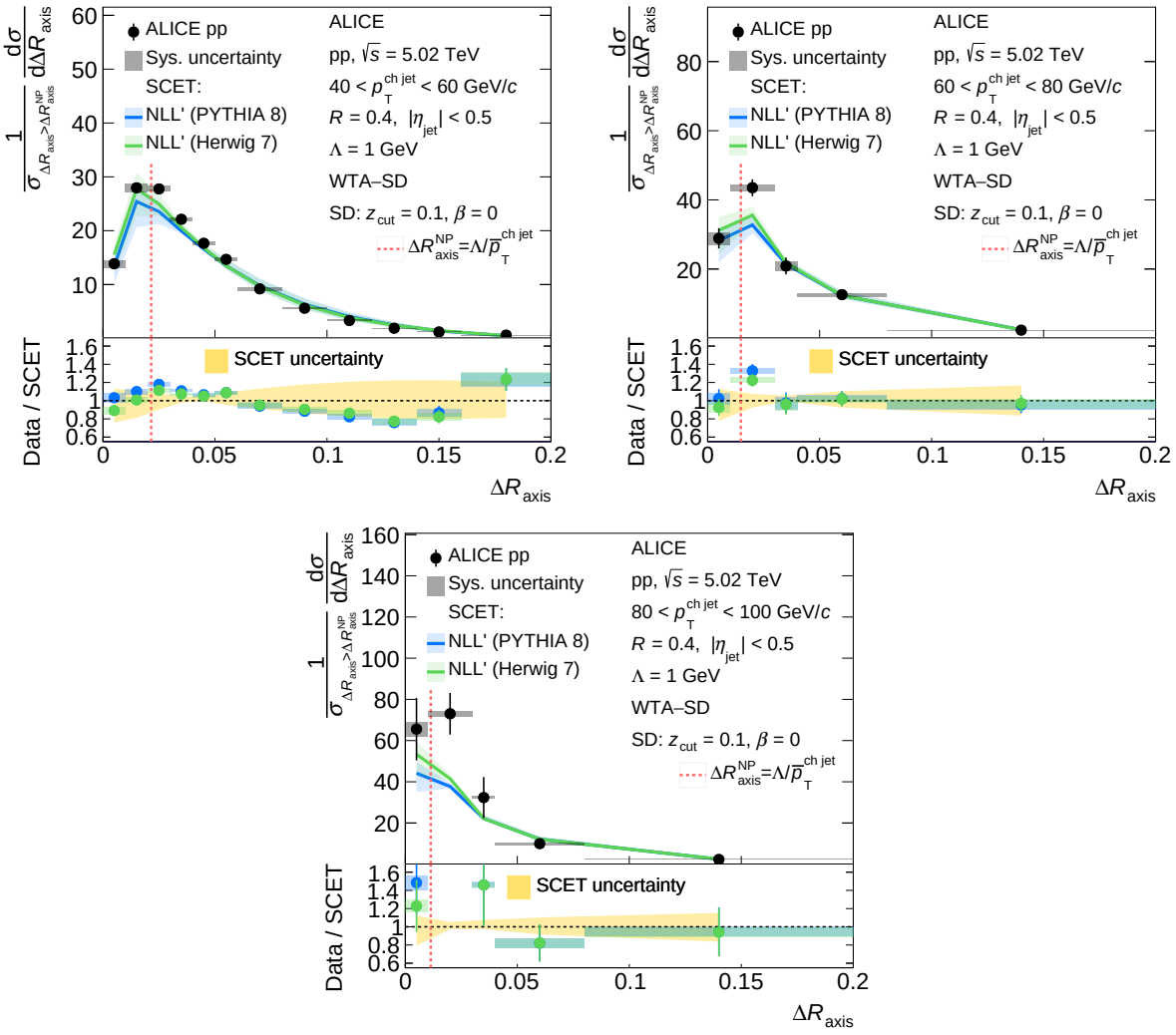


Figure B.2: Same as Figure B.1, for WTA−SD ($z_{\mathrm{cut}} = 0.1$, $\beta = 0$) for $R = 0.4$.

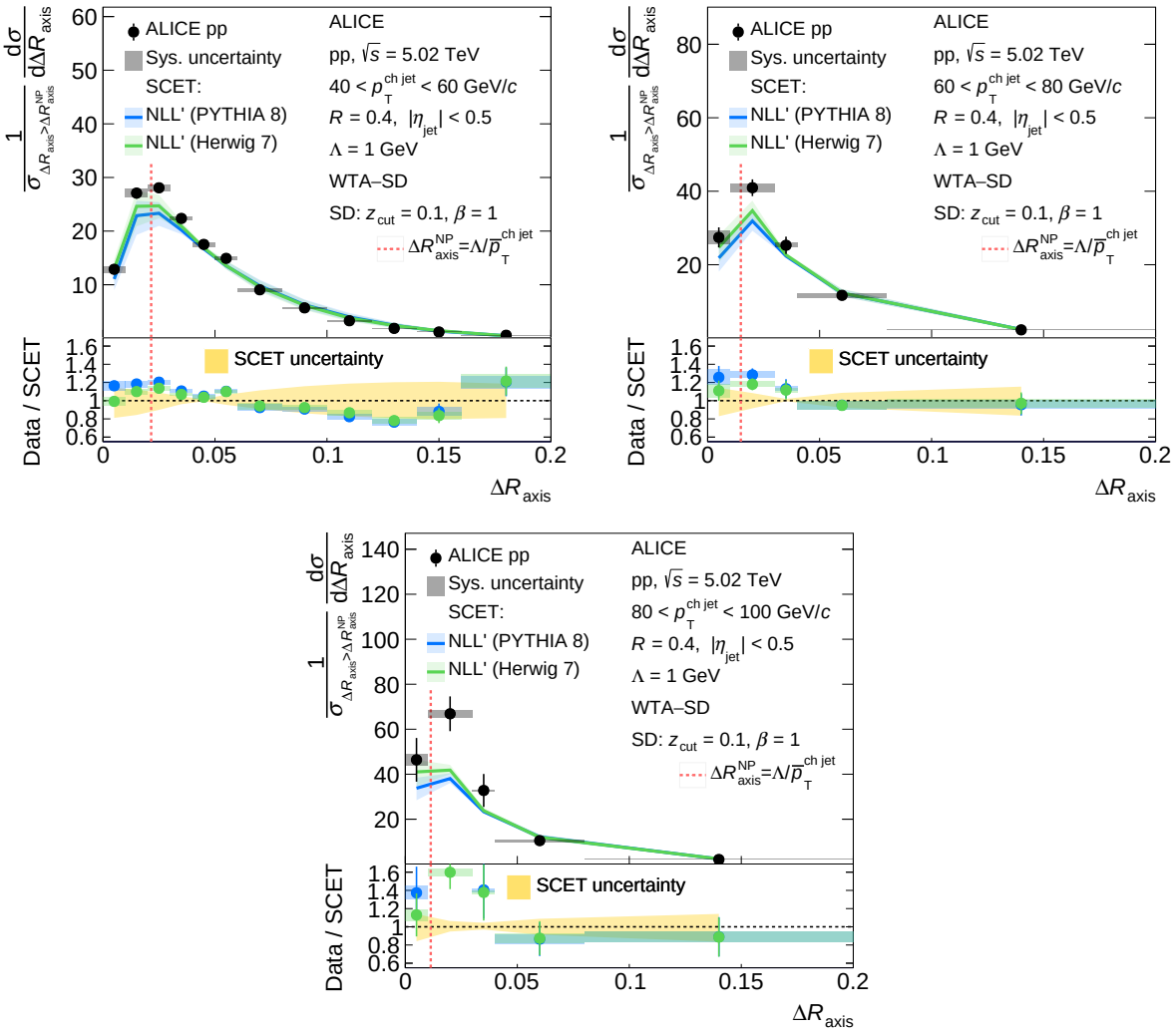


Figure B.3: Same as Figure B.1, for WTA−SD ($z_{\mathrm{cut}} = 0.1$, $\beta = 1$) for $R = 0.4$.

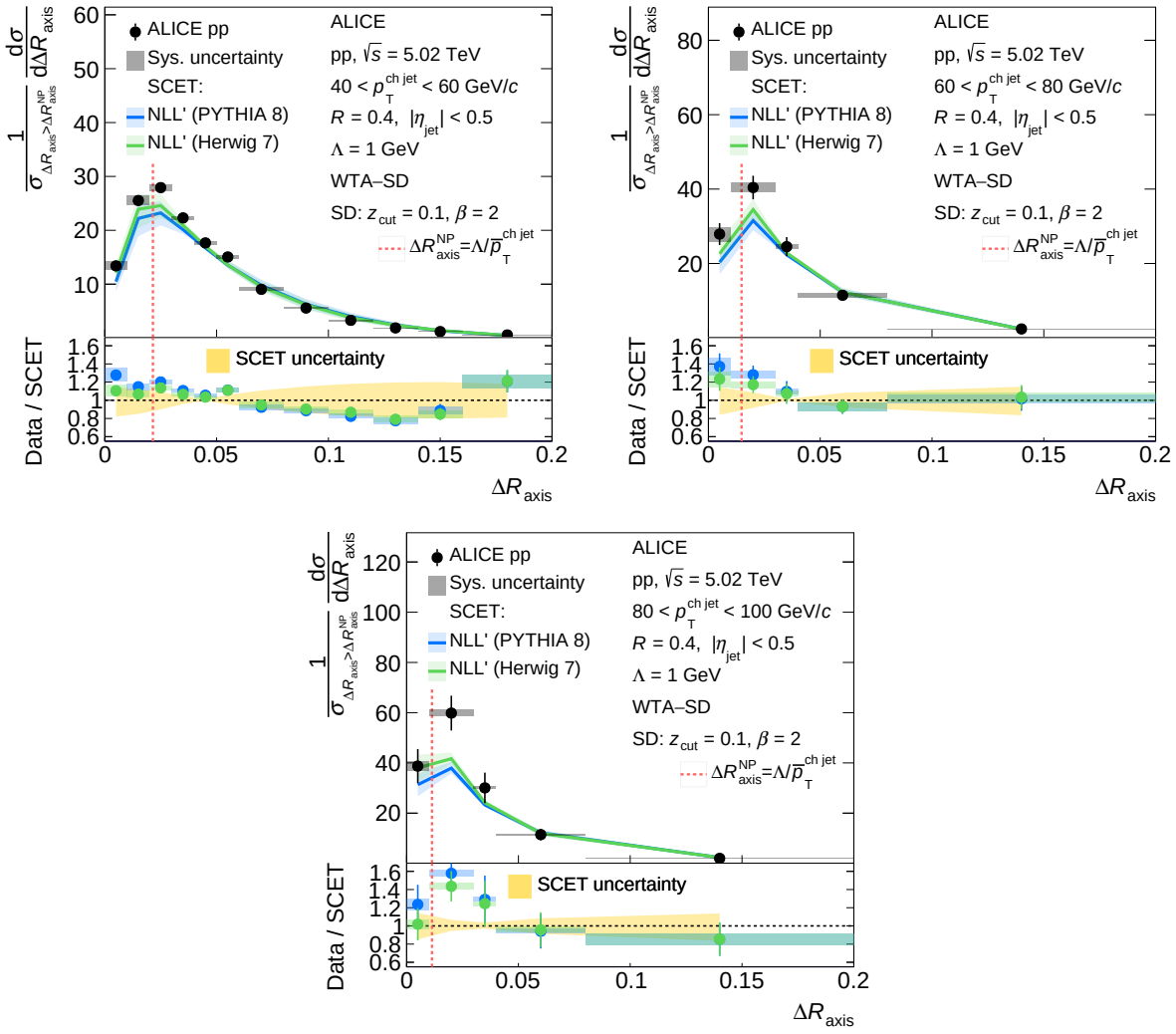


Figure B.4: Same as Figure B.1, for WTA−SD ($z_{cut} = 0.1$, $\beta = 2$) for $R = 0.4$.

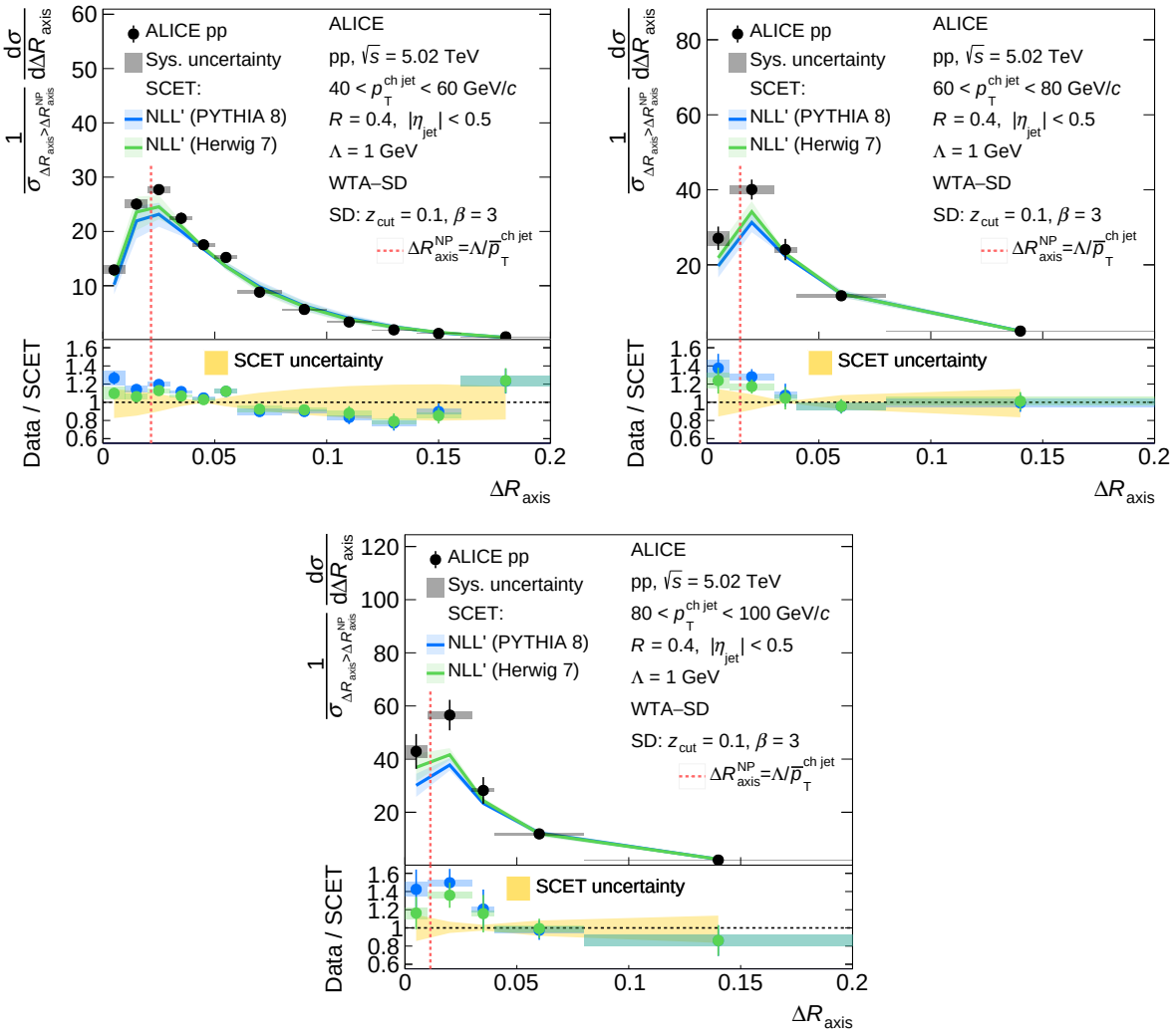


Figure B.5: Same as Figure B.1, for WTA−SD ($z_{\mathrm{cut}} = 0.1$, $\beta = 3$) for $R = 0.4$.

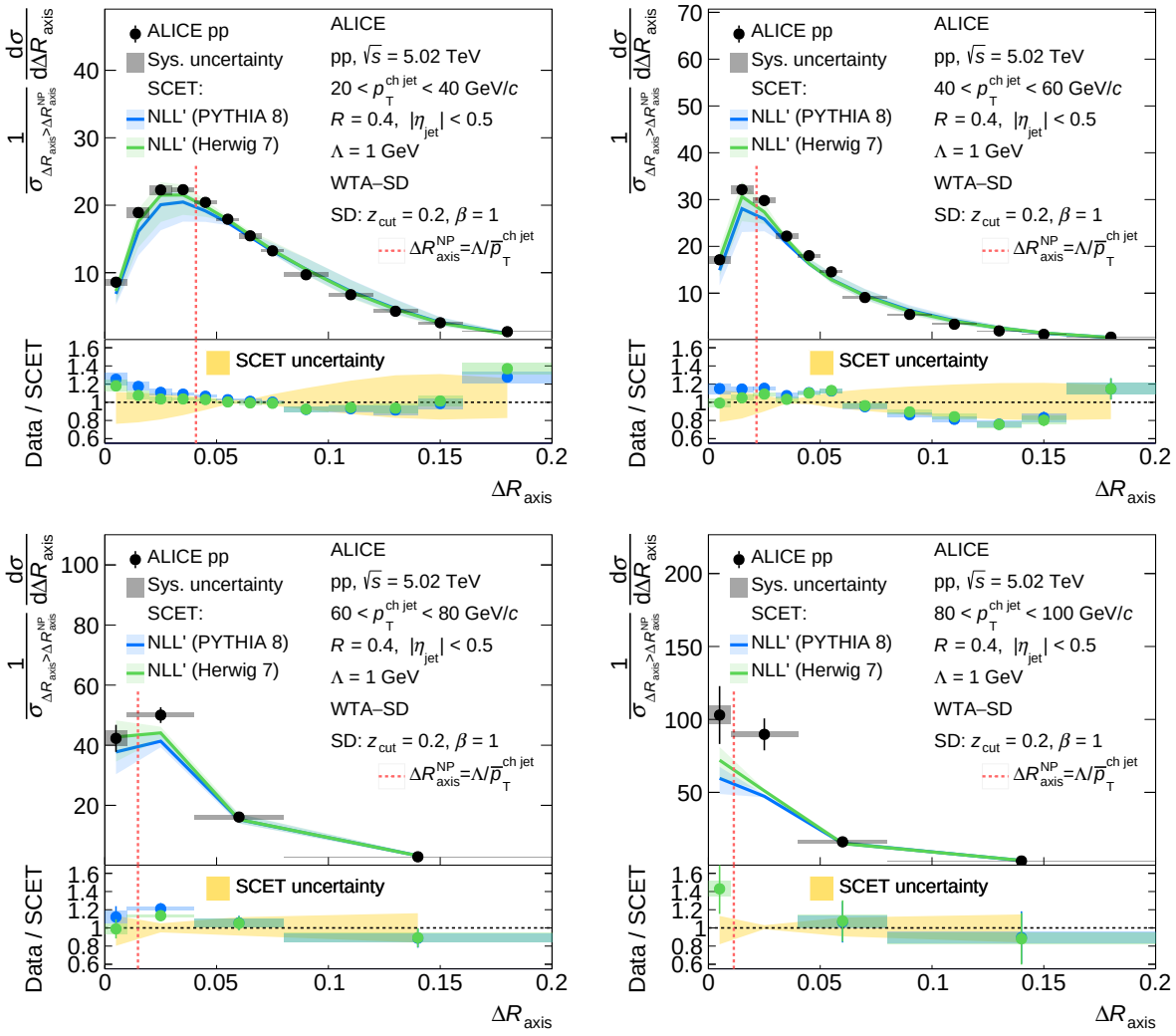


Figure B.6: Same as Figure B.1, for WTA−SD ($z_{\mathrm{cut}} = 0.2$, $\beta = 1$) for $R = 0.4$.

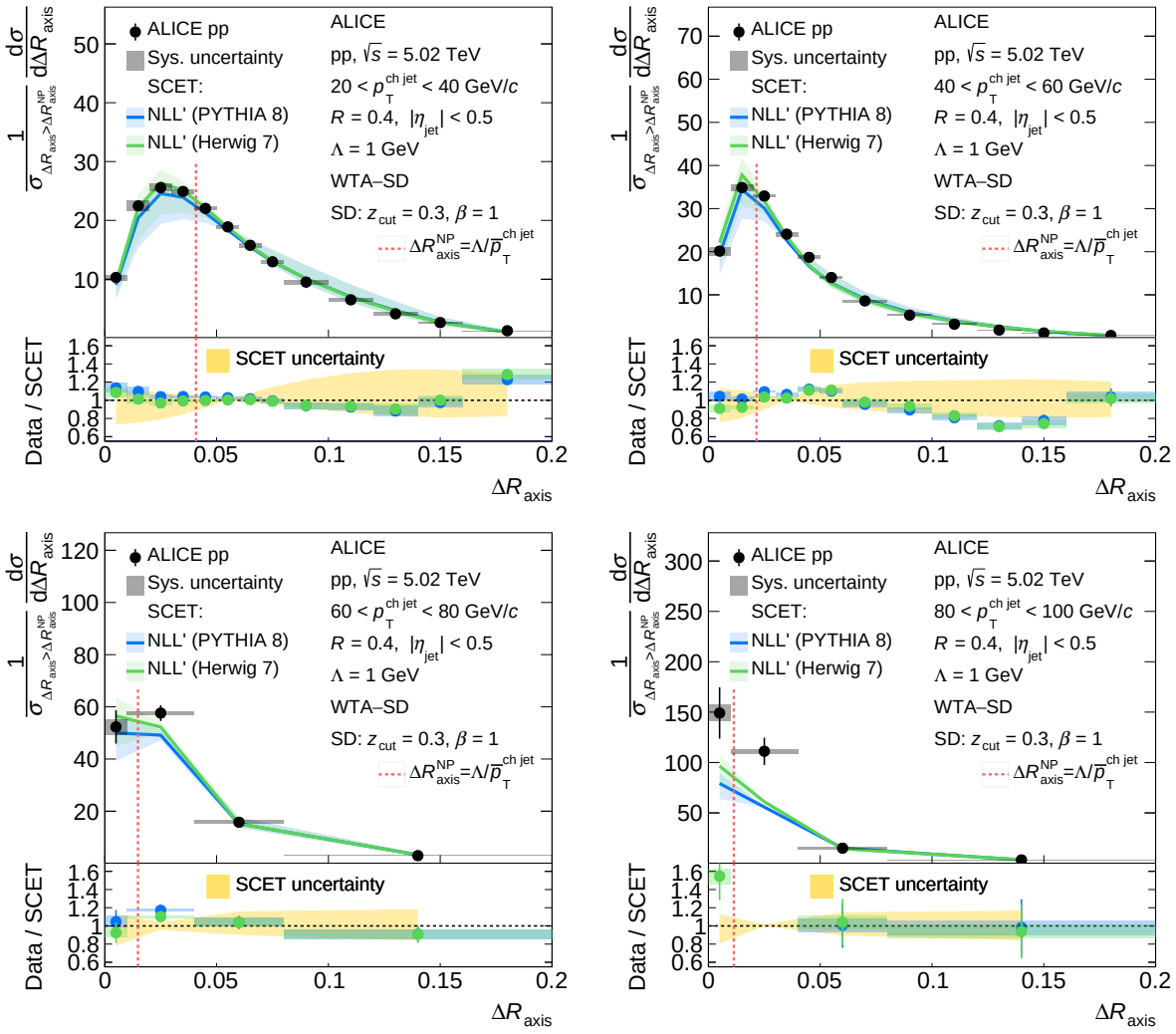


Figure B.7: Same as Figure B.1, for WTA−SD ($z_{\mathrm{cut}} = 0.3$, $\beta = 1$) for $R = 0.4$.

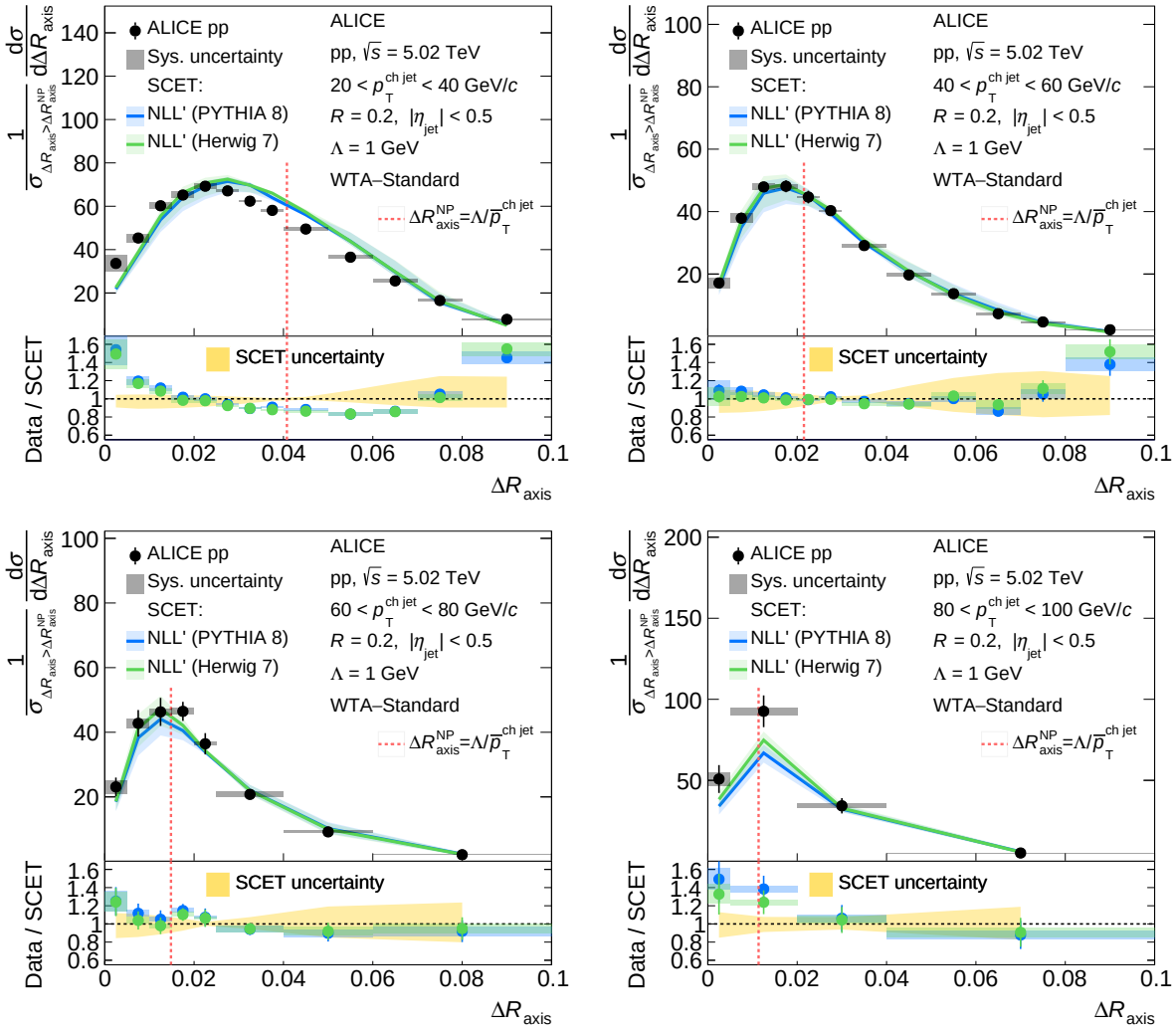


Figure B.8: Same as Figure B.1, for Standard - WTA for $R = 0.2$.

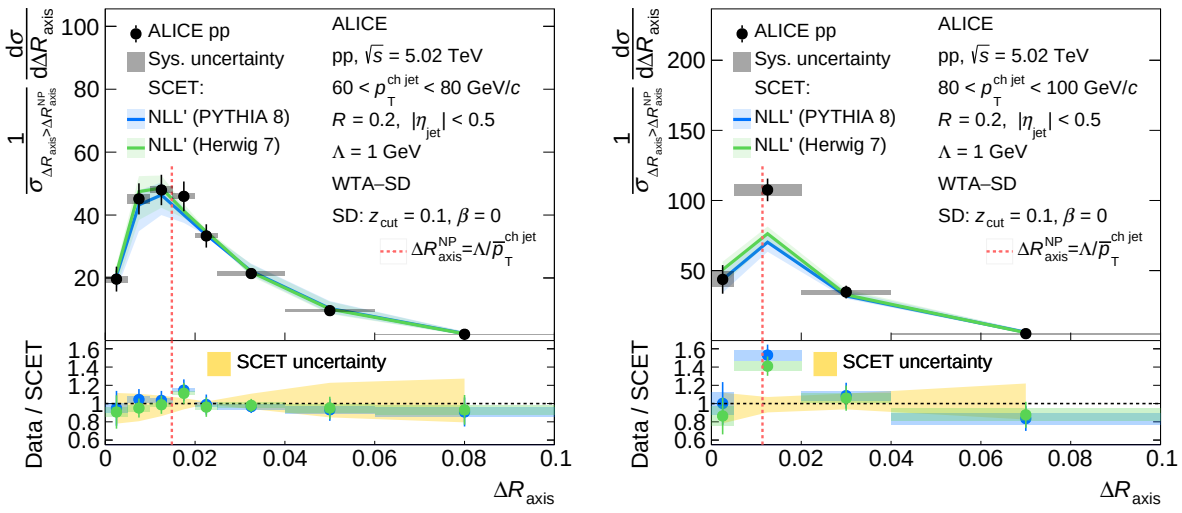


Figure B.9: Same as Figure B.1, for WTA−SD ($z_{\rm cut} = 0.1$, $\beta = 0$) for $R = 0.2$.

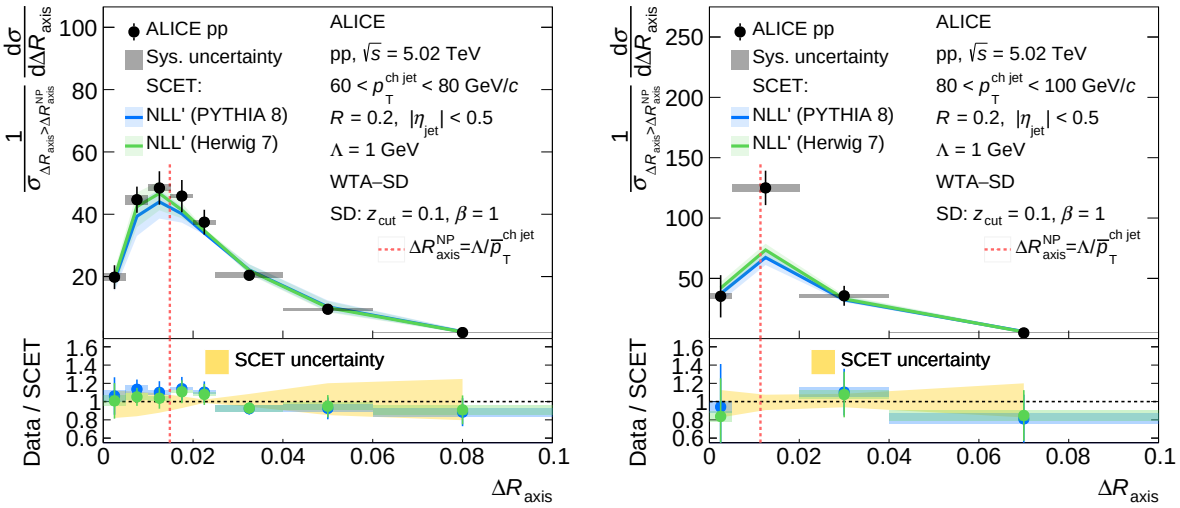


Figure B.10: Same as Figure B.1, for WTA−SD ($z_{\rm cut} = 0.1$, $\beta = 1$) for $R = 0.2$.

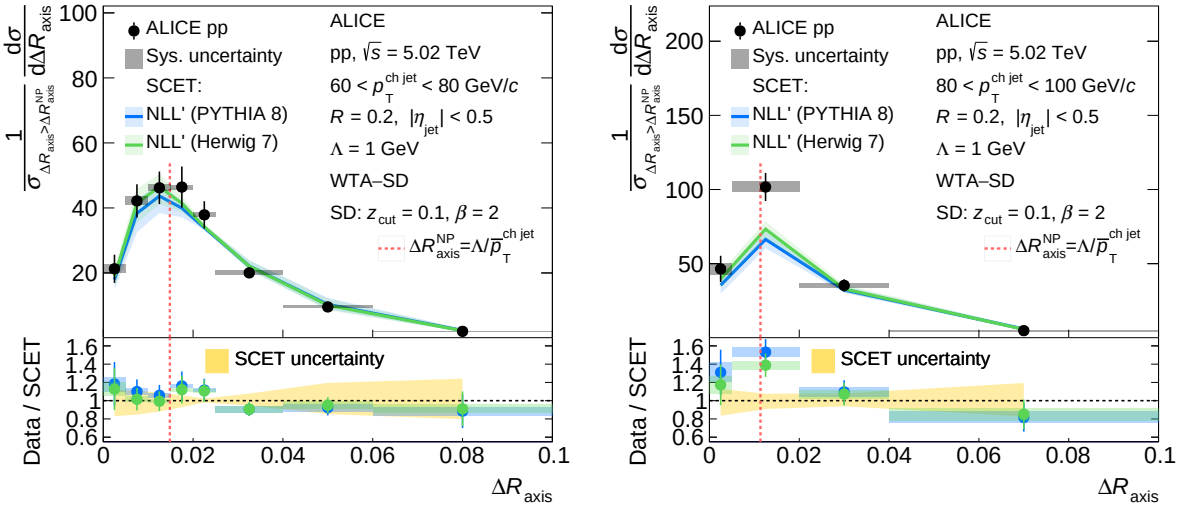


Figure B.11: Same as Figure B.1, for WTA−SD ($z_{\rm cut} = 0.1$, $\beta = 2$) for $R = 0.2$.

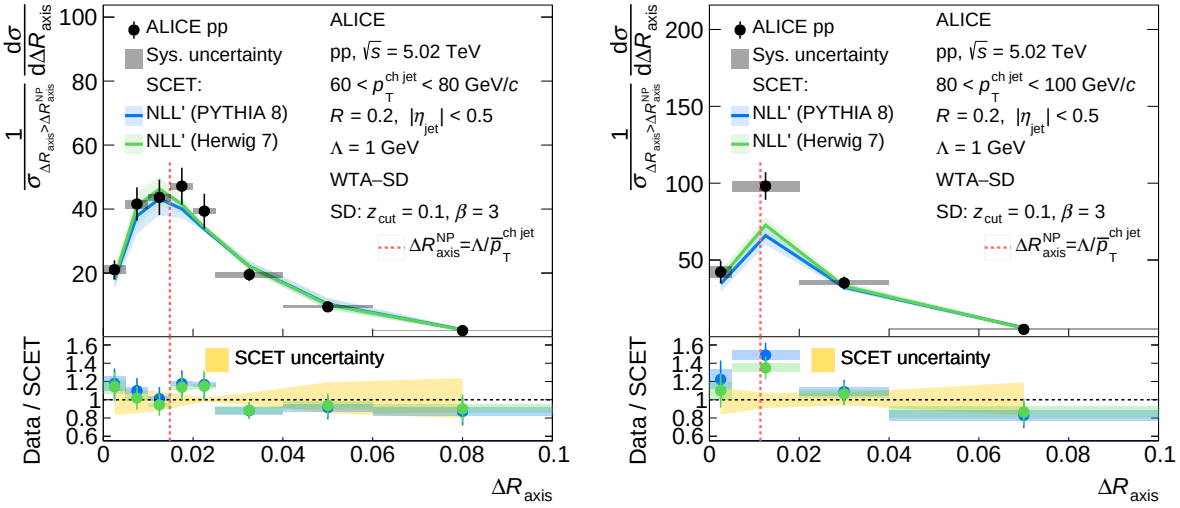


Figure B.12: Same as Figure B.1, for WTA−SD ($z_{\rm cut} = 0.1$, $\beta = 3$) for $R = 0.2$.

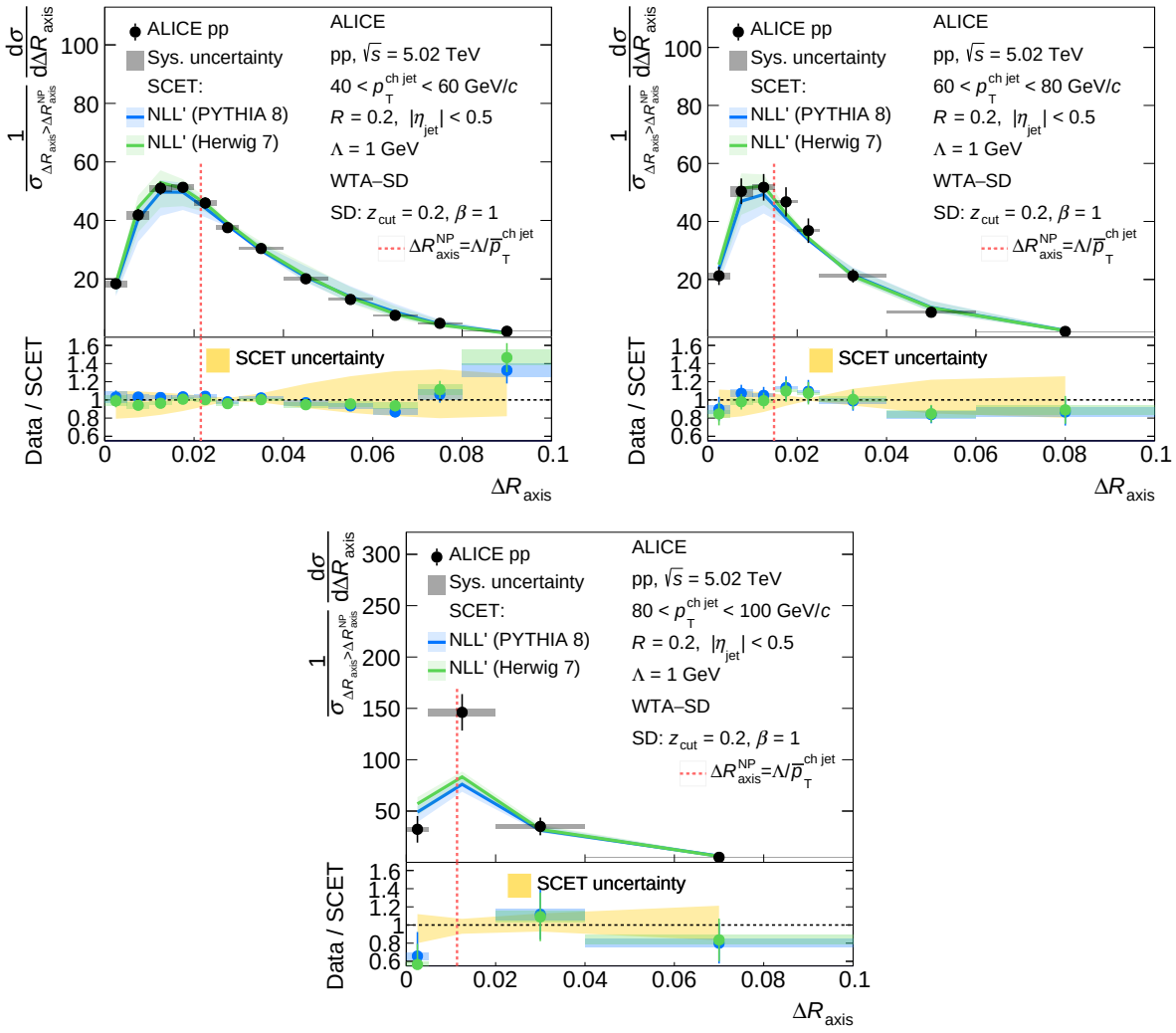


Figure B.13: Same as Figure B.1, for WTA−SD ($z_{\mathrm{cut}} = 0.2$, $\beta = 1$) for $R = 0.2$.

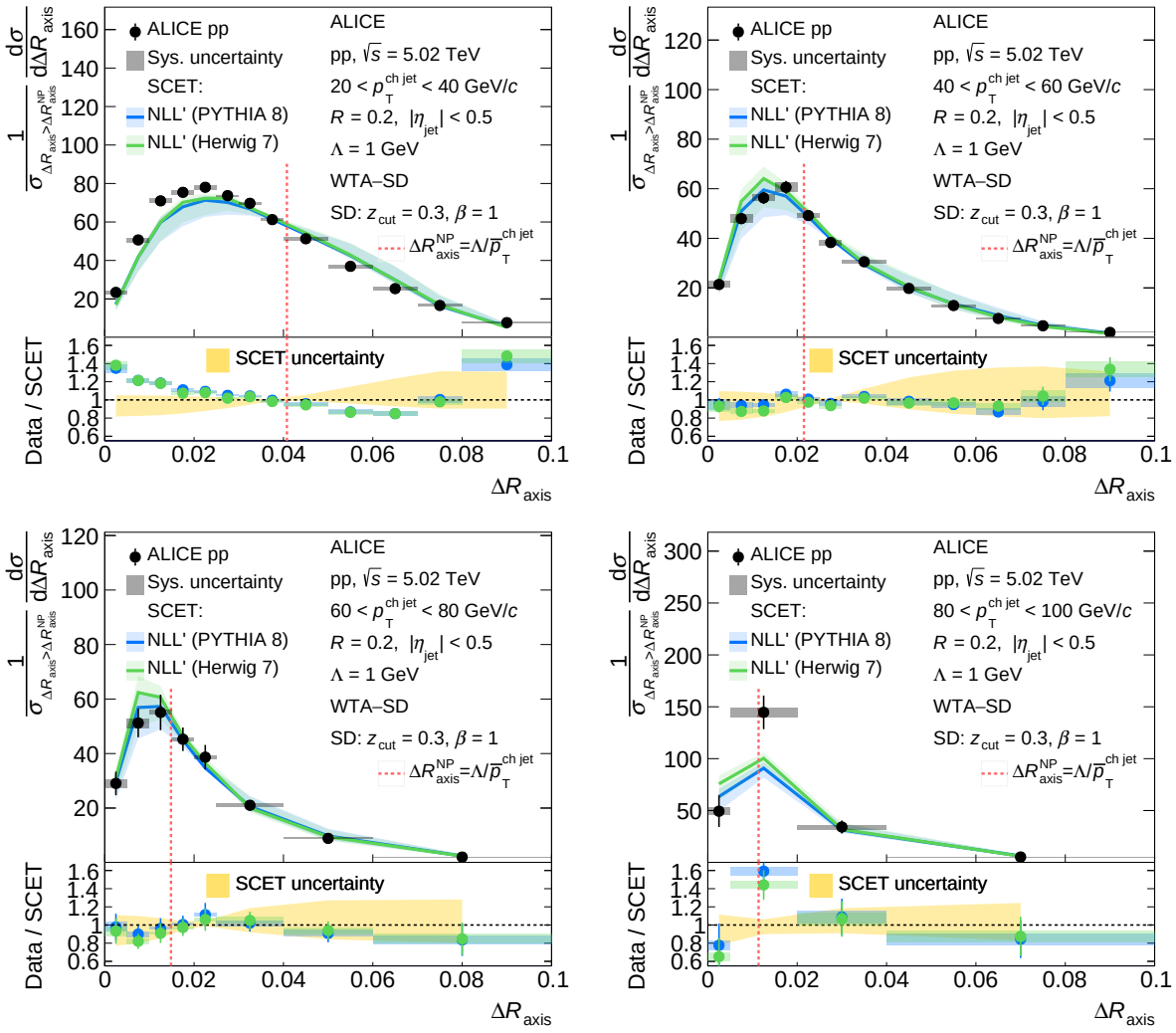


Figure B.14: Same as Figure B.1, for WTA−SD ($z_{\mathrm{cut}} = 0.3$, $\beta = 1$) for $R = 0.2$.