Measurements of jet substructure in pp and Pb-Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV with ALICE

by

Ezra D. Lesser

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Barbara V. Jacak, Chair
Professor Marjorie D. Shapiro
Professor Karl A. van Bibber

Fall 2023

CERN-THESIS-2023-218
22/10/2023



Measurements of jet substructure in pp and Pb-Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV with ALICE

Copyright 2023
by
Ezra D. Lesser

Abstract

Measurements of jet substructure in pp and Pb-Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV with ALICE

by

Ezra D. Lesser

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Barbara V. Jacak, Chair

Jet substructure observables are powerful tools to search for new physics and test theoretical descriptions of perturbative and non-perturbative processes in QCD. In heavy-ion collisions, jet substructure observables are used to elucidate the structure and dynamics of the quark-gluon plasma. One substructure observable is jet mass, which probes the virtuality of hard-scattered partons and their modified fragmentation. Additionally, generalized jet angularities allow differential measurements of the jet shower and its modification, as two parameters independently vary the weight of the jet constituents' relative angle and transverse momentum. Previous measurements of the jet mass and jet angularities show conflicting deviations when compared with models. This thesis presents new measurements of the jet mass and jet angularities to resolve this conflict, using charged-particle tracks in pp and Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV and jet resolution parameters $R = 0.2$ and 0.4. The results from this work are compared to ALICE measurements of heavy-flavor jets, which provide a high-powered probe of perturbative QCD at low transverse momentum. Jet angularities in jets containing a charm meson are compared to the inclusive measurements of this thesis, revealing a significant narrowing due to the QCD dead cone modifying jet fragmentation. Jet angularity results are also compared to QCD predictions using both folding and shape function corrections for nonperturbative effects. Jet grooming can be used to isolate specific splittings inside the jet fragmentation history, and ALICE measurements of the groomed-jet splitting angle and momentum fraction with soft drop and dynamical grooming are also compared to QCD predictions to test parton branching. The high-precision tracking system of ALICE enables these measurements over a broad range in transverse momentum, with a low-transverse momentum reach that is unique at the LHC.

Draw from the well of unchanging
And its union nourishes on
In the right rearranging
Till the last confusion is gone
Water brothers trust in the ultimust
Of the always-singing song they pass along

James Thomas Hall, 1967

If I have seen further, it is by standing on the shoulders of giants.

Sir Isaac Newton, 1675

Contents

Contents **ii**

List of Figures **iv**

List of Tables **xvii**

1 Introduction **1**
1.1 Quantum chromodynamics 2
1.2 Scattering experiments 7
1.3 Jets 15
1.4 Relativistic heavy-ion collisions 26
1.5 Statement of purpose 39

2 Experimental apparatus **43**
2.1 The Large Hadron Collider (LHC) 43
2.2 A Large Ion Collider Experiment (ALICE) 46

3 Analysis procedure **56**
3.1 Data sets 56
3.2 Jet reconstruction 59
3.3 Unfolding 63
3.4 Systematic uncertainties 83

4 Theoretical corrections **90**
4.1 Parameters of obtained predictions 90
4.2 Folding-based corrections 93
4.3 Nonperturbative shape function 101

5 Results **105**
5.1 Groomed jet splitting observables in vacuum 106
5.2 Jet angularities in vacuum 109
5.3 Quenched jet mass and angularities 129

6 Discussion and conclusions **151**
6.1 Discussion 151
6.2 Outlook 155

Bibliography **157**

A Additional kinematic efficiency checks **182**
A.1 Generator resilience 182
A.2 Test at higher jet transverse momentum 184

B Heavy-flavor tagged jet analysis **187**
B.1 Data sets 187
B.2 HF meson and jet reconstruction 188
B.3 Jet corrections 189
B.4 Systematic uncertainties 193

List of Figures

1.1 Feynman diagrams for the gluon self-interaction terms in QCD. 3
1.2 Experimental measurements of α_s at varying energy scale Q. The perturbative accuracy of the extraction is given in parentheses for each observable. For details, see Ref. [3]. 4
1.3 Leading-order (two-vertex) Feynman diagrams for $2 \to 2$ scattering in QFT, labelled by particle four-momenta $\mathbf{p}_i$. These diagrams are read from left to right, with $i = 1, 2$ corresponding to incoming particles, and $i = 3, 4$ to outgoing particles. 8
1.4 Structure function $F_2(x, Q^2)$ for the proton, as measured by different experiments for varying x and Q^2 [3]. The structure function varies only weakly with respect to Q^2, despite the horizontal axis spanning five orders of magnitude. 10
1.5 Unpolarized parton distribution functions $xf(x)$ at the energy scales of $\mu^2 = 10$ GeV^2 (left) and 10^4 GeV^2 (right), as evaluated at NNLO [3, 30]. The horizontal axis is the momentum fraction (Bjorken) x. The valence up and down quarks ($f = \mathrm{u_v}$ and $\mathrm{d_v}$, respectively) dominate at large values of x, while the probability of scattering off gluons (g) and sea quarks (s, c, b, antiquarks) dominates in the low-x regime. 11
1.6 Example s-channel (quark-antiquark annihilation) Feynman diagrams corresponding to a three-jet event. The addition of an extra vertex for the outgoing gluon suppresses these matrix elements by a factor of g_s with respect to the leading-order $2 \to 2$ ones. 16
1.7 Experimental results from TASSO [48] depicting both dijet (left) and 3-jet (right) events, with projections onto three orthogonal planes. Reconstructed tracks are drawn as lines with length corresponding to their momentum. 17
1.8 Comparison of jet sizes and shapes for a sample simulated event, as reconstructed with the k_T, Cambridge/Aachen, SISCone, and anti-k_T algorithms [66]. 19
1.9 Proposed procedure for tagging Higgs-initiated jets via the decay into a $\mathrm{b\bar{b}}$ pairs [71]. A large-radius ($R \approx 1$) "fat" jet is first reconstructed using the CA algorithm, containing the boosted $\mathrm{b\bar{b}}$ pair along with any soft and collinear radiation. The fat jet is then split into subjets using the mass drop procedure, which searches for a sufficiently symmetric and heavy splitting in the jet clustering history. The radius of these two subjets is then reduced to $R_\mathrm{filt} \approx 0.4$ in order to filter out contamination from the UE. 20

1.10 Depictions of the grooming procedure. *Left:* radiations are iteratively removed from a jet of radius R until a splitting satisfying the grooming condition is satisfied, from which jet substructure observables can be calculated. *Right:* SD grooming removes soft (small z) and wide-angle (large θ) radiation, while retaining some soft-collinear information, as determined by the scaling of β. 22
1.11 Diagram representing the procedure for the direct measurement of the dead cone [119]. At the top, the CA algorithm reclusters the jet in order of smallest angles first, reconstructing the angularly-ordered parton shower. After the reclustering is complete, the jet is then declustered, with the heavy quark energy E_{Radiator} and momentum axis adjusted with each subsequent emission. The radiation angle θ_i and the heavy quark energy $E_{\text{Radiator},i}$ is recorded at each declustering step i. 24
1.12 Experimental results for the direct dead cone measurement of D^0-tagged jets in increasing bins of E_{Radiator} [119]. Pink shaded areas represent the kinematic region $\theta < \theta_0$, where gluon radiation is expected to be suppressed. The charm quark mass is assumed to be $m_{\text{c}} = 1.275$ GeV/c^2. Data are compared to the PYTHIA 8 and SHERPA MC event generators, both in the nominal charm quark case and in the light quark (LQ) / inclusive limit, where no dead cone effect is expected. 25
1.13 *Left:* first experimental observation of the difference between structure functions for iron (Fe) and deuterium (D), known as the EMC effect [128]. The enhancement at small x was not predicted, as it was expected that the ratio would increase or be approximately consistent with unity. It was later shown [131] that the modification increases with nuclear size. *Right:* sketch of the nucleus-nucleon structure function ratio $R_{F_2}^{\text{A}} = F_2^{\text{A}}(x, Q^2)/[\text{A}F_2^{\text{nucleon}}(x, Q^2)]$ for a generic nucleus and nucleon [132]. See text for explanation of different regions in x. 27
1.14 Hyperon enhancement as observed by the WA97 experiment [150]. Hyperon production is found to scale faster than the average number of binary particle interactions $\langle N_{\text{part}} \rangle$, with a 1:1 scaling given by the solid line. Instead, Ω production is found to be proportional to $\langle N_{\text{part}} \rangle^{1.72}$ (dashed line). 29
1.15 Lattice QCD predictions of the pressure p, energy density ϵ, and entropy density s for the QGP at varying temperatures [157, 158]. The ideal gas (non-interacting / Stefan-Boltzmann) limit is shown as a black dashed line. The phase transition is a continuous crossover, with smoothly rising spectra above the critical temperature $T_{\text{c}} \approx 150$ MeV. The Hadron Resonance Gas (HRG) model [159] is effective at describing the hadronic equation of state at low temperatures, but diverges above the QGP phase transition. 30
1.16 Cartoon of a non-central AA collision [160], with the two outgoing nuclear remnants colored purple, and the overlapping region forming a QGP (orange). The horizontal plane of symmetry formed by the collision is known as the *reaction plane* (green). The QGP is oblong or almond-like in shape, and its internal pressure gradient causes a nonuniform expansion as predicted by relativistic hydrodynamics. 31

1.17 Measurement of the event multiplicity distribution with the ALICE detector using Pb–Pb data at $\sqrt{s_{\rm NN}} = 5.02$ TeV [164, 165]. Data from the ALICE V0 detectors (see Sect. 2.2.4) is fit with Glauber MC coupled to a negative binomial distribution (NBD) particle production model. Integrated centrality regions are shown below the curve. 32
1.18 Measurement of the elliptic flow coefficient $\bar{v}_2$ in Pb–Pb collisions with centrality ranging from 0 to 60% and for several particle species as a function of their $p_{\rm T}$, as calculated using the Scalar Product (SP) method [167]. The flow increases for more peripheral collisions, where events produce a more anisotropic QGP. The mass ordering for $p_{\rm T} < 3$ GeV is given by an interplay between the radial and elliptic flow which modifies $\bar{v}_2$ according to mass. After some threshold value, increases of particle $p_{\rm T}$ decrease the effect of flow, as the particles tend to rapidly emerge from the fluid. 34
1.19 Schematic view of the phase diagram of QCD as a function of temperature and baryon density [171]. The existence of a QCD critical point has not been confirmed and is currently being searched for by the Beam Energy Scan, Phase II (BES-II) at the Relativistic Heavy Ion Collider (RHIC) on Long Island in New York. Similarly, the proposed color superconductor phase has not experimentally been probed, and is still under investigation. 35
1.20 Recent measurements of the jet cross section in pp and Pb–Pb collisions (left) and the corresponding nuclear modification factor $R_{\rm AA}$ (right) [180]. Results are shown for centrality ranging from 0 to 80%, in bins with width 10%. Significant modification is observed for all reported centralities, with drastically enhanced quenching observed for the more central collisions. 37
1.21 Experimental measurements resulting in the girth-mass puzzle [73, 223]. *Left:* measurement of the jet girth $g = \lambda_1 R$ in Pb–Pb data and compared to pp models, as given by PYTHIA. Significant quenching is observed in the Pb–Pb/pp ratios, with an approximate factor of 2 enhancement at small angles, with a corresponding suppression at large angles. *Right:* measurement of the jet mass $m_{\rm jet}$ in Pb–Pb data, theoretically related to the girth through its connection to the thrust (Eq. 1.28). No significant quenching effect is observed within uncertainties. 40

2.1 Diagram of the CERN accelerator complex [245]. 44
2.2 Cross-sectional diagram of ALICE at the CERN LHC during Run 2 (2015-2018) [255]. Different subdetector systems are labelled, and the inner ITS region is shown in an expanded view in the upper right. See text for details. 47
2.3 Distribution of the specific energy loss ($\mathrm{d}E/\mathrm{d}x$) in the ITS (left) and TPC (right) [257]. The solid black curves depict the per-particle parametrizations of the expected mean energy loss. The slow "relativistic rise" regime is visible at low p. 50

2.4 *Left:* A diagram showing charge amplification and collection for the MWPCs of the ALICE TPC, leading to a three-dimensional reconstruction of the original charged particle trajectory. *Right:* an example Pb–Pb event in the ALICE TPC [261]. 1582 positively-charged (darker tracks) and 1579 negatively-charged (lighter tracks) particles are produced; about 80 percent are π mesons. 51
2.5 *Left:* example mass probability density functions f_i for π, K, and p hadrons, with Gaussian fits to guide the eye [268]. The probability $P_{\mathrm{K}}(m)$ of a track with measured mass m to correspond to a K meson is also shown. Tracks are required to match in the TPC and ITS, and have $0.5 \leq p \leq 2.5$ GeV/c. Note that the y-scale is logarithmic. *Right:* speed $\beta = v/c$ distribution for tracks entering the TOF detector. 53
2.6 Inverse track p_{T} resolution $\sigma_{1/p_{\mathrm{T}}}$ for tracks found at various stages in the ALICE fitting procedure [257]. $\sigma_{1/p_{\mathrm{T}}}$ is related to the p_{T} resolution $\sigma_{p_{\mathrm{T}}}$ by $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = p_{\mathrm{T}}\,\sigma_{1/p_{\mathrm{T}}}$. 54

3.1 Diagram of the CS method [287]. *Left:* an initial hard-scattering produces a jet (blue) inside an uncorrelated background (red). Ghost particles (green) are deployed uniformly in the event. *Right:* after CS, the p_{T} of most tracks is reduced. A small number of ghost particles also remain (circled), which are now included in the event. 61
3.2 *Left:* all possible matchings between combined- and detector-level $R = 0.2$ jets in Pb–Pb MC. A dashed line is plotted at the matching condition of $\Delta R < 0.6R$. *Right:* $p_{\mathrm{T}}^{\text{ch jet}}$ smearing between combined- and detector-level $R = 0.2$ jets, using CS with $R_{\mathrm{max}} = 0.1$. [286] 63
3.3 PYTHIA 8 + GEANT3 + embedding + background subtraction characterizations of $R = 0.2$ charged-jet reconstruction in Pb–Pb collisions, using CS with $R_{\mathrm{max}} = 0.1$ [286]. All distributions are shown per $p_{\mathrm{T,truth}}^{\text{ch jet}}$ interval (labelled $p_{\mathrm{T}}^{\mathrm{gen}}$). The value of $p_{\mathrm{T}}^{\mathrm{det}}$ corresponds to combined, background-subtracted, and matched jets. See text for details. *Upper left:* jet-by-jet JES. *Upper right:* mean JES shift. *Lower left:* JER. *Lower right:* jet reconstruction efficiency. 64
3.4 PYTHIA 8 distributions of λ_1 for inclusive $R = 0.2$ jets in Pb–Pb collisions, in $p_{\mathrm{T}}^{\text{ch jet}}$ bins ranging from 40 to 150 GeV/c [286]. The distributions are shown both at truth level (in blue, before the particles undergo interactions with the detector) and at detector level (in red, after events are reconstructed by the ALICE GEANT3 detector simulation, embedded in Pb–Pb data, and the CS procedure is applied). 65
3.5 Example detector/truth matching. *Left:* the $p_{\mathrm{T}}^{\text{ch jet}}$ matching is shown for pp truth- and detector-level jets, as calculated via the anchored MC simulation [289]. *Right:* a projection of the Pb–Pb RM onto the truth- and detector-level (combined and subtracted) observable axes for λ_1, binned according to the measured data [286]. In both plots, a large diagonal component is observed, with smearing due to detector effects. 66

3.6 Raw spectrum of λ_1 versus $p_{\mathrm{T,det}}^{\mathrm{ch\,jet}}$ for $R = 0.2$ charged-particle jets in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV [286]. The number of counts in each bin is displayed. To make the unfolding procedure stable, the binning is optimized to limit the number of bins having $N < 10$ counts. 67
3.7 Distributions of detector-level (embedded, background-subtracted) λ_1 for $R = 0.2$ jets in simulated Pb–Pb collisions from PYTHIA8 + GEANT3 compared to raw data [286]. Reasonable agreement is observed for all $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$ bins, indicating that unfolding can be used. Shape biases are considered as a systematic uncertainty; see Sect. 3.4. 68
3.8 Example shape closure tests (+0.5 variation) for $\lambda_{1.5}$ using the anchored MC production, and scaled by Eq. 3.5, as measured in Pb–Pb collisions for $R = 0.2$ jets in the lowest $40 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 60$ GeV/c bin [286]. The solutions have increasing n_{iter} from 5 (left) to 10 (middle) to 15 (right), demonstrating a simultaneous improvement in the closure. Note that the highest λ_α bin is truncated from the reported distributions, but it is included in the unfolding for efficiency considerations. . . 70
3.9 Example convergence tests λ_1 in Pb–Pb collisions for $R = 0.2$ jets [286]. Note that the last bin is truncated from the reported distributions, but it is included in the unfolding for efficiency considerations. 71
3.10 Example refolding tests projected onto the $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$ and observable axes for λ_1 in Pb–Pb collisions [286]. Each plot is for one of the respective $p_{\mathrm{T,det}}^{\mathrm{ch\,jet}}$ bins, with the binning used for the lowest $40 \leq p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 60$ GeV/c reported distribution. As can be seen, the test is successful for the relevant range in each case. *Upper left:* $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$ projection for all $p_{\mathrm{T,det}}^{\mathrm{ch\,jet}}$ bins; *upper right:* λ_α projection for $40 \leq p_{\mathrm{T,det}}^{\mathrm{ch\,jet}} < 45$ GeV/c; *bottom left:* λ_α projection for $45 \leq p_{\mathrm{T,det}}^{\mathrm{ch\,jet}} < 50$ GeV/c; *bottom right*: λ_α projection for $50 \leq p_{\mathrm{T,det}}^{\mathrm{ch\,jet}} < 60$ GeV/c. 72
3.11 Example statistical closure tests for λ_1 using the anchored MC production smeared by data uncertainties as measured in Pb–Pb collisions for $R = 0.2$ jets [286]. Each plot is given for one of the $p_{\mathrm{T,truth}}^{\mathrm{ch\,jet}}$ bins, each using a different binning (and hence a different RM) to account for statistical and shape differences. Note that the last bin is truncated from the reported distributions, but it is included in the unfolding for efficiency considerations. 73
3.12 Example shape closure tests (+0.5 variation) for λ_1 using the anchored MC production, and scaled by Eq. 3.5, as measured in Pb–Pb collisions for $R = 0.2$ jets [286]. Each plot is given for one of the $p_{\mathrm{T,truth}}^{\mathrm{ch\,jet}}$ bins, each using a different binning (and hence a different RM) to account for statistical and shape differences. Note that the last bin is truncated from the reported distributions, but it is included in the unfolding for efficiency considerations. 74

3.13 Example prior closure tests (+0.5 variation) for λ_1 using the anchored MC production, and scaled by Eq. 3.5, as measured in Pb–Pb collisions for $R = 0.2$ jets [286]. Each plot is given for one of the $p_{\mathrm{T,truth}}^{\mathrm{ch\ jet}}$ bins, each using a different binning (and hence a different RM) to account for statistical and shape differences. Note that the last bin is truncated from the reported distributions, but it is included in the unfolding for efficiency considerations. 75
3.14 Example statistical uncertainty divergence for the unfolded λ_1 distributions in Pb–Pb collisions using $R = 0.2$ jets as a function of n_{iter} [286]. Note that the last bin is truncated from the reported distributions, but it is included in the unfolding for efficiency considerations. 76
3.15 Example thermal closure tests of λ_1 for $R = 0.2$ jets in MC embedded into simulated thermal background for Pb–Pb, then subtracted with CS [286]. Any non-closure is taken as a systematic uncertainty. Note that the last bin is truncated from the reported distributions, but it is included in the unfolding for efficiency considerations. 78
3.16 Sketch of the regions of phase space covered a response matrix, as projected onto $p_{\mathrm{T}}^{\mathrm{jet}}$ [293]. 79
3.17 Example 2D test plots of the kinematic efficiency in pp for inclusive $R = 0.2$ charged jets in the truncated range of $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ and λ_α, and corresponding to the detector-level binning cuts described in the text [289]. The color scale represents the MC-estimated kinematic efficiency, which ranges from 0 (worst) to 1 (best). Plots are (left to right, then top to bottom) for $\alpha = [1, 1.5, 2, 3]$. The λ_α binning used in these distributions is the binning choice utilized for $p_{\mathrm{T}}^{\mathrm{ch\ jet}} \in [20, 40]$ GeV/c; the top λ_α bin is not reported. 81
3.18 Example kinematic efficiency distributions of λ_1 for inclusive $R = 0.2$ charged jets in Pb–Pb collisions, using several $p_{\mathrm{T,truth}}^{\mathrm{ch\ jet}}$ selections, as evaluated using the anchored PYTHIA 8 + GEANT3 MC simulation [286]. The efficiency is worst for the lowest $40 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 60$ GeV/c bin, due to the requirement that $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}} > 40$ GeV/c to remove combinatorial jets from the unfolding procedure and thus ensure good thermal closure. 82
3.19 Example unfolding systematic uncertainties of λ_1 for $R = 0.2$ charged-particle jets in Pb–Pb collisions [286]. See text for details on each variation. 85
3.20 Generator systematic uncertainty of λ_1 for $R = 0.2$ charged-particle jets in Pb–Pb collisions [286]. Generator 1 is Herwig 7 default tune. Generator 2 is JEWEL Pb-Pb with recoils off. Each are taken as a ratio to PYTHIA 8 Monash 2013 through the same fast simulation. 86
3.21 Distributions of δp_{T} in Pb–Pb collisions for $R = 0.2$ charged-particle jets [286]. The top plot is "undersubtracted", with $R_{\mathrm{max}} = 0.05$; the middle plot is the nominal case, with $R_{\mathrm{max}} = 0.1$; the bottom row is "oversubtracted" with $R_{\mathrm{max}} = 0.5$. 88
3.22 Total systematic uncertainty of λ_1 for $R = 0.2$ charged-particle jets in Pb–Pb collisions, along with individual constituents. The total uncertainty is taken as the quadrature sum of all contributions. [286] 89

4.1 $p_{\mathrm{T}}^{\mathrm{jet}}$ residuals between parton and charged-hadron level for $R = 0.2$ (left) and $R = 0.4$ (right) charged-particle jets in PYTHIA 8 [289]. 94
4.2 $p_{\mathrm{T}}^{\mathrm{jet}}$ residuals between parton and charged-hadron level for $R = 0.2$ (left) and $R = 0.4$ (right) charged-particle jets in Herwig 7 [289]. 94
4.3 $p_{\mathrm{T}}^{\mathrm{jet}}$ response matrix projection for $R = 0.2$ (left) and $R = 0.4$ (right) charged-particle jets in PYTHIA 8 between parton and charged-hadron levels [289]. . . . 95
4.4 $p_{\mathrm{T}}^{\mathrm{jet}}$ response matrix projection for $R = 0.2$ (left) and $R = 0.4$ (right) charged-particle jets in Herwig 7 between parton and charged-hadron levels [289]. 95
4.5 Comparison of the $\lambda_{1.5}$ distributions according to parton-level theory predictions with the selected event generators for $R = 0.2$ jets the 10-90 GeV/c range [289]. 97
4.6 Comparison of the $\lambda_{1.5}$ distributions according to parton-level theory predictions with the selected event generators for $R = 0.2$ jets the 90-200 GeV/c range [289]. 98
4.7 Example ratios depicting the size of $\lambda_{1.5}$ correction between parton, hadron, charged-hadron, and charged-hadron with MPI levels, for default PYTHIA 8 (top), PYTHIA 8 with lower `TimeShower` cutoff (middle), and Herwig 7 (bottom). The matched jets use radius parameter $R = 0.2$ and are projected for $p_{\mathrm{T}}^{\mathrm{ch\ jet}} \in [60, 80]$ GeV/c. 100
4.8 Comparison of the parton-level theory predictions for $\lambda_{1.5}$ with nonperturbative shape function convolution versus PYTHIA 8 for $R = 0.2$ jets in the 10-90 GeV/c range. 103
4.9 Comparison of the parton-level theory predictions for $\lambda_{1.5}$ with nonperturbative shape function convolution versus PYTHIA 8 for $R = 0.2$ jets in the 90-200 GeV/c range. 104

5.1 ALICE measurements of z_{g} (left) and θ_{g} (right) distributions in pp collisions at $\sqrt{s} = 5.02$ TeV with SD for three values of the grooming parameter β, compared with PYTHIA 8 Monash 2013 [37, 38] calculations [300]. 106
5.2 ALICE measurements of θ_{g} distributions with SD, compared with pQCD predictions including nonperturbative corrections. The integral of the perturbative region, $\theta_{\mathrm{g}} > \theta_{\mathrm{g}}^{\mathrm{NP}}$ (to the right of the dashed vertical blue line; see Sect. 4.1), is set at unity. The nonperturbative scale is taken to be $\Lambda = 1$ GeV/c [300]. 107
5.3 ALICE measurements of z_{g} (left) and θ_{g} (right) distributions in pp collisions at $\sqrt{s} = 5.02$ TeV with dynamical grooming for three values of the grooming parameter a, compared with PYTHIA 8 Monash 2013 [37, 38] calculations [300]. 107
5.4 ALICE measurements of z_{g} (top) and θ_{g} (bottom) distributions in pp collisions at $\sqrt{s} = 5.02$ TeV with dynamical grooming for two values of the grooming parameter a as compared with pQCD calculations [300]. 108
5.5 Comparison of ungroomed jet angularities λ_α in pp collisions for $R = 0.4$ to MC predictions using PYTHIA 8 and Herwig 7, as described in the text. Four equally-sized $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ intervals are shown, with edges ranging between 20 and 100 GeV/c. The distributions are normalized to unity [301]. 111

5.6 Comparison of ungroomed jet angularities λ_α in pp collisions for $R = 0.2$ to MC predictions using PYTHIA 8 and Herwig 7, as described in the text. Four equally-sized $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ intervals are shown, with edges ranging between 20 and 100 GeV/c. The distributions are normalized to unity [301]. 112
5.7 Comparison of groomed jet angularities $\lambda_{\alpha,\mathrm{g}}$ in pp collisions for $R = 0.4$ to MC predictions using PYTHIA 8 and Herwig 7, as described in the text. Four equally-sized $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ intervals are shown between 20 and 100 GeV/c. The distributions are normalized to the groomed jet tagging fraction [301]. 113
5.8 Comparison of groomed jet angularities $\lambda_{\alpha,\mathrm{g}}$ in pp collisions for $R = 0.2$ to MC predictions using PYTHIA 8 and Herwig 7, as described in the text. Four equally-sized $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ intervals are shown between 20 and 100 GeV/c. The distributions are normalized to the groomed jet tagging fraction [301]. 114
5.9 Comparison of ungroomed jet angularities λ_α in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions with MC hadronization corrections in the range $60 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 80$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_\alpha > \lambda_\alpha^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region. 115
5.10 Comparison of groomed jet angularities $\lambda_{\alpha,\mathrm{g}}$ in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions with MC hadronization corrections in the range $60 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 80$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_{\alpha,\mathrm{g}} > \lambda_{\alpha,\mathrm{g}}^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region. 116
5.11 Comparison of ungroomed jet angularities λ_α in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions with MC hadronization corrections in the range $80 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 100$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_\alpha > \lambda_\alpha^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region. 117
5.12 Comparison of groomed jet angularities $\lambda_{\alpha,\mathrm{g}}$ in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions with MC hadronization corrections in the range $80 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 100$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_{\alpha,\mathrm{g}} > \lambda_{\alpha,\mathrm{g}}^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region. 118
5.13 Comparison of ungroomed jet angularities λ_α in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions using $F(k)$ convolution in the range $40 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 60$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_\alpha > \lambda_\alpha^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region. . 119

5.14 Comparison of groomed jet angularities $\lambda_{\alpha,\mathrm{g}}$ in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions using $F(k)$ convolution in the range $40 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 60$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_{\alpha,\mathrm{g}} > \lambda_{\alpha,\mathrm{g}}^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region. . 120
5.15 Comparison of ungroomed jet angularities λ_{α} in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions using $F(k)$ convolution in the range $60 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 80$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_{\alpha} > \lambda_{\alpha}^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region. . 121
5.16 Comparison of groomed jet angularities $\lambda_{\alpha,\mathrm{g}}$ in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions using $F(k)$ convolution in the range $60 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 80$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_{\alpha,\mathrm{g}} > \lambda_{\alpha,\mathrm{g}}^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region. . 122
5.17 Comparison of ungroomed jet angularities λ_{α} in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions using $F(k)$ convolution in the range $80 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 100$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_{\alpha} > \lambda_{\alpha}^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region. . 123
5.18 Comparison of groomed jet angularities $\lambda_{\alpha,\mathrm{g}}$ in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions using $F(k)$ convolution in the range $80 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 100$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_{\alpha,\mathrm{g}} > \lambda_{\alpha,\mathrm{g}}^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region. . 124
5.19 Comparison of D^0-tagged jet angularities λ_{α} to semi-inclusive baseline in pp collisions for $R = 0.4$ charged-particle jets and $\alpha = 1$ (top) or 1.5 (bottom). Results are compared to PYTHIA MC predictions for both results. Jets are reported in $10 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 20$ GeV/c with $p_{\mathrm{T}}^{\mathrm{D}^0} > 5$ GeV/c, and the distributions are normalized to unity [303]. 126
5.20 Comparison of D^0-tagged jet angularities λ_{α} to semi-inclusive baseline in pp collisions for $R = 0.4$ charged-particle jets and $\alpha = 2$ (top) or 3 (bottom). Results are compared to PYTHIA MC predictions for both results. Jets are reported in $10 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 20$ GeV/c with $p_{\mathrm{T}}^{\mathrm{D}^0} > 5$ GeV/c, and the distributions are normalized to unity [303]. 127
5.21 Comparisons of the inclusive quark-, gluon-, and D^0-initiated charged-particle jet angularities in PYTHIA 8 at $\sqrt{s} = 5.02$ TeV. Generated distributions require a unique angular matching between a hard-scattered parton and the charged-particle jet [303]. 128

5.22 ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_1 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $40 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 60$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 131
5.23 ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_1 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $60 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 80$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 132
5.24 ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_1 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $80 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 100$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 133
5.25 ALICE measurement of ungroomed (left) and SD groomed (right) λ_1 for $R = 0.2$ charged-particle jets in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $100 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 150$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 134
5.26 ALICE measurement of ungroomed (top) and SD groomed (bottom) $\lambda_{1.5}$ for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $40 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 60$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 135
5.27 ALICE measurement of ungroomed (top) and SD groomed (bottom) $\lambda_{1.5}$ for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $60 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 80$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 136
5.28 ALICE measurement of ungroomed (top) and SD groomed (bottom) $\lambda_{1.5}$ for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $80 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 100$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 137
5.29 ALICE measurement of ungroomed (left) and SD groomed (right) λ_1 for $R = 0.2$ charged-particle jets in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $100 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 150$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 138

5.30 ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_2 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $40 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 60$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 139
5.31 ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_2 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $60 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 80$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 140
5.32 ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_2 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $80 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 100$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 141
5.33 ALICE measurement of ungroomed (left) and SD groomed (right) λ_2 for $R = 0.2$ charged-particle jets in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $100 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 150$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 142
5.34 ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_3 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $40 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 60$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 143
5.35 ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_3 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $60 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 80$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 144
5.36 ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_3 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $80 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 100$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 145
5.37 ALICE measurement of ungroomed (left) and SD groomed (right) λ_3 for $R = 0.2$ charged-particle jets in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $100 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 150$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 146

5.38 ALICE measurement of ungroomed (top) and SD groomed (bottom) m_{jet} for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $40 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 60$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 147
5.39 ALICE measurement of ungroomed (top) and SD groomed (bottom) m_{jet} for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $60 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 80$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 148
5.40 ALICE measurement of ungroomed (top) and SD groomed (bottom) m_{jet} for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $80 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 100$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 149
5.41 ALICE measurement of ungroomed (left) and SD groomed (right) m_{jet} for $R = 0.2$ charged-particle jets in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $100 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 150$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching. 150

6.1 Measurement of the jet girth $g = \lambda_1^1 R$ in Pb–Pb data at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV from LHC Run 1 (left) and at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV from Run 2 (right). On the left figure, the bottom ratio panel compares data to PYTHIA baselines, as does the central ratio panel on the right figure. However, the comparison to a proper pp baseline in the bottom ratio panel significantly pushes the ratio closer to unity. 153
6.2 Measurement of the jet mass m_{jet} (left) and thrust λ_2^1 (right) in Pb–Pb and pp data at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV, as presented in Sect. 5.3. Systematic uncertainties are assumed to be totally uncorrelated between pp and Pb–Pb. Comparisons are made to several models [192, 195, 197–201, 305], which show varying behavior with respect to measured data. 154

A.1 Kinematic efficiency of λ_1 for $R = 0.2$ jets in Pb–Pb collisions for $40 < p_{\mathrm{T,truth}}^{\mathrm{ch\,jet}} < 60$ GeV/c with four different generator selections, each combined with the same fast simulation [286]. Top left: PYTHIA8 Monash 2013. Top right: Herwig7 default tune. Bottom left: JEWEL Pb–Pb, recoils off. Bottom right: JEWEL Pb–Pb, recoils on, without recoil subtraction. The final bin is truncated from the reported distributions. 183
A.2 Kinematic efficiency of λ_α for $R = 0.2$ charged-particle jets for $60 < p_{\mathrm{T,truth}}^{\mathrm{ch\,jet}} < 80$ GeV/c with a stricter $p_{\mathrm{T,det}}^{\mathrm{ch\,jet}}$ cut, as simulated using the anchored PYTHIA 8 + GEANT3 production embedded in Pb–Pb data and background-subtracted with CS [286]. The final bin is truncated from the reported distributions. All four plots are computed using the same jets, with $\alpha = 1$, 1.5, 2, and 3. 185

A.3 Example closure tests for λ_1 with $R = 0.2$ charged-particle jets using a stricter $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}}$ cut, as simulated using the anchored PYTHIA 8 + GEANT3 production embedded in Pb–Pb data and background-subtracted with CS [286]. The final bin is truncated from the reported distributions. The test is considered successful in each case. 186

B.1 Two diagrams of a $\mathrm{D}^0 \to \mathrm{K}^- \pi^+$ decay [303]. *Left:* a "prompt" D^0 meson, originating from the hard scattering, decays at a secondary vertex $\mathrm{V_{sec}}$, nominally 122 μm away from the IP in its rest frame. *Right:* a "non-prompt" D^0 meson, originating from the "feed-down" decay of a B meson, decays with slightly different kinematics. 188

B.2 *Left:* example D^0 invariant mass spectra with signal and background fits, using $8 \leq p_{\mathrm{T}}^{\mathrm{D}^0} < 13$ GeV/c candidates inside of jets with $10 \leq p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 20$ GeV/c. *Right:* the corresponding correction of the jet angularity λ_1 (girth) raw yield. 190

B.3 *Left:* prompt (red) and non-prompt (blue) D^0-meson jet acceptance $\times$ efficiency as a function of detector-level $p_{\mathrm{T}}^{\mathrm{D}^0}$, as evaluated with PYTHIA 8 + GEANT3 [317]. *Right:* fraction of jet angularities λ_1 coming from the feed-down of a B meson, as calculated with POWHEG + PYTHIA 8 and folded to detector level, versus the number of inclusive, background-subtracted D^0-tagged jets, as measured in data [303]. Systematic uncertainties (shaded bands) are discussed in Sect. B.4. . 191

B.4 Total systematic uncertainty of D^0-tagged jet angularities λ_α for $R = 0.4$ charged-particle jets in pp collisions, along with individual sources. The total uncertainty is taken as the quadrature sum of all contributions. [303] 194

List of Tables

2.1 Layer type, dimensions, precision, active area, and channels for the ALICE ITS during LHC Run 2 [256]. 49

3.1 Estimated $R = 0.2$ jet matching inefficiency, requiring a unique match between detector and truth levels, as calculated in PYTHIA 8 + GEANT3 for the pp and Pb–Pb data samples. In Pb–Pb, this percentage is calculated by the number of matches which satisfy all selection cuts divided by those which have at least one detector-level match (to eliminate combinatorial jets); note that no Pb–Pb data is used below $p_{\mathrm{T}}^{\mathrm{ch\ jet}} = 40$ GeV/c. 62
3.2 Estimated tracking efficiency uncertainty as calculated in PYTHIA 8 + GEANT3 for the relevant Pb–Pb data samples per $p_{\mathrm{T,track}}$ bin [297]. 83

B.1 List of the topological cuts used to reconstruct D^0 mesons across various $p_{\mathrm{T}}^{\mathrm{D}^0}$ intervals in pp collisions at $\sqrt{s} = 5.02$ TeV [303]. 189
B.2 Parameters of the POWHEG + PYTHIA 8 simulations which are varied to estimate uncertainty on the D^0-tagged jet contribution from feed-down decays of B mesons [303]. 193

Acknowledgments

I would like to acknowledge several people who have made sizable contributions to my academic career. First thanks goes to my mother and father, for supporting and encouraging me in these endeavors for the better part of three decades. Second goes to Christine Aidala, who gave me a first opportunity in experimental research as an undergraduate, guiding my path considerably; it was through her that I discovered this passion. Numerous hours were also spent learning with Joe Osborn, an excellent teacher and true friend to me. I similarly thank Bing Zhou, Tom Schwarz, Jean Krisch, and Tiesheng Dai for offering me the awesome chance to perform hardware work for ATLAS and join the detector crew at CERN. This incredible opportunity influenced my life and goals immensely.

I owe a considerable amount of gratitude to my Ph.D. advisor Barbara Jacak, who saw enough potential in me to offer an opportunity to join her group, and to support my education and research for half a decade. It would be impossible to ask for a better supervisor, and our frequent check-ins provided essential guidance, both academically and personally. I similarly thank James Mulligan for spending copious time working with me on my first jet analyses, for being a model researcher and collaborator, and for answering my many questions when no one else could. Thanks also to the other staff at LBNL, especially Mateusz Płoskoń, Peter Jacobs, and Nikki Apadula, who have always prioritized strong science from our group, and Preeti, Minjung, Wenqing, and Emma for reviewing an early draft of this document.

It cannot be overemphasized how grateful I am to the ALICE Collaboration for several opportunities over these years which have shaped me considerably. Marta Verweij and Filip Krizek served as phenomenal working group conveners and offered me the chance to speak at my first international conferences. A huge debt of thanks is owed to Andrea Dainese for his kindness and encouragement and for seeing my potential. Many thanks also go to Ivan Ravasenga, Hartmut Hillemanns, Federico Ronchetti, Sarah Porteboeuf, Edith Buthelezi, and Sabyasachi "Gouda" Siddhanta for trusting and guiding me during my first shift blocks for the collaboration, including being Shift Leader. Driving this experiment while collecting real physics data provided an ethereal excitement for me as a junior collaborator.

Deep thanks goes to the staff, faculty, and especially the members of my committee at UC Berkeley who have pushed me forwards. Marjorie Shapiro has always offered helpful guidance as an academic advisor, and gave encouragement when it was needed most. Similar thanks goes to Karl van Bibber and Lawrence Hall for altruistically sharing their humbling level of expertise. I am also grateful for the guidance and support of Joelle Miles, a flawless expert of the physics department: this institution would surely crumble without her.

Finally I must express appreciation to all the friends and family who have endured these years by my side: Preeti Dhankher, Chi Chung (Alvin) Li, Ben and Molly Cunningham, Kyle Lee, Clark Faville, Mike Somavilla, Cooper Wagner, Ethan Cannaert, Kenny Vetter, Jay Polt, Jeroen Tenkink, Brent Marley, and others. Thanks to Mike Joseph, Ethan Hawker, Andrew Kartsounes, and Mike Letovsky for stimulating me in high school, and similarly Roky, Tommy, Stacy, John Ike, Ronnie, Benny, and Danny 1 & 2 for continual inspiration. To everyone else who has seen me through to become what I am, thanks, love and mercy.

Chapter 1

Introduction

Mankind has looked to the heavens pondering its existence for millennia. From the ancient philosophers through modern civilization today, an aggregation of scientific discoveries has revolutionized human life and understanding. In recent times, mankind has harnessed the power of electricity, elucidated the secrets of the atom, mastered the ability of flight, and watched as living beings visited another celestial body, the moon, and safely returned home. While mere dreams and musings of generations past, these modern accomplishments and more are attributed in no small part to mankind's continuing research of the physical sciences.

Yet many fundamental mysteries still remain. The origins of creation, and the enduring question *why*, still plague the human condition. Theories are born, survive awhile, then crumble, and give birth to a new generation of ideas. The determinism of Newtonian mechanics, for instance, now comes with an asterisk, given by the apparent randomness of quantum theory. With such developments comes a new array of deep questions, such as the interconnectedness of all physical bodies, an entanglement beyond human senses.

It is impossible to know where modern research will eventually lead: through these studies of physics man strives to understand God, and with such understanding the possibilities are truly limitless. One can only hope that with such knowledge life will improve for all mankind, and the appreciation of God's many mysteries will be made ever stronger.

This thesis details some recent advances that have been made in the field of high-energy particle and nuclear physics. The following introductory section gives an overview of the modern theory which governs our understanding of the universe at the tiniest scales currently known. In particular, special attention is paid to the strong nuclear interaction, and the mechanisms used to study it at high energies. This chapter is concluded with a statement of purpose for the research of interest (Sect. 1.5), which is the main subject of this writing.

The following chapters give details of the particular physics analyses which were completed during the tenure of this appointment. Chapter 2 details the experimental apparatus, while Ch. 3 enumerates the relevant data sets, procedures, and systematic checks used for the analyses. Ch. 4 explains the corrections that are applied to pQCD predictions before they are compared to data. Results are shown in Ch. 5, with discussion and closing remarks delivered in Ch. 6.

1.1 Quantum chromodynamics

The physical theory which governs interactions of the strong nuclear force is called quantum chromodynamics, or QCD. A detailed review is given in Refs. [1–6]. QCD is a quantum field theory with dynamics dictated by its gauge-invariant Lagrangian,

$$\mathcal{L}_{\rm QCD} = \sum_q \left\{ \bar{\psi}_{q,i} \left[i\gamma^\mu (D_\mu)_{ij} - m_q \delta_{ij} \right] \psi_{q,j} \right\} - \frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a, \tag{1.1}$$

where δ_{ij} is the Kronecker delta function and γ^μ are the Dirac gamma matrices.

The QCD Lagrangian bears a striking resemblance to that of quantum electrodynamics (QED), the quantum field theory which governs electromagnetic interactions, and the resulting physics is comparable in several ways. The spin-1/2 (fermion) fields $\psi_{q,i}$ are Dirac spinors which represent fundamental particles called quarks, each with a type (*flavor*) q and charge (*color*) i. The Dirac adjoints (conjugates) $\bar{\psi}_{q,i}$ represent the corresponding anti-particles, called anti-quarks. The flavor q correlates to the lepton flavor of QED, while the color charge i compares to QED's electrical charge. Charge in QCD is further complicated, however, by the local SU(3) color gauge symmetry of the theory, allowing the index i to run from 1 to 3.

The gauge covariant derivative D_μ is composed of an ordinary partial derivative ∂_μ term along with an additional term which preserves local gauge invariance,

$$(D_\mu)_{ij} = \partial_\mu \delta_{ij} - i g_s (T_a)_{ij} \mathcal{A}^a_\mu. \tag{1.2}$$

The spin-1 fields $\mathcal{A}^a_\mu$ represent the gauge bosons that mediate interactions in the theory, which are called gluons. The gluon fields are connected to the Dirac quark fields via T_a, infinitesimal generators of rotation in SU(3). A popular representation of these generators is $T_a = \lambda_a/2$, where λ_a correspond to the Gell-Mann matrices, with $a \in [1, 8]$. Unlike the single photon which mediates QED, these eight generators signify eight different gluons which can connect the incoming and outgoing quark colors i and j. This term also includes the constant g_s, which sets the scale of the fundamental coupling strength between quarks and gluons. The s subscript is used to explicitly denote that this is the coupling constant for the strong force (QED has a similar coupling constant g, sometimes written as g_e or just e).

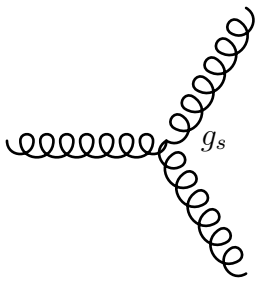
The second term inside the sum of Eq. 1.1, which bears the fundamental constants m_q, directly connects $\bar{\psi}_{q,i}$ to $\psi_{q,j}$, and therefore can be interpreted as terms which give the quarks their mass. There are no such terms for the gauge fields $\mathcal{A}^a_\mu$, rendering them massless.

The final term in Eq. 1.1 involves the (gauge invariant) gluon field strength tensor,

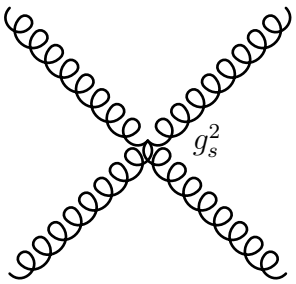
$$G^a_{\mu\nu} = \partial_\mu \mathcal{A}^a_\nu - \partial_\nu \mathcal{A}^a_\mu + g_s f^a{}_{bc} \mathcal{A}^b_\mu \mathcal{A}^c_\nu, \tag{1.3}$$

where f^{abc} are c-number commutators of the λ_a matrices, known as the *structure constants*,

$$[\lambda_a, \lambda_b] = 2i f^{abc} \lambda_c, \quad \text{with} \quad f^{abc} = \text{number} = f_{abc} = f^a{}_{bc}. \tag{1.4}$$



(a) Vertex corresponding to $(\partial\mathcal{A})(\mathcal{A}\mathcal{A})$.



(b) Vertex corresponding to $(\mathcal{A}\mathcal{A})(\mathcal{A}\mathcal{A})$.

Figure 1.1: Feynman diagrams for the gluon self-interaction terms in QCD.

Equation 1.3 is the analog of the electromagnetic field strength tensor $F_{\mu\nu}$ in QED. The third term in Eq. 1.3, however, directly connects the gauge fields, which does not appear in QED. This term is a mathematical result of the required SU(3) color symmetry, as SU(3) is a non-Abelian group with noncommuting generators T_a, whereas the U(1) symmetry of QED has c-number generators, which commute. This results in several consequential features of the theory. The first are cross-terms in Eq. 1.1 of the form $(\partial\mathcal{A})(\mathcal{A}\mathcal{A})$ and $(\mathcal{A}\mathcal{A})(\mathcal{A}\mathcal{A})$, which correspond to 3- and 4-gluon interaction vertices, respectively, as depicted in Fig. 1.1. In other words, the gluons themselves carry color charge, unlike the chargeless photon of QED.

Another interesting feature which arises from the non-Abelian gauge symmetry is the requirement that the strong coupling is universal, meaning that all quarks must couple to gluons with the same strength. This is untrue for QED, where charged leptons carry a larger electric charge than the fractional one carried by quarks.

1.1.1 The running of α_s

A third consequence of QCD's non-Abelian nature, asymptotic freedom, was a surprise which resulted in Frank Wilczek, David Gross, and David Politzer being awarded the 2004 Nobel Prize in Physics for their 1973 discovery [7, 8]. The *bare* coupling strength g_s is not directly observed in nature, with its strength varying according to the energy scale at which an interaction occurs. This *running* of the coupling is caused by the screening from virtual particles created via outgoing flux near a charged source. The variation of a coupling constant g with respect to energy scale μ is given by the renormalization group equation

$$\beta(g) = g\frac{\partial g}{\partial \mu} = \frac{\partial g}{\partial \ln \mu}, \tag{1.5}$$

which is called the beta function. At the order of one-loop level calculations in QCD,

$$\beta(g_s) = -\left(11 - \frac{n_q}{6} - \frac{2n_i}{3}\right)\frac{g_s^3}{16\pi^2}, \tag{1.6}$$

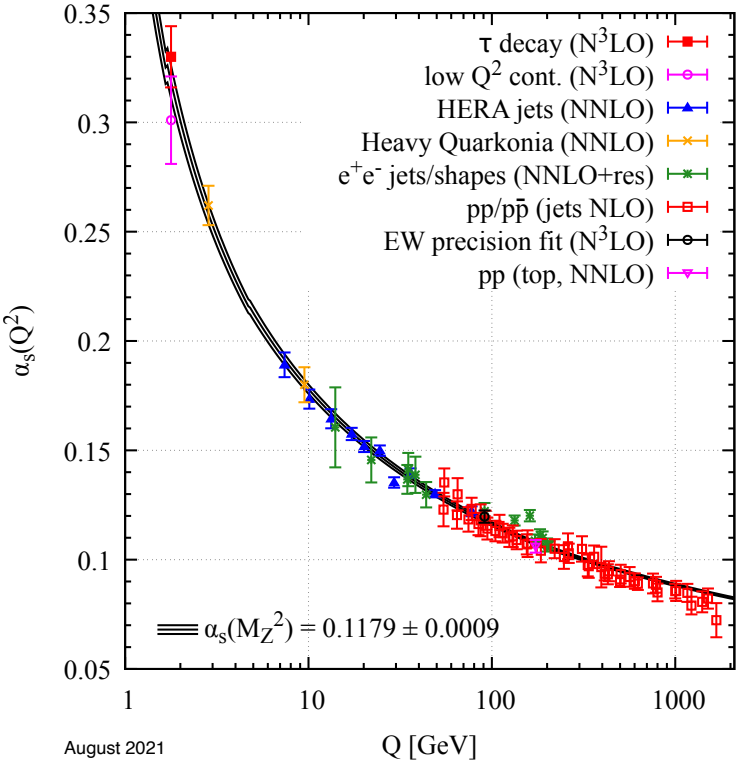


Figure 1.2: Experimental measurements of α_s at varying energy scale Q. The perturbative accuracy of the extraction is given in parentheses for each observable. For details, see Ref. [3].

where n_q is the number of quark flavors and n_i is the number of scalar color bosons. The global negative sign is unique for non-Abelian theories, and implies that the coupling goes to infinity as the scale μ decreases, with the divergent scale labelled $\Lambda_{\text{QCD}} \approx 200$ MeV. It also means that the theory becomes asymptotically free as μ increases, or equivalently that the strong force becomes weaker at large energies, which is the opposite behavior of QED. Furthermore, the weakening of the coupling is logarithmic,

$$\alpha_s(\mu^2) \equiv \frac{g_s^2(\mu^2)}{4\pi} \approx \frac{1}{\beta_0 \ln(\mu^2/\Lambda^2)}, \tag{1.7}$$

where β_0 is a constant. It is therefore necessary to specify the energy scale μ when describing experimentally-derived values for the coupling constant α_s (or g_s).

Figure 1.2 shows measurements of the running of α_s using data from various observables at several different experiments. The value of α_s is extracted by comparing these experimental results to theoretical predictions from perturbative QCD (pQCD), with accuracy ranging from Next-to-Leading Order (NLO) to N^{3}LO, depending on the observable. These results show excellent agreement with the predicted theoretical scaling and give a globally extracted value of $\alpha_s(m_{\text{Z}}^2) = 0.1179(9)$, where $m_{\text{Z}} \approx 2.5$ GeV/c^2 is the mass of the Z electroweak vector

boson. It is worth noting that α_s is the only free parameter in QCD apart from the quark masses, and thus by specifying this parameter at some scale, it can be renormalized to all scales, and the theory is complete.

Compared to other fundamental constants, such as the electron charge or mass, the level of precision to which α_s has been measured is rudimentary, at only four significant digits. This relative imprecision is due to the complexity of QCD, as experimental measurements must be compared to complicated theoretical predictions to extract α_s. Though significantly more involved to compute than its electromagnetic cousin, QCD can be used to calculate differential cross sections using perturbation theory; however, due to the fact that α_s is roughly an order of magnitude larger than the low-energy value for the electromagnetic coupling constant $\alpha_{\text{EM}} \approx 1/137$, perturbation theory proves to be less precise for QCD than for QED at identical orders, with LO pQCD often having significant corrections at the level of a factor of 2. These calculations are therefore notably improved by including higher-order corrections – which, in turn, improve the extraction of α_s.

1.1.2 Nonperturbative consequences

The divergence of α_s at low energy scales provides further complications. While high-energy interactions in QCD can be described by quarks and gluons, the degrees of freedom under a lower-energy microscope are quite different. As the energy of a probe changes, so does its Compton wavelength, with lower energies corresponding to larger distance scales following the usual inverse (de Broglie) relation $\lambda = h/p$. At the scale of Λ_{QCD}, then, the length scales are large enough that the interaction strength between quarks and gluons, dictated by a diverging α_s, must be enormous. However, plugging in numbers to the Compton formula, one expects this divergence to happen at $hc/\Lambda_{\text{QCD}} \approx 6 \times 10^{-15}$ m $= 6$ fm, much smaller than the length scales of everyday life.

The physical manifestation of this observation is *color confinement* [9], which enforces color-charged particles to be confined inside of color-neutral objects. From a phenomenological perspective, the strong force is mediated between two colored particles via the exchange of gluons, which form "color flux tubes" between the quarks. If the color-charged particles are pulled apart, the force between them grows, and as the color flux tubes grow stronger, more energy is required to pull them further. At some scale, the energy inside the color flux tubes is so great that they break and form a new quark-antiquark pair. The divergence of α_s dictates this breaking scale, which then determines the size of color-neutral objects.

Quarks and gluons then are always confined inside of composite objects, called *hadrons*. The proton is one example, as is the neutron, with both inheriting their quantum properties from three *valence* quarks (i.e., one red, one blue, and one green, which together are colorless). These are examples of a wider class of *baryons*, which all have three such valence quarks. It is also possible to have hadrons containing only two valence quarks (e.g., one red and one anti-red), which are called *mesons*. Quarks and gluons are sometimes collectively referred to as partons, as they are parts (constituents) of their hadronic wholes (the name *partons* was coined by Richard Feynman when the quark model was still not well-verified).

Hadrons, however, are significantly more complex than a mere 2- or 3-quark structure. These valence quarks constantly exchange gluons, which can interact with themselves and one another, or split into quark-antiquark pairs which exist for a while before interacting with other partons or annihilating back into a gluon. These extra quarks and gluons form a *sea* of ever-evolving parton distributions inside the hadron. Sea quark production via gluon splitting can be generally understood via normal perturbative evolution, although there is a high density of soft (meaning low-energy), nonperturbative gluon radiation, and experimental probes have similarly revealed significant nonperturbative effects [10, 11]. There have also been recent advances in calculating the general parton distribution functions (PDFs) inside of hadrons by simulating QCD on a finite, discretized lattice [12]. Despite much larger error bands than current experimental extraction, advances in computing (including, for example, quantum processors) may improve the precision of these theoretical predictions in coming years. For an introduction to lattice QCD in this context, see Ref. [13–15].

In addition to lattice calculations, chiral perturbation theory (ChPT) can be used for calculations in the low-energy regime of QCD. ChPT is an effective field theory (EFT) which utilizes the approximate chiral symmetry of QCD, along with other symmetries such as charge conjugation and parity, to construct a model Lagrangian for low-energy QCD interactions. Such calculations have proven successful for modeling interactions of π mesons with nucleons, as well as some other hadronic interactions. For an introduction to phenomenological Lagrangians and ChPT, see Ref. [16, 17].

1.1.3 Hadronization

While regular nuclear matter is composed of hadrons, the way in which partons form these complex structures is unclear. High-energy partons radiate energy, primarily in the form of gluons, until their energy scale falls to the point where $\alpha_s \geq 1$, after which perturbation theory is no longer a useful expansion of their dynamics. Since the partition function of QCD has never been solved for hadrons, it is not possible at present to describe hadronization from first-principles QCD. Therefore, phenomenological models have been introduced to model this nonperturbative process.

One common example is the Lund string model [18, 19]. Lattice calculations show that the color flux tubes formed between a quark-antiquark pair contain an approximately proportional amount of energy to the distance between them. This linear relationship is modelled by the introduction of relativistic strings, confined to the longitudinal axis between the two partons. Hadron spectroscopy measurements suggest a string constant of $\kappa \approx 1$ GeV/fm, which corresponds to an effective mass density of the one-dimensional string. As the partons stretch apart, energy in the string increases, until it has sufficient energy to break, with two new quarks produced at the broken string edges ($\mathrm{q\bar{q}} \rightarrow \mathrm{q\bar{q}'} + \mathrm{q'\bar{q}}$, where the quark is labelled as q and the antiquark as $\bar{\mathrm{q}}$). This string-breaking procedure continues until only on-mass-shell hadrons remain. (It should be noted that these phenomenological strings, which replace color flux tubes, are fundamentally different than the objects of string theory, which instead replace fundamental particles.)

Another approach is cluster hadronization [20–22]. In this model, all final-state gluons (meaning those which survive after successive branching/emissions off of hard-scattered partons) are split into quark-antiquark pairs, which are then connected by color lines to form color-singlet clusters. These clusters are taken to have small mass and length, with those above a given mass threshold allowed to fission into progressively smaller clusters. All clusters then fragment into hadrons, either by splitting into two hadrons, or by exchanging energy with other local clusters on-mass-shell for the hadron of their corresponding flavor.

A third candidate mechanism is statistical hadronization [23–25], which assumes that the clusters produced by pQCD have a finite spatial extent, and that all multi-hadronic states arising from those clusters (compatible with their quantum numbers) are equally likely. Along with the coalescence model [26], statistical hadronization has been used to explain signatures of the quark-gluon plasma [27], the production of which is discussed in Sect. 1.4.2.

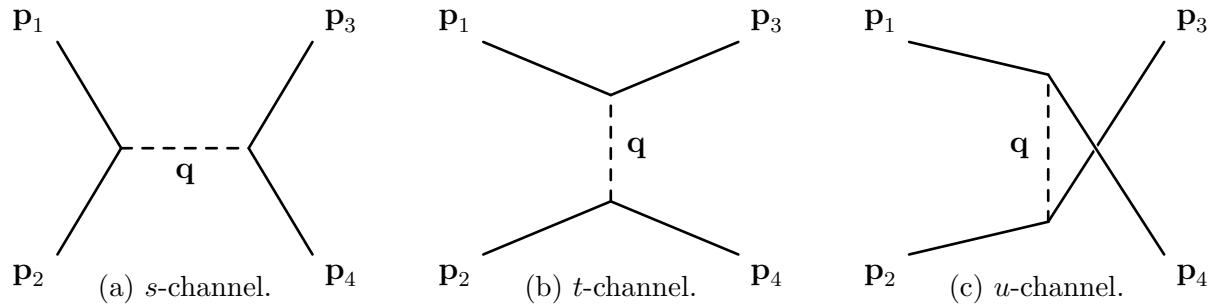
Popular implementations of phenomenological models such as these typically have several different parameters which must be tuned to experimental data in order to provide accurate predictions of experiment and a consistent picture across different observables and energy regimes. The number of parameters inside such models can also vary with time, as inconsistencies are discovered and resolved, or as improvements are made to the underlying physical assumptions. Providing measurements which directly test these models, or which can be used to tune them to higher precision, is an active area of experimental research.

It is also possible to measure so-called fragmentation functions [28, 29], which describe the probability of a particular partonic species fragmenting (hadronizing) into a specific hadron. These will be discussed in Sect. 1.2 below.

1.2 Scattering experiments

In the past 50 years since the discovery of QCD, one of the most productive ways to experimentally study the theory has been via high-energy *scatterings*, or collisions, of one particle off another. These experiments are typically designed in one of two ways. In the first design concept, called a *fixed-target* experiment, particles of some known species are accelerated to high energies and directed toward a *target* of stationary material. Particles from the beam scatter off of those in the target, and the resulting particles can then be measured and compared to the predicted interactions from QCD. These experiments are asymmetric, as the center-of-mass frame for the colliding particles is boosted with respect to the lab frame.

The second experimental design can be tuned to be more symmetric in the lab frame. Two beams of particles – not necessarily of the same species – are accelerated in opposing directions, until they are crossed at an *interaction point* (IP). Particles from the two beams can interact with each other at the IP, producing events which are then observed experimentally. These collider experiments are capable of reaching higher interaction energies than the fixed-target approach, as the secondary beam contributes significantly to the center-of-mass energy. One experimental approach to measuring these events, the ALICE experiment at



(a) s-channel. (b) t-channel. (c) u-channel.

Figure 1.3: Leading-order (two-vertex) Feynman diagrams for $2 \to 2$ scattering in QFT, labelled by particle four-momenta $\mathbf{p}_i$. These diagrams are read from left to right, with $i = 1, 2$ corresponding to incoming particles, and $i = 3, 4$ to outgoing particles.

the CERN LHC, will be the primary subject of Ch. 2, and this latter experimental design will be the focus in this thesis.

Generally a standard set of parameters are defined to describe the kinematics of a collider experiment. The *Mandelstam variables* are convenient parameters to describe the kinematics of $2 \to 2$ scattering, whose leading-order Feynman diagrams are given in Fig. 1.3. In terms of the incoming and outgoing particle four-momenta $\mathbf{p}_i$, the three Mandelstam variables are typically defined

$$s \equiv (\mathbf{p}_1 + \mathbf{p}_2)^2 = (\mathbf{p}_3 + \mathbf{p}_4)^2, \tag{1.8}$$

$$t \equiv (\mathbf{p}_1 - \mathbf{p}_3)^2 = (\mathbf{p}_4 - \mathbf{p}_2)^2, \quad \text{and} \tag{1.9}$$

$$u \equiv (\mathbf{p}_1 - \mathbf{p}_4)^2 = (\mathbf{p}_3 - \mathbf{p}_2)^2, \tag{1.10}$$

using natural units where the speed of light $c = 1$. These parameters correspond to the squared four-momenta of the interactions $\mathbf{q}^2$ as given in Fig. 1.3. It is worth noting that the center-of-mass energy of the interaction can be identified as $\sqrt{s}$, which is a frequently-used shorthand at collider experiments.

One of the simplest interactions which probes QCD is that of an electron (e^-) scattering off a proton (p), mediated by a virtual photon (γ). At energies low enough that the Compton wavelength of the γ is larger than or comparable in magnitude to the size of the proton, the scattering is elastic, and the proton remains intact. At low energies, these interactions can be modelled by Rutherford scattering (non-relativistic) or Mott scattering (relativistic, but proton recoil neglected), which at first-order in perturbation theory yields

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\mathrm{Mott}} = \frac{\alpha_{\mathrm{EM}}^2}{4E^2 \sin^4(\theta/2)} \cos^2\left(\frac{\theta}{2}\right) \left|F(\mathbf{q}^2)\right|^2, \tag{1.11}$$

where α_{EM} is the electromagnetic fine-structure constant and $F(\mathbf{q}^2)$ is the proton *form factor*,

$$F(\mathbf{q}^2) = \int \rho(\vec{r}) \exp(i\vec{q} \cdot \vec{r}) \mathrm{d}^3\vec{r}, \tag{1.12}$$

where $\vec{q}$ is the three-momentum component of $\mathbf{q}$. The fraction $\mathrm{d}\sigma/\mathrm{d}\Omega$ is known as a *differential cross section*, which experimentally describes the fractional probability of a particular interaction per solid angle (Ω, with $\mathrm{d}\Omega = \sin\theta\ \mathrm{d}\theta\ \mathrm{d}\phi$). While Eq. 1.11 results from a purely QED interaction, the form factor $F(\mathbf{q}^2)$ is the Fourier transform of the proton's charge distribution, signifying that measurement of elastic scattering probabilities directly probes the charge distribution – and thus the internal quark structure – of the proton. For similar reasons as explained in Sect. 1.1 above, the functional form of $F(\mathbf{q}^2)$ has not been derived from first principles, and therefore measurement provides an invaluable look into proton structure.

1.2.1 Structure and scaling

It is useful to describe the kinematics of $\mathrm{e}^-\mathrm{p}$ scattering using Lorentz-invariant quantities (for frame-independent quantification) which can be expressed in terms of the kinematics of the e^-, which is typically measured experimentally, unlike the outgoing proton. Taking the lines $i = 1, 3$ to correspond to the incoming and outgoing e^- and $i = 2, 4$ to that of the proton, the variable

$$Q^2 \equiv -q^2 = -\mathbf{q}\cdot\mathbf{q} \approx 4E_1E_3\sin^2\frac{\theta}{2} > 0 \tag{1.13}$$

is often used to describe the scale of four-momentum exchanged in the interaction; indeed, this variable Q is also what appears as the scaling term in Fig. 1.2.

At higher interaction energies, where the virtual photon has a wavelength smaller than the size of the proton, the photon no longer scatters elastically off of the proton as a whole, but rather off the partons inside it. In such cases, the dimensionless parameter x (sometimes called "Bjorken x" after theoretician James D. Bjorken),

$$x \equiv \frac{Q^2}{2\mathbf{p}_2\cdot\mathbf{q}} = \frac{Q^2}{2m_\mathrm{p}\nu} \quad \text{with} \quad \nu \equiv \frac{\mathbf{p}_2\cdot\mathbf{q}}{m_\mathrm{p}}, \tag{1.14}$$

is used to describe the elasticity of the scattering processes, with $x \in [0, 1]$. The extreme case $x = 1$ describes a fully elastic collision with the composite proton, whereas the opposite case $x \to 0$ occurs as the γ probes smaller and smaller length scales inside the proton, in which case the proton typically breaks apart. This *deep inelastic scattering* (DIS) in high-energy $\mathrm{e}^-\mathrm{p}$ collisions is used to probe the internal structure of the proton, which lives below the nonperturbative limit of QCD.

The fractional energy lost by the e^- in such collisions, as evaluated in the rest frame of the proton, is given by the *inelasticity* y,

$$y \equiv \frac{\mathbf{p}_2\cdot\mathbf{q}}{\mathbf{p}_2\cdot\mathbf{p}_1} = \left(\frac{2m_\mathrm{p}}{s - m_\mathrm{p}^2}\right)\nu, \tag{1.15}$$

where $y \in [0, 1]$. Combining with the definition of x from Eq. 1.14, the variables x and y are related to one another by

$$Q^2 = (s - m_\mathrm{p}^2)xy, \tag{1.16}$$

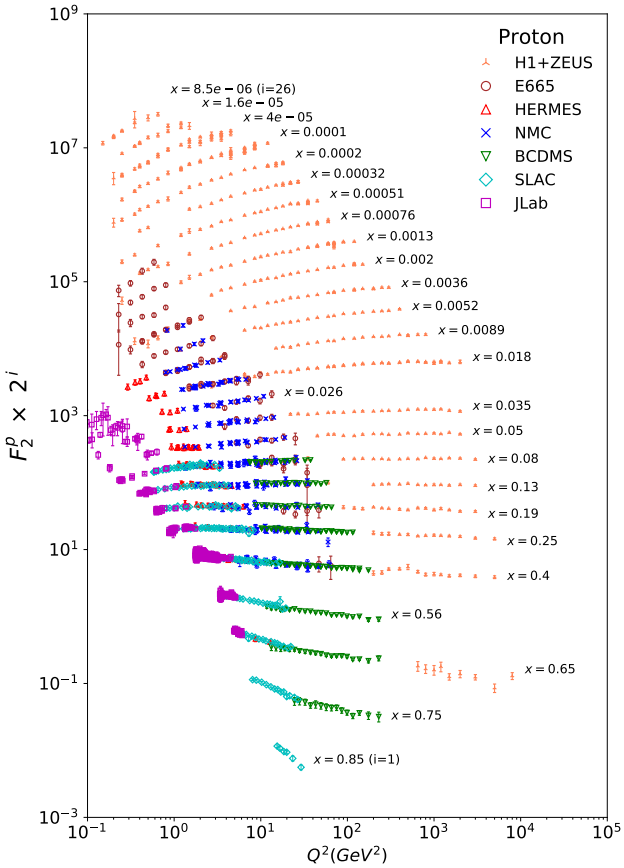


Figure 1.4: Structure function $F_2(x, Q^2)$ for the proton, as measured by different experiments for varying x and Q^2 [3]. The structure function varies only weakly with respect to Q^2, despite the horizontal axis spanning five orders of magnitude.

such that the kinematics of DIS at fixed $\sqrt{s}$ can be uniquely described by any two of x, Q^2, and y (or ν).

Experimental data from DIS has been instrumental for illuminating the partonic structure of the proton. It can be shown that the differential cross section for DIS (where $Q^2 \gg x^2 m_\mathrm{p}^2 y^2$) is [6]

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}x\,\mathrm{d}Q^2} \approx \frac{4\pi\alpha^2}{Q^4}\left[(1-y)\frac{F_2(x,Q^2)}{x} + y^2 F_1(x,Q^2)\right], \tag{1.17}$$

where $F_1(x, Q^2)$ and $F_2(x, Q^2)$ are called the *structure functions* of the proton. Systematic studies of these structure functions at the Stanford Linear Accelerator Center (SLAC) in Stanford, California revealed that they are nearly Q^2 independent, allowing their approximation as

$$F_1(x, Q^2) \to F_1(x) \quad \text{and} \quad F_2(x, Q^2) \to F_2(x). \tag{1.18}$$

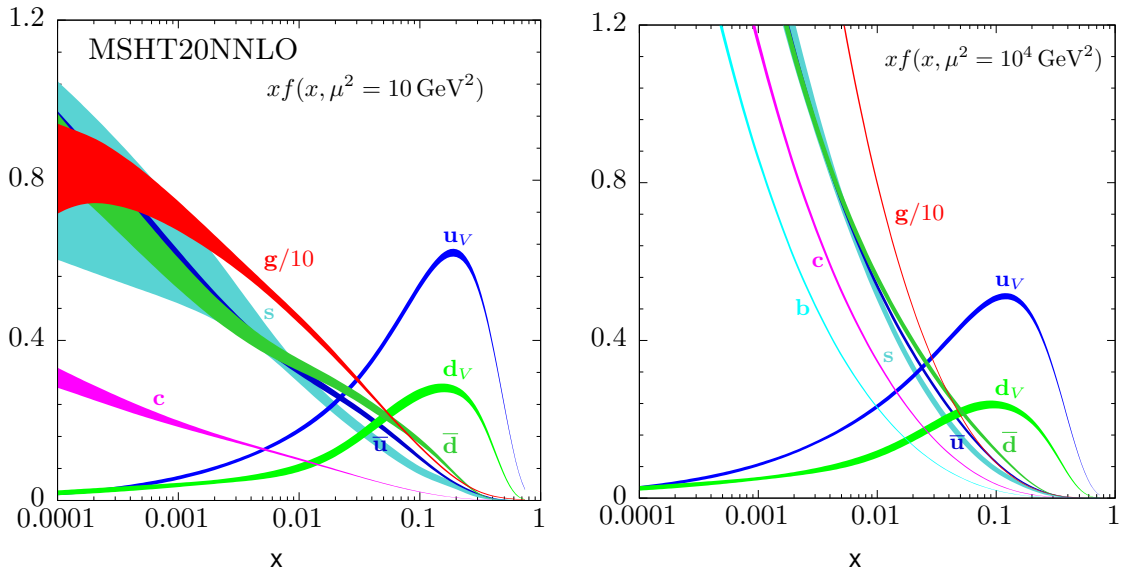


Figure 1.5: Unpolarized parton distribution functions $xf(x)$ at the energy scales of $\mu^2 = 10$ GeV^2 (left) and 10^4 GeV^2 (right), as evaluated at NNLO [3, 30]. The horizontal axis is the momentum fraction (Bjorken) x. The valence up and down quarks ($f = \mathrm{u_v}$ and $\mathrm{d_v}$, respectively) dominate at large values of x, while the probability of scattering off gluons (g) and sea quarks (s, c, b, antiquarks) dominates in the low-x regime.

This phenomenon is known as *Bjorken scaling*; see Fig. 1.4. Another observation for DIS (where Q^2 is greater than a few GeV^2) was that the structure functions are related,

$$F_2(x) = 2xF_1(x), \tag{1.19}$$

which is called the *Callan-Gross relation*. This observation signifies that the electric and magnetic contributions to Eq. 1.17, which underlie $F_1(x)$ and $F_2(x)$, are related by the magnetic moment of a single Dirac particle, implying that DIS is actually the elastic scattering of a virtual boson off a point-like parton inside the proton.

If one considers the scattered parton to carry momentum $\mathbf{p}_\mathrm{q} = \xi\mathbf{p}_2$, where ξ is the fractional momentum of the proton carried by the struck quark, it can be shown [6] that, when the proton energy is sufficiently large ($E_2 \gg m_\mathrm{p}$),

$$\xi = Q^2/(2\mathbf{p}_2 \cdot \mathbf{q}) \equiv x. \tag{1.20}$$

The x variable is therefore typically also called the parton momentum fraction.

Since quarks and gluons are constantly interacting with each other inside the proton, and since these dynamics live beneath the perturbative limit of QCD, x is not known *a priori*, and its probability distribution must be measured. As the quark and gluon content of hadrons are also dynamic, these measurements must be carried out for each flavor, and

are generally known as *Parton Distribution Functions* (PDFs) $f(x)$. In the parton model, these PDFs are related to the structure functions by

$$F_1(x, Q^2) = \frac{1}{2} \sum_i Q_i^2 f_i(x), \tag{1.21}$$

where Q_i is the charge carried by quark flavor i. Figure 1.5 shows the most recent constraints on the partonic PDFs for the unpolarized proton, where unpolarized refers to the proton spin. By assuming isospin symmetry of the up- and down-type quarks, the similar valence structure of protons and neutrons, it can be extracted from experimental data that the up- and down-type valence quarks only carry about half of the proton's momentum, with the rest carried by gluons (which are electrically neutral, and therefore do not contribute to DIS) and sea quarks.

Comparing the two panels of Figure 1.5 also reveals that, despite the approximate scaling behavior given by Eq. 1.18, the energy scale independence of f_i is violated at very low and very high values of x. At larger values of Q^2, corresponding to shorter wavelengths of the virtual photon, DIS resolves smaller length scales inside the proton and is therefore sensitive to small-length gluon emissions ($q \to qg$). This energy scale dependence is understood from QCD, and can be directly calculated by using the DGLAP (Dokshitzer-Gribov-Lipatov-Altarelli-Parisi) evolution equations, which are derived using fixed-order calculations in pQCD. For more information, see Ref. [31–33].

1.2.2 Universality and factorization

The application of data-extracted functions such as f_i are not limited to e^-p collisions. A fundamental assumption of particle physics is that the proton appearing in e^-p collisions should be the same as the one collided in proton-proton (pp) or even proton-nucleus (pA) collisions, if probed at identical x and with identical Q^2. This assumption is called *universality*, i.e. the PDFs measured in e^-p should be *universal* in all proton collisions.

Nonetheless, the total scattering amplitudes for different collision systems will still be distinct. In pp collisions, the most probable interaction between two incident partons of the opposing protons is mediated via the strong force, due to the greatly enhanced strength of the QCD interaction over the QED one ($g_s \gg e$). This means that QFT calculations now involve sums over colored partons mediated by 8 different gluons, rather than a single virtual photon. The probability of encountering a particular parton is described by the PDF of the nucleon, but the scattering itself is described by pQCD.

The total cross section for interactions produced in collider experiments must therefore depend on both PDFs and cross sections as calculated in perturbation theory. It turns out that the total cross section for DIS has the form [34]

$$\sigma_{\mathrm{DIS}}(x, Q^2) = \sum_i \int_x^1 f_i(\xi)\ \sigma_{\mathrm{B}}(x/\xi, Q^2)\ \mathrm{d}\xi, \tag{1.22}$$

where $\sigma_{\mathrm{B}}(\xi, Q^2)$ is the standard Born-level (i.e., leading-order) cross section. Equation 1.22 has the nontrivial feature that PDFs and perturbative cross sections can be *factorized*, where f_i needs no knowledge of the particular $2 \to 2$ scattering encapsulated by σ_{B}, and vice-versa (σ_{B} is independent of the PDFs). This factorization has been theoretically proven [34] at leading power of Q for DIS, electron-positron annihilation ($\mathrm{e^+e^-} \to$ hadron + X), the Drell-Yan process ($\mathrm{q\bar{q}} \to \gamma^* \to$ lepton-antilepton pair), as well as jet and heavy quark production in hadronic collisions, though its extension to all-orders in QFT has not been rigorously proven. Experimental searches for cases where factorization might be broken are ongoing.

1.2.3 Fragmentation functions

In DIS experiments, the outgoing $\mathrm{e^-}$ is typically the simplest particle to measure experimentally; leptons are easier to identify, and since they are unaffected by hadronization, they directly probe the scattering kinematics. In some cases, however, a parton (or partons) from the incoming proton are kicked with sufficiently high Q^2 that they hadronize and traverse detectable regions outside the beam pipe. The measurement of a leading hadron in what is known as *semi-inclusive* DIS (or SIDIS) adds additional information about the kinematics of the scattering. This can be used in the extraction of transverse momentum dependent PDFs (TMDs) [35, 36], or to gain other details such as for the identity of the struck quark.

The mechanisms for producing specific hadrons, however, are not simple to describe from QCD, as hadronization is nonperturbative; see Sect. 1.1.3. The production of hadrons from initiating partons is described statistically using probability distributions known as *fragmentation functions*. Fragmentation functions are typically written as $D_i^h(x, \mu^2)$, where i is the initiating (anti)quark or gluon, h is the produced hadron, and μ is the energy scale (interchangeable with Q, as above). Much like PDFs, fragmentation functions must be extracted experimentally, which is typically done using cross section measurements from $\mathrm{e^+e^-}$ collisions, where the lack of (known) electron substructure eliminates PDFs from the calculations and leaves only the perturbative QED amplitude and the (factorized) fragmentation function $D_i^h(x, \mu^2)$.

After extracting these functions from $\mathrm{e^+e^-}$ data, they can be used universally in theoretical predictions for other measurements. Despite a resounding success in applying these fragmentation functions to more complicated collision systems such as pp, their general universality has never been proven from first principles, and recent experimental results in the production of hadrons containing heavy-flavor valence quarks have raised questions about whether this universality may have been broken; see Sect. 1.3.4 for details.

1.2.4 Monte Carlo event generators

With measurements of the nonperturbative PDFs and fragmentation functions completing the perturbative picture, predictions can in principle be made for arbitrary experimental observables. However, direct calculations may be prohibitively difficult, or frameworks may have not been developed for their computation, such as generic observables which are IRC-

unsafe (see Sect. 1.3.1). To obtain predictions for such observables, a statistics-based approach is used, where events are generated probabilistically according to the PDFs of the scattered hadrons, the perturbative cross section, and phenomenological hadronization models (see Sect. 1.1.3) which are tuned to reproduce measured fragmentation functions. The observable of interest is then "measured" from these simulated events, which over a sufficient number of simulated events becomes a distribution which can be compared to experimental data. Using this repeated random sampling to obtain a numerical result is known as the Monte Carlo (MC) method.

Since a MC simulation has no knowledge of the observable it will be asked to predict, it is an extremely flexible method to generate predictions for arbitrary distributions. Nonetheless, this approach has a few drawbacks. It is not a fully-analytical procedure, so interpreting results in terms of QCD is challenging. Any discrepancies between MC predictions and experimental data often cannot be traced to a single part of the generator, meaning that it is difficult to place quantitative limits on a specific aspect of QCD using a MC generator. Similarly, MC generators have no clear level of perturbative accuracy. The cross section for a hard scattering is calculated to fixed order in pQCD, but this cross section is then followed by a phenomenological parton shower, which modifies the perturbative accuracy. In the PYTHIA event generator [37–39], for example, the parton shower is based on LO evolution equations in DGLAP and is estimated to approximate Leading Logarithmic (LL) accuracy, though this estimation is not rigorous. At the same time, the MC model can be used to generate arbitrary statistics, meaning that statistical error bands on the results do not necessarily represent the (much larger) uncertainty on the theory, given the limited perturbative accuracy and uncertainty on the extracted PDF and fragmentation behavior. MC generators therefore cannot replace rigorous, direct QCD predictions.

MC event generators can also be used to simulate events for other purposes. When coupled with an experimental detector model, as for example implementable in the GEometry ANd Tracking (GEANT) software package [40], MC events can be used to simulate material interactions that occur in experimental measurements and to correct for these inefficiencies. Simulations can also typically be modified to enhance rare processes, to study the behavior of intermediary states (e.g. the event before hadronization), or to include species of undiscovered particles in searches for new physics. The usefulness and flexibility of MC models makes them essential for many modern studies in scattering experiments.

The PYTHIA event generator, as mentioned above, is a general-purpose MC event generator for generating events in the high-energy scattering of leptons, hadrons, and heavy nuclei. Underlying the hadronic scattering model is LO pQCD [41], combined with phenomenological approaches to simulate the *Underlying Event* (UE), undesirable contamination that occurs in some scattering experiments due to *pileup*, multiple simultaneous pp collisions within the same bunch crossing, interactions of the beam with residual gas inside the beam pipe, or Mutli-Parton Interactions (MPI) occurring within the same event. After the shower, partons are hadronized using the string fragmentation approach. PYTHIA has many parameters which can be modified to adjust the model predictions; these are set by *tuning* the model to experimental data, and several tunes are freely available with the software.

Different tunes can produce substantially different predictions, and alternate tunes should therefore be considered separately from one another. Tune variations can also be used as an aid to assess systematic uncertainties on the choice of physics model.

Herwig [22, 42, 43] is another multi-purpose MC event generator for use in high-energy collider experiments. Some types of events, such as those including the Drell-Yan process, can be generated at higher perturbative accuracy (NLO) than in PYTHIA. Herwig events use an angularly-ordered parton shower based on QCD emissions from a dipole. In addition to nonperturbative models of UE effects, the generator uses cluster hadronization to convert final-state partons into hadrons, before simulating hadronic decays to the desired timescale.

There are several other MC generators, including POWHEG, an NLO generator which can be matched with the PYTHIA or Herwig shower and hadronization to produce predictions with better perturbative accuracy [44]. This improvement can produce significant changes for processes that are suppressed at leading order, such as heavy-quark production via gluon splitting (i.e., $g \to b\bar{b}$). For a theoretical review of MC event generator phenomenology and a list of some popular models, see Ref. [45].

1.3 Jets

The $2 \to 2$ scattering described above is a leading-order picture of the partonic scattering which occurs in QCD-dominated collisions such as pp. Partons have divergent branching probabilities in pQCD which are described by the DGLAP evolution equations (see Sect. 1.2.1). In the phenomenological picture, hard-scattered partons, which often carry a large component of momentum transverse to the colliding beams, then fragment into softer partons via consecutive gluon emissions, which is sometimes referred to as the *parton shower*. The shower eventually reaches a low enough virtuality scale that the partons are confined into hadrons, which then carry onward the initiating parton's transverse momentum. Therefore, the partons which carry a large virtuality from an initial hard scattering are eventually detected as collimated sprays of particles, known as *jets*.

Experimental evidence for jets was first seen at SPEAR [46]. Since then, jet production has been observed as one of the most prolific occurrences in proton and nuclear collider experiments. Over the last several decades, they have also proven to be one of the most useful; for a theoretical review, see Ref. [47]. Their formation is impacted by QCD at several scales, including the PDFs of the initial nucleons, the high-Q^2 scattering amplitude, the parton shower down to low virtuality, and the culminating nonperturbative fragmentation into hadrons. The highest-probability event topology is the simplest: dijets, where two hard-scattered partons produce two jets back-to-back in their center-of-mass frame, whose boost is dependent on the transverse momentum fraction carried by those partons.

One of the hard-scattered partons can also emit a hard gluon, as depicted in Fig. 1.6. If the gluon is sufficiently hard it will form its own jet, creating a three-jet event. Searches for

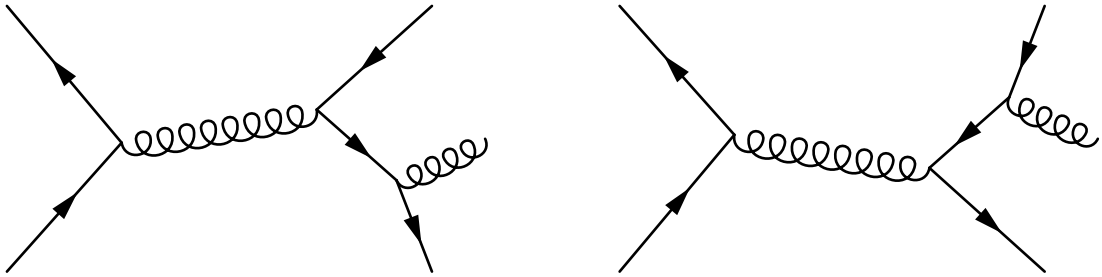


Figure 1.6: Example s-channel (quark-antiquark annihilation) Feynman diagrams corresponding to a three-jet event. The addition of an extra vertex for the outgoing gluon suppresses these matrix elements by a factor of g_s with respect to the leading-order $2 \to 2$ ones.

this event topology were carried out by evaluating event *sphericity*,

$$S = \frac{3}{2}\frac{\langle p_{\mathrm{T}}^2 \rangle}{\langle p^2 \rangle}, \tag{1.23}$$

where p_{T}^2 is particle transverse momentum squared, p^2 is particle total momentum squared, and the angle brackets designate averages over all particles in the event. S measures the isotropy of an event, where high-sphericity events correspond to highly isotropic topologies, which are more likely to have three-jet (or more) configurations. The initial discovery of three-jet events using this high-sphericity search was the first strong experimental evidence for the existence of the gluon, as purely dijet predictions could not account for such topologies at sufficiently high energy [48].

1.3.1 Jet reconstruction and IRC safety

While identifying simple jet topologies like those depicted in Fig. 1.7 might be easy to approximate by eye, the general case is more ambiguous, and the jet object as a mathematical construct can be defined in several different ways. The ideal approach would have the benefit of being both convenient experimentally as well as useful theoretically. From an experimental perspective, the specific partonic interactions that created the jets are not known, so jets must be reconstructed independently of QCD. Several different approaches exist, and if two different algorithms are used to identify the jets from some event, they could trace the same reconstructed tracks to different jets; the number of jets themselves could also change.

From a theoretical perspective, the jets which are reconstructed experimentally should correspond to objects which can be calculated perturbatively. One way to ensure calculability is to require that the jet reconstruction algorithms satisfy the conditions of Infrared and Collinear (IRC) safety [49]. IRC-safe jet reconstruction algorithms satisfy two conditions:

1. *Infrared (IR) safety:* The reconstructed jets in an event will not change if an infinitely soft (infrared) particle is added to the event.

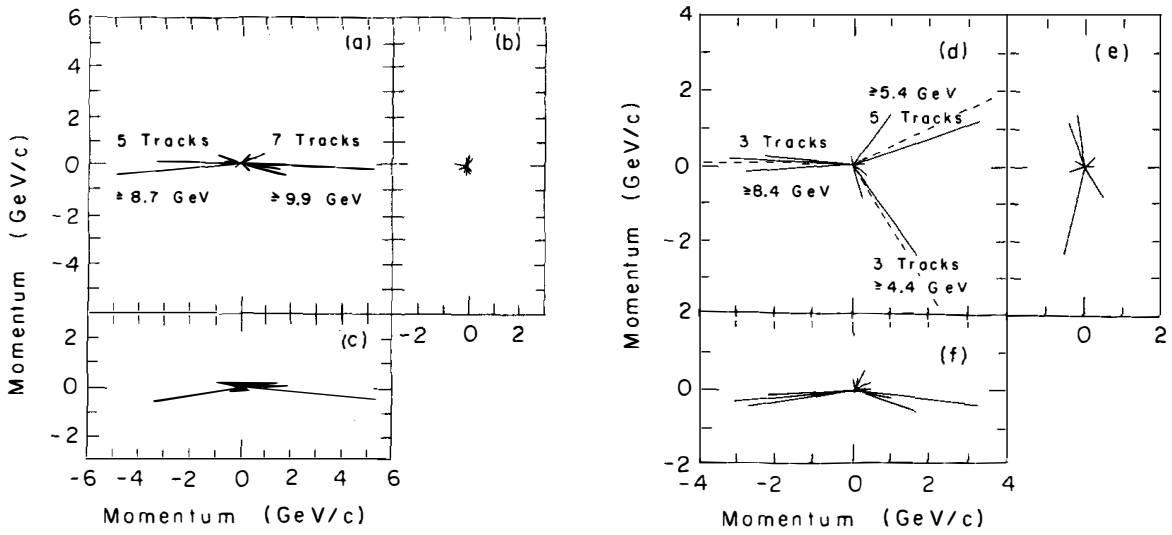


Figure 1.7: Experimental results from TASSO [48] depicting both dijet (left) and 3-jet (right) events, with projections onto three orthogonal planes. Reconstructed tracks are drawn as lines with length corresponding to their momentum.

2. *Collinear (C) safety:* The reconstructed jets in an event will not change if a track is split into two collinear tracks which together carry the same four-momentum of the original track.

These requirements remove divergences in jet calculations and ensure theoretical calculations can be done [50]. Specifically, in Soft Collinear Effective Theory (SCET) [51–55], these requirements cancel singularities that appear between real and virtual emissions, permitting Taylor expansions around $\alpha_s = 0$. Though IRC safety is sufficient to ensure a method of calculability, it is not a necessary requirement for jets to be theoretically calculable: for example, ratios of IRC-safe observables are not themselves IRC-safe, but are "Sudakov safe" [56], meaning calculations can be done by Taylor expanding around finite values of α_s [57].

Early experimental measurements often employed *cone algorithms*, which required constituent tracks to fall within a cone of some predefined radius. These algorithms produce jets which trace out rigid circles in pseudorapidity and azimuthal angle (in the $\eta - \phi$ plane). Examples include the Iterative Cone algorithms with Progressive Removal (IC-PR) or the Split Merge procedure (IC-SM), which are collinear and infrared unsafe, respectively. The Seedless Infrared Safe Cone (SISCone) algorithm [58] is an IRC-safe jet reconstruction algorithm with $\mathcal{O}\left[Nn\log(n)\right]$ time complexity, where N is the number of particles in an event and n is the number of particles in a cone. At the time, this efficiency was very attractive for experimentalists. Furthermore, cone algorithms offer resilience to effects from the UE.

The other primary class of jet reconstruction algorithms are known as *sequential recom-*

bination algorithms, or (more uncommonly) *hierarchical agglomerative clustering algorithms*. Rather than defining the jet based solely on particle distributions in position space, sequential recombination algorithms group tracks together using a similarity ("distance") metric combining information from momentum space. Early examples include the JADE algorithm [59, 60] and the k_{T} (or $k_{\perp}$) algorithm [61–63], the latter of which comes in several variants. The generalized k_{T} algorithm defines its particle-pair distance metric as

$$d_{ij} = \min(p_{\mathrm{T},i}^{b},\ p_{\mathrm{T},j}^{b}) \left(\frac{\Delta R_{ij}}{R}\right)^{2}, \tag{1.24}$$

where the indices i and j label the particle pair, p_{T} is the transverse momentum carried by the particle, R is a user-defined parameter (typically between 0.1 and 1) called the *jet resolution* or *jet radius*, and $\Delta R_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$ is the particle-pair distance in the rapidity-azimuthal angle (y-ϕ) plane. The fractional distance $\Delta R_{ij}/R$ can be interpreted as a unitless angle. The exponent b is the generalized parameter which determines the relative weighting of track momentum with respect to particle distance; the original logitudinally-invariant k_{T} algorithm uses $b = 2$.

Along with its distance to every other particle in the event, each particle is also assigned a distance to the beam,

$$d_{i\mathrm{B}} = p_{\mathrm{T},i}^{b}. \tag{1.25}$$

After computing these distance metrics for all particles, the algorithm finds the minimum of all d_{ij} and $d_{i\mathrm{B}}$ in the event. If the smallest distance is given by some d_{ij}, particles (jets) i and j are combined into a single jet, with their combined jet four-momenta defined using one of various *recombination schemes*. The theoretically most straightforward approach [64] is the E-scheme recombination, which simply adds the four-momenta together. If $d_{i\mathrm{B}}$ is the smallest distance, the jet i is considered final and removed from the list of jets. This process repeats until all jets are finalized.

The exponent b has a significant effect on the shape of reconstructed jets and the order in which particles are clustered. The nominal k_{T} case of $b = 2$ tends to reconstruct jets with large areas and rough edges, which can be sensitive to soft effects. The case of $b = 0$, which is completely p_{T}-independent, clusters together particles based on their closeness in the y-ϕ plane, leading to jets which are angularly-ordered. This ordering reflects the theoretically expected angular ordering of emissions from DGLAP. This parametrization is known as the Cambridge/Aachen (CA) reconstruction algorithm [65].

The case of $b = -2$ is known as the anti-k_{T} algorithm [66]. Contrary to the k_{T} algorithm, which tends to combine softer particles first, the anti-k_{T} algorithm tends to cluster together harder particles first. Unlike the nonuniform boundaries of k_{T} and CA jets, the jets reconstructed with the anti-k_{T} algorithm are cone-like and produce consistent jet areas, as shown in Fig. 1.8. This results in the jets being resilient to soft, nonperturbative effects such as the UE [66], which makes the anti-k_{T} algorithm preferable in collision systems with higher levels of background effects, such as collision systems containing QCD interactions.

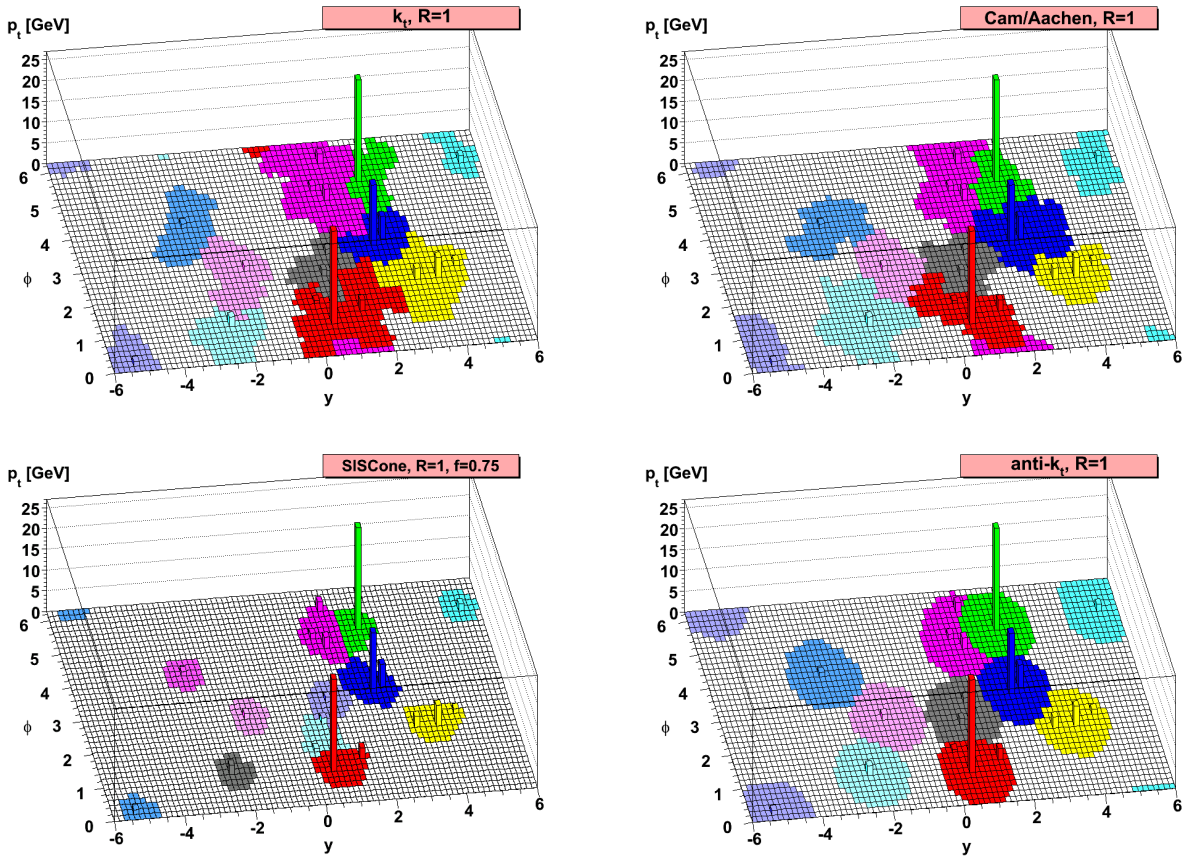


Figure 1.8: Comparison of jet sizes and shapes for a sample simulated event, as reconstructed with the k_{T}, Cambridge/Aachen, SISCone, and anti-k_{T} algorithms [66].

These k_{T}-based algorithms are IRC-safe. For many years, however, it was believed that these algorithms were significantly more complicated to implement, and suffered from the relative time inefficiency of $\mathcal{O}(N^3)$ [67]. Since the release of the open-source FastJet software package [68], accessible implementations are freely available with jet reconstruction time efficiencies outdoing the cone-based approaches by roughly two orders of magnitude. These factors have made sequential recombination algorithms, and the anti-k_{T} algorithm in particular, popular in experimental measurements over the last decade.

1.3.2 Jet substructure

Early studies of jets in scattering experiments were classified by using global event-shape observables, such as the sphericity S, as defined in the introduction to Sect. 1.3 above. Event-shape observables connect the perturbative short-distance dynamics of QCD to the nonperturbative long-distance ones [69], providing tests for QCD at opposing scales.

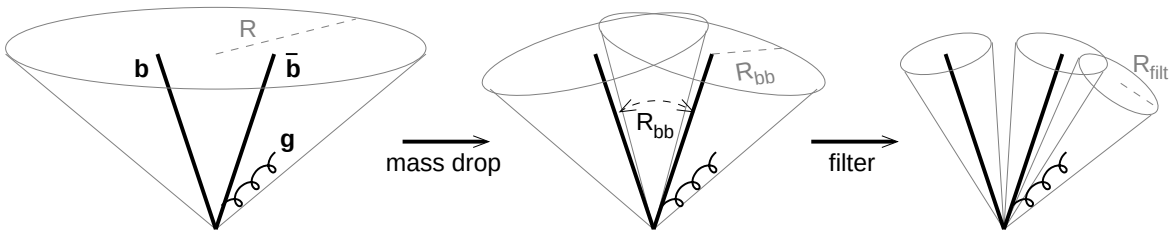


Figure 1.9: Proposed procedure for tagging Higgs-initiated jets via the decay into a $b\bar{b}$ pairs [71]. A large-radius ($R \approx 1$) "fat" jet is first reconstructed using the CA algorithm, containing the boosted $b\bar{b}$ pair along with any soft and collinear radiation. The fat jet is then split into subjets using the mass drop procedure, which searches for a sufficiently symmetric and heavy splitting in the jet clustering history. The radius of these two subjets is then reduced to $R_{\text{filt}} \approx 0.4$ in order to filter out contamination from the UE.

It was soon discovered that the internal structure of jets carries complementary information to event shapes. Individual jets were explored through experimental study of fragmentation functions, as introduced in Sect. 1.2.3, as well as jet multiplicity – the number of particles in a jet. Due to the differing color factors, the probability for a QCD emission from a hard-scattered quark ($C_{\text{F}} = 4/3$) is lower than that of a hard-scattered gluon ($C_{\text{A}} = 3$), leading gluon-initiated jets to have higher multiplicities and broader shapes than quark-initiated jets of similar momentum. Measuring jet multiplicity and breadth, then, was suggested as a way to differentiate between jets that came from a hard-scattered quark versus gluon [70].

Jet substructure garnered a broader interest after it was suggested that looking inside jets could be used to tag rare particle decays, such as the (then-undiscovered) Higgs boson decaying to a bottom $q\bar{q}$ pair [71], as depicted in Fig. 1.9. In addition to jet substructure's newly discovered ability to tag rare decays, observables such as the angularities and energy flow correlations, which had originally been proposed as event shapes [69, 72], were reformulated as probes of QCD using jet substructure.

One of the most widely-studied jet substructure observables, both experimentally [73–87] as well as theoretically [88–93], is the *jet invariant mass*,

$$m_{\text{jet}} = ||\mathbf{p}_{\text{jet}}|| = \sqrt{E_{\text{jet}}^2 - p_{\text{jet}}^2}, \tag{1.26}$$

where E_{jet} and p_{jet} are the jet energy and momentum, respectively. The jet mass is a proxy for the virtuality Q of the hard-scattered parton, which is larger for jets with broader fragmentation. Closely related are the *jet angularities* [69, 94–97],

$$\lambda_\alpha^\kappa \equiv \sum_{i \in \text{jet}} \left(\frac{p_{\text{T},i}}{p_{\text{T}}^{\text{jet}}} \right)^\kappa \left(\frac{\Delta R_i}{R} \right)^\alpha = \sum_{i \in \text{jet}} z_i^\kappa \theta_i^\alpha \tag{1.27}$$

where i runs over constituents in the jet, p_{T} designates transverse momentum, R is the jet resolution parameter, and $\Delta R_i \equiv \sqrt{(y_{\mathrm{jet}} - y_i)^2 + (\phi_{\mathrm{jet}} - \phi_i)^2}$ gives the distance between the jet and its ith constituent in the y-ϕ plane. The continuous, tunable parameters α and κ define the specific angularity observable.[1] The second equality is equivalent by convention, where $z_i \equiv p_{\mathrm{T},i}/p_{\mathrm{T}}^{\mathrm{jet}}$ is the *jet fragmentation function*, and $\theta_i \equiv \Delta R_i/R$. (Note that this definition of "fragmentation function" refers to perturbative emissions occurring during the parton shower, and is fundamentally different than the functions described in Sect. 1.2.3.)

The jet angularities encompass a variety of observables which have been studied independently from one another. The case $\kappa = \alpha = 0$ returns the familiar jet multiplicity, while setting $\kappa = 2$ and $\alpha = 0$ gives the square of the *momentum dispersion* (p_{T}D), which measures the second moment of the jet constituent p_{T} distribution. These observables have been suggested to be sensitive to the differences between quark and gluon jets, and have therefore been employed in such taggers. The jet angularities are IRC-safe for $\kappa = 1$ and $\alpha > 0$ [49], which encompasses both the *jet girth* $g = \lambda_1^1 R$ [98] and *jet thrust* λ_2^1 [72]. The latter has a direct theoretical relation to m_{jet} [99],

$$\text{jet thrust } \lambda_2^1 = \left(\frac{m_{\mathrm{jet}}}{p_{\mathrm{T}}^{\mathrm{jet}} R}\right)^2 + \mathcal{O}[(\lambda_2^1)^2]. \tag{1.28}$$

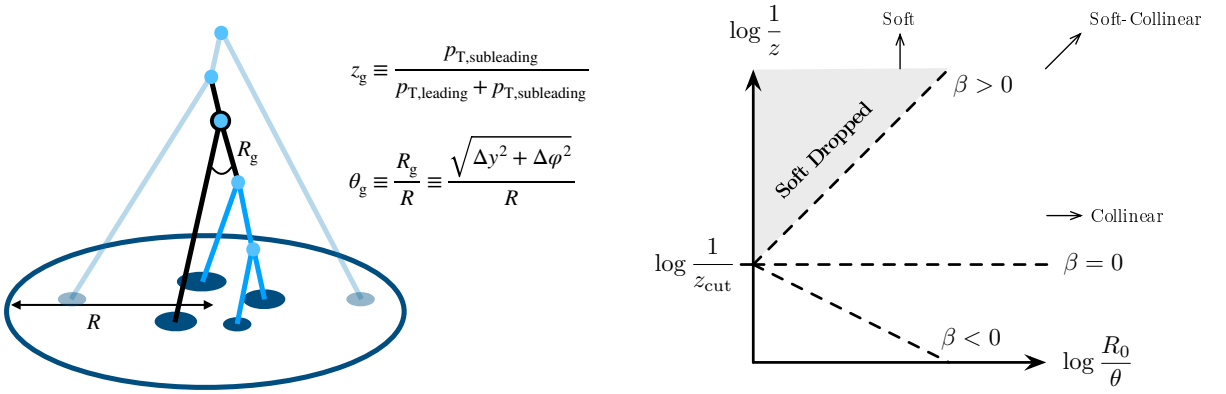
There are many other jet substructure observables which are outside the scope of this thesis. For a recent theoretical review, see Ref. [100]; for a review of some recent experimental results, see Ref. [101] and [102], Sect. 6.

1.3.3 Jet grooming

Jet substructure techniques can also refine the jet object to reduce both experimental and theoretical uncertainties. This procedure, referred to as *jet grooming*, aims to remove soft (small z) and wide-angle (large θ) radiation from a jet, isolating its more central and hard components. Early grooming techniques included trimming [105], pruning [106, 107], and mass drop tagging (MDT) [71], as debuted for the example shown in Fig. 1.9. It was soon discovered that this latter approach could be modified (mMDT) [88] to simplify theoretical calculations by removing Sudakov double logarithms [108] (from the incomplete cancellation of soft divergences) and non-global logarithms (NGLs) [109] (those that are sensitive to radiation in only a part of phase space) from field theoretic resummations, while also significantly reducing sensitivity to nonperturbative effects.

The Soft Drop (SD) algorithm [104] generalizes the mMDT approach. In the SD method, jets are first reclustered using the CA algorithm to an angularly ordered structure. This means that jets can be reconstructed from a sample of particles using any algorithm (e.g.

[1] The notation λ_α is employed to represent the jet angularities instead of the commonly-used notation λ_β in order to avoid conflict with the letter β, which is also used to denote the angular parameter of the soft drop grooming algorithm; see Sect. 1.3.3.



(a) Groomed jet cartoon with two observables (z_g and R_g) defined by the tagged splitting [103].

(b) Plot of the radiation phase space removed by SD grooming [104].

Figure 1.10: Depictions of the grooming procedure. *Left:* radiations are iteratively removed from a jet of radius R until a splitting satisfying the grooming condition is satisfied, from which jet substructure observables can be calculated. *Right:* SD grooming removes soft (small z) and wide-angle (large θ) radiation, while retaining some soft-collinear information, as determined by the scaling of β.

anti-k_T), and the reclustering will preserve the shape of those jets while reordering the splittings inside them into a structure which follows the DGLAP equations. The SD algorithm then iteratively steps through the CA clustering history and evaluates the SD condition at each splitting,

$$\frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} \overset{?}{>} z_{\text{cut}} \left(\frac{\Delta R_{12}}{R}\right)^{\beta}, \tag{1.29}$$

where $p_{T,i}$ is the transverse momentum carried by branch i at the splitting and R_{12} is the distance in the y-ϕ plane, as defined in Sect. 1.3.1 above. The user-defined parameters z_{cut} and β are used to tune the strength and angular- versus momentum-dependence of the grooming. If the SD condition is not satisfied by the splitting, the softer of the two branches is removed from the jet ("groomed away"), and the algorithm continues on to evaluate the next (subleading) splitting in the harder branch. If the condition is satisfied, the procedure is concluded. If no such splitting satisfies the SD condition, then the removal of all successive splittings means that the jet is removed from the groomed sample entirely.

From a theoretical standpoint, the SD algorithm has several useful properties. SD-groomed jets are IRC-safe, though soft-collinear radiation can be maintained and controlled inside the jet by varying β. The case of $\beta = 0$ is roughly equivalent to the mMDT tagger, but is Sudakov safe [56] to all orders in α_s. Furthermore, properties of the tagged splitting itself, such as the relative transverse momentum carried by the branches and the angle at which the splitting occurred, are both experimentally observable and theoretically calculable,

providing tests of the expected fragmentation structure and scaling from QCD [110]. Such measurements have been carried out by several experimental collaborations, who observe good agreement with pQCD within uncertainties [75, 103, 111–114].

One recent extension of grooming techniques is *dynamical grooming* [115–118], which reclusters jets using the CA algorithm similarly to the SD algorithm, but then searches the entire jet clustering history for the "hardest" splitting, defined by

$$\text{“hardness”} = \kappa^{(a)} = \frac{1}{p_{\mathrm{T}}^{\mathrm{jet}}} \max_{i \in \text{CA seq.}} \left[z_i(1-z_i)\, p_{\mathrm{T},i} \left(\frac{\theta_i}{R} \right)^a \right], \tag{1.30}$$

where a is a free parameter used to select on splittings ordered in different ways:

1. Setting $a = 2$, referred to as TimeDrop, selects the splitting with the shortest formation time $t_f^{-1} \sim \kappa^{(2)} p_{\mathrm{T}}$;
2. Setting $a = 1$, referred to as k_{T}Drop, selects the splitting with the largest relative transverse momentum $k_{\mathrm{T}} \sim \kappa^{(1)} p_{\mathrm{T}}$;
3. Setting $a = 0$, referred to as zDrop, selects the splitting with the most symmetric momentum sharing between the two branches. This case is collinear-unsafe, and is hence typically approximated with $a \approx 0.1$, which maintains calculability.

Dynamical grooming therefore benefits from enhanced interpretability of its free parameter with respect to the SD approach, where z_{cut} and β must be fine-tuned and do not necessarily correlate to any meaningful splitting in QCD.

1.3.4 Heavy-flavor jets

The study of heavy-flavor hadrons – those initiated with a c or b quark (there are no hadrons with a top valence quark, which decay before they can hadronize) – offers valuable insight into QCD. A detailed study of heavy quark dynamics was carried out in [120]. The invariant mass for a hard-scattered quark which undergoes gluon emission ($\mathrm{q} \to \mathrm{q}'\mathrm{g}$) can be expressed as

$$m_{\mathrm{q}}^2 \approx 2E_{\mathrm{q}'}E_{\mathrm{g}}(1-\cos\theta), \tag{1.31}$$

where θ is the angle between the 3-momenta of q′ and g. In the collinear (small-θ) limit, valid within the jet cone, a Taylor expansion of the cosine in Eq. 1.31 yields

$$\theta \approx \frac{m_{\mathrm{q}}}{\sqrt{E_{\mathrm{q}'}E_{\mathrm{g}}}} = \frac{1}{\sqrt{\zeta(1-\zeta)}} \frac{m_{\mathrm{q}}}{E_{\mathrm{q}}}, \tag{1.32}$$

where $\zeta = E_{\mathrm{g}}/E_{\mathrm{q}}$ is the energy fraction carried by the gluon. Taking the initiating parton to be relativistic ($E_{\mathrm{q}} \gg m_{\mathrm{q}}$) and defining $\theta_0 = m_{\mathrm{q}}/E_{\mathrm{q}}$, the radiation probability for gluon emission at a small opening angle $\theta < \theta_0 \ll 1$ is [121]

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}E_{\mathrm{g}}} \right)_{\mathrm{q}\to\mathrm{q}'\mathrm{g}} = \frac{\alpha_s}{\pi} C_{\mathrm{F}} \frac{(2\sin\theta/2)^2\, \mathrm{d}(2\sin\theta/2)^2}{[(2\sin\theta/2)^2 + \theta_0^2]^2} \frac{1}{E_{\mathrm{g}}} [1 + \mathcal{O}(\theta_0, E_{\mathrm{g}})] \sim \frac{1}{E_{\mathrm{g}}} \frac{\theta^2 \mathrm{d}\theta^2}{(\theta^2+\theta_0^2)^2}, \tag{1.33}$$

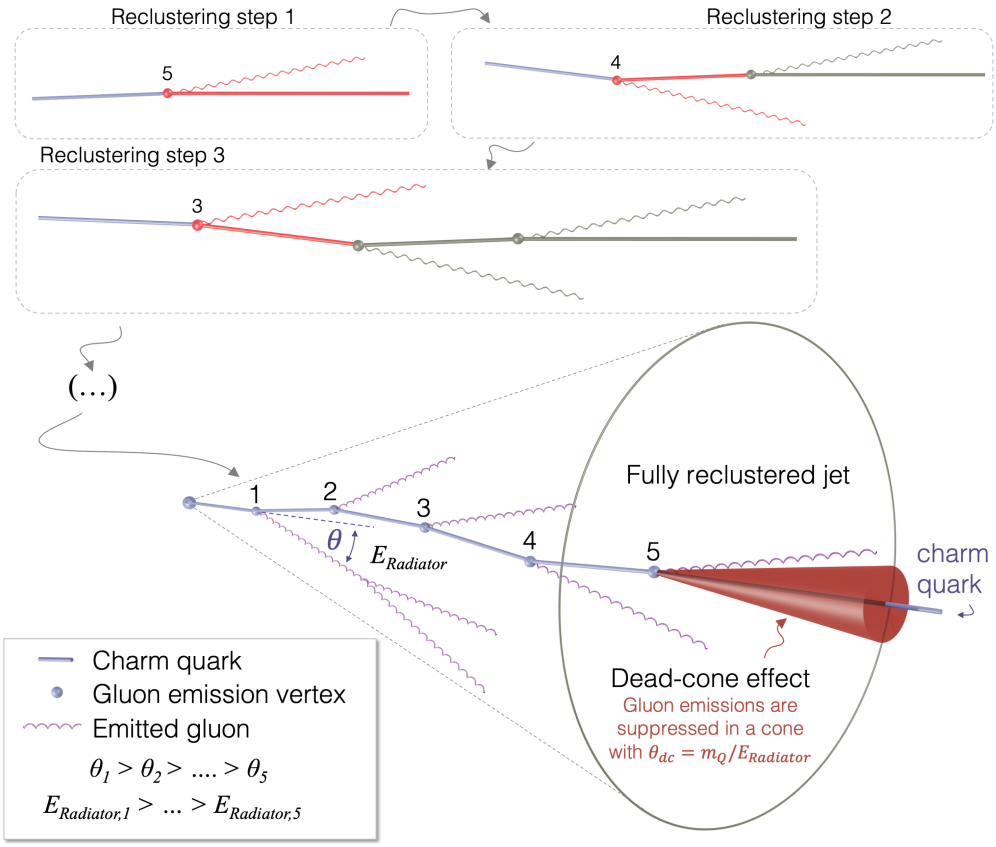


Figure 1.11: Diagram representing the procedure for the direct measurement of the dead cone [119]. At the top, the CA algorithm reclusters the jet in order of smallest angles first, reconstructing the angularly-ordered parton shower. After the reclustering is complete, the jet is then declustered, with the heavy quark energy E_{Radiator} and momentum axis adjusted with each subsequent emission. The radiation angle θ_i and the heavy quark energy $E_{\mathrm{Radiator},i}$ is recorded at each declustering step i.

where $\mathcal{O}(\theta_0, E_{\mathrm{g}})$ contains small power correction terms of the order of θ_0. Comparing with the double-logarithmic approximation (DLA) [121], the angular integration is no longer logarithmic in θ^2 for the kinematic region $\theta < \theta_0$, and the particle yield from this region contributes minutely ($\sim\sqrt{\alpha_s}$) to the total multiplicity. This effect is known as *angular screening*, and the kinematic region $\theta < \theta_0$ where gluon radiation is suppressed is called the QCD *dead cone* [121].

Due to the significant mass ordering between 'light' and 'heavy' quarks – with $m_{\mathrm{u,d}} \sim$ MeV/c^2, while $m_{\mathrm{c,b}} \sim$ GeV/c^2 – the dead cone angle θ_0 is significantly enhanced for heavy quarks with respect to lighter ones [122], an effect which is experimentally observable. To capture this radiation pattern, jets are reconstructed around heavy-flavor hadrons, which themselves are typically reconstructed from the heavy hadron's decay daughters. Jet sub-

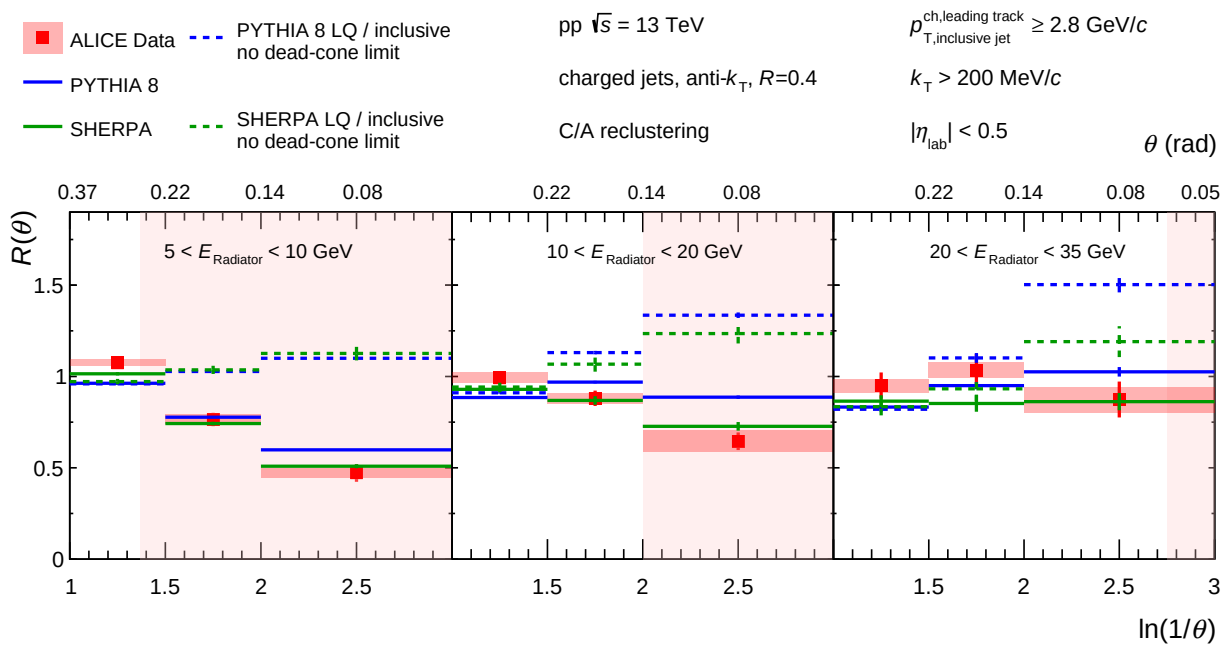


Figure 1.12: Experimental results for the direct dead cone measurement of D^0-tagged jets in increasing bins of E_{Radiator} [119]. Pink shaded areas represent the kinematic region $\theta < \theta_0$, where gluon radiation is expected to be suppressed. The charm quark mass is assumed to be $m_c = 1.275$ GeV/c^2. Data are compared to the PYTHIA 8 and SHERPA MC event generators, both in the nominal charm quark case and in the light quark (LQ) / inclusive limit, where no dead cone effect is expected.

structure can then be used to unravel the dead cone effect.

Such measurements of the dead cone, however, eluded experimental observation for decades after its initial prediction, due to the challenging nature of tagging heavy-flavor jets with sufficient statistics and efficiency, while untangling the perturbative emissions of the heavy quark from a dynamically changing quark axis, along with the nonperturbative corrections from hadronization, which smear the kinematic region $\theta < \theta_0$. The first experimental studies observed a relative suppression in the multiplicities for particles near the jet axis [123], which then could be combined with MC models and the expected dead cone suppression to extract a value for the mass of the heavy quark [124].

The first direct observation of the dead cone [119] employed CA reclustering to trace the heavy-flavor quark along its trajectory following successive emissions. The procedure is sketched in Fig. 1.11, and results are shown in Fig. 1.12, where a significant suppression in radiation is observed within the dead cone limit for D^0-tagged (charm) jets. This trend is reproduced by MC event generators, as mass scaling is included in LO pQCD cross sections. Turning on light quarks in these models eliminates the relative suppression, confirming the mass dependence of the effect.

Though the dead cone effect has now been directly observed, theoretical understanding of heavy-flavor jet substructure is still incomplete. Probing the jet fragmentation of heavy quarks at all angles provides stringent tests of pQCD, as heavy quarks are perturbatively well-described: since the quark mass is near the scale of Λ_{QCD}, the production of heavy-quarks can be described down to lower p_{T} using perturbation theory than for light quarks. However, formalisms for calculating heavy-flavor jet substructure observables in perturbative QCD have only recently appeared. One outstanding mystery is that the number of charm baryons produced relative to charm mesons at low transverse momentum has been observed to be roughly four times larger than expected from the "cleaner" hadronization in $\mathrm{e^+e^-}$ collisions [125]. This signals the possibility of hadronization universality breaking for charm hadrons, for which jet substructure studies could illuminate the source.

1.4 Relativistic heavy-ion collisions

Examples of universal nonperturbative effects constrained from $\mathrm{e^+e^-}$ and $\mathrm{e^-p}$ collisions have been presented above, with connections to their use in more complicated QCD-dominated scattering experiments, such as pp. In a similar way, QCD constraints from these data are used to understand the baseline for more complicated heavy-ion (nuclear) collisions. These systems are sometimes designated as AA, for a generic nucleus species A.

1.4.1 Cold nuclear matter effects

The introduction of a heavy nucleus into a collider experiment modifies several different factors with respect to a binary nucleon-nucleon scattering. When a proton or neutron is embedded into a nucleus, potential energy is converted into binding energy, and the nucleon begins to exchange color-charged particles with other nucleons. While the nucleus has a net positive electric charge, this *residual strong interaction*, historically called the *strong nuclear force*, is enough to overpower the electric repulsion between protons. While quarks and gluons are confined inside color-neutral objects, nucleons can still interact via QCD at close enough distances, mediated by color-bound particles such as the π^0 meson. At small enough distances ($\lesssim 0.7$ fm), the interaction is repulsive, leading the nucleons to sit in a potential well phenomenologically described by the Reid potential [126].

As a newly bound nucleon undergoes QCD interactions with other nucleons in its vicinity, the PDF of the nucleon also changes [127]. While the nucleon binding energy was initially considered to be too small with respect to the energy exchanged in DIS to see any such fluctuations, this modification has been known since the 1970s, but its underlying mechanisms are not well understood. The structure function modification observed by the European Muon Collaboration in 1982 [128] was opposite to the expected trend; the result is shown with respect to Bjorken x in Fig. 1.13. This was a surprising discovery, with the current leading explanations being mean-field modifications or short-ranged correlated pairs [129, 130]. The modification of nuclear PDFs is an example of a so-called *cold nuclear matter effect*,

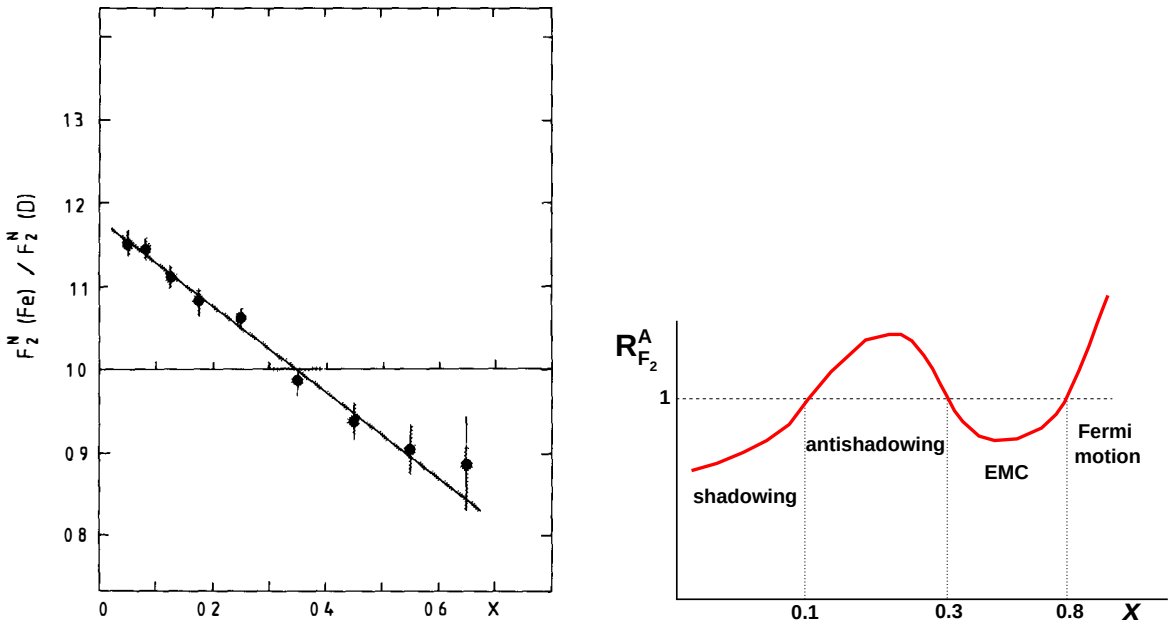


Figure 1.13: *Left:* first experimental observation of the difference between structure functions for iron (Fe) and deuterium (D), known as the EMC effect [128]. The enhancement at small x was not predicted, as it was expected that the ratio would increase or be approximately consistent with unity. It was later shown [131] that the modification increases with nuclear size. *Right:* sketch of the nucleus-nucleon structure function ratio $R_{F_2}^{\mathrm{A}} = F_2^{\mathrm{A}}(x,Q^2)/[\mathrm{A}F_2^{\mathrm{nucleon}}(x,Q^2)]$ for a generic nucleus and nucleon [132]. See text for explanation of different regions in x.

meaning that it is a function of the nuclear structure, and a high-energy (hot) interaction is not required to create it (though may be used to probe it).

The nuclear PDF (nPDF) tends to have several regions of relative enhancement and suppression as compared to unbound nucleons, as depicted in Fig. 1.13. At very low x (with $x \lesssim 0.1$), the distributions are suppressed, in an effect called *nuclear shadowing*. This suppression has been attributed to different physics phenomena in phenomenological approaches, including photon-to-vector-meson fluctuations resulting in hadronic flux being absorbed by nucleons on the outside of the nucleus [127], multiple scatterings inside the nucleus [132], as well as a higher probability for gluon-gluon fusions in a nuclear environment where gluon distributions are broadened [133]. As x is increased to $0.1 \lesssim x \lesssim 0.3$, the shadowing is reversed into *antishadowing*, a corresponding enhancement required by the relative suppression at lower x. The EMC effect is observed for $0.3 \lesssim x \lesssim 0.8$, with Fermi motion of the nucleon inside of the nucleus dominating at $0.8 \lesssim x \leq 1$.

Cold nuclear matter effects are typically uncovered by high-energy scattering of leptons or hadrons off of atomic nuclei. In pA collisions, for example, it was observed that the hadrons

typically produced at low $p_{\mathrm{T}} \lesssim 1$ GeV/c) were boosted to higher p_{T} than in pp collisions [134], which is called the Cronin effect. This observation has historically been attributed to multiple scattering inside the nucleus, though it has only been explained phenomenologically [135]. In a similar way, it is possible that the existence of a nuclear environment could modify the (otherwise perturbative) parton shower. Experimental results are unclear, showing no evidence of jet fragmentation modification at midrapidity for both large and small values of $\sqrt{s}$ [136–138], while at forward rapidities (sensitive to lower x) modification has been observed in several collision systems at small $\sqrt{s}$ [139, 140], but not at large $\sqrt{s}$ [141], where measurements at lower values of $p_{\mathrm{T}}^{\mathrm{jet}}$ may be required.

1.4.2 Quark-gluon plasma

Due to their reduced charge-to-mass ratio, heavy ions can only be accelerated to around half of the per-nucleon energies as protons using the same collider machine; however, since heavy nuclei are also much more massive than single protons, the heavy-ion collisions themselves carry much more total kinetic energy than pp collisions alone. In the late 1970s and early 1980s, it was predicted that, due to QCD being an asymptotically free gauge theory, these large temperatures and densities would create a state of matter where quarks and gluons are decoupled, behaving as a plasma [142–144]. According to lattice QCD calculations, the transition to this state of matter, termed the *quark-gluon plasma* (QGP) or sometimes referred to more generically as the *QCD medium*, should occur around a critical temperature of $T_{\mathrm{c}} \approx 150$ MeV [145]; this corresponds to energy densities of about $\epsilon \approx 1$ GeV/fm^3, roughly seven times that of normal nuclear matter [146]. Prior to experimental creation in the laboratory, the last time such energy densities were known to have existed were within the first few microseconds after the Big Bang, though it has been suggested that such matter could exist at the cores of neutron stars at ultrahigh density and lower temperature [142, 147].

Perturbative predictions can be made in the high-density $\epsilon \to \infty$ limit [142], where small-distance QCD interactions are asymptotically free. Similarly, thermal perturbation theory can be used at large temperatures $T \gg T_{\mathrm{c}}$ [144], while lattice calculations can be run for small QCD systems, where calculations are feasible on modern computers. At the temperature and density of the phase transition, the theory is nonperturbative, and at present must be described by models. Experimental data is therefore essential to probe the formation and structure of the QGP, where many mysteries remain. An overview will be detailed below, prefaced by a historical recap. For a recent detailed review, see Ref. [148].

One predicted effect in QGP production is *strangeness enhancement*. Within ordinary nuclei, s quarks come predominantly from the gluon sea (see Fig. 1.5), and the production of *hyperons* (strange hadrons: those hadrons which contain one or more s valence quarks, but no c, b, or t) is therefore suppressed relative to production of hadrons that contain only u or d valence quarks. In a QGP, where the temperature is larger than the mass of the s quark $m_{\mathrm{s}} \approx 100$ MeV/c^2, the system comes to a chemical equilibrium for the u, d, and s quarks via thermal production, which is mass-ordered by m_{q}/T. The heavier quarks,

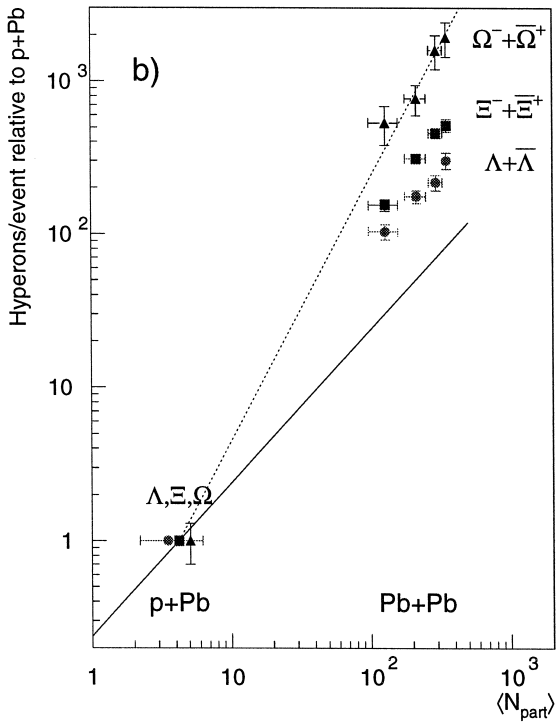


Figure 1.14: Hyperon enhancement as observed by the WA97 experiment [150]. Hyperon production is found to scale faster than the average number of binary particle interactions $\langle N_{\rm part}\rangle$, with a 1:1 scaling given by the solid line. Instead, Ω production is found to be proportional to $\langle N_{\rm part}\rangle^{1.72}$ (dashed line).

whose masses are significantly above the QGP thermal energy density, are not sufficiently produced to reach the same equilibrium. As a result, hyperons and antihyperons should be produced much more abundantly after a QGP has been created than a corresponding superposition of binary nucleon-nucleon collisions.[2] Strangeness enhancement was clearly observed in the late 1990s in lead-lead (Pb–Pb) collisions by experiments on the Super Proton Synchrotron (SPS) [150–153], with the heaviest hyperons being produced roughly 100 times more frequently in Pb–Pb collisions than they would have been in superimposed pp collisions. Example results are shown in Fig. 1.14.

After its creation, the QGP expands to roughly 30-50 times its original size before hadrons are set free [146]. This was studied by measurements of two-particle correlations at the same SPS experiments, via a technique known as *Bose-Einstein interferometry* [154, 155]. Using a method dating back to the stellar interferometry developed by Hanbury Brown and Twiss (HBT) [156], two-particle correlations in heavy-ion collisions use measurements of the angular hadron intensity to characterize the QGP at the point of chemical freeze-out, at $T = T_{\rm c}$.

[2]Strange-antistrange pairs can also be produced in a hot hadron gas; however, the energy for production in a QGP ($\sim 2m_{\rm s} \approx 200$ MeV/c^2) is lower than the hadronic channels (≈ 670 MeV/c^2) [149].

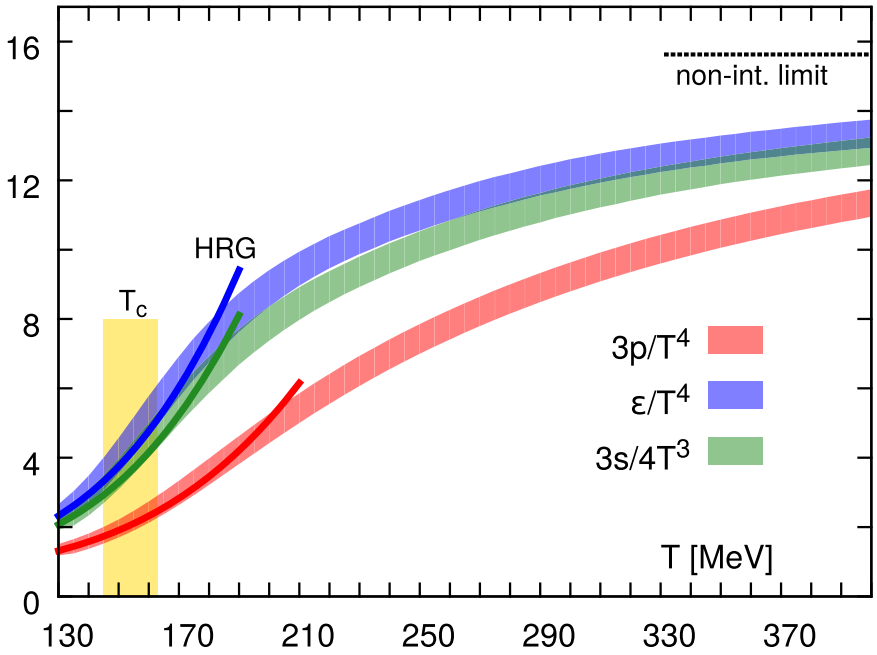


Figure 1.15: Lattice QCD predictions of the pressure p, energy density ϵ, and entropy density s for the QGP at varying temperatures [157, 158]. The ideal gas (non-interacting / Stefan-Boltzmann) limit is shown as a black dashed line. The phase transition is a continuous crossover, with smoothly rising spectra above the critical temperature $T_c \approx 150$ MeV. The Hadron Resonance Gas (HRG) model [159] is effective at describing the hadronic equation of state at low temperatures, but diverges above the QGP phase transition.

Hadrons emitted from the QGP then interact as a hadron gas, until the system cools to the kinetic freeze-out of $T \approx 100$ MeV (with energy density $\epsilon \approx 50$ MeV/c), at which point the hadrons continue their trajectories, continue through normal hadronic decays, and finally enter the experimental detectors.

1.4.3 Hydrodynamic flow

Because of asymptotic freedom, it was originally predicted that the quarks and gluons produced in hot, dense heavy-ion collisions would approach the weak coupling limit and have properties of an ideal gas. This was found, however, to not be the case [157, 158]; see Fig. 1.15. The phase transition from nuclear matter to a QGP is continuous, similar to the second-order phase transition which occurs during the ionization of an atomic plasma, and rises smoothly above T_c but remains below the Stefan-Boltzmann limit. There are different possibilities for why ideal gas behavior might not occur. It is possible that, since the quarks and gluons are asymptotically free, they do not interact enough to reach thermal equilibrium. On the other hand, it is possible that the energies reached experimentally are

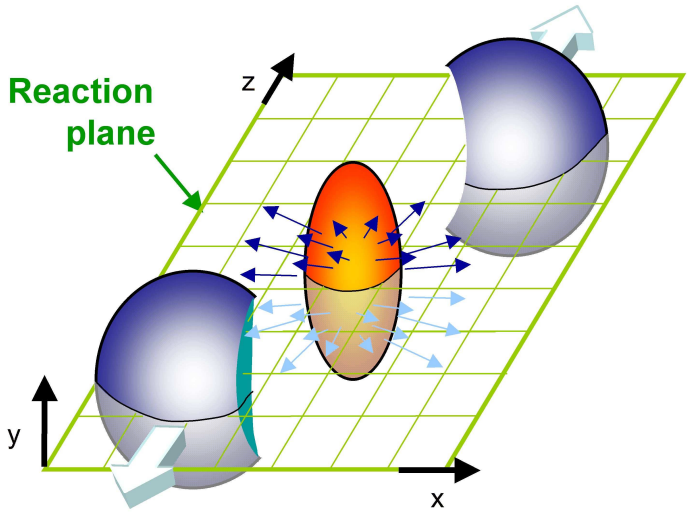


Figure 1.16: Cartoon of a non-central AA collision [160], with the two outgoing nuclear remnants colored purple, and the overlapping region forming a QGP (orange). The horizontal plane of symmetry formed by the collision is known as the *reaction plane* (green). The QGP is oblong or almond-like in shape, and its internal pressure gradient causes a nonuniform expansion as predicted by relativistic hydrodynamics.

not high enough to completely decouple the partons, leaving some remnant strong coupling. This could result in complex partonic states carrying the energy difference between the non-interacting limit; in other words, the degrees of freedom for a "quark-gluon" plasma could include complex color states, rather than free quarks and gluons.

Experimental studies reveal that relativistic hydrodynamics [161] can be used to predict and extract properties of the QGP with striking detail, implying that the QGP behaves like a fluid. One revealing area of study into QGP fluid dynamics is that of *collective* or *hydrodynamic flow* [162]. Due to the finite size of atomic nuclei, AA collisions occur with some impact parameter b which separates the centers of the two nuclei. The overlapping region is not spherically symmetric, but instead has an oblong shape as depicted in Fig. 1.16. As the QGP expands, its dynamics are driven by both its anisotropy and its viscosity, as hydrodynamics converts spatial anisotropies into momentum anisotropies [148]. If particles in the QGP were free-streaming, the motion of particles inside the QGP would be random, which would wipe out any observable anisotropy in the final state; similarly, if the QGP were a fluid with large viscosity, these momentum anisotropies would be dampened.

The azimuthal flow of particles in the final state can be quantified from a Fourier transform of the angular distribution of hadrons in the final state [163],

$$\frac{\mathrm{d}\bar{N}}{\mathrm{d}\theta} = \frac{\bar{N}}{2\pi}\left(1 + 2\sum_{n=1}^{\infty} \bar{v}_n \cos[n(\theta - \bar{\Psi}_n)]\right), \tag{1.34}$$

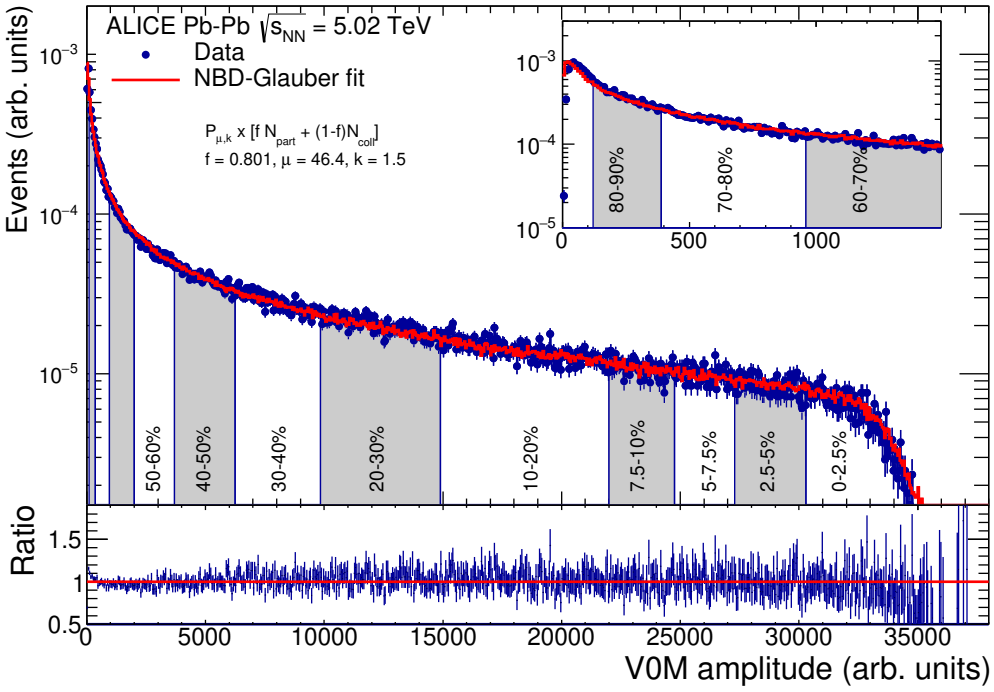


Figure 1.17: Measurement of the event multiplicity distribution with the ALICE detector using Pb–Pb data at $\sqrt{s_{\rm NN}}$ = 5.02 TeV [164, 165]. Data from the ALICE V0 detectors (see Sect. 2.2.4) is fit with Glauber MC coupled to a negative binomial distribution (NBD) particle production model. Integrated centrality regions are shown below the curve.

where $\bar{N}$ is the average number of hadrons per event, θ is the azimuthal angle transverse to the reaction plane, $\bar{v}_n$ are the *anisotropic flow coefficients*, and $\bar{\Psi}_n$ are the *event plane angles*. The size of $\bar{v}_n$ describe the strength of the nth harmonic component in the angular distribution, while $\bar{\Psi}_n$ is the angle of maximum multiplicity.

The extracted values of these parameters will depend not only on the properties of the fluid but also on b. For the most *central* collisions, where $b \approx 0$, the shape of the initial fluid is more spherical, with the primary anisotropies coming solely from the "lumpiness" of the colliding nuclei. As b is increased, the collisions become more *peripheral*, and the created fluid becomes more oblong. From an experimental perspective, however, b is fundamentally unknown, so in order to approximate this dependence, collisions are classified based on the multiplicity of the event. High-multiplicity collisions are typical of smaller b, not only since the increased number of participating nucleons increases the yield of final-state particles, but because the high energy densities increase thermal production. Larger b collisions create a smaller volume of QGP with less total energy, and thus fewer particles are created in the final state.

Using the assumption that nuclei are non-interacting bags of nucleons – the so-called

Glauber model, named after Roy Glauber – which are sampled from measured density distributions [166], the number of interactions and the produced particle yields per event can be predicted. These yields are experimentally correlated with the *centrality* of the collision, which is typically reported as a percentage range of the total cross section, with 0% being the most central and 100% being the most peripheral collisions. Note that this is an integrated percentage, and that experiments can never measure the full range of the interaction cross section. An example centrality determination is shown in Fig. 1.17.

A measurement of the coefficient for the second Fourier coefficient $\bar{v}_2$, also called the *elliptic flow*, in high-energy ($\sqrt{s}$ = 2.76 TeV) Pb–Pb collisions is given in Fig. 1.18. ($\bar{v}_1$ and $\bar{v}_3$ are similarly known as the coefficients of *radial* and *triangular* flow, respectively.) The flow is found to be nonzero for all centralities, and to significantly increase with more peripheral collisions, which matches the geometrical picture.

These measurements of flow can be carried out for different particle species in various kinematic regimes and at different centralities. When compared to hydrodynamic predictions, these measurements probe fluid properties of the QGP. Recent extractions [148] give a QGP formation time of $\tau_0 \approx 0.2$ fm/c, and a shear viscosity η (not to be confused with pseudorapidity η) to entropy density s ratio (*specific viscosity*) $\eta/s \approx 1/4\pi$, which is lower than any other known fluid. For this reason, the QGP is sometimes called a "nearly-perfect" fluid. This exact value of η/s also leads to a corresponding gravitational description of a black hole horizon in 4+1 dimensional Anti-de Sitter space [168, 169] whose oscillations correspond to the motion of the plasma [170]. Calculations in this formalism are tractable and would potentially be equivalent to the intractable ones of QCD. Extracting η/s with higher precision is therefore highly useful to probe the veracity of this correspondence.

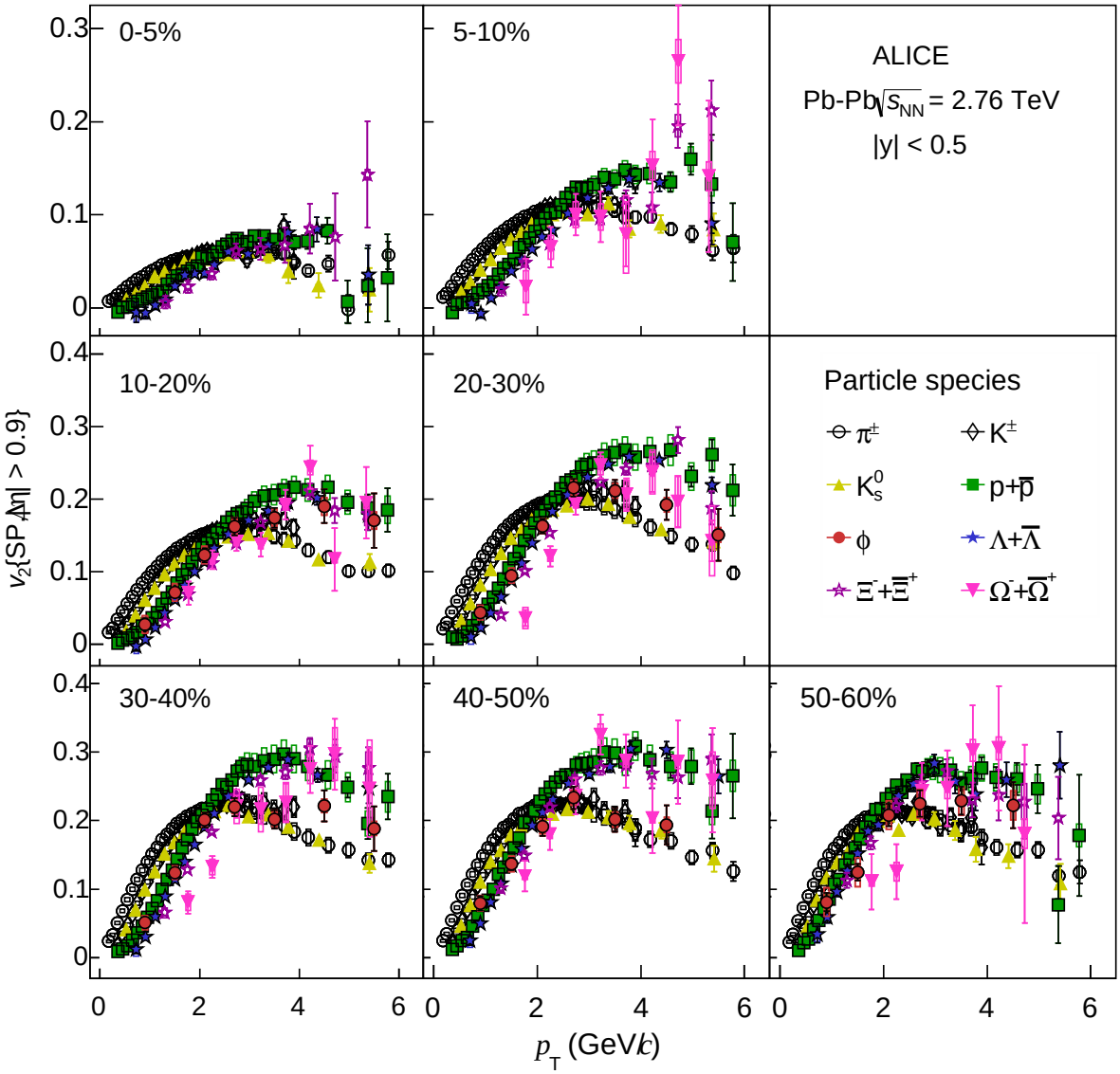


Figure 1.18: Measurement of the elliptic flow coefficient $\bar{v}_2$ in Pb–Pb collisions with centrality ranging from 0 to 60% and for several particle species as a function of their p_{T}, as calculated using the Scalar Product (SP) method [167]. The flow increases for more peripheral collisions, where events produce a more anisotropic QGP. The mass ordering for $p_{\mathrm{T}} < 3$ GeV is given by an interplay between the radial and elliptic flow which modifies $\bar{v}_2$ according to mass. After some threshold value, increases of particle p_{T} decrease the effect of flow, as the particles tend to rapidly emerge from the fluid.

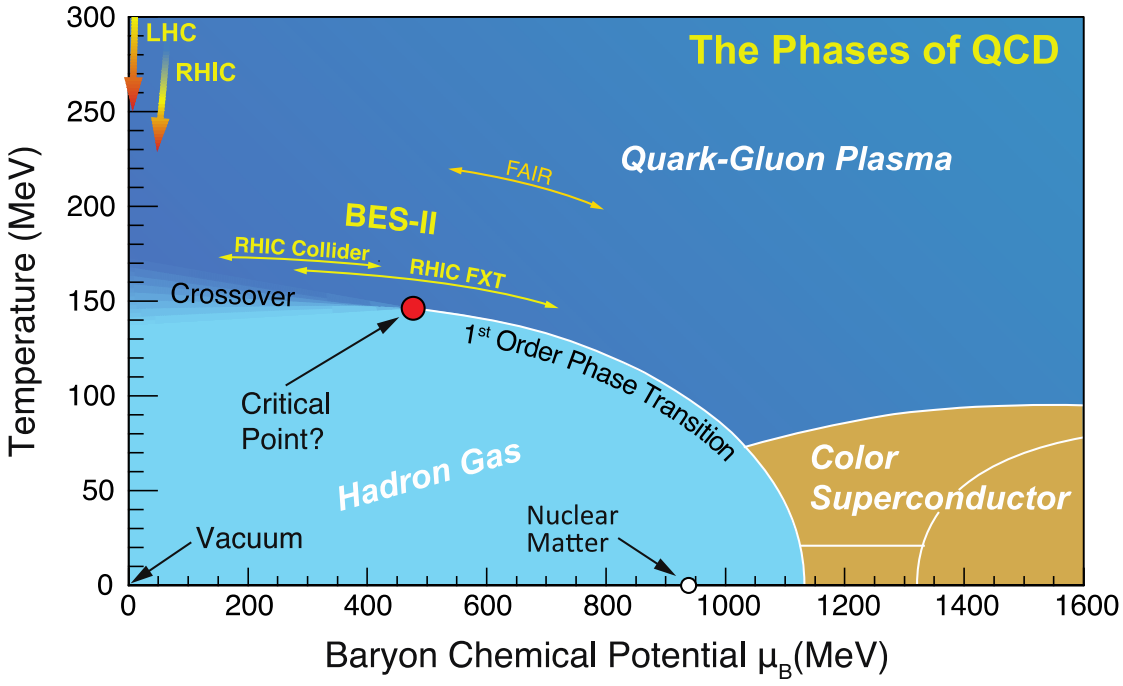


Figure 1.19: Schematic view of the phase diagram of QCD as a function of temperature and baryon density [171]. The existence of a QCD critical point has not been confirmed and is currently being searched for by the Beam Energy Scan, Phase II (BES-II) at the Relativistic Heavy Ion Collider (RHIC) on Long Island in New York. Similarly, the proposed color superconductor phase has not experimentally been probed, and is still under investigation.

1.4.4 Phase diagram

It is worth briefly mentioning the extensive recent (and ongoing) effort to map out the QCD phase diagram. By varying the number of baryons relative to antibaryons produced in collisions (by varying $\sqrt{s}$) and by recording the temperature of the medium produced, one can search for the possible existence of a QCD critical point, above which the hadron gas and QGP phases are separated by a first-order phase transition. Some models have predicted that in the *chiral limit*, where QCD has two massless quarks, a sharp second-order phase transition at low baryon density (which restores chiral symmetry) becomes a first-order phase transition at a tricritical point with $\mu_B > 0$ [148, 172]. Some lattice calculations have been performed in this regime, but suffer from a "sign problem" [173] at finite μ_B which has (thus far) prevented these calculations from being taken from a discretized lattice to the continuum limit [148].

At ultrahigh baryon density and lower temperature, QCD matter is expected to enter a color superconductor regime, which may be found at the centers of neutron stars [142, 147]. On a microscopic level, a QCD superconductor is a "degenerate Fermi gas of quarks with a

condensate of Cooper pairs near the Fermi surface that induces color Meissner effects" [174], emulating the behavior of an ordinary (electromagnetic) superconductor. Analytic calculations can be performed at the highest densities to study properties of a color-flavor locked (CFL) superfluidity phase, which may break down at lower densities.

For more information on studies of the QCD phase diagram, see Ref. [174–177].

1.4.5 Jet quenching

As jets are naturally sensitive to QCD at various momentum scales Q, jet observables offer a unique probe of the QGP at its various stages of evolution. After a hard scattering, occurring before (or perhaps during) the QGP formation, dijets traverse the fluid with different path lengths depending on the location of the scattering within the colliding nuclei. The normal probability for a gluon radiation in perturbative QCD is enhanced by medium-induced emissions, which affect both the jet momentum and substructure. These medium-induced effects are referred to as *jet quenching*.

The energy loss or gain of jets (or other probe i) through the QGP produced in AA collisions can be expressed via the *nuclear modification factor*,

$$R^i_{\mathrm{AA}}(p_{\mathrm{T}}) = \frac{\mathrm{d}^2 N^i_{\mathrm{AA}}/(\mathrm{d}p_{\mathrm{T}}\,\mathrm{d}\eta)}{\langle T_{\mathrm{AA}}\rangle\,\mathrm{d}^2\sigma^i_{\mathrm{nn}}/(\mathrm{d}p_{\mathrm{T}}\,\mathrm{d}\eta)}, \tag{1.35}$$

where N^i_{AA} is the jet (or particle i) yield in AA collisions, σ^i_{nn} is the corresponding cross section for that species in nucleon-nucleon (e.g. pp) collisions, and $\langle T_{\mathrm{AA}}\rangle$ is the average number of binary nucleon-nucleon collisions per AA collision scaled by the inelastic nucleon-nucleon cross section,

$$\langle T_{\mathrm{AA}}\rangle = \langle N_{\mathrm{coll}}\rangle/\sigma^{\mathrm{inel}}_{\mathrm{nn}}. \tag{1.36}$$

The value of $\langle N_{\mathrm{coll}}\rangle$ is typically calculated in the Glauber approach. By Eq. 1.35, the value of $R_{\mathrm{AA}} = 1$ would correspond to AA being a superposition of nucleon-nucleon collisions. This observable was first defined for single hadrons [178, 179]. A similar modification factor can also be constructed by comparing central to peripheral collisions, which is typically called R_{CP}, or for mixed collision systems such as pA (called R_{pA}).

Measurements of the jet cross section in AA and pp collisions have consistently revealed a suppression in the R_{AA} that is not seen for electromagnetic probes which are largely unquenched by the plasma nor for jets in smaller collision systems (i.e., in R_{pA}) [181]. A recent systematic measurement of the jet R_{AA} is given in Fig. 1.20. The centrality dependence of the quenching is phenomenologically understood as more central events creating a larger volume of fluid which increases the jet mean free path through the QGP and thus its modification. As $p^{\mathrm{jet}}_{\mathrm{T}}$ increases, R_{AA} heads closer to unity, as jets are able to escape the QGP with less (but still significant) quenching.

The path-length dependence of jet quenching can lead to an imbalance in the dijet energy yield as measured in heavy-ion collisions. This effect, known as dijet asymmetry, has also been experimentally measured [182, 183]. However, the effect of jet quenching on the

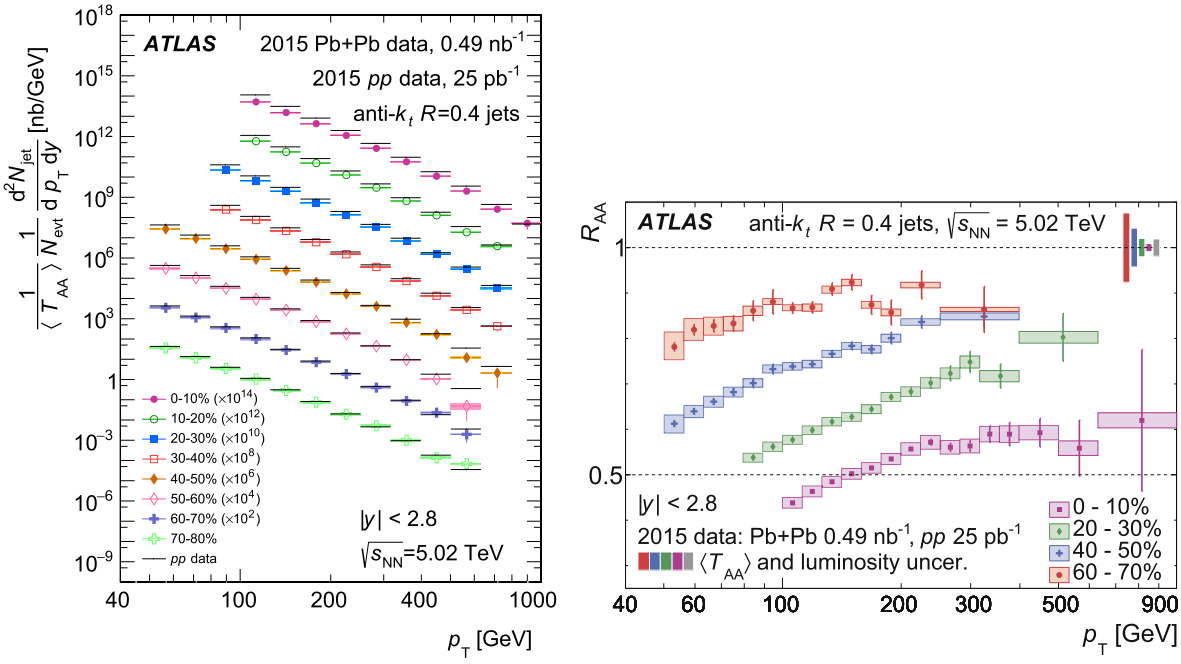


Figure 1.20: Recent measurements of the jet cross section in pp and Pb–Pb collisions (left) and the corresponding nuclear modification factor R_{AA} (right) [180]. Results are shown for centrality ranging from 0 to 80%, in bins with width 10%. Significant modification is observed for all reported centralities, with drastically enhanced quenching observed for the more central collisions.

fragmentation of jets is still under experimental study, and these studies form a portion of the results in this thesis.

1.4.6 Heavy-ion MC models

Much like in pp collisions (see Sect. 1.2.4), MC approaches can be used to predict the dynamics of heavy-ion collisions. In addition to the standard techniques of pp models, AA models must have additional parameters to describe the collision geometry and predict event modifications due to the effect of the QGP. This is implemented in different ways, depending on the preferred phenomenological model of QGP structure and interactions.

There are several approaches to calculate medium-induced gluon radiation based on pQCD [184]. The *multiple soft scattering approximation* (MSSA), also known as the BDMPS-Z (Baier-Dokshitzer-Mueller-Peigne-Schiff and Zakharov) approximation, takes the mean free path to be small, which allows resummation of high-density QGP effects to all orders; this is the basis for medium-induced parton showers [185–187]. The *single hard scattering approximation* (SHSA), as first proposed by GLV (Gyulassy-Levai-Vitev) [188] and Wiedemann [189], instead assumes that the medium is dilute, i.e. that the mean free path is much

larger than the size of the medium. This is extended in the *Higher-Twist* (HT) approximation, which further assumes that the gluon p_{T} is much larger than the energy scale Q of medium interactions [190, 191].

The JEWEL framework [192] is currently the only fully open-source implementation of jet evolution in heavy-ion collisions available.[3] Based on the underlying PYTHIA model, JEWEL uses user-defined PDFs (e.g. nPDFs from LHAPDF6 [193]) to produce the $2 \to 2$ hard-scattered partons, which then propagate through a parton shower with the addition of extra medium-induced radiation. JEWEL assumes that jets interact with partons inside of the plasma, and, depending on whether enabled by the user, can create medium *recoil* scattering centers, which then must be treated like background and "subtracted" from the final observables before comparing to data [194]. Comparisons to JEWEL are often shown with this recoil effect disabled, as well as enabled with observable-specific subtraction procedures.

The Hybrid model [195] is also based on PYTHIA and adds medium-induced emissions to the shower. This approach assumes a strongly-coupled medium and uses holographic calculations obtained via gauge/gravity duality (as briefly mentioned in Sect. 1.4.3) to model the extra energy loss, obtaining a stopping distance that is comparable to the result from $\mathcal{N} = 4$ supersymmetric Yang-Mills (SYM) theory [168, 196]. This approach can be combined with other effects such as elastic Molière scattering inside the medium [197] or the addition of a "wake" around the high-momentum parton [198]. Toggling these options and comparing the results to experimental data allows indirect tests of the physical existence of these phenomena inside the QGP.

The Jet Energy-loss Tomography with a Statistically and Computationally Advanced Program Envelope (JETSCAPE) framework [199] is a recent MC generator with a modular approach, able to incorporate different pre-included as well as third-party models to build a simulation of heavy-ion collisions. Typically included in reported predictions are the Modular All-Twist Transverse and Elastic scattering-induced Radiation (MATTER) [200] and Linear Boltzmann Transport (LBT) [201] models for partonic energy loss. One of JETSCAPE's unique features is a full hydrodynamic simulation, which, though computationally expensive, is theoretically advantageous as compared to tuned heuristic approaches. Viscous hydrodynamics can be included using a model based on the Kurganov-Tadmor method [202], called MUScl for Ion Collisions (MUSIC) [203, 204], or a model written in Open Computer Language called CLVisc [205, 206]. JETSCAPE also offers an implementation of Three-dimensional Extended Reduced Thickness Event-by-event Nuclear Topology ($\mathrm{T_R}$ENTo 3D) [207] to model the initial-state geometry of nuclear collisions, and the Simulating Many Accelerated Strongly-interacting Hadrons (SMASH) [208] model of hadronic transport as an afterburner. For implementation specifics of these, see Ref. [209].

[3] JETSCAPE is also open-source, but a properly-tuned version using Bayesian inference from a wide set of experimental measurements to set its various parameters is still internal-only.

1.5 Statement of purpose

The purpose of this thesis is to present novel studies of QCD, both in vacuum and in the presence of a QGP, via jet substructure measurements in both pp and Pb–Pb collisions.

In studies of data from pp collisions, measurements of jet substructure enable stringent tests of pQCD and facilitate studies of nonperturbative effects which are not yet under satisfactory theoretical control [210]. Jet substructure observables offer both flexibility and rigor: they can be constructed to be theoretically calculable from first-principles pQCD while simultaneously maintaining sensitivity to jet radiation in regions of phase-space selected by the observable. Jet grooming algorithms, such as SD and dynamical grooming (see Sect. 1.3.3), can additionally be used to remove soft wide-angle radiation via well-controlled approaches, reducing nonperturbative effects. This defines two families of jet substructure observables: one that can be constructed from all jet constituents and one based on a subset of jet constituents which remain after grooming procedures.

In heavy-ion collisions, pp studies of jet substructure provide a baseline for studies of jet quenching (see Sect. 1.4.5), where the jet formation is modified by the hot, dense QGP (see Sect. 1.4.2 for overview). Several aspects of the QGP are unknown or not understood from first-principles QCD. For example, it is not clear how a nearly-perfect, strongly-interacting fluid arises from a gauge theory which is asymptotically free. The degrees of freedom of the medium may be quarks and gluons, or more complex correlated states, depending on the coupling strength inside the plasma and whether or not it is thermally equilibriated (as discussed in Sect. 1.4.3). Similarly, hard probes which traverse the fluid may scatter incoherently off of individual quarks and gluons inside the QGP, or could scatter coherently off of several quarks and gluons at once; the distance scale with which hard-scattered partons can resolve the fluid is referred to as the *resolution length* L_{res} in phenomenological models [211]. It is also possible that some of these scatterings could be sufficiently hard as to cause wide-angle Rutherford-like effects via elastic Molière scattering [197].

From the jet perspective, it is not obvious how the enhanced QCD emissions from medium interactions should affect the shape of observed jets, which are both broadened by the additional gluon radiation and narrowed by thermal absorption of low-p_{T} constituents [212]. To complicate matters further, it is possible that the existence of a QGP could break factorization, meaning that the framework used to calculate observables in pp collisions (see Sect. 1.2.2) is not applicable in a heavy-ion environment. To determine model validity and differentiate between these competing effects, systematic studies of jet substructure are required. Grooming techniques in particular have contributed to studies of heavy-ion collisions, in order to explore whether the QGP modifies the hard substructure of jets [118, 213–222].

Several measurements of these jet substructure observables have been performed across multiple high-energy collider experiments. Groomed jet observables have previously been studied in pp and heavy-ion collisions [75, 86, 87, 103, 112, 114, 224, 225] as well as in $\mathrm{e^+e^-}$ collisions [226]. The benefits of different jet grooming algorithms remain a topic of ongoing study, since different grooming algorithms have different perturbative structure and offer different flexibility via grooming parameters that can be adapted to specific physics goals

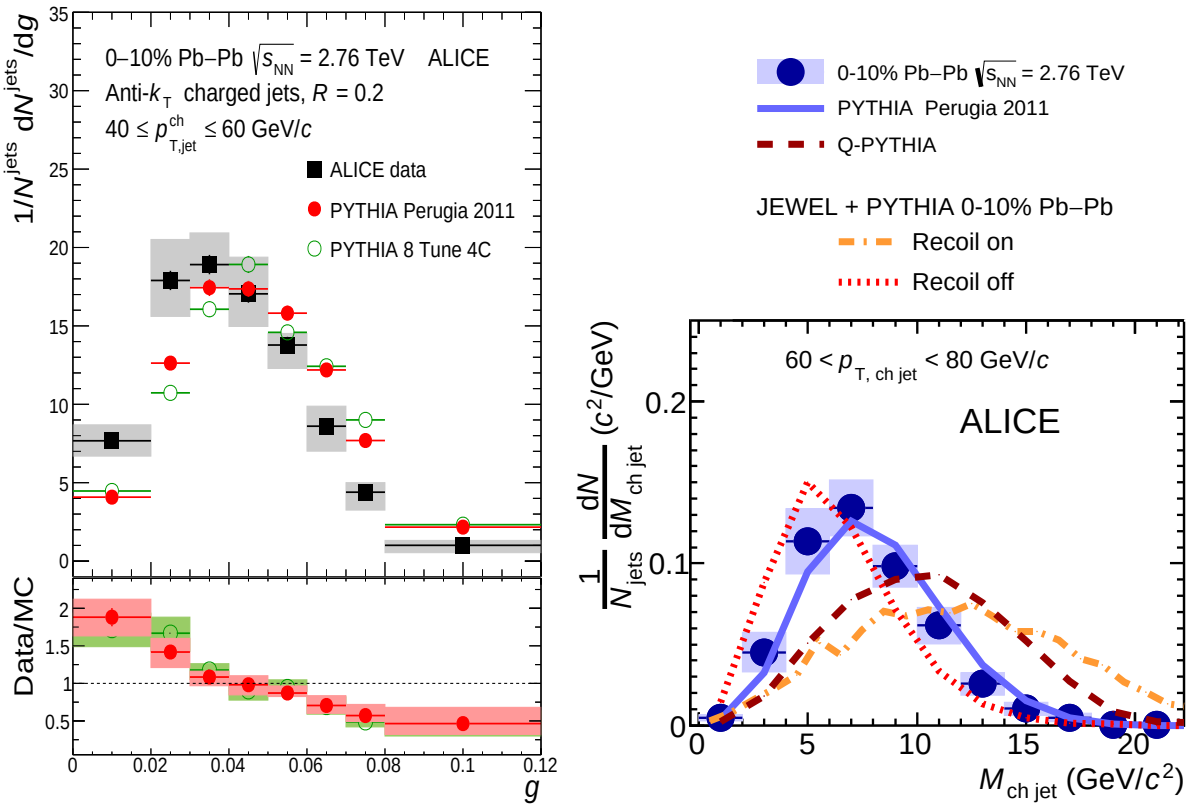


Figure 1.21: Experimental measurements resulting in the girth-mass puzzle [73, 223]. *Left:* measurement of the jet girth $g = \lambda_1 R$ in Pb–Pb data and compared to pp models, as given by PYTHIA. Significant quenching is observed in the Pb–Pb/pp ratios, with an approximate factor of 2 enhancement at small angles, with a corresponding suppression at large angles. *Right:* measurement of the jet mass m_{jet} in Pb–Pb data, theoretically related to the girth through its connection to the thrust (Eq. 1.28). No significant quenching effect is observed within uncertainties.

in either proton–proton or heavy-ion collisions (see e.g. Refs. [118, 214, 222]). This thesis presents studies using the SD and dynamical grooming algorithms, and tests the ability of pQCD calculations to describe their behavior.

Measurements are presented for a subsample of the IRC-safe jet angularities (Eq. 1.27), as well as the same jet angularities with SD grooming, in which the sum runs only over the constituents of the groomed jet. These include the jet girth [98] and the jet thrust [72], the latter of which is related to the jet mass m_{jet} by Eq. 1.28; the thrust, however, is more robust against nonperturbative effects than m_{jet} since it does not depend explicitly on the hadron masses. Additional values of α are chosen to probe the transition region between the girth and thrust, and the large-angle radiation region beyond both. Similar measurements of the inclusive and groomed jet mass, m_{jet} and $m_{\text{jet,g}}$, are undertaken using an identical

sample of jets. Finally, ALICE measurements of the groomed jet radius R_{g} and momentum splitting fraction z_{g} (see Fig. 1.10), using both SD and dynamical grooming, are used to further interpret the results of this thesis. Comparing these ALICE measurements to direct pQCD predictions explores the dynamics of splittings in the jet fragmentation.

Due to their IRC-safe nature, the jet angularities, R_{g}, and z_{g} are all calculable from pQCD. In particular, the IRC-safe jet angularities offer the possibility to systematically vary the observable definition in a way that is theoretically calculable and therefore provide a rich opportunity to study both perturbative and nonperturbative QCD [227–230]. Jet angularities were recently calculated in pp collisions both in the ungroomed [99] and groomed [231] cases, as well as for jets produced in association with a Z boson [232]. These calculations use all-order resummation of large logarithms up to next-to-leading-logarithmic (NLL′) accuracy [233]. Measurements of λ_α and $\lambda_{\alpha,\mathrm{g}}$ serve to test these analytical predictions, in particular the role of resummation effects and power corrections. Moreover, by measuring multiple values of α, one can test the predicted scaling of nonperturbative shape functions that are used to model hadronization, which depend only on a single nonperturbative parameter for all values of α [234, 235].

This thesis presents jet angularities constructed from charged-particle jets. While charged-particle jets are IRC-unsafe [236], comparisons to these theoretical predictions can nonetheless be carried out by following a nonperturbative correction procedure, as will be outlined in Ch. 4. In order to account for these nonperturbative effects including hadronization (see Sect. 1.1.2) as well as to correct for the missing neutral (uncharged) particles in the charged-particle jet measurement, two correction methods are used. First, a purely MC-based approach is used with both PYTHIA [37–39] and Herwig [42, 43], testing the agreement of direct pQCD calculations as combined with different phenomenological approaches (see Sect. 1.1.3). Secondly, a shape-function based approach is used, which is a theoretically-derived approach dependent on one nonperturbative parameter, which is expected to be universal (see Sect. 1.2.2). Different regions of the measured distributions probe these different underlying aspects of QCD, while varying the observable definition changes the region of QCD phase space being explored.

The ungroomed jet angularity λ_1 has been measured in both pp [82, 112, 223] and $\mathrm{p\overline{p}}$ [80] collisions. The ungroomed jet thrust λ_2 has also been measured in pp collisions [112]. The closely related ungroomed and groomed jet mass have been extensively measured in pp collisions [74–79, 81, 82, 84–87], and the ungroomed mass was also studied in $\mathrm{p\overline{p}}$ [80] and p–Pb [73] collisions. Many of these measurements have focused on using jet substructure for tagging objects at high p_{T}, rather than for fundamental studies of QCD, and with the exception of the jet mass there were no comparisons of jet angularities to analytical calculations prior to these results, nor were any such comparisons made for charged-particle jets.

In previous measurements carried out for the girth g and m_{jet} in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV and compared to pp models [73, 223], significant quenching modification was observed for g, while no significant modification was seen for m_{jet}; see Fig. 1.21. The origin of this discrepancy is unclear: differing values of R and $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$, which vary the relative

energy loss and nonperturbative dependence, versus varied values of α, which change sensitivity to medium recoil effects and angular broadening, could both account for the difference. This conundrum has become known as the *girth-mass puzzle*.

Using pp measurements as a baseline, new measurements of the jet angularities in Pb–Pb collisions at identical center-of-mass energy can systematically quantify in-medium substructure modifications. These angularities, compared with new measurements of m_{jet} in the same data, are presented using equivalent jet definitions for the first time to address the girth-mass puzzle. Background-subtracted charged-particle jets are measured with $40 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 150$ GeV/c, extending the kinematic reach of previous measurements and probing the strength of jet-medium interactions at varying energy scales. Theoretical comparisons are given, which provide discrimination between models and inform future quenched jet substructure studies.

Concurrently, jet substructure measurements for heavy-flavor tagged jets, such as charmed jets containing a D^0 meson, allow direct probes of how the QCD dead cone affects jet shapes (see Sect. 1.3.4). The ability to make perturbative predictions down to lower $p_{\mathrm{T}}^{\mathrm{jet}}$ enhances tests of perturbative dynamics such as the DGLAP evolution equations and provides a baseline for future studies in AA, where the dead cone may be filled in by medium-induced radiation [237]. The inclusive jet angularities of this thesis are compared with ALICE measurements of the heavy-flavor tagged jet angularities to study these effects [238].

The following chapters give details regarding the performance of these measurements and their subsequent interpretation. Chapter 2 details the experimental apparatus, while Ch. 3 enumerates the relevant data sets, procedures, and systematic checks used for the analyses. Ch. 4 explains the corrections that are applied to pQCD predictions before they are compared to data. Results are shown in Ch. 5, with discussion and closing remarks delivered in Ch. 6.

Chapter 2

Experimental apparatus

This chapter describes the experimental apparatus and the specific detector subsystems which are most relevant for the analyses presented in this thesis. Section 2.1 introduces the Large Hadron Collider (LHC), the collider on which the experimental apparatus was built. For a more detailed report on the LHC design, see Ref. [239–241]. Section 2.2 discusses A Large Ion Collider Experiment (ALICE) and some of its performance characteristics.

2.1 The Large Hadron Collider (LHC)

The LHC is the largest collider in the world, straddling the Franco-Swiss border just west of Geneva, Switzerland. Constructed between 1998 and 2008 by the *Conseil européen pour la recherche nucléaire* (CERN), or the European Organization for Nuclear Research, the collider ring is 27 km in diameter and runs though a small tunnel roughly 150 m underground. This tunnel was originally used for the Large Electron-Positron Collider (LEP), which was operated from 1989 to 2000 before being dismantled. The LHC is composed of two separate beam pipes held at ultra-high vacuum (UHV), roughly 10^{-10} to 10^{-11} mbar, where two different beams of particles (one in each beam pipe) can be accelerated in opposing directions before being made to collide with one another.

The machine works by accelerating bunches of particles through successive electromagnetic kicks along a circular path. The *cyclotron*, first invented by Ernest Lawrence at the University of California Berkeley in 1929, was the first type of accelerator to accomplish this task; the 1939 Nobel Prize in Physics was awarded to Lawrence for his discovery [242]. Cyclotrons use a radio frequency (RF) alternating current to accelerate particles along a spiral path, starting from a small radius of curvature and moving outwards. The curvature of the particle trajectory is controlled by imposition of a uniform magnetic field. The cyclotron design has been largely superseded by that of the *synchrotron*, which was first proposed in the mid-1940s [243, 244]. The synchrotron design, of which the LHC is an example, uses a fixed-radius ring with varying magnetic field to accelerate particles as they circulate. The frequency of the RF currents are also adjusted to synchronize with the arrival times of parti-

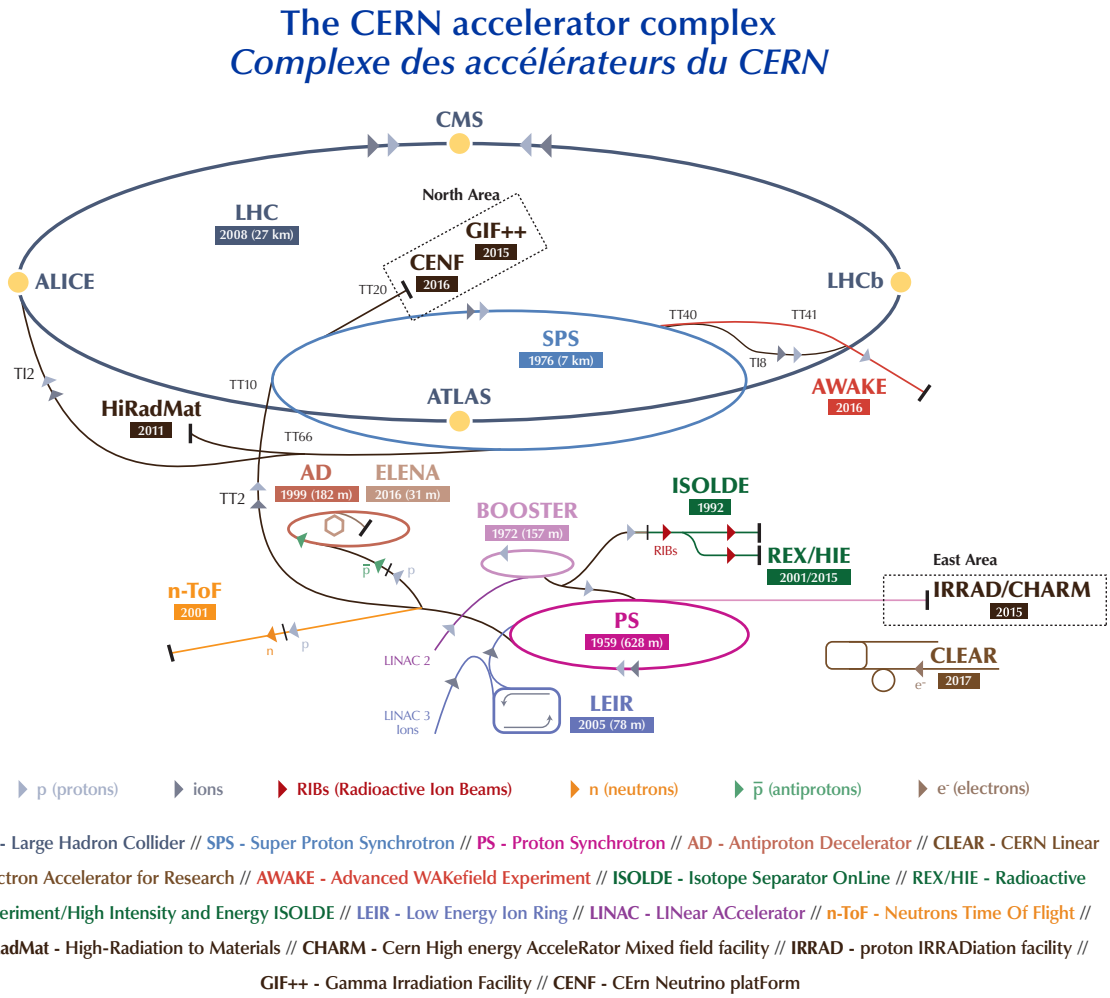


Figure 2.1: Diagram of the CERN accelerator complex [245].

cles on each successive rotation. The maximum attainable energy in the synhcrotron design is limited by the achievable magnetic field strength, particle energy loss due to *synchrotron (magnetobremsstrahlung) radiation*, and the maximum frequency which is attainable for the RF fields.

The particles which are injected into the LHC first go through a series of smaller accelerators before they can be boosted to LHC energies; see Fig. 2.1. Protons and heavy ions are first accelerated by a Linear Accelerator (LINAC) into the Low Energy Ion Ring (LEIC), which separates the long injection pulse into shorter *bunches* (with about 1.2×10^{11} particles each), and accelerates them from energies of 4.2 to 72 MeV. These bunches are then injected into the Proton Synchrotron (PS), which was constructed in 1959 as CERN's first

synchrotron. The PS originally served to accelerate 10 GeV protons for fixed-target experiments, some embedded inside of bubble chambers such as the 81 cm Saclay [246]. With several upgrades over the years, the PS now accelerates protons to 25 GeV before injecting them into the Super Proton Synchrotron (SPS), housed in a circular tunnel 2.2 km in diameter. Originally called Sp$\bar{\text{p}}$S, the SPS operated as a proton-antiproton collider from 1981 to 1991, before being repurposed as an injector for LEP [247]. SPS beams provided data for the UA1 and UA2 experiments, which discovered the W and Z electroweak bosons [248]. Protons injected into the SPS are accelerated to 450 GeV before being injected into the LHC, which can then accelerate these bunches up to 6.8 TeV for a maximum pp center-of-mass collision energy of $\sqrt{s} = 13.6$ TeV.

Particles are injected into the LHC at one of eight insertion points, where they can also be extracted, accelerated, or collided. Sixteen RF cavities (eight for each beam) are used to accelerate the particles, with each cavity delivering 2 MV at 400 MHz. The oscillatory nature of the electric field preserves the bunch structure of the beam, as particles on the front of the bunch are accelerated less than those in the back. The accelerator ring is controlled by superconducting electromagnets which are liquid-Helium cooled to 1.9 K, colder than the 2.7 K of outer space. There are 1232 main dipole magnets, each 14.3 m in length and weighing 35 metric tons, which produce magnetic field strengths of approximately 8.3 T in order to bend the particle beams around the ring. This requires a current of approximately 11 to 12 kA through the superconducting coils.

2.1.1 Luminosity

Before reaching one of four *interaction points* (IP) where experimental detectors are located, the beams are focused by a system of three quadrupole magnets, called an inner triplet. The inner triplets narrow the beams from about 200 to 16 μm, increasing the particle density to correspondingly boost the *instantaneous luminosity* $\mathcal{L}$,

$$\mathcal{L} = \frac{1}{\sigma}\frac{\mathrm{d}N}{\mathrm{d}t}, \tag{2.1}$$

where σ is the collision cross section and $\mathrm{d}N$ is the number of events per unit time $\mathrm{d}t$. Note that $\mathcal{L}$ has units of [Area]$^{-1}$ [Time]$^{-1}$, which is often quoted in cm^{-2}s^{-1}. The value of $\mathcal{L}$ primarily depends on the number of particles per bunch, the size and density of the bunches, and how much the beams overlap. The LHC can circulate up to 2808 bunches at a time, with a bunch crossing happening at each IP every 25 ns. This corresponds to 11,245 rotations in the LHC per second, at a speed of about $0.99999999c$.

In general, $\mathcal{L} = \mathcal{L}(t)$ varies as a function of time. As bunches circulate and cross at each IP, the quality of the beams degrade due to particle losses and broadening of the bunches. This corresponds to a decreasing $\mathcal{L}(t)$, which reduces the number of observable interactions. Therefore the particle beams are regularly *dumped* (ejected) after sufficient degradation, allowing new bunches to be injected from the SPS and accelerated to LHC energies. The total luminosity delivered and collected by an experiment during a period of runs is referred

to as the *integrated luminosity* $\mathcal{L}_{\text{int}}$,

$$\mathcal{L}_{\text{int}} = \int \mathcal{L}(t)\,\mathrm{d}t, \tag{2.2}$$

which has units of $[\text{Area}]^{-1}$, typically reported in inverse *barns* b^{-1} (or, equivalently, femtobarns fb^{-1}, microbarns μb^{-1}, etc.), where 1 b = 100 fm^2.

2.1.2 Pileup

Since the LHC collides bunches of particles, it is possible for several binary particle interactions to happen in each bunch crossing. This occurrence, referred to as *pileup*, presents a challenge to experimental measurements, as uncorrelated secondary interactions can form a significant background for rare processes that randomly occur in some collisions. Due to the high speed at which collisions occur, sometimes particles produced in one bunch crossing can also be misidentified as coming from another bunch crossing, which is referred to as *out-of-bunch* pileup. Despite the high number of individual particles per bunch, the spacing between the particles is still much higher than the size of the particles themselves; nonetheless, at peak LHC luminosity, the number of binary interactions per bunch crossing, called μ, can average as high as $\langle\mu\rangle \approx 60$ [249, 250].

To account for this effect, experiments opt for one of two different approaches. The first approach is to identify each individual interaction by tracing the produced particles back to their originating vertex, and/or to remove background on a track-by-track level. Various algorithms and subtraction strategies have been proposed for this, such as Particle Flow [251], Pileup Per Particle Identification (PUPPI) [252], and Pileup Mitigation with Machine Learning (PUMML) [253]. Another approach to pileup mitigation, which is the one employed by the analyses in this thesis, is to purposefully reduce $\mathcal{L}$ by slightly detuning the beams until pileup effects are negligible. While the first approach offers the ability to capture higher statistics (important for studying rare interactions), the second approach benefits from reduced systematic uncertainty on event reconstruction, which is beneficial for precision measurements of QCD.

2.2 A Large Ion Collider Experiment (ALICE)

ALICE is one of the four major (and one of eight total) experiments positioned on the LHC, located at IP2 ("Point 2") on the French side of the border near the commune of Saint-Genis-Pouilly. From its original technical proposal [254], ALICE was designed as a general-purpose heavy-ion experiment, constructed to handle the large number of charged-particle tracks ($\mathrm{d}N_{\text{ch}}/\mathrm{d}y \approx 8000$) predicted to be produced in AA collisions at LHC energies (TeV scale per-nucleon). It is the only LHC experiment explicitly designed for nuclear collisions, with its primary physics goals being study of QCD bulk matter and the QGP. The detector is concentrated at midrapidity, where the baryon density is minimized, while the energy

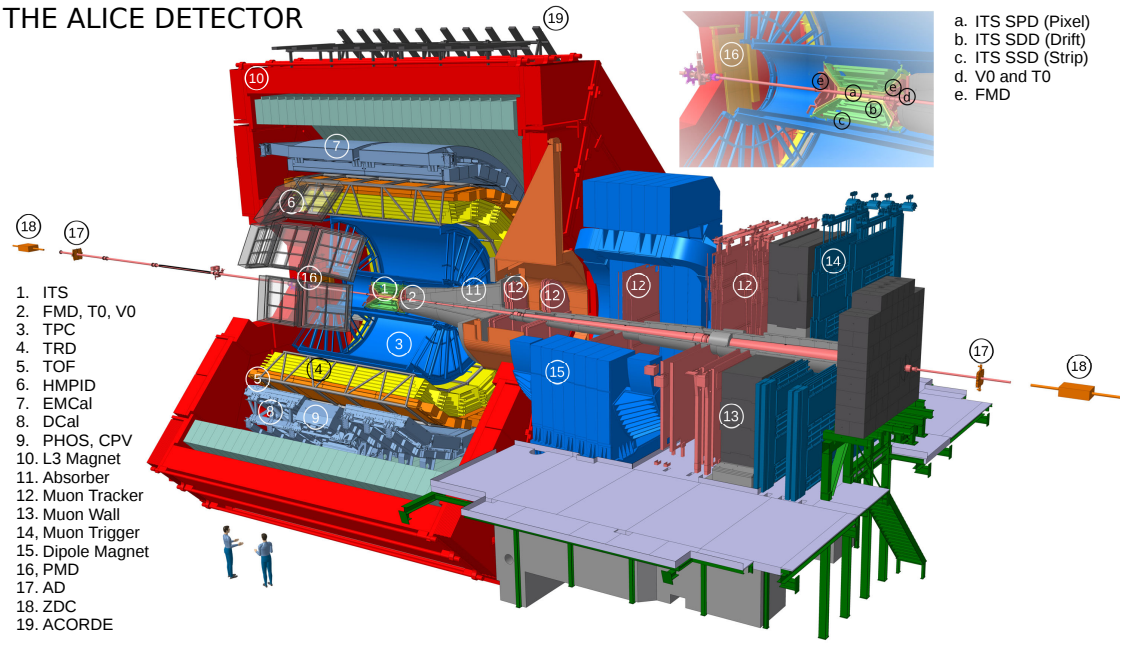


Figure 2.2: Cross-sectional diagram of ALICE at the CERN LHC during Run 2 (2015-2018) [255]. Different subdetector systems are labelled, and the inner ITS region is shown in an expanded view in the upper right. See text for details.

density is maximized. This provides complementary information to the lower-energy QCD studies provided by the SPS and RHIC. ALICE is capable of measuring events from pp and p–Pb collisions using the same apparatus.

The ALICE detector consists of three main parts:

1. a central barrel (covering the psuedorapidity range $-0.9 < \eta < 0.9$),
2. a forward muon spectrometer ($-4.0 < \eta < -2.4$), and
3. a set of multiplicity detectors in the forward regions used for triggering ($2.8 < \eta < 5.1$ and $-3.7 < \eta < -1.7$).

The central barrel is enclosed inside a large solenoid magnet, originally used by the L3 experiment at LEP, which creates a magnetic field strength of 0.5 T. The muon spectrometer contains an additional dipole magnet of strength 0.67 T to bend $\mu^{\pm}$ produced at forward rapidity, from (for example) J/ψ decays. The multiplicity detectors are arrays of scintillator counters constructed in both forward regions.

A diagram depicting the ALICE detector and its various subdetectors is given in Fig. 2.2. The following subsections overview the primary detector systems which are used for results presented in this thesis; for a full review of the ALICE detector and its performance, see Ref. [256, 257].

2.2.1 Inner Tracking System (ITS)

At the heart of the central barrel, directly surrounding the beam pipe at the nominal IP, is the ALICE Inner Tracking System (ITS) [258]. The ITS serves several purposes:

1. One of the primary goals of the ITS is to determine the location of the *primary vertex*, the point where the collision occurred, to a very high precision (within a resolution of 100 μm).

2. The ITS must similarly be able to reconstruct secondary vertices which come from the decays of heavy particles, such as hadrons containing a b or c valence quark.

3. In some cases, the ITS should be able to contribute to event selection (triggering).

4. Finally, the ITS should provide a good constraint on the p_{T} and angle of the particles which traverse it.

The ITS is a silicon-based detector composed of six cylindrical layers, ranging from radii of 4 to 43 cm and covering a pseudorapidity range of $|\eta| < 0.9$ for all vertices within ± 5.3 cm of the nominal IP along the beam direction. The two innermost ITS layers are based off of Silicon Pixel Detectors (SPD). Silicon pixels offer the best resolution for resolving high track densities near the IP, which can reach 50 tracks per cm^2 in heavy-ion events, while maintaining high speed capabilities, with a chip clock frequency of 10 MHz.

Like all silicon detectors, SPD pixels work by measuring ionization electrons / holes left by the ionizing radiation which traverse them. These charges then drift across the pixel due to an applied electric field, controlled via a *reverse-bias voltage*; the 'reversal' in this name refers to the direction of the applied voltage difference with respect to the p-n junction in the semiconductor. In particular, if a diode is reverse-biased, the cathode has a higher voltage than the anode, so that no (or little) current will flow unless free charges are introduced (e.g. by ionization in the bulk) or the diode breaks down due to a high enough voltage. As the drift charges enter the silicon circuitry, the charges create current pulses, which are then amplified. If produced above a particular threshold value (set to reduce noise), a hit is recorded by the detector. The SPD pixels are arranged in arrays of 256 by 160 cells each measuring 50 μm in the azimuthal direction ($r\phi$) and 425 μm in the longitudinal direction (z), with a thickness of 200 μm. In total, there are nearly 10 million SPD channels in the inner two layers of the ITS, covering an active area of only 0.21 m^2.

By design, silicon detectors are primarily sensitive to charged particles, which can ionize the atomic constituents of the semi-conductor bulk via electromagnetic interactions with the atomic electrons. Neutrally-charged particles require much larger radiation length to scatter, and thus typically do not interact in the silicon ITS.

The next two layers of the ITS, situated outside of the SPD layers, use Silicon Drift Detectors (SDD), manufactured from homogenous Neutron Transmutation Doped (NTD) silicon with a thickness of 300 μm. Each sensor has a sensitive area of 70.17 ($r\phi$) by 75.26 (z) mm^2 split into two drift regions by a central cathode strip carrying a HV bias of -2.4 kV,

Layer	Type	r (cm)	$\pm z$ (cm)	$r\phi$ prec. (μm)	z prec. (μm)	Area (m^2)	Channels
1	pixel	3.9	14.1	12	100	0.07	3 276 800
2	pixel	7.6	14.1	12	100	0.14	6 553 600
3	drift	15.0	22.2	35	25	0.42	43 008
4	drift	23.9	29.7	35	25	0.89	90 112
5	strip	38.0	43.1	20	830	2.20	1 148 928
6	strip	43.0	48.9	20	830	2.80	1 459 200

Table 2.1: Layer type, dimensions, precision, active area, and channels for the ALICE ITS during LHC Run 2 [256].

though performance is largely unaffected from HV degradation to -1.65 kV. These voltages correspond to drift speeds of 8.1 (5.6) μm/ns, with the signal at each anode being sampled at roughly 40 MHz. This corresponds to roughly 100k cells per detector, each read out by 512 electronic channels. The SDD are also the slowest detectors in ALICE, and hence are excluded from the *trigger clusters*, or groups of detectors that read out together, which are used in this analysis.

The exterior two ITS layers are double-layered Silicon Strip Detectors (SSD). Each of the 1698 modules are 300 μm thick and have 768 strips on each side, covering an area of 75 by 42 mm^2. These are the lowest-resolution detectors in the ITS, but cover the largest active area. A summary of the dimensional characteristics for the six ALICE ITS layers is given in Table 2.1. In total, it has 6.28 m^2 of active detector area, and creates an approximate 7.2% radiation length for perpendicular charged particles.

The SDD and SSD layers are additionally equipped with an analog readout, where the pulse peaks can be directly observed. The integral of the pulse is proportional to the amount of ionization energy deposited in the silicon, which over some known thickness of material is referred to as the *specific energy loss* $\mathrm{d}E/\mathrm{d}x$. This is quite a useful measurement, as $\mathrm{d}E/\mathrm{d}x$ can be used in conjunction with the *Bethe-Bloch formula* [6],

$$\frac{\mathrm{d}E}{\mathrm{d}x} \approx -4\pi\hbar^2\alpha_{\mathrm{EM}}^2 \frac{n\mathrm{Z}}{m_{\mathrm{e}}\beta^2}\left\{\ln\left[\frac{2(\beta\gamma c)^2 m_{\mathrm{e}}}{I_{\mathrm{e}}}\right] - \beta^2\right\}, \tag{2.3}$$

to extract the speed $\beta = v/c$ of the particle traversing a medium with atomic number Z, number density n, and effective ionization potential I_{e}. Combining measurement of β with p allows extraction of the particle mass; and, thus, the type of particle itself. This procedure is typically called *particle identification* (PID). In the nonrelativistic regime ($p \lesssim 2$ GeV/c) where mass effects are significant, $\mathrm{d}E/\mathrm{d}x$ can typically perform PID with a high degree of certainty. Example $\mathrm{d}E/\mathrm{d}x$ curves from the ITS can be seen in the left panel of Fig. 2.3.

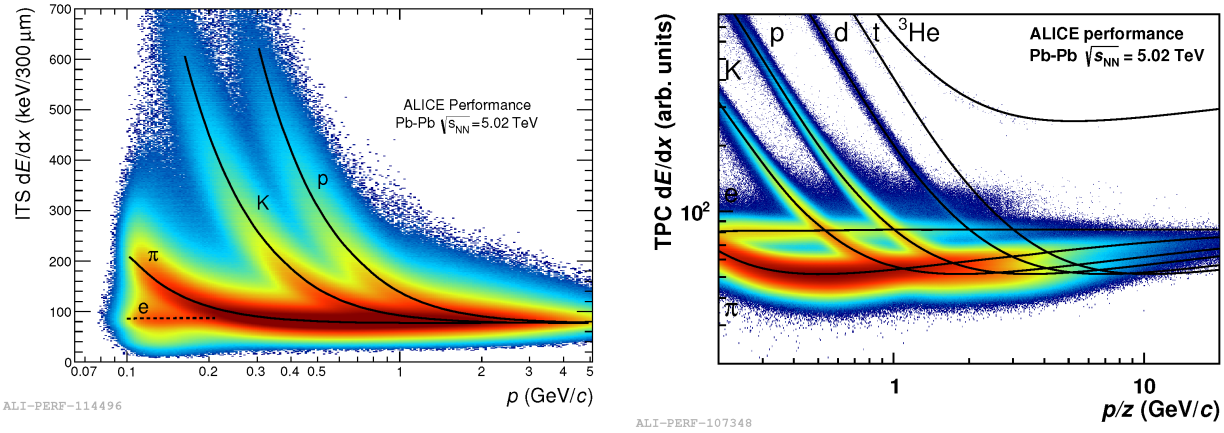


Figure 2.3: Distribution of the specific energy loss ($\mathrm{d}E/\mathrm{d}x$) in the ITS (left) and TPC (right) [257]. The solid black curves depict the per-particle parametrizations of the expected mean energy loss. The slow "relativistic rise" regime is visible at low p.

2.2.2 Time Projection Chamber (TPC)

The TPC [259] is the main tracking apparatus of ALICE, which is a unique feature among LHC experiments. Originally developed in the mid-1970s by David Nygren at Lawrence Berkeley National Laboratory (LBNL) [260], the TPC is a cylindrical container of liquid or gas which is ionized by radiation that traverses it, in a similar fashion to silicon. The charges are then drifted to the ends of the cylinder and collected, with their drift time across a known electric potential "projecting" out the longitudinal position. By measuring the transverse position of the ionization charges (x, y) and the time difference between a known event and the charge collection (t, projecting z), a three-dimensional picture can be reconstructed. Furthermore, the amount of charge collected can be correlated with the energy deposited in the detector, yielding an effective measure of $\mathrm{d}E/\mathrm{d}x$ and thus PID; see the right panel of Fig. 2.3. Since particles traversing the TPC create many space-points which each provide a measurement of $\mathrm{d}E/\mathrm{d}x$ (as opposed to the single measurement per silicon layer provided by the ITS), the PID capabilities of the ALICE detector are greatly enhanced by the $\mathrm{d}E/\mathrm{d}x$ samples from the TPC.

While a TPC could in principle have continuous measurement of tracks, the resolution of a TPC is limited by the *diffusion constant* $d = lv/3$ of its gas mixture, where l is the mean free path of a free electron inside the gas, and v is its speed. For a free charge moving in the presence of an electric field, its root mean square normal distance due to diffusion, both parallel ($\parallel$) and perpendicular ($\perp$) to the field, is given by [260]

$$\sigma_{\parallel} = \sqrt{2dt} \quad \text{and} \quad \sigma_{\perp} = \frac{\sigma_{\parallel}^2}{1 + \omega^2\tau^2}, \tag{2.4}$$

where $\omega = eB/m_{\mathrm{e}}$ is the cyclotron frequency, t is the drift time, and τ is the mean collision

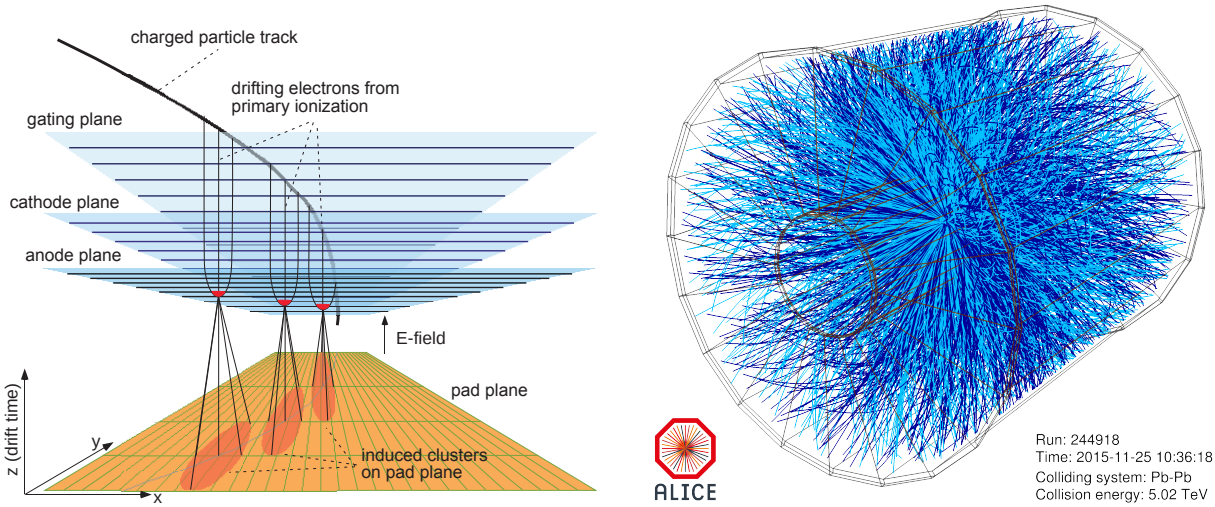


Figure 2.4: *Left:* A diagram showing charge amplification and collection for the MWPCs of the ALICE TPC, leading to a three-dimensional reconstruction of the original charged particle trajectory. *Right:* an example Pb–Pb event in the ALICE TPC [261]. 1582 positively-charged (darker tracks) and 1579 negatively-charged (lighter tracks) particles are produced; about 80 percent are π mesons.

time. The resolution can also be adversely affected by the presence of nonuniformity in the electric field, typically caused by other charges in the TPC; these are referred to as *space-charge fluctuations*.

The ALICE TPC has an inner (outer) radius of about 0.85 m (2.5 m) and an overall length along the beam direction of 5 m, yielding a volume of around 90 m^3. The detector is filled with a gaseous mixture of neon (Ne) doped with carbon dioxide (CO_2) and nitrogen (N_2), a mixture chosen for its drift speed, low diffusion, low radiation length, small space-charge effect, and ageing and stability properties [256, 262]. The outer walls of the TPC are composed of a field cage, with a central high-voltage electrode dividing it into two longitudinal halves, and creating a static electric field in its volume with high uniformity. The central electrode is a stretched, aluminized sheet of Mylar, 22 μm in thickness. The endcaps of the TPC are each covered by Multi-Wire Proportional Chambers (MWPCs) with an active area of 32.5 m^2. The cathode pad readout, combining a grid of anode wires above the pad plane, a cathode wire plane, and a gating grid, is depicted in Fig. 2.4. The TPC also has a complicated cooling system for temperature stability and a laser calibration system to precisely calibrate the position of the readout chambers and allow online monitoring of temperature and space-charge distortions.

Since particles drifting across the TPC take approximately 90 μs to reach the readout planes, the TPC is the limiting factor for the delivered luminosity. In pp collisions, where space-charge fluctuations are roughly one order of magnitude lower than in Pb–Pb, an interaction rate of 350 kHz ($\mathcal{L} = 5 \times 10^{30}$ cm^{-2} s^{-1}, or equivalently 5 μb^{-1} s^{-1}) delivers triggered

events with about $\langle\mu\rangle \approx 60$ pp interactions, of which about 30 satisfy *minimum bias* (MB) criteria (see Sect. 2.2.4). As discussed in Sect. 2.1.2, this pileup can be reduced during event reconstruction by eliminating tracks which lead to a different primary vertex. However, this leads to worsening $p_{\mathrm{T}}^{\mathrm{jet}}$ and observable resolution due to inaccuracies in the vertex tagger, space-charge effects in the TPC, and track reconstruction uncertainties.

To reduce this effect, the luminosity delivered to ALICE is decreased by detuning the beams. Luminosity determination is accomplished via the van der Meer scan (vdM) technique, as discussed in Ref. [263–265]. The low speed of the TPC readout was a major motivation for the ALICE Collaboration to upgrade the detector for upcoming data collection during LHC Runs 3 and 4 by replacing the MWPCs with endcaps based on Gas Electron Multipliers (GEMs) with an improved readout chip; for more information, see Ref. [266].

2.2.3 Time-Of-Flight (TOF) detector

Surrounding the TPC is the Transition Radiation Detector (TRD) [267] of ALICE, which is primarily used for identification of electrons with momenta above 1 GeV/c. Outside of the TRD is the TOF detector [268], which uses double-stacked Multi-gap Resistive Plate Chambers (MRPCs) to measure the arrival time for particles produced in the event. This arrival time is compared to a start time delivered by the T0 detector, consisting of two arrays of Cherenkov counters positioned on opposite sides of the IP [257]. The difference in time between these two detectors yields a measurement of particle velocity which, in combination with $\mathrm{d}E/\mathrm{d}x$ from the tracking detectors, can be used to constrain PID. The detector is designed to work in the intermediate momentum range, below about 2.5 GeV/c for $\pi^{\pm}$ and $\mathrm{K}^{\pm}$ and up to 4 GeV/c for protons, with a π/K and K/p separation better than 3σ [256].

Similar to the TPC, the electric field within the TOF must be kept high and uniform over its whole volume. Ionization charges formed within the TOF are accelerated towards the collecting electrodes and, once they gain enough energy, ionize other atoms along their path. This addition of *secondary ionization* leads to an analog signal amplification known as a *gas avalanche*. This same effect is utilized in the MWPCs of the TPC, as depicted in Fig. 2.4. Separation of the mass peaks for π, K, and p are shown in Fig. 2.5.

2.2.4 VZERO (V0) scintillator detectors

The ALICE V0 detectors [269] are two small-angle arrays of scintillators, called V0A and V0C, situated on either side of the IP. These detectors are multi-functional, as they provide minimum-bias (MB) triggers for the central barrel detectors as well as multiplicity counters to indicate centrality in Pb–Pb collisions. This is possible since the number of particles detected on the V0 arrays is monotonically dependent on the number of primary emitted particles [256]. The V0 detectors are therefore quintessential for MB, multiplicity, semi-central, and central triggers on ALICE. For details on the particular MB trigger used in this thesis, see Sect. 3.1.

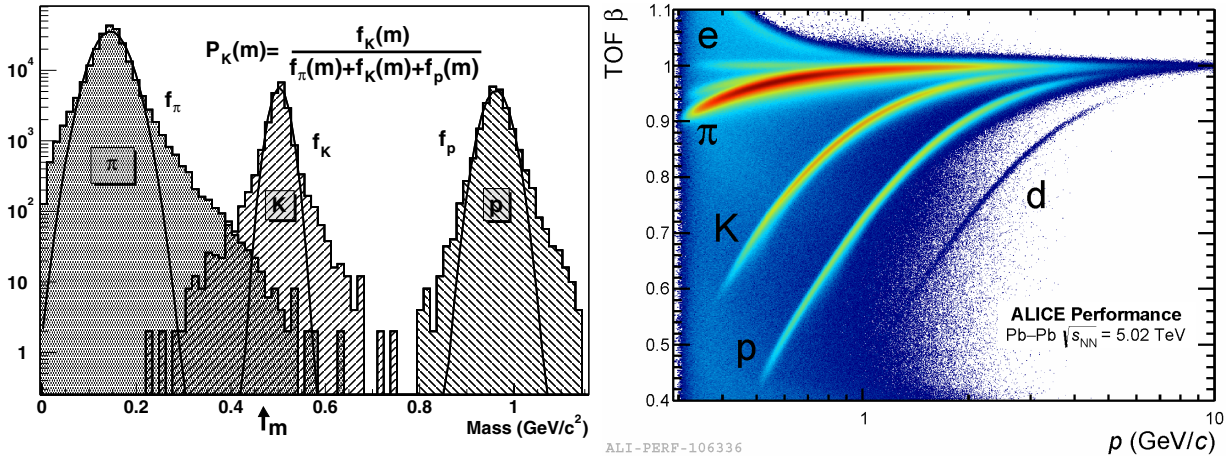


Figure 2.5: *Left:* example mass probability density functions f_i for π, K, and p hadrons, with Gaussian fits to guide the eye [268]. The probability $P_{\mathrm{K}}(m)$ of a track with measured mass m to correspond to a K meson is also shown. Tracks are required to match in the TPC and ITS, and have $0.5 \leq p \leq 2.5$ GeV/c. Note that the y-scale is logarithmic. *Right:* speed $\beta = v/c$ distribution for tracks entering the TOF detector.

In pp collisions, the efficiency for the detection of at least one charged particle detected in either side of the V0 is about 84%, including environmental effects. Rejecting events that do not have a hit in the V0 detector significantly reduces unneeded data flow due to interactions of protons with the residual gas of the vacuum chamber, or the high background rate of the forward $\mu^{\pm}$ spectrometer trigger chambers. The V0 also participates in the luminosity determination of ALICE with a good precision of about 10%; for more details, see Ref. [270].

2.2.5 Electromagnetic calorimeter (EMCal)

ALICE is equipped with an EMCal [271] to measure highly energetic photons and electrons, which deposit their entire energy inside the detector. The EMCal also collects the energy of neutral π^0 mesons though their photonic decay modes. The detector uses a *Shashlik* layout of alternating layers of lead and plastic scintillator, with wavelength-shifting fibers running longitudinally through the stack to collect scintillation light [272]. The basic modules consist of 2x2 arrays of optically-isolated *towers*, which are each read out individually.

The ALICE EMCal does not cover the entire acceptance of the central barrel tracking detectors, instead spanning only about 1.9 radians (less than one third of 2π) in the azimuthal direction across the pseudorapidity range $0.7 \leq \eta \leq 0.7$. Information from the EMCal allows improved energy resolution on the measurement of "full" jets (including neutral particles); however, the EMCal cannot offer the same spatial or momentum resolution as the ITS and TPC, which are necessary for precision measurements of jet substructure. For that reason,

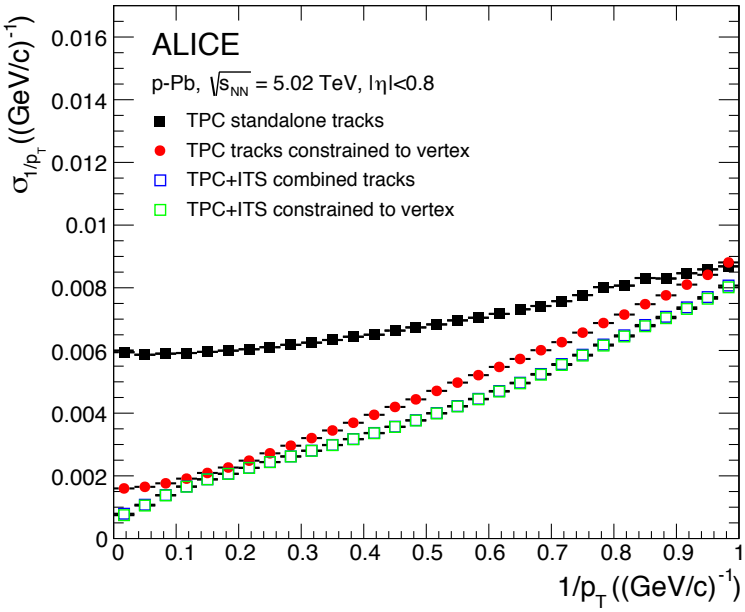


Figure 2.6: Inverse track p_{T} resolution $\sigma_{1/p_{\mathrm{T}}}$ for tracks found at various stages in the ALICE fitting procedure [257]. $\sigma_{1/p_{\mathrm{T}}}$ is related to the p_{T} resolution $\sigma_{p_{\mathrm{T}}}$ by $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = p_{\mathrm{T}}\,\sigma_{1/p_{\mathrm{T}}}$.

the EMCal is not used for the analyses in this thesis, which instead rely soley upon "charged" jets (including only central barrel tracking).

2.2.6 Central barrel track finding

After information from an event is collected, data from the various ALICE subdetector systems must be combined in order to determine what particles were observed in the event, along with their corresponding momenta, angle, rapidity, and PID, if required. A detailed description is given in Ref. [257]. Track finding in ALICE begins with a clusterization step, where the information from all included subdetectors, including positions, signal amplitudes, signal times, and their associated errors, are converted into "clusters." The clusterization is performed separately for each detector, depending on its parameters and function. The SPD is then used to preliminarily determine the primary interaction vertex, assumed to be the vertex connecting the largest number of contributing tracklets. If such a space point is not found, the algorithm searches along the beam axis for the maximum of the points of closest approach (PCA) of tracklets to the nominal beam axis.

Track finding and fitting is then performed using the Kalman filter technique [273]. This procedure starts at the outside of the TPC. Track "seeds" are first built using two TPC clusters along with the preliminary vertex point, followed by three TPC clusters with no vertex constraint. The seeds are propagated inwards, adding the nearest cluster that fulfils a proximity cut. The TPC readout chambers have 159 tangential pad rows, meaning that a

(nearly-)perpendicular track can ideally produce 159 clusters within the TPC volume. Tracks are required to have at least 20 clusters in the TPC and no less than half of the clusters that would be expected for the track trajectory (though additional quality cuts are added later; see Ref. [274] for details). The TPC track-finding efficiency is approximately 80% for tracks with $p_{\mathrm{T}} \gtrsim 0.5$ GeV/c, limited primarily by gaps between the readout chambers, with a drop-off at lower momentum caused by energy loss and multiple scattering in the detector material. It does not significantly change with detector occupancy.

Reconstructed TPC tracks are then propagated inwards to the ITS, which is used to create a tree of combined track hypotheses for each TPC track, both with and without a constraint on the primary vertex. The highest quality candidates from each tree are selected and added to the event. Due to the TPC efficiency drop-off at low-p_{T}, a standalone ITS reconstruction is performed with those clusters that were not used in the ITS–TPC tracks. Seeds are constructed using clusters from the three innermost ITS layer, and propagated to the other layers, similar as before. All of the track hypotheses are refitted by a Kalman filter and the track with the best fit χ^2 is accepted, with its clusters being removed from further searches. This procedure is repeated a few times to increase the efficiency, enabling tracking of particles with p_{T} down to approximately 80 MeV/c [257].

After calculating their PCA to the preliminary primary vertex, tracks are then refit using the Kalman filter in the outgoing direction using the previously-determined clusters. The track length integral and time-of-flight is updated at each step, and tracks using TRD and TOF information are matched to clusters in these detectors, where such calculation ends. In the final step, the tracks are refit one final time, again inwards from the outside of the detector. The position, direction, inverse curvature, and associated covariance matrix are determined for each track at this stage. The repetition of these inward-outward and outward-inward reconstructions improves the track p_{T} resolution, which is given in Fig. 2.6.

Tracks reconstructed in TPC and ITS, referred to as *global tracks*, are finally used to determine the location of the primary interaction vertex with a higher precision than initial estimation by SPD tracklets. A precise vertex fit is performed by extrapolating global tracks to their PCA to the beamline, using track weighting to suppress the contribution of any remaining outliers. The majority of tracks reconstructed in this procedure originate from the primary interaction vertex, though some tracks come from particle decays and material interactions with the detector. These can be reduced by applying cuts on the distance of closest approach (d_0) to the primary vertex.

Alternatively, tracks with $d_0 > 0.5$ mm (in pp) or 1 mm (in Pb–Pb) can be selected to reconstruct secondary vertices from photon conversions and particle decays. The reconstruction of complex heavy-flavor decays close to the interaction point, e.g. $\mathrm{D}^0 \to \mathrm{K}^-\pi^+$ and charge conjugates, is searched for by considering unlike-sign track pairs and selecting those passing a set of topological cuts [275], including the decay length in the transverse plane $L_\perp/\sigma_{L_\perp}$ and the transverse *pointing angle* $\cos(\theta_{\mathrm{pointing}})$, the angle between the flight line of the candidate (i.e., the vector connecting the primary to secondary vertex) and the direction of its reconstructed momentum. The reconstruction of D^0 mesons is given in Sect. B.2.

Chapter 3

Analysis procedure

This chapter details the data analysis procedure for the studies motivated in Ch. 1 and presented in Ch. 5. The relevant data sets are first summarized in Sect. 3.1, and a list of the relevant MC productions are enumerated. Following the track finding procedure (for which an overview was given in Sect. 2.2.6), the details of jet reconstruction are given in Sect. 1.3.1, and the jet substructure observables are then computed. These measurements are fully unfolded to account for detector effects, and for background effects in Pb–Pb; this procedure is described in Sect. 3.3. Various unfolding checks are summarized, and systematic uncertainties are then calculated in Sect. 3.4.

3.1 Data sets

The data for these analyses were collected during LHC Run 2 (2015-2018) using the ALICE detector (see Sect. 2.2). The Pb–Pb data was collected at (nucleon-nucleon) center-of-mass energy $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV, which is the highest energy achievable for Pb ions at the LHC. In order to make the pp data suitable for comparison, the selected pp data was also collected at center-of-mass energy $\sqrt{s} = 5.02$ TeV.

3.1.1 pp data selection

The pp data were collected during 2017; specifically, the data sets used are from ALICE LHC17p and LHC17q reconstruction pass 1 Analysis Object Data (AOD) files, which contain filtered event information parsed from the raw Event Summary Data (ESD) files recorded during data taking, which also include track reconstruction and vertexing information. The `FAST` trigger cluster is combined with the `CENT_woSDD` trigger cluster [276], both of which are reconstructed using the central barrel tracking ("`CENT`") but without the SDD; see Sect. 2.2.1. The run list is the TPC "good" list (with globally good tracking detectors) as defined by the ALICE Data Preparation Group (DPG):

- **LHC17p** (41 runs): 282343, 282342, 282341, 282340, 282314, 282313, 282312, 282309, 282307, 282306, 282305, 282304, 282303, 282302, 282247, 282230, 282229, 282227, 282224, 282206, 282189, 282147, 282146, 282127, 282126, 282125, 282123, 282122, 282120, 282119, 282118, 282099, 282098, 282078, 282051, 282050, 282031, 282025, 282021, 282016, 282008
- **LHC17q** (3 runs): 282367, 282366, 282365

The pp events are selected using a MB trigger (called `kINT7`, requiring a coincidence of hits in both V0 scintillator arrays; see Sect. 2.2.4) according to the following requirements:

1. During the primary vertex reconstruction, the number of vertex contributors is required to be greater than zero, i.e., there is at least one successfully-reconstructed vertex from the event.
2. The primary vertex position relative to the nominal IP is required to be within -10 cm $< z_{\mathrm{vtx}} < 10$ cm (along the beam direction), and within 1 cm in the transverse plane.

Beam-induced background events are removed using two neutron Zero-Degree Calorimeters (ZDCs) [277] located at ± 112.5 m along the beam axis from the center of the detector. Events where there are multiple reconstructed primary vertex candidates are rejected, and track quality selection criteria ensure that tracks used in the analysis are from only one vertex. Out-of-bunch pileup is rejected using V0 and SPD cuts, as implemented in `AddTaskPhysicsSelection(isMC, kTRUE)` [278]. Events are acquired at instantaneous luminosities between approximately 10^{30}-10^{31} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$, corresponding to a low level of pileup with approximately $0.004 < \mu < 0.03$ events per bunch crossing. The pp data sample contains 870 million events and corresponds to an integrated luminosity of $\mathcal{L}_{\mathrm{int}} = 18.0(4)$ nb^{-1} [264].

From the selected events, "hybrid" tracks [274] are selected for the analyses, in which two classes of tracks are included:

- Global tracks (see Sect. 2.2.6), requiring $d_0 < 2.4$ cm for the transverse direction and 3.2 cm for the longitudinal direction, and including at least one hit in an SPD layer.
- *Complementary tracks*, which have the same requirements as good global tracks, but do not require an SPD hit. Instead, the track is refit with a constraint to the primary vertex, which improves its p_{T} resolution.

The inclusive jet analysis also uses a PYTHIA 8 MC production (Monash 2013 tune) [37, 38] with a full GEANT3 [40] ALICE detector simulation, which simulates how produced particles are transported through the ALICE detector, undergo material interactions, and are picked up by the various detector subsystems. In order to create a MC data sample with adequate statistics at higher values of $p_{\mathrm{T}}^{\mathrm{jet}}$ – where the rapidly falling differential jet cross section makes events extremely rare – the production is generated in 20 $p_{\mathrm{T,hard}}$ (or $\hat{p}_{\mathrm{T}}$) bins,

which correspond to the invariant p_{T} of the simulated $2 \to 2$ interactions [279]. Each bin is populated with approximately 6M events, with bin edges

$$\begin{aligned}&[5,\ 7,\ 9,\ 12,\ 16,\ 21,\ 28,\ 36,\ 45,\ 57,\ 70,\ 85,\\ &\ 99,\ 115,\ 132,\ 150,\ 169,\ 190,\ 212,\ 235,\ \infty]\ \mathrm{GeV}/c.\end{aligned} \tag{3.1}$$

This MC sample, generated in 2018 and named LHC18b8, is anchored run-by-run to LHC17pq runs, and so the good runlist is the same as those in data:

- **LHC18b8** (44 runs): 282343, 282342, 282341, 282340, 282314, 282313, 282312, 282309, 282307, 282306, 282305, 282304, 282303, 282302, 282247, 282230, 282229, 282227, 282224, 282206, 282189, 282147, 282146, 282127, 282126, 282125, 282123, 282122, 282120, 282119, 282118, 282099, 282098, 282078, 282051, 282050, 282031, 282025, 282021, 282016, 282008, 282367, 282366, 282365

The data sets used for the D^0-tagged jet analysis are given in Sect. B.1. For general jet Quality Assurance (QA) checks for these pp run periods and for the associated MC, see Ref. [280, 281].

3.1.2 Pb–Pb data selection

The measured Pb–Pb data set for this analysis was collected during 2018, specifically LHC18q and LHC18r pass 3 AODs. The run list is the central barrel tracking "good" run list, defined by the DPG:

- **LHC18q** (136 runs): 296623, 296622, 296621, 296619, 296618, 296616, 296615, 296594, 296553, 296552, 296551, 296550, 296548, 296547, 296516, 296512, 296511, 296510, 296509, 296472, 296433, 296424, 296423, 296420, 296419, 296415, 296414, 296383, 296381, 296380, 296379, 296378, 296377, 296376, 296375, 296312, 296309, 296307, 296304, 296303, 296280, 296279, 296273, 296270, 296269, 296247, 296246, 296244, 296243, 296242, 296241, 296240, 296198, 296197, 296196, 296195, 296194, 296192, 296191, 296143, 296142, 296135, 296134, 296133, 296132, 296123, 296074, 296068, 296066, 296065, 296063, 296062, 296060, 296016, 295947, 295945, 295943, 295942, 295941, 295937, 295936, 295913, 295910, 295909, 295908, 295881, 295861, 295860, 295859, 295856, 295855, 295854, 295853, 295831, 295829, 295826, 295825, 295822, 295819, 295818, 295816, 295791, 295788, 295786, 295763, 295762, 295759, 295758, 295755, 295754, 295725, 295723, 295721, 295719, 295718, 295717, 295714, 295712, 295677, 295676, 295675, 295673, 295671, 295668, 295667, 295666, 295665, 295615, 295612, 295611, 295610, 295589, 295588, 295587, 295586, 295585

- **LHC18r** (91 runs): 297595, 297590, 297588, 297558, 297544, 297542, 297541, 297540, 297537, 297512, 297483, 297479, 297452, 297451, 297450, 297446, 297442, 297441, 297415, 297414, 297413, 297406, 297405, 297380, 297379, 297372, 297367, 297366, 297363, 297336, 297335, 297333, 297332, 297317, 297315, 297312, 297311, 297310,

297278, 297222, 297221, 297218, 297196, 297195, 297193, 297133, 297132, 297129, 297128, 297124, 297123, 297119, 297118, 297117, 297085, 297035, 297031, 296966, 296941, 296938, 296935, 296934, 296932, 296931, 296930, 296903, 296900, 296899, 296894, 296852, 296851, 296850, 296848, 296839, 296838, 296836, 296835, 296799, 296794, 296793, 296790, 296787, 296786, 296785, 296784, 296781, 296752, 296694, 296693, 296691, 296690

The event selection is similar to pp, using a MB (`kINT7`) trigger, asserting the same vertexing requirements, and rejecting out-of-bunch pileup. The Pb–Pb analyses in this theses are performed only for the 0-10% centrality region, for which the total number of accepted Pb–Pb events is approximately 91M. The same hybrid track selection criteria are used as in pp.

The analysis also uses a PYTHIA 8 MC production (Monash 2013 tune) with the full GEANT3 detector simulation (LHC20g4). The production is similar to the pp case, but the detector simulation was modified corresponding to the different run period. The production consists of the same 20 $p_{\mathrm{T,hard}}$ bins as in pp, each populated with approximately 8M events. The MC is anchored to selected LHC18qr runs, with run numbers:

- **LHC20g4** (16 runs): 297588, 295612, 295788, 295819, 296191, 296244, 296380, 296415, 296550, 296690, 296935, 297132, 297317, 297379, 297479, 297544

For general jet QA of the Pb–Pb run period and associated MC, see Ref. [281, 282]. (Note that some of the QA/analysis plots shown are from LHC18qr pass 1 and the accompanying MC LHC19f4, but all were verified to be insignificantly different from those from LHC18qr and LHC20g4.)

Since a pp model is used to generate the MC for the Pb–Pb analyses, additional corrections are required. First, the tracking efficiency of ALICE is slightly degraded in Pb–Pb compared to pp [283]. The centrality-dependence of the tracking efficiency is determined by comparing its magnitude in central versus peripheral collisions, as generated by HIJING detector simulations [284, 285]. Since the tracking efficiency is observed to be about 2% worse in central Pb–Pb data, 2% of tracks coming from the Pb–Pb MC, after the GEANT3 detector simulation, are randomly rejected before the events are analyzed.

Secondly, detector-level MC events must be embedded into raw data to simulate an uncorrelated (but QCD-driven) background from the creation of the QGP. The treatment of this embedding is described in Sect. 3.2.3 below.

3.2 Jet reconstruction

3.2.1 Background subtraction

Before jets are reconstructed in Pb–Pb collisions, background is subtracted for each event. For the data collected in pp collisions, the background is not subtracted from the data, as

the contribution from pileup is small (see Sect. 3.1.1 above). In order to make consistent comparisons between the data and the theoretical calculations, underlying event effects such as MPI (see Sect. 1.2.4), along with other nonperturbative effects including hadronization, are included in model corrections, as described in Ch. 4.

For Pb–Pb data, the uncorrelated thermal background from the QGP causes a much more significant bias than the underlying event in pp [286]. To account for this, an event-by-event constituent subtraction (CS) [287] is applied, which is able to correct both the total $p_{\mathrm{T}}^{\mathrm{jet}}$ and substructure simultaneously. In this approach, the event background density ρ is first estimated using the FastJet grid-median method [288], with grid size = 1.0. This is accomplished by projecting the event onto a y-ϕ grid, with the scalar p_{T} sum of the particles then being computed for each "cell." The value of ρ is obtained by taking the median of $\sum_{i\in\mathrm{cell}} p_{\mathrm{T},i}/A_{\mathrm{cell}}$ over all cells, where A_{cell} is the cell area.

Using this value, the CS algorithm deploys "ghost" particles uniformly in the y-ϕ plane, with each ghost carrying transverse momentum $p_{\mathrm{T},g} = \rho A_g$. The parameter A_g represents the "area" of the ghost, and its value is chosen by the user, with the nominal suggested value being $A_g = 0.01$ [287]. Each particle i is then paired with each ghost k, and their weighted distance $D_{i,k} = p_{\mathrm{T},i}^{\alpha}\Delta R_{i,k}$ is calculated, where $\Delta R_{i,k}$ is the same y-ϕ angular distance as in Eq. 1.27. The free parameter α has a nominal suggested value of $\alpha = 0$ (i.e., no p_{T}-dependence). Then, starting from the lowest $D_{i,k}$, the hardest track of i and k is reassigned $p_{\mathrm{T}} = |p_{\mathrm{T},i} - p_{\mathrm{T},k}|$, and the softer track is removed from the event. This subtraction procedure continues until the sorted list $D_{i,k}$ is exhausted or $\Delta R_{i,k} > R_{\mathrm{max}}$, another free parameter. The effect of this CS procedure is depicted in Fig. 3.1.

For this analysis, the nominal values of $A_g = 0.01$ and $\alpha = 0$ are used. Of the three user-defined parameters, R_{max} has the strongest affect on the subtraction, with $R_{\mathrm{max}} = 0.1$ being the best parameter to center the $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ resolution for $R = 0.2$ jets (see right panel of Fig. 3.2). Variations of this parameter are also used for calculation of a background subtraction systematic uncertainty; see Sect 3.4.4 for details.

3.2.2 Jet selection

Jets are reconstructed from charged-particle tracks with the FastJet 3.3.2 software package [68] using the anti-k_{T} algorithm [66] (see Sect. 1.3.1) with the E-recombination scheme.[1] in the The value of $p_{\mathrm{T,track}}$ is directly obtained from the track finding, as described in Sect. 2.2.6. All tracks are reconstructed assuming that they are charged π mesons, using $m_{\pi^{\pm}} = 0.1396$ GeV/c^2. (A variation on this track mass assumption is taken as a systematic uncertainty for λ_α – see Sect. 3.4.). In pp data, both the radii $R = 0.2$ and 0.4 are used; in Pb–Pb, where there is a large uncorrelated background from particles not associated with a single hard scattering, only the radius $R = 0.2$ is explored, as the background is enhanced for larger R.

[1] These charged-particle (track) jets, which do not contain information from neutrally-charged particles, are sometimes abbreviated "ch. jets" in the text; e.g., these jets carry transverse momentum $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$.

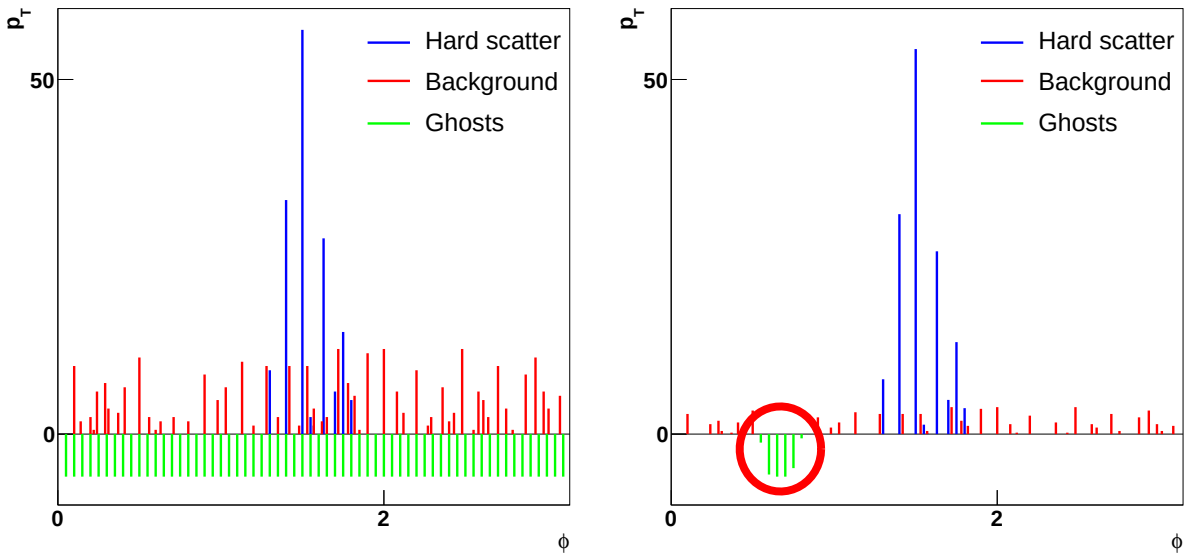


Figure 3.1: Diagram of the CS method [287]. *Left:* an initial hard-scattering produces a jet (blue) inside an uncorrelated background (red). Ghost particles (green) are deployed uniformly in the event. *Right:* after CS, the p_{T} of most tracks is reduced. A small number of ghost particles also remain (circled), which are now included in the event.

Although the reported transverse momentum $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ range is under 100 GeV/c in pp, all reconstructed charged-particle jets in the range $5 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 200$ GeV/c are collected, in order to maximize statistics in the unfolding procedure. In Pb–Pb, jets have a higher $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ cut in order to eliminate purely combinatorial ("fake") jets from the thermal background, which contaminate the unfolding; this is described below in Sect. 3.3. Each jet axis is required to be within the fiducial volume of the TPC, $|\eta_{\mathrm{jet}}| < 0.9 - R$. Jets containing a track with $p_{\mathrm{T}} > 100$ GeV/c are removed from the collected data sample, due to limited momentum resolution. In the reported $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ range, this effect is negligible (see Sect. 3.2.3 below).

3.2.3 Jet matching

In order to account for detector inefficiencies and to understand the jet reconstruction performance, a sample of jets is produced using PYTHIA 8 [38] and run through a full simulation of the ALICE detector using GEANT3 [40]; details of the event generation are listed in Sect. 3.1. Jets are reconstructed using the original PYTHIA 8 event sample before the particles undergo interactions with the detector ("truth" level), as well as at after the particles are reconstructed by the ALICE GEANT3 detector simulation ("detector" level).

For the MC used in pp collisions, truth-level jets are then matched the detector-level jets using geometrical criteria. The closest truth-level to detector-level jet pairs are found, and the matched jets are required to be within an angular distance of $\Delta R < 0.6R$ from one another. Furthermore, the match is required to be unique; if more than one suitable match

$p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ bin (GeV/c)	0 - 10	10 - 20	20 - 30	30 - 40	40 - 50	50 - 60	60 - 70
LHC17pq (pp) (%)	2.4	0.4	0.1	0.04	0.02	0.01	0.005
LHC18qr (Pb–Pb) (%)	16	5.7	1.6	0.8	0.6	0.4	0.4

Table 3.1: Estimated $R = 0.2$ jet matching inefficiency, requiring a unique match between detector and truth levels, as calculated in PYTHIA 8 + GEANT3 for the pp and Pb–Pb data samples. In Pb–Pb, this percentage is calculated by the number of matches which satisfy all selection cuts divided by those which have at least one detector-level match (to eliminate combinatorial jets); note that no Pb–Pb data is used below $p_{\mathrm{T}}^{\mathrm{ch\ jet}} = 40$ GeV/c.

is found, the jets are removed from the event sample. The percentage of jets having a unique match is $p_{\mathrm{T}}^{\mathrm{jet}}$-dependent, and is given in Table 3.1. The minimum transverse momentum cut $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}} \geq 5$ GeV/c is also enforced at this level.

In Pb–Pb collisions, the background subtraction procedure adds additional smearing which must be corrected. To quantify this effect, the detector-level MC events are embedded into real events randomly chosen from 0-10% centrality Pb–Pb raw data. The embedding is performed at the level of reconstructed tracks, with the PYTHIA particle list appended by tracks from raw data. Three collections of jets are then constructed: pp truth jets (using PYTHIA truth information), pp detector jets (using PYTHIA with detector simulation), and "combined" jets using the embedded sample, consisting of a jet constructed from MC detector-level tracks and real Pb–Pb tracks. Combined-level events also undergo the CS procedure before jet reconstruction, as done with the raw data. A truth-level jet from PYTHIA must then be assigned with an associated combined jet. To define such a match, the same geometrical matching criteria are used to connect a combined jet geometrically to its nearest pp detector jet, while additionally requiring that the combined jet contains at least 50% of the tracks of the pp detector jet, as measured by p_{T} (implicitly enforcing uniqueness). The pp detector-level jet is matched to its corresponding pp truth jet with the same criteria.

The large uncorrelated background in Pb–Pb collisions creates a significant smearing effect which is larger than that of the regular pp detector smearing. As an example of this matching procedure, Fig. 3.2 shows distributions of all possible combined- to detector-level jet matchings along with the effective smearing on $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ from the matching and background subtraction procedure.

In order to verify general quality of the generated, selected, and matched jets, several checks are performed. Figure 3.3 shows these quality checks for the Pb–Pb charged-particle jets used in these analyses. Distributions are evaluated for the *jet energy scale* (JES),

$$\mathrm{JES} = \frac{p_{\mathrm{T,det}}^{\mathrm{ch\ jet}} - p_{\mathrm{T,truth}}^{\mathrm{ch\ jet}}}{p_{\mathrm{T,truth}}^{\mathrm{ch\ jet}}}, \tag{3.2}$$

which describes the shift in $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ between detector- (combined-) and truth-level jets for pp

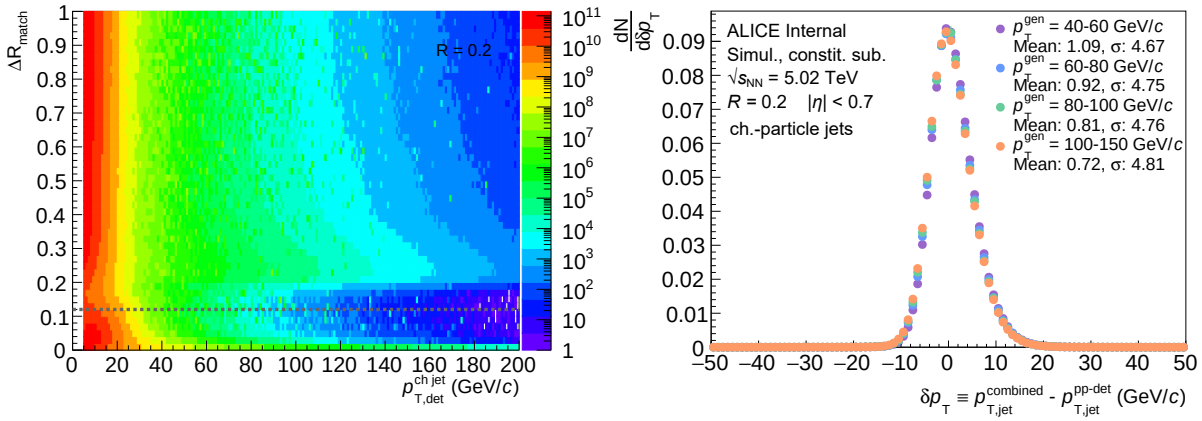


Figure 3.2: *Left:* all possible matchings between combined- and detector-level $R = 0.2$ jets in Pb–Pb MC. A dashed line is plotted at the matching condition of $\Delta R < 0.6R$. *Right:* $p_{\mathrm{T}}^{\text{ch jet}}$ smearing between combined- and detector-level $R = 0.2$ jets, using CS with $R_{\max} = 0.1$. [286]

(Pb–Pb) collisions. In particular, tracking inefficiency can cause the JES to be negative, while "fake" tracks (from detector noise or combinatorial clusters) or detector smearing of real tracks can cause the JES to slightly increase, as anti-k_{T} jets are reconstructed according the smeared detector-level track distributions. The former tracking inefficiency effects tend to be more substantial, and the latter effects becomes less significant at larger values of $p_{\mathrm{T}}^{\text{ch jet}}$. Jet-by-jet JES distributions therefore typically have a large negative tail, with mean values shifting to negative values at larger values of $p_{\mathrm{T}}^{\text{ch jet}}$. This is enhanced for Pb–Pb events, where the CS approach can target both real and uncorrelated jet constituents.

The *jet energy resolution* (JER) is also evaluated,

$$\mathrm{JER} = \frac{\sigma(p_{\mathrm{T,truth}}^{\text{ch jet}})}{p_{\mathrm{T,truth}}^{\text{ch jet}}}, \tag{3.3}$$

which describes the detector smearing on the reconstructed $p_{\mathrm{T}}^{\text{ch jet}}$ due to material interactions and multiple scattering within the detector volume. Finally, the *jet reconstruction efficiency* is evaluated, which is simply the number of jets reconstructed at detector level corresponding to a truth-level jet, versus the actual number of truth-level jets. In Pb–Pb collisions, the thermal background and CS procedure greatly reduce this efficiency at low $p_{\mathrm{T}}^{\text{ch jet}}$.

3.3 Unfolding

As alluded to in the section above, the reconstructed jets are not identical representations of the jets produced in QCD. The experimental measurement of jets and their substructure is affected by several detector effects:

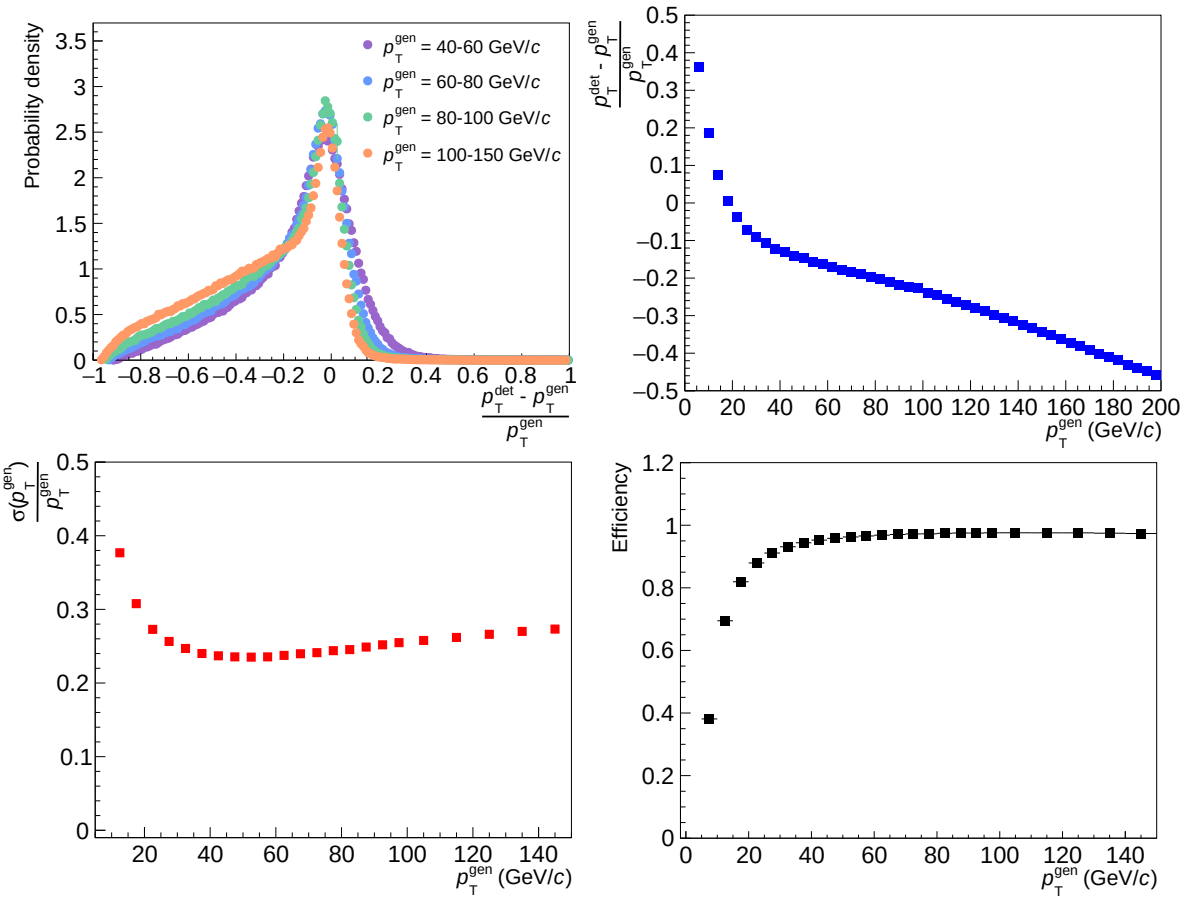


Figure 3.3: PYTHIA 8 + GEANT3 + embedding + background subtraction characterizations of $R = 0.2$ charged-jet reconstruction in Pb–Pb collisions, using CS with $R_{\max} = 0.1$ [286]. All distributions are shown per $p_{\mathrm{T,truth}}^{\mathrm{ch\ jet}}$ interval (labelled $p_{\mathrm{T}}^{\mathrm{gen}}$). The value of $p_{\mathrm{T}}^{\mathrm{det}}$ corresponds to combined, background-subtracted, and matched jets. See text for details. *Upper left:* jet-by-jet JES. *Upper right:* mean JES shift. *Lower left:* JER. *Lower right:* jet reconstruction efficiency.

- **Tracking inefficiency:** as mentioned in Sect. 2.2.6, the ALICE detector cannot capture all physical particles with perfect accuracy. Some number of tracks are lost in the track-finding procedure, traverse through gaps in the detector active area, or do not deposit enough energy to be measured.

- **Particle-material interactions:** as particles traverse the detector, they will interact with it, and lose some energy in the process. Sometimes the particle-detector scatterings can be significant, and must be corrected.

- **Track p_{T} resolution**: apart from multiple scattering effects, the ALICE detector has

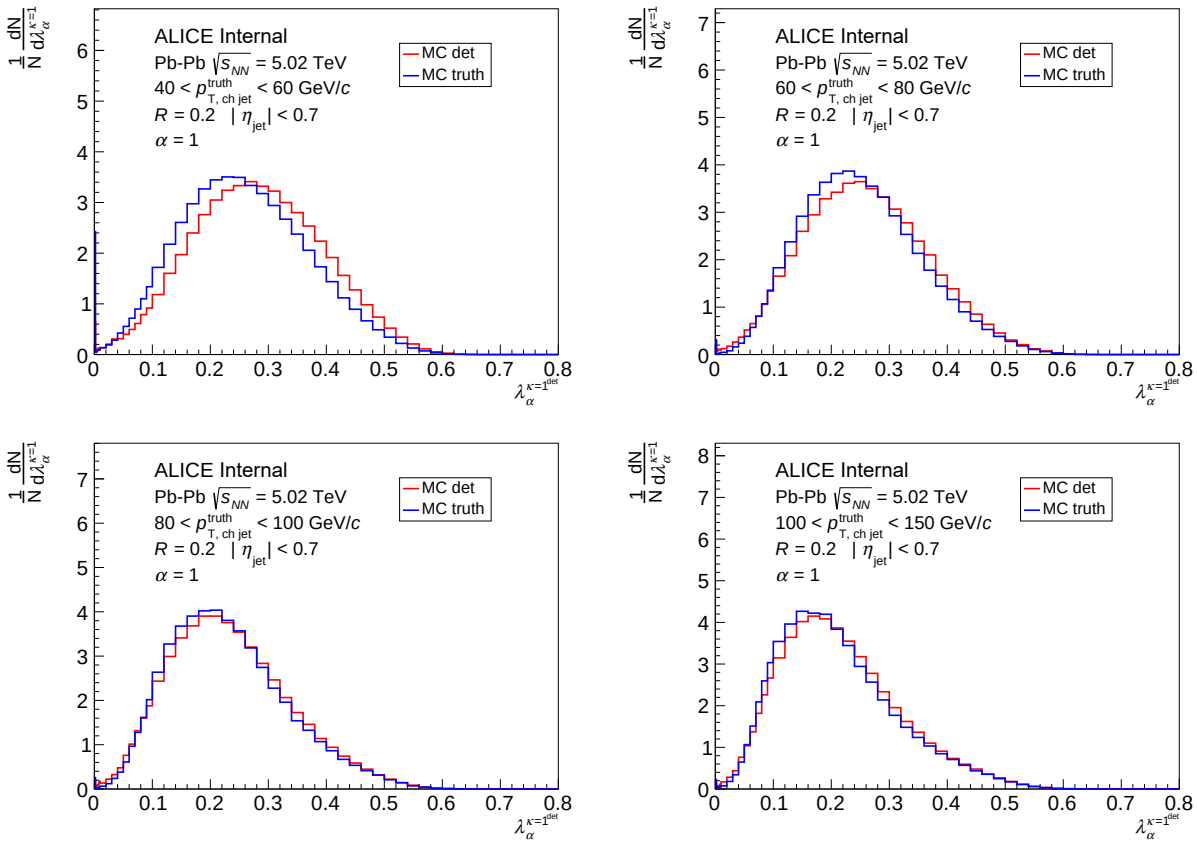


Figure 3.4: PYTHIA 8 distributions of λ_1 for inclusive $R = 0.2$ jets in Pb–Pb collisions, in $p_{\text{T}}^{\text{ch jet}}$ bins ranging from 40 to 150 GeV/c [286]. The distributions are shown both at truth level (in blue, before the particles undergo interactions with the detector) and at detector level (in red, after events are reconstructed by the ALICE GEANT3 detector simulation, embedded in Pb–Pb data, and the CS procedure is applied).

a finite resolution defined by the size of the individual silicon detectors and MWPC readout for the TPC. This resolution causes a smearing on the reconstructed track momentum; see Fig. 2.6.

These effects are typically $p_{\text{T}}^{\text{ch jet}}$ dependent, with larger corrections being needed at low-$p_{\text{T}}^{\text{ch jet}}$. Example distributions of λ_α showing the magnitude of these detector effects is given in Fig. 3.4

To account for these detector effects, a *response matrix* (RM) is constructed to describe the detector response for a particular observable. A four-dimensional RM is constructed for each observable O, which depends simultaneously in $p_{\text{T}}^{\text{ch jet}}$ and O,

$$\text{R}\left(p_{\text{T,det}}^{\text{ch jet}}, p_{\text{T,truth}}^{\text{ch jet}}, O_{\text{det}}, O_{\text{truth}}\right), \tag{3.4}$$

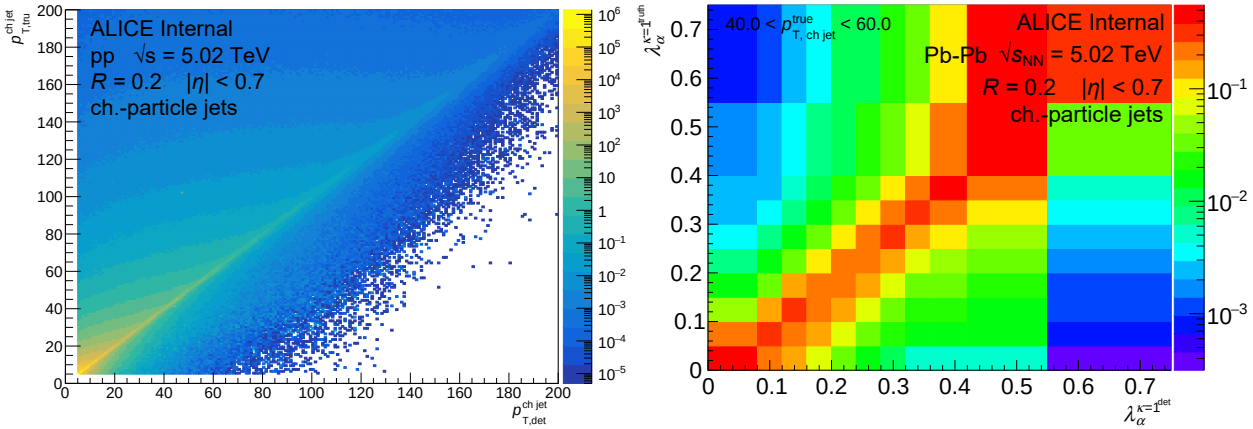


Figure 3.5: Example detector/truth matching. *Left:* the $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ matching is shown for pp truth- and detector-level jets, as calculated via the anchored MC simulation [289]. *Right:* a projection of the Pb–Pb RM onto the truth- and detector-level (combined and subtracted) observable axes for λ_1, binned according to the measured data [286]. In both plots, a large diagonal component is observed, with smearing due to detector effects.

which is filled using the anchored MC productions described in Sect. 3.1. Some projections of an example RM are shown in Fig. 3.5. In principle, the RM could be inverted and applied to the detector-level distribution in order to obtain truth-level information. However, due to the finite statistics of the MC approach, the RM is not exact, and inverting it causes significant statistical fluctuations which make the resulting distributions unreasonable.

To properly account for MC imprecision, an *unfolding* technique is used in conjunction with the constructed RM. Several unfolding algorithms exist; for these analyses, the iterative Bayesian approach [290] is used, which uses Bayes' theorem to iteratively apply corrections in a multidimensional way, as implemented in the RooUnfold [291] software package. Every unfolding method injects a small amount of bias into the result in the form of *regularization*, ensuring that the resulting distributions do not have unreasonable fluctuations. In the Bayesian method, the regularization is controlled by the number of iterations n_{iter} through the algorithm. If n_{iter} is too small, the algorithm will not have converged, and small changes in n_{iter} will result in substantial corrections to the observable, an undesirable bias on the choice n_{iter}. If n_{iter} is instead too large, statistical fluctuations will begin to dominate.

The ideal n_{iter} is therefore the smallest number such that the solution has sufficiently converged (i.e. the ratio of the unfolded distributions $\frac{O(n_{\mathrm{iter}})}{O(n_{\mathrm{iter}}-1)}$ is within a few percent), but where the statistical uncertainties have not grown too large. Therefore, for each observable in these analyses, n_{iter} is chosen such that sufficient *closure* (described in Sect. 3.3.1 below) is observed, while ensuring that the combined statistical and systematic (see Sect. 3.4) uncertainties are reasonable.

In order for the unfolding procedure to be stable, it is necessary for the data to be binned such that statistical variations are minimized. Due to the wide variation in shapes for

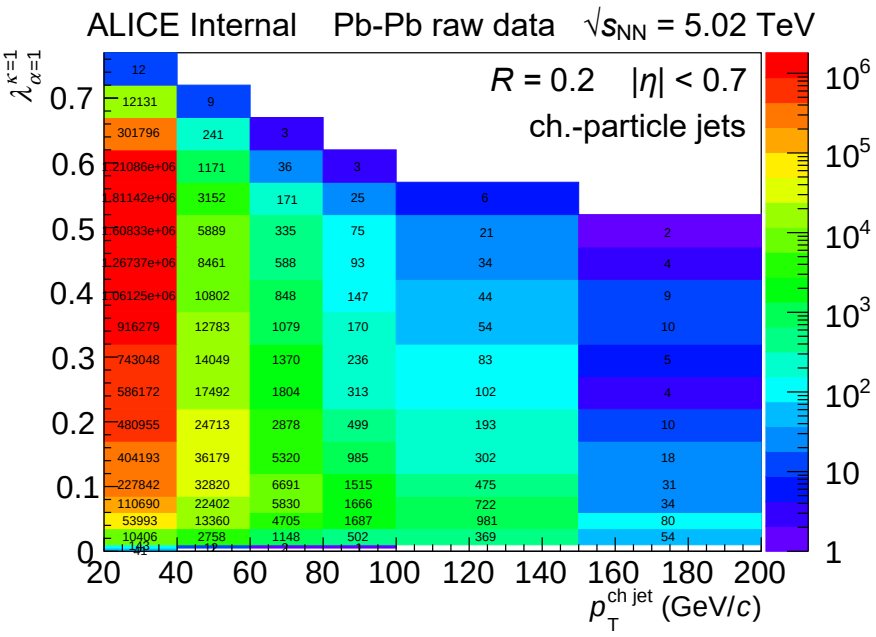


Figure 3.6: Raw spectrum of λ_1 versus $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}}$ for $R = 0.2$ charged-particle jets in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV [286]. The number of counts in each bin is displayed. To make the unfolding procedure stable, the binning is optimized to limit the number of bins having $N < 10$ counts.

different $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ ranges of the λ_α and m_{jet} distributions, a different binning and thus separate unfolding is carried out for each $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ bin in addition to each R and α setting. An example of the raw data binned to the appropriate setting for $40 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 60$ GeV/c is given for λ_α in Fig. 3.6.

While the iterative Bayesian approach is remarkably resilient to the shape of the input distributions [290], the detector-level distributions from PYTHIA 8 + GEANT3 must still reasonably approximate the raw distributions from data in order to reduce shape bias being introduced from the assumed MC *prior distribution* which begins the iterative Bayesian unfolding procedure [290]. An example of this check is given in Fig. 3.7.

3.3.1 Unfolding tests

In order to verify that the unfolding procedure is convergent and reasonable, five unfolding validation tests are performed:

- **Convergence test**: the unfolding process is required reasonably converge as the number of iterations increases. The deviations in the distribution of $\frac{O(n_{\mathrm{iter}})}{O(n_{\mathrm{iter}}-1)}$ should tend towards unity as n_{iter} is increased.

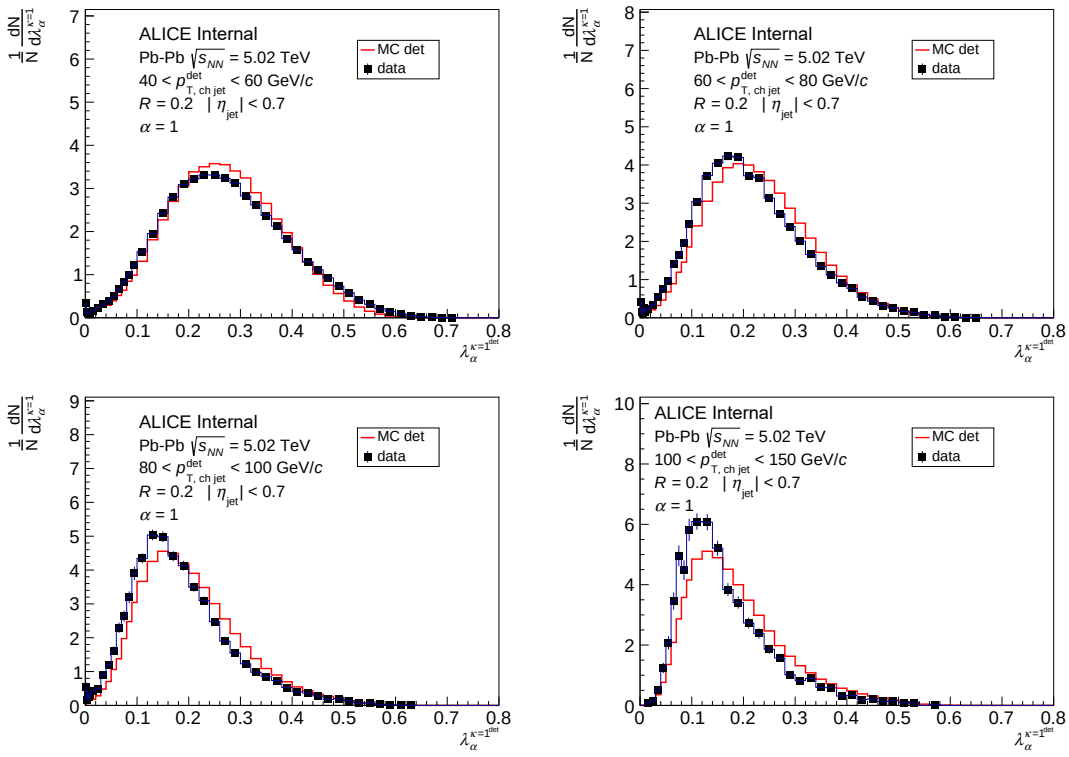


Figure 3.7: Distributions of detector-level (embedded, background-subtracted) λ_1 for $R = 0.2$ jets in simulated Pb–Pb collisions from PYTHIA8 + GEANT3 compared to raw data [286]. Reasonable agreement is observed for all $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ bins, indicating that unfolding can be used. Shape biases are considered as a systematic uncertainty; see Sect. 3.4.

- **Refolding test**: this test is used to verify that the unfolding procedure is self-consistent. The unfolded solution is multiplied (forward-folded) by the original RM, and the resulting *refolded* distribution is compared to the original detector-level spectrum. The result should be similar (but not identical), partly due to the bias injected during unfolding, and also due to the fact that a larger number of bins at detector level as compared to truth level can introduce fluctuations during the refolding, which were correspondingly suppressed during the unfolding.

- **Statistical closure test**: this test demonstrates that the unfolding procedure is reasonably robust to the statistical fluctuations of the measured distribution. The process is as follows:

  1. the MC detector-level spectrum is smeared by statistical errors on measured data;
  2. the smeared MC detector-level spectrum is then unfolded using the same parameters as in data;

3. the unfolded MC spectrum is compared to the truth-level MC spectrum.

- **Shape closure test**: this test demonstrates that the unfolding procedure is reasonably robust to the shape of the input distribution. The procedure is similar to the statistical closure test, with an additional modification to the RM:

  1. the RM is scaled in some reasonable way to reflect uncertainty on the shape of the MC, by e.g. the differences observed in Fig. 3.7. This scaled RM is then projected onto its truth- and detector-level axes to obtain scaled MC spectra at both truth and detector level;
  2. the scaled MC detector-level distribution is scaled by statistical errors on measured data;
  3. the scaled, smeared MC detector-level spectrum is then unfolded using the original, unscaled RM;
  4. the resulting unfolded, scaled MC spectrum is compared to scaled MC truth-level distribution.

- **Prior closure test**: this test demonstrates that the unfolding procedure is reasonably robust to the shape of the prior distribution. The procedure is similar to the shape closure test, but instead the scaled RM and unscaled spectra are used:

  1. the MC detector-level spectrum is smeared by statistical errors on measured data;
  2. the prior distribution is scaled. In practice, this is accomplished by scaling the RM that is fed into RooUnfold, using the same scaling criteria as mentioned above;
  3. the smeared MC detector-level spectrum is unfolded using the scaled prior (RM);
  4. the resulting unfolded MC spectrum is compared to MC truth-level distribution.

Note that a successful closure test is consistent with unity given the uncertainties; e.g., 68% of the points should be within 1σ of unity, 95% should be within 2σ, etc. For the shape and prior closure tests, the scaling function which is used is

$$p_{\mathrm{T}}^{\pm 0.5} \times [1 \pm 0.5 * (2O - 1)] \tag{3.5}$$

for the $O = \lambda_\alpha$, $\lambda_{\alpha,\mathrm{g}}$, and θ_{g} analyses, or by

$$p_{\mathrm{T}}^{\pm 0.5} \times z_{\mathrm{g}}^{\pm 0.5} \tag{3.6}$$

for the z_{g} analysis, or by

$$p_{\mathrm{T}}^{\pm 0.5} \times N_{\mathrm{counts}}^{\pm 0.1} \tag{3.7}$$

for the m_{jet} and $m_{\mathrm{jet,g}}$ analyses. These scaling functions were determined by comparing the observable shapes using jets from the anchored MC production at detector level to those from the raw data (e.g., Fig. 3.7). The exponential drop-off in the $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ spectrum is varied

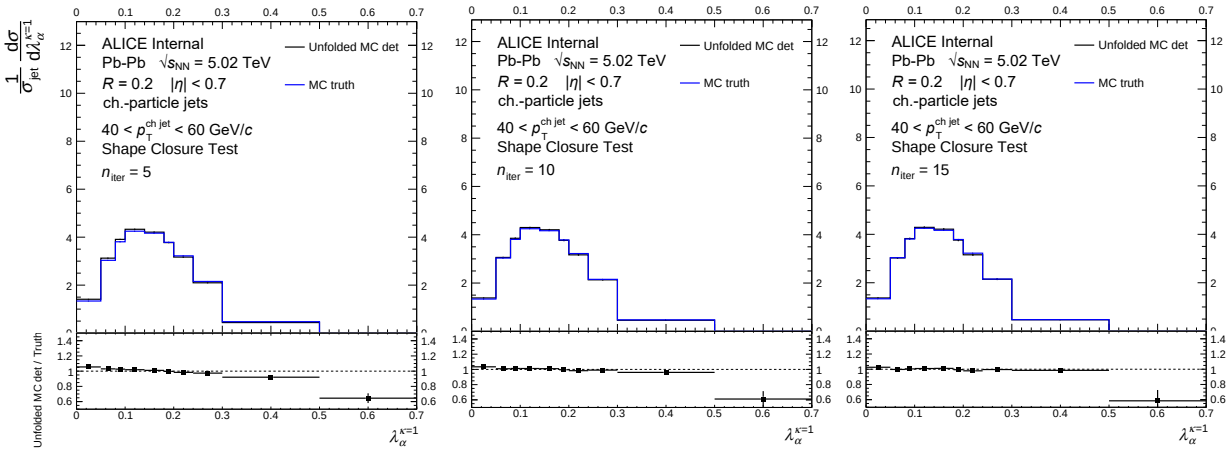


Figure 3.8: Example shape closure tests (+0.5 variation) for $\lambda_{1.5}$ using the anchored MC production, and scaled by Eq. 3.5, as measured in Pb–Pb collisions for $R = 0.2$ jets in the lowest $40 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 60$ GeV/c bin [286]. The solutions have increasing n_{iter} from 5 (left) to 10 (middle) to 15 (right), demonstrating a simultaneous improvement in the closure. Note that the highest λ_α bin is truncated from the reported distributions, but it is included in the unfolding for efficiency considerations.

by a power of ± 0.5, which is a liberal estimate on the inaccuracy of the $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ spectrum as predicted by PYTHIA 8. Along the observable axes, λ_α varies approximately linearly, so a linear variation is given with a $\pm 50\%$ variation in the extreme tails of the distribution; m_{jet} instead mostly undergoes a sharpening or smoothing of the distribution, reflected by a variation of power ± 0.1 on the number of counts.

For the jet angularities, n_{iter} is chosen to be 3 for pp and 5 for Pb–Pb, as this is where good closure is observed in shape and statistical closure tests. However, the lowest $p_{\mathrm{T}}^{\mathrm{jet}}$ bin in Pb–Pb is an exception, where $n_{\mathrm{iter}} = 10$ and 15 is used for the groomed and ungroomed distributions, respectively, as the extra iterations are helpful to reduce the stronger bias introduced by the shape of the prior, and thus improve the observed closure, as well as reduce the systematic uncertainties. An example of this improvement in the shape closure test can be seen in Fig. 3.8. For the jet mass, n_{iter} is similarly chosen based upon the most acceptable closure under these unfolding tests.

Examples of these closure tests are shown in Fig. 3.9 for the convergence test, Fig. 3.10 for the refolding test, Fig. 3.11 for the statistical closure test, Fig. 3.12 for the shape closure test, and Fig. 3.13 for the prior closure test. Note that for many of these figures, the full range shown in the unfolding is not actually reported in the final result; many distributions are truncated, and the highest bin is removed. To see the final binnings used in the reported distributions, see Ch. 5.

In the Bayesian unfolding approach, some smearing of the statistical uncertainty on the unfolded distributions is performed with each successive iteration. The result is that the

statistical uncertainties on the final unfolded result will be larger than on the raw data. Some example distributions showing the evolution of these uncertainties on the unfolded result as a function of the n_{iter} is shown for λ_α in Fig. 3.14.

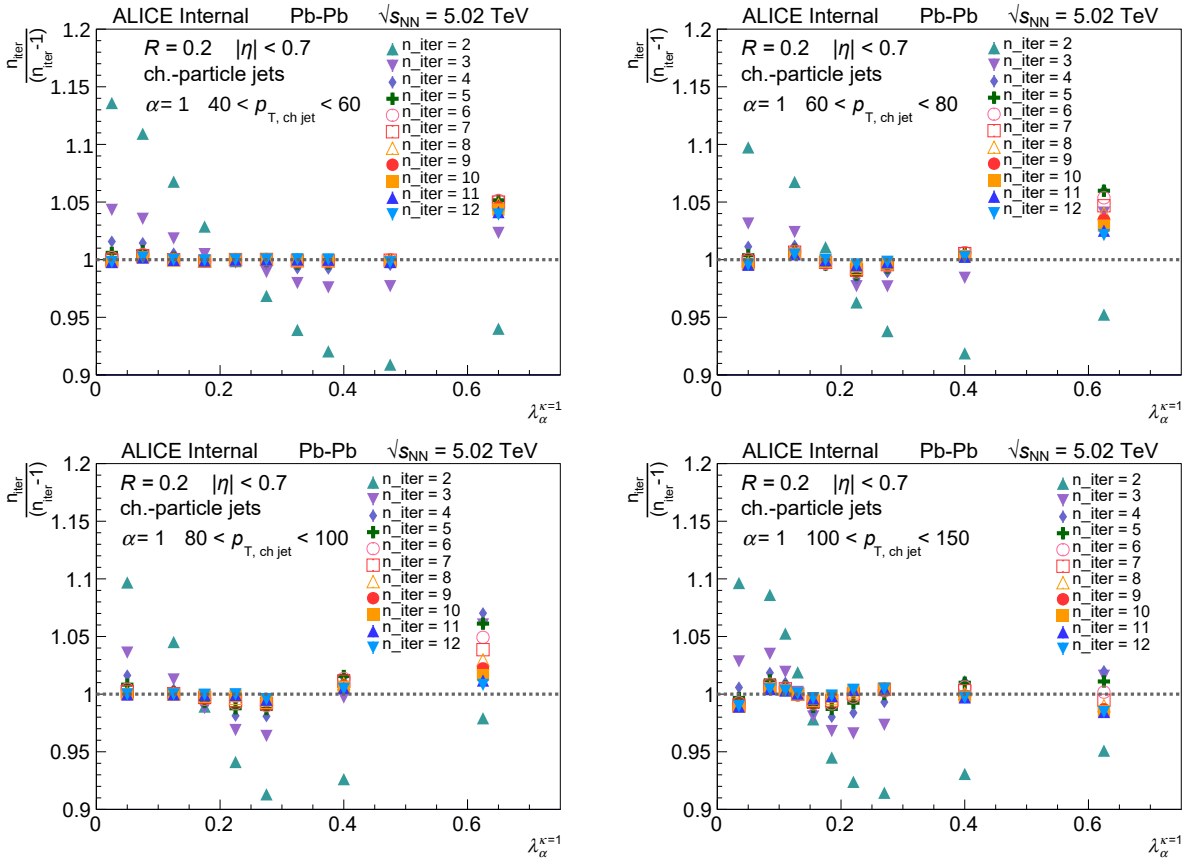


Figure 3.9: Example convergence tests λ_1 in Pb–Pb collisions for $R = 0.2$ jets [286]. Note that the last bin is truncated from the reported distributions, but it is included in the unfolding for efficiency considerations.

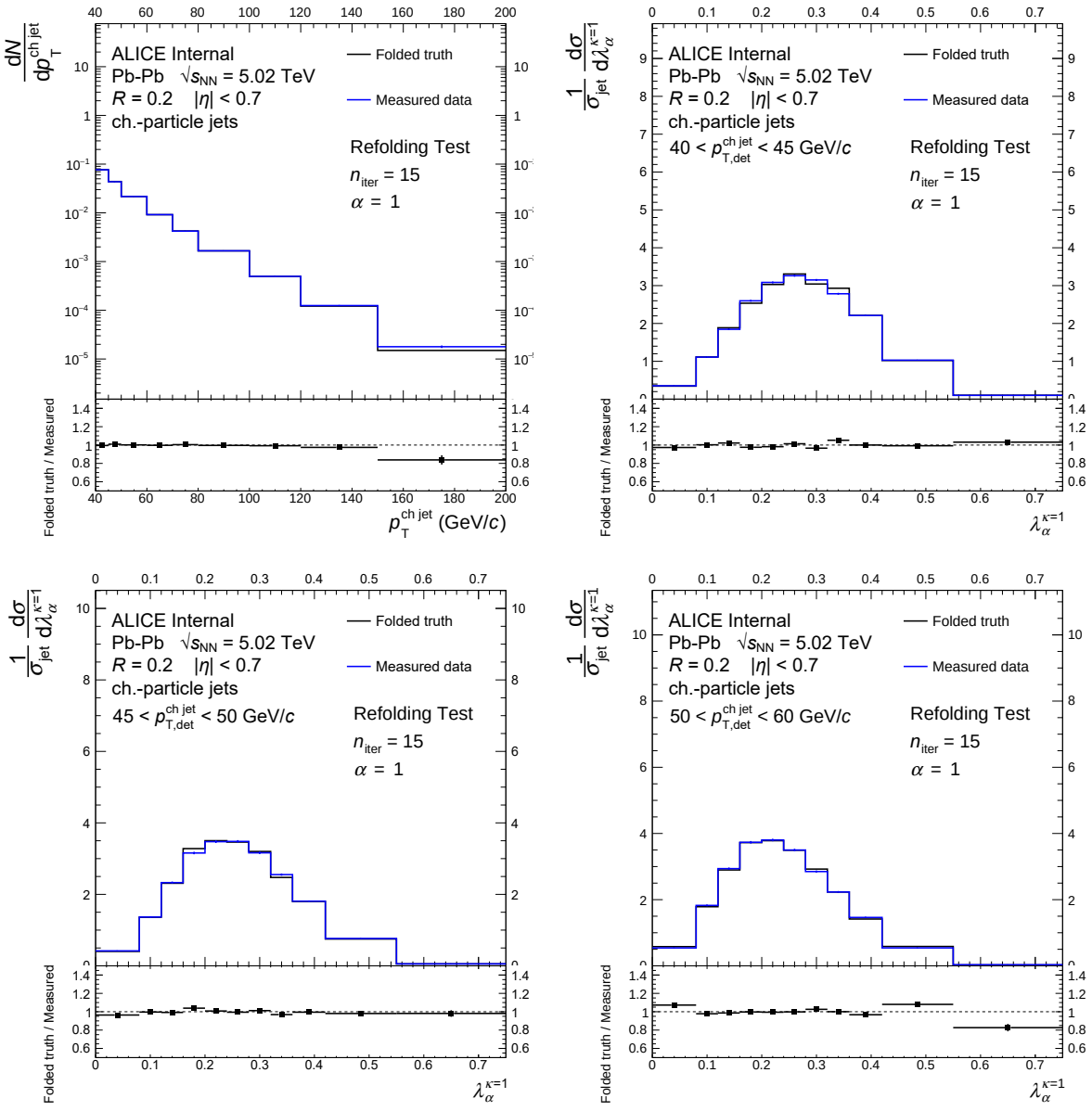


Figure 3.10: Example refolding tests projected onto the $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ and observable axes for λ_1 in Pb–Pb collisions [286]. Each plot is for one of the respective $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}}$ bins, with the binning used for the lowest $40 \leq p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 60$ GeV/c reported distribution. As can be seen, the test is successful for the relevant range in each case. *Upper left:* $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ projection for all $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}}$ bins; *upper right:* λ_α projection for $40 \leq p_{\mathrm{T,det}}^{\mathrm{ch\ jet}} < 45$ GeV/c; *bottom left:* λ_α projection for $45 \leq p_{\mathrm{T,det}}^{\mathrm{ch\ jet}} < 50$ GeV/c; *bottom right*: λ_α projection for $50 \leq p_{\mathrm{T,det}}^{\mathrm{ch\ jet}} < 60$ GeV/c.

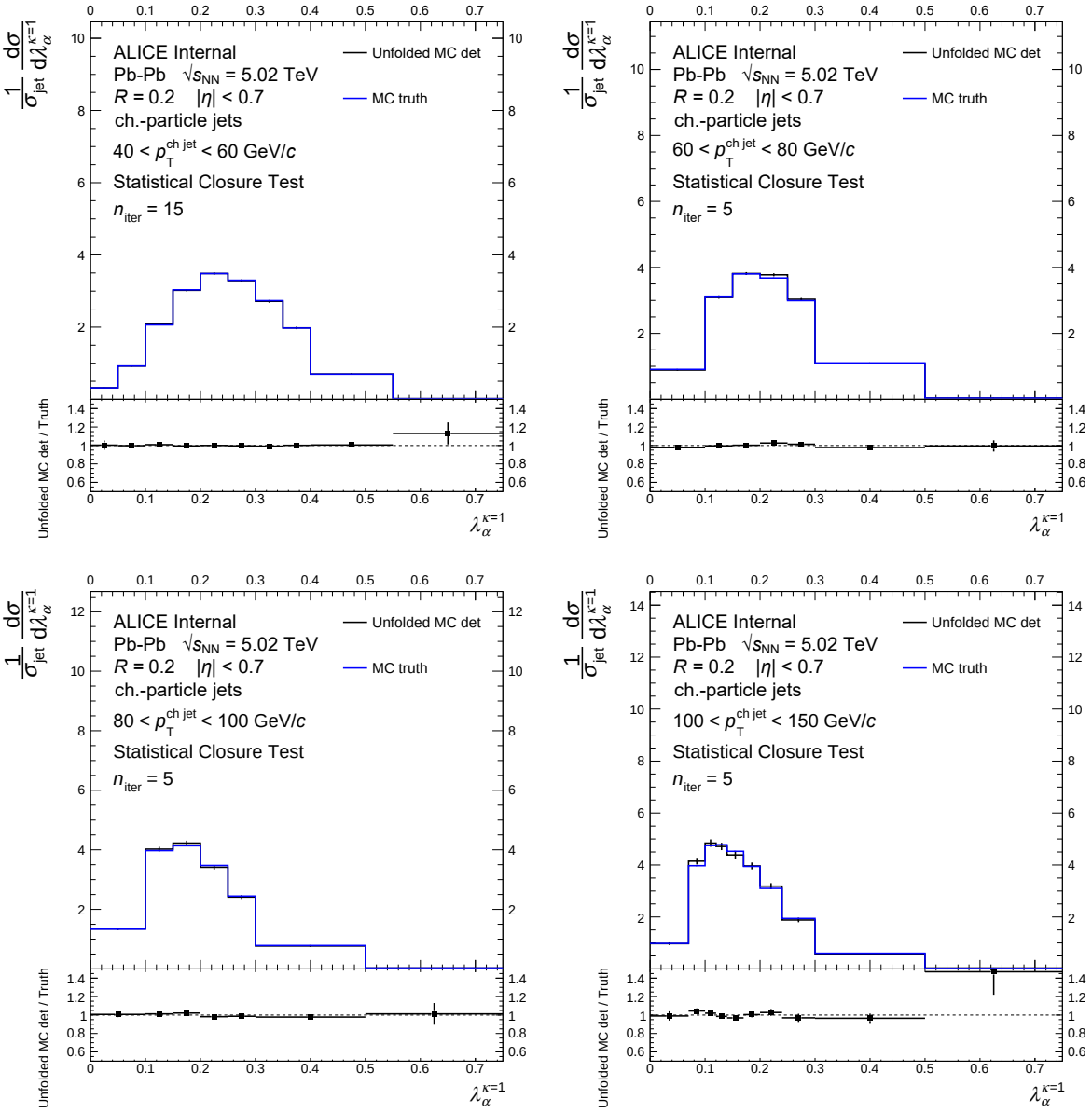


Figure 3.11: Example statistical closure tests for λ_1 using the anchored MC production smeared by data uncertainties as measured in Pb–Pb collisions for $R = 0.2$ jets [286]. Each plot is given for one of the $p_{\mathrm{T,truth}}^{\mathrm{ch\ jet}}$ bins, each using a different binning (and hence a different RM) to account for statistical and shape differences. Note that the last bin is truncated from the reported distributions, but it is included in the unfolding for efficiency considerations.

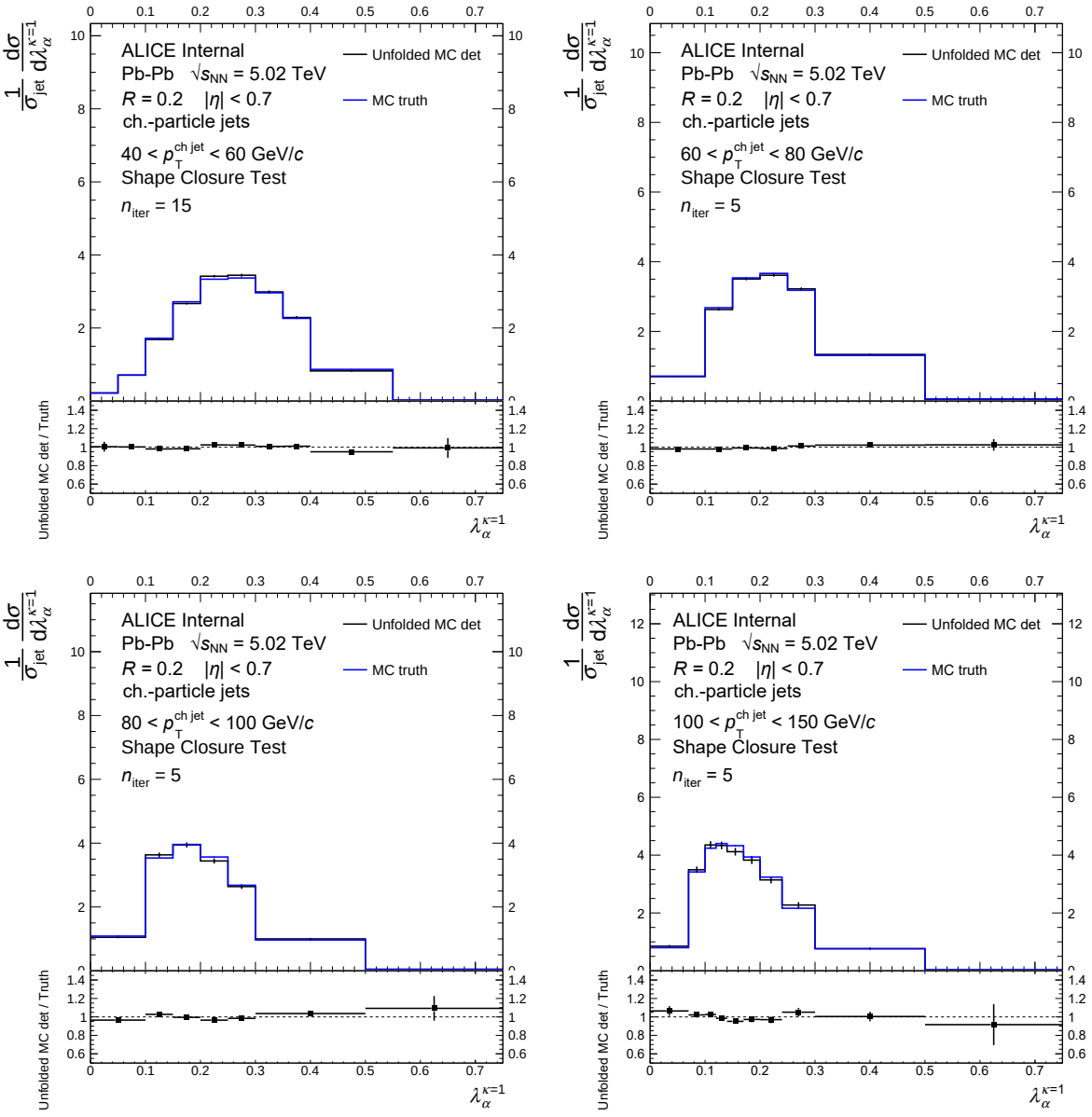


Figure 3.12: Example shape closure tests (+0.5 variation) for λ_1 using the anchored MC production, and scaled by Eq. 3.5, as measured in Pb–Pb collisions for $R = 0.2$ jets [286]. Each plot is given for one of the $p_{\mathrm{T,truth}}^{\mathrm{ch\,jet}}$ bins, each using a different binning (and hence a different RM) to account for statistical and shape differences. Note that the last bin is truncated from the reported distributions, but it is included in the unfolding for efficiency considerations.

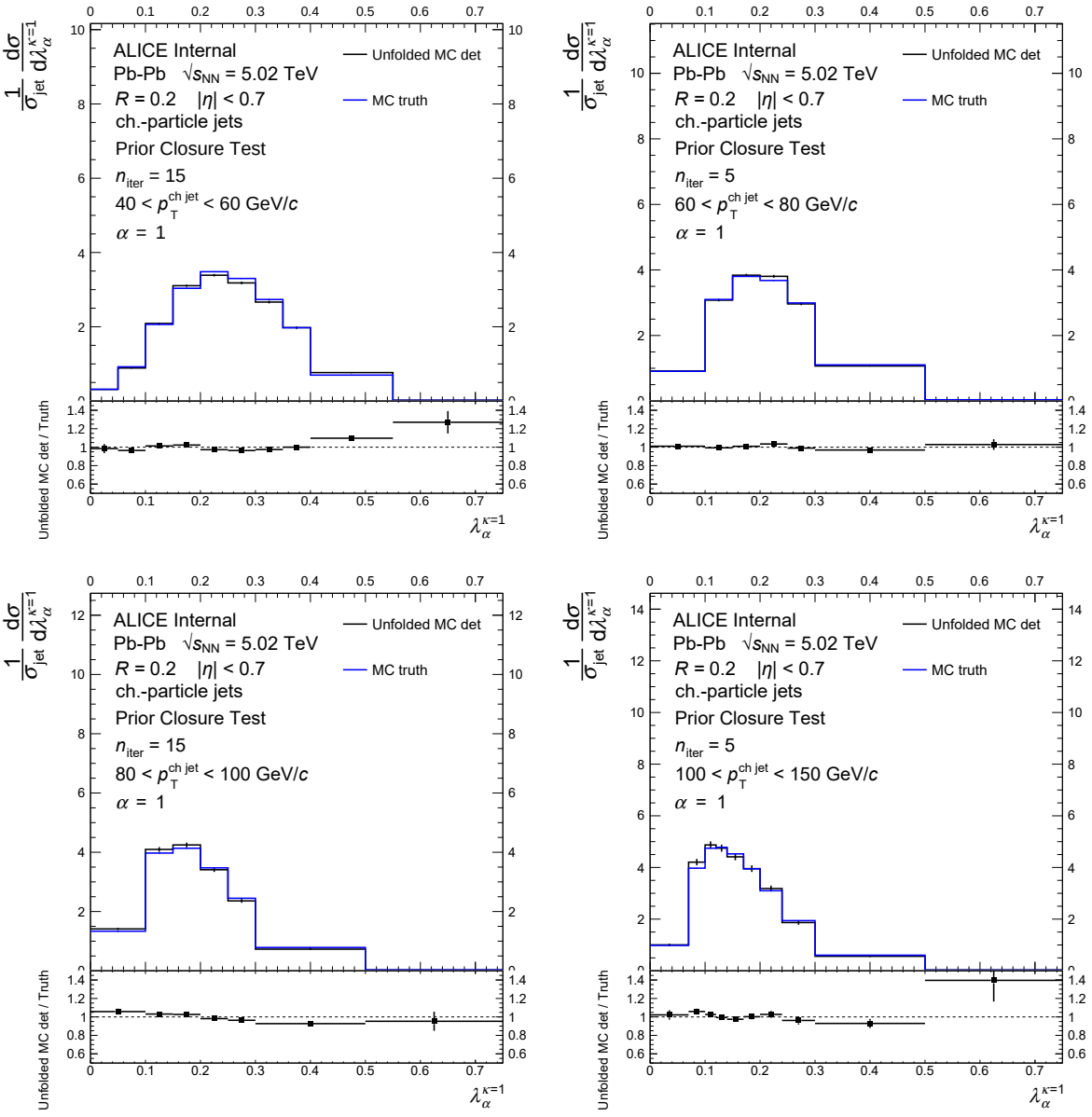


Figure 3.13: Example prior closure tests (+0.5 variation) for λ_1 using the anchored MC production, and scaled by Eq. 3.5, as measured in Pb–Pb collisions for $R = 0.2$ jets [286]. Each plot is given for one of the $p_{\mathrm{T,truth}}^{\mathrm{ch\,jet}}$ bins, each using a different binning (and hence a different RM) to account for statistical and shape differences. Note that the last bin is truncated from the reported distributions, but it is included in the unfolding for efficiency considerations.

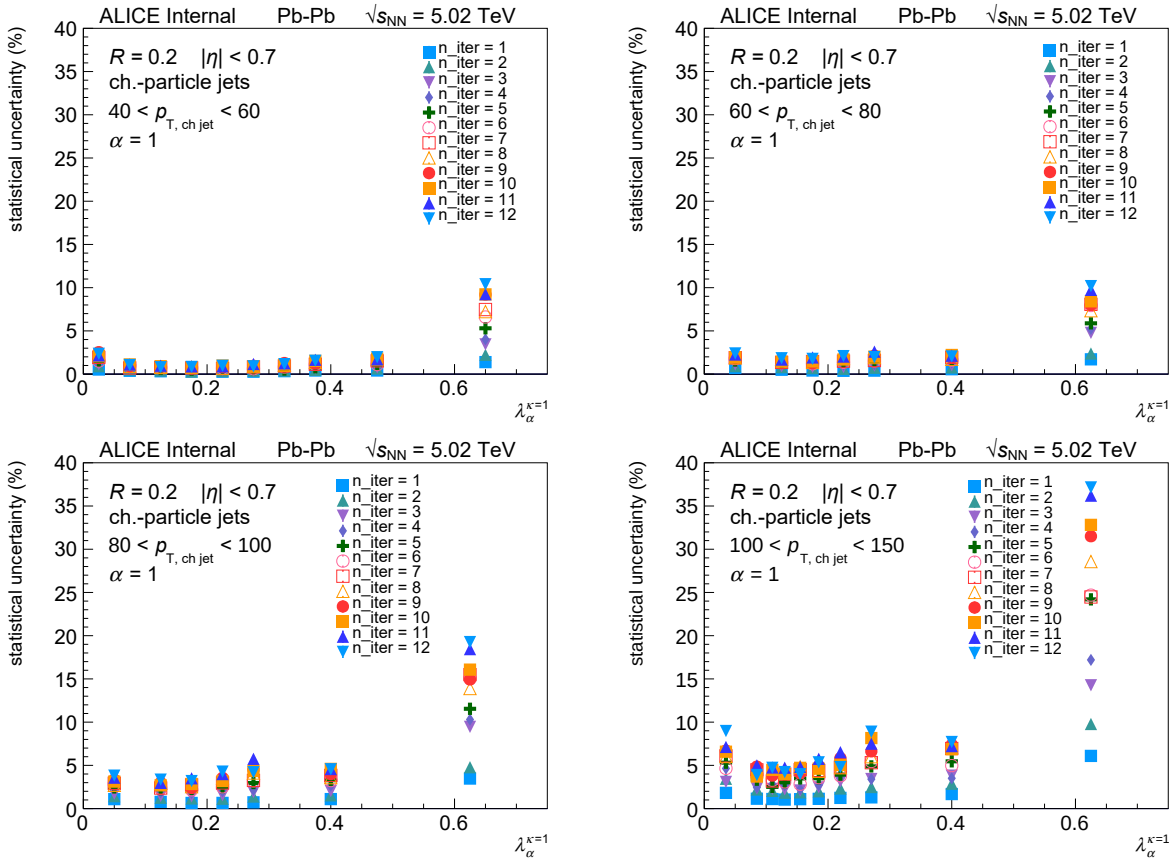


Figure 3.14: Example statistical uncertainty divergence for the unfolded λ_1 distributions in Pb–Pb collisions using $R = 0.2$ jets as a function of n_{iter} [286]. Note that the last bin is truncated from the reported distributions, but it is included in the unfolding for efficiency considerations.

3.3.2 Thermal closure tests

The refolding and self-closure tests check the mathematical consistency of the unfolding framework, but they do not test whether the unfolded solution is physically correct. In particular, the rejection of combinatorial jets or background splittings in the final result must be verified for Pb–Pb collisions. To do so, a *thermal closure test* is performed using MC events in a simulated thermal background. The entire analysis is repeated using "combined" jets clustered from the combination of PYTHIA 8 + GEANT3 detector-level particles along with the generated thermal background particles, in a similar procedure to the nominal case:

1. construct the combined-level jet spectrum by embedding detector-level MC into the thermal background,
2. apply CS, build the RM, and
3. unfold the combined-level jets to compare the result to the truth-level MC spectrum.

Since the background does not have any jet component, this test is able to verify whether the analysis procedure indeed recovers the correct solution. That is, the test ensures that if there is a jet signal and a background, the analysis procedure will correctly produce the jet spectrum, and not be contaminated by the background.

In order to create the thermal background, N particles are generated with their p_{T} sampled from a Gamma distribution:

$$f_{\Gamma}\left(p_{\mathrm{T}};\alpha,\beta\right) \propto p_{\mathrm{T}}^{\alpha-1} e^{-p_{\mathrm{T}}/\beta} \tag{3.8}$$

N is modelled as a Gaussian, and the parameter $\alpha = 2$ is selected. The free parameters $\overline{N}$, σ_N, and β are then chosen so that, when jets are reconstructed from the data, their $p_{\mathrm{T}}^{\text{ch jet}}$ smearing fits the δp_{T} distribution (see Fig. 3.2) for $R = 0.2$ jets in 0-10% centrality Pb–Pb data. The chosen values must also be compatible with observed values of N and $\langle p_{\mathrm{T}} \rangle$ in data. The parameters $\overline{N} = 2500$, $\sigma_N = 500$, and $\beta = 0.4$ are selected, which gives a reasonable description of δp_{T} as shown in Sect. 5.2.5 of Ref. [292]. The jet-finding procedure is then performed on the hybrid event to obtain a combined-level jet spectrum, and a response matrix is filled,

$$\mathrm{R}\left[\left(p_{\mathrm{T,det}}^{\text{ch jet}}\right)^{\text{PYTHIA+Thermal}}, \left(p_{\mathrm{T,truth}}^{\text{ch jet}}\right)^{\text{PYTHIA}}, O_{\text{det}}^{\text{PYTHIA+Thermal}}, O_{\text{truth}}^{\text{PYTHIA}}\right], \tag{3.9}$$

for each observable O.

Some example thermal closure tests for λ_α are given in Fig. 3.15. These plots generally demonstrate successful closure within uncertainties. Note that we also take any non-closure as an additional systematic uncertainty; see Sect. 3.4.

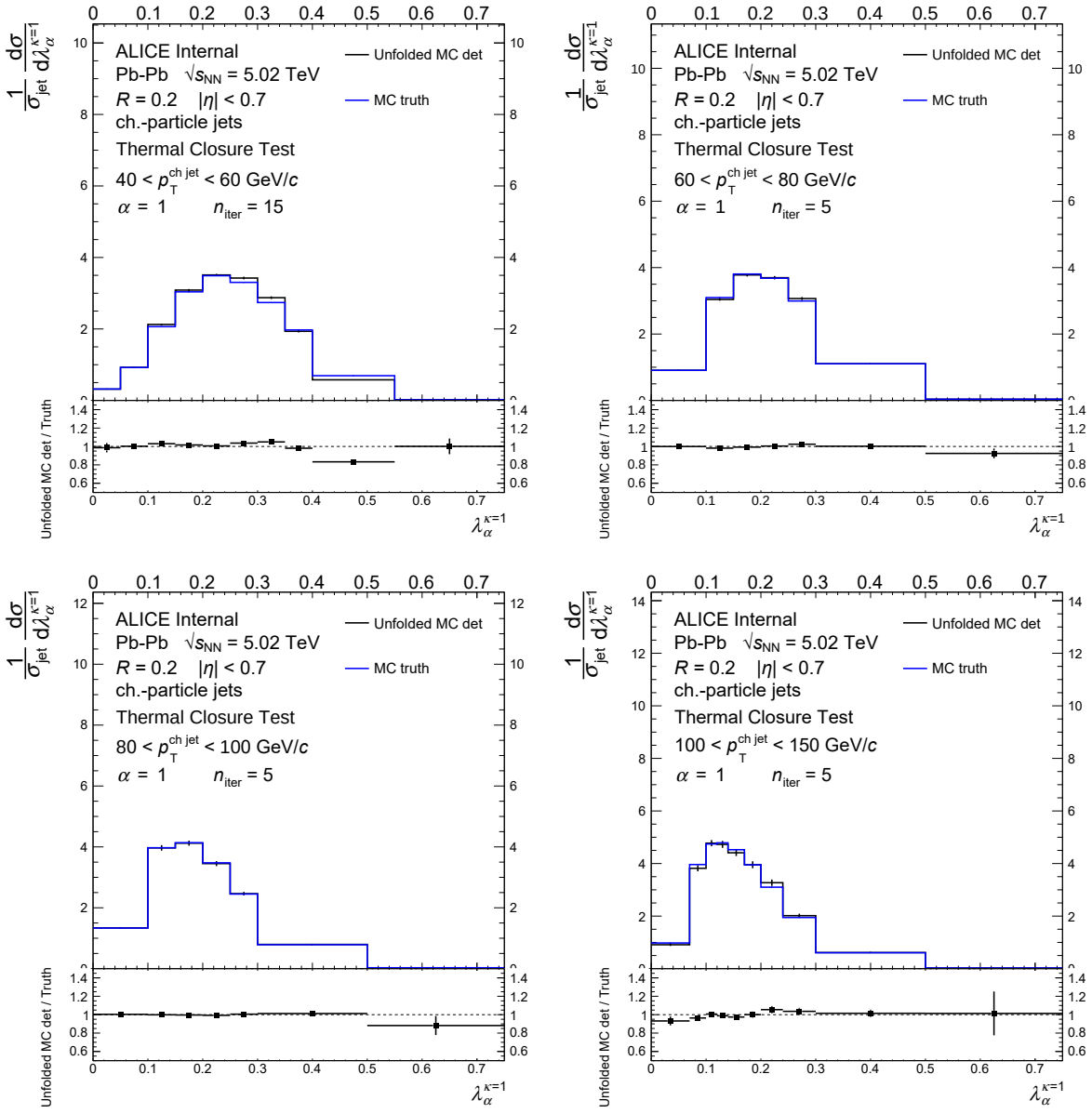


Figure 3.15: Example thermal closure tests of λ_1 for $R = 0.2$ jets in MC embedded into simulated thermal background for Pb–Pb, then subtracted with CS [286]. Any non-closure is taken as a systematic uncertainty. Note that the last bin is truncated from the reported distributions, but it is included in the unfolding for efficiency considerations.

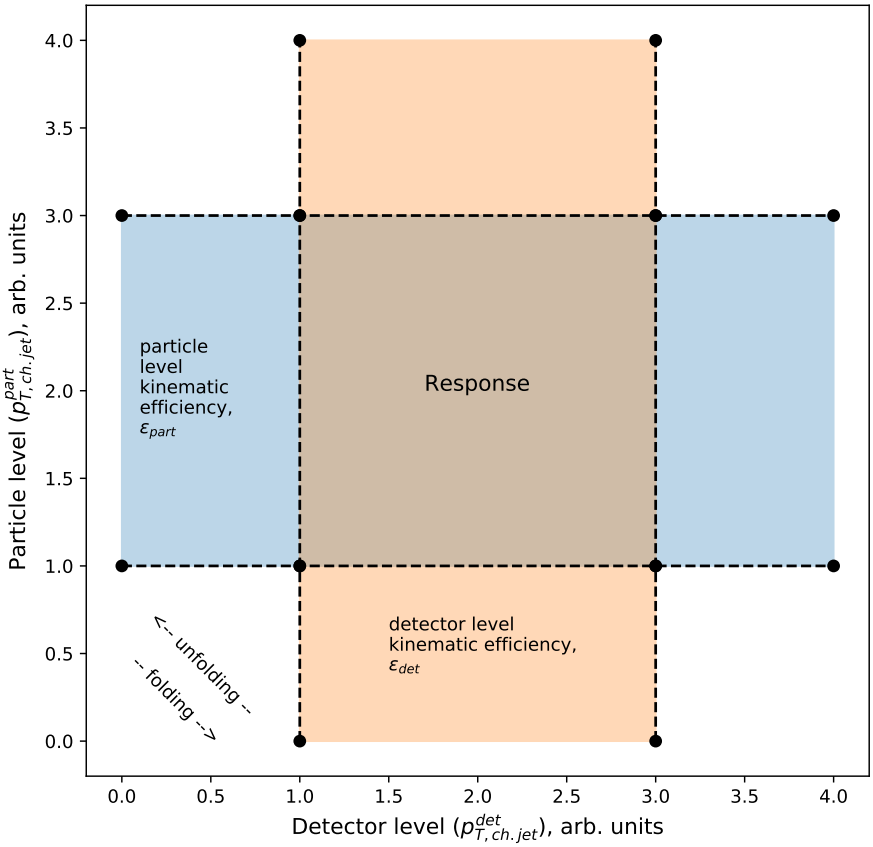


Figure 3.16: Sketch of the regions of phase space covered a response matrix, as projected onto $p_{\mathrm{T}}^{\mathrm{jet}}$ [293].

3.3.3 Kinematic efficiency

When applying the unfolding procedure, the input range of the measured $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}}$ and detector-level observable only constrains the unfolded solution to a certain extent, as visualized in Fig. 3.16. In particular, since the RM only covers a finite area of phase space, parts of the distribution which should be mapped inside or outside of the target range during the unfolding (or, during folding, as it occurs for the feed-down D^0-tagged jets in Sect. B.3) may not always be covered. This lends to some bias in the shape of the RM, which is characterized by the *kinematic efficiency* $\varepsilon_{\mathrm{kin}}$:

$$\varepsilon_{\mathrm{kin}}(\lambda_{\alpha,\mathrm{truth}}) \equiv \frac{\frac{dN}{d\lambda_{\alpha,\mathrm{truth}}}(\lambda_{\alpha,\mathrm{det}} \in [\lambda_{\alpha,\mathrm{det}}^{\mathrm{min}}, \lambda_{\alpha,\mathrm{det}}^{\mathrm{max}}], p_{\mathrm{T,det}}^{\mathrm{ch\ jet}} \in [p_{\mathrm{T,det,min}}^{\mathrm{ch\ jet}}, p_{\mathrm{T,det,max}}^{\mathrm{ch\ jet}}])}{\frac{dN}{d\lambda_{\alpha,\mathrm{truth}}}(\lambda_{\alpha,\mathrm{det}} \in [0,1], p_{\mathrm{T,det}}^{\mathrm{ch\ jet}} \in [0,\infty])}, \tag{3.10}$$

for a fixed value of $p_{\mathrm{T,truth}}^{\mathrm{ch\ jet}}$. In practice, the $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}}$ spectrum cannot be known down to 0 and up to ∞, so the values $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}} \in [5, 200]$ GeV/c are used in the denominator for pp,

and $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}} \in [5, 300]$ GeV/c for Pb–Pb, where better statistics in data permit the creation of a larger RM for the unfolding. For m_{jet}, z_{g}, and θ_{g}, the kinematic efficiency is defined analogously to Eq. 3.10.

The imperfect efficiency will be corrected fully by MC in the final result. In the RooUnfold software package [291], this is performed using the 'miss' (truth entries with no corresponding measurement) and 'fake' (measurements with no corresponding truth entry) functionality. In order to minimize the direct MC bias from this effect, $\varepsilon_{\mathrm{kin}}$ should be high. In pp, the range $[p_{\mathrm{T,det,min}}^{\mathrm{ch\ jet}}, p_{\mathrm{T,det,max}}^{\mathrm{ch\ jet}}]$ is also set to [5, 200] GeV/c, with the bottom limit being set to maximize data without entering too much into a non-perturbative regime (where the RM behavior is unverified), and the top limit being set where statistics are so low that counts above are negligible. The range $[\lambda_{\alpha,\mathrm{det}}^{\mathrm{min}}, \lambda_{\alpha,\mathrm{det}}^{\mathrm{max}}]$ is chosen based on the statistics of the distribution of $\lambda_{\alpha,\mathrm{det}}$ versus $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}}$, which must be rebinned so as to ensure stability during the unfolding procedure. With these choices, the kinematic efficiency is 1 for all bins.

In order to validate the results of this choice, and that any kinematic efficiency correction from truncation effects are minimal, the binning is varied and the observables are recalculated with the altered binning, while the $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}}$ range is also truncated to [10, 120] GeV/c as a systematic check. Figure 3.17 shows 2D kinematic efficiencies using this variation, which reveals that the calculated values of $\varepsilon_{\mathrm{kin}}$ are within approximately 5% of unity in the reported range $p_{\mathrm{T}}^{\mathrm{ch\ jet}} \in [20, 100]$ GeV/c for λ_α. This implies that the effect is even smaller (power-suppressed in $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$) for the utilized range. This falls within the systematic uncertainties assigned to the unfolding, as discussed in Sect. 3.4. (The largest λ_α bin in these plots are included in the unfolding steps, but are dropped from the reported ranges.)

In Pb–Pb, the kinematic efficiency is typically reduced relative to pp due to background effects. In particular, due to the high background, a minimum $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}}$ cut is required at 40 GeV/c (for inclusive $R = 0.2$ jets) in order to ensure thermal closure. Further, the $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ can also be smeared out of the selected $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ range by background effects. An example of the bin-by-bin $\varepsilon_{\mathrm{kin}}$ for λ_α in Pb–Pb is shown in Fig. 3.18. For the $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ bins adjacent to the $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}}$ cut we describe above, the kinematic efficiencies are rather low, which points to large corrections from MC. Note however that since the final distributions are normalized to unity, the absolute level of the kinematic efficiency does not enter as an overall MC correction, but rather affects the shape. Uncertainties on this shape correction are addressed by considering the generator-dependence of the result, as discussed in Sect. 3.4. Additional checks for this lowest $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ bin are given in Appendix A.

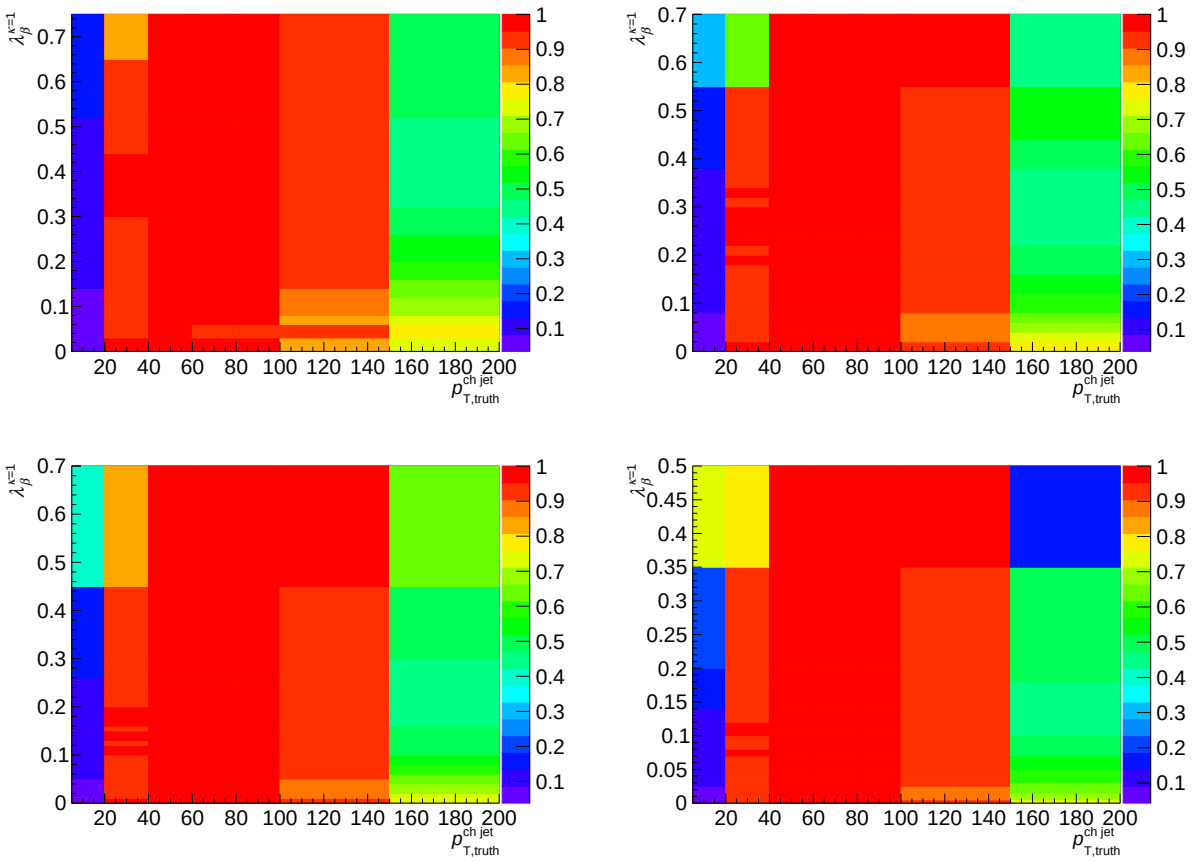


Figure 3.17: Example 2D test plots of the kinematic efficiency in pp for inclusive $R = 0.2$ charged jets in the truncated range of $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ and λ_α, and corresponding to the detector-level binning cuts described in the text [289]. The color scale represents the MC-estimated kinematic efficiency, which ranges from 0 (worst) to 1 (best). Plots are (left to right, then top to bottom) for $\alpha = [1, 1.5, 2, 3]$. The λ_α binning used in these distributions is the binning choice utilized for $p_{\mathrm{T}}^{\mathrm{ch\ jet}} \in [20, 40]$ GeV/c; the top λ_α bin is not reported.

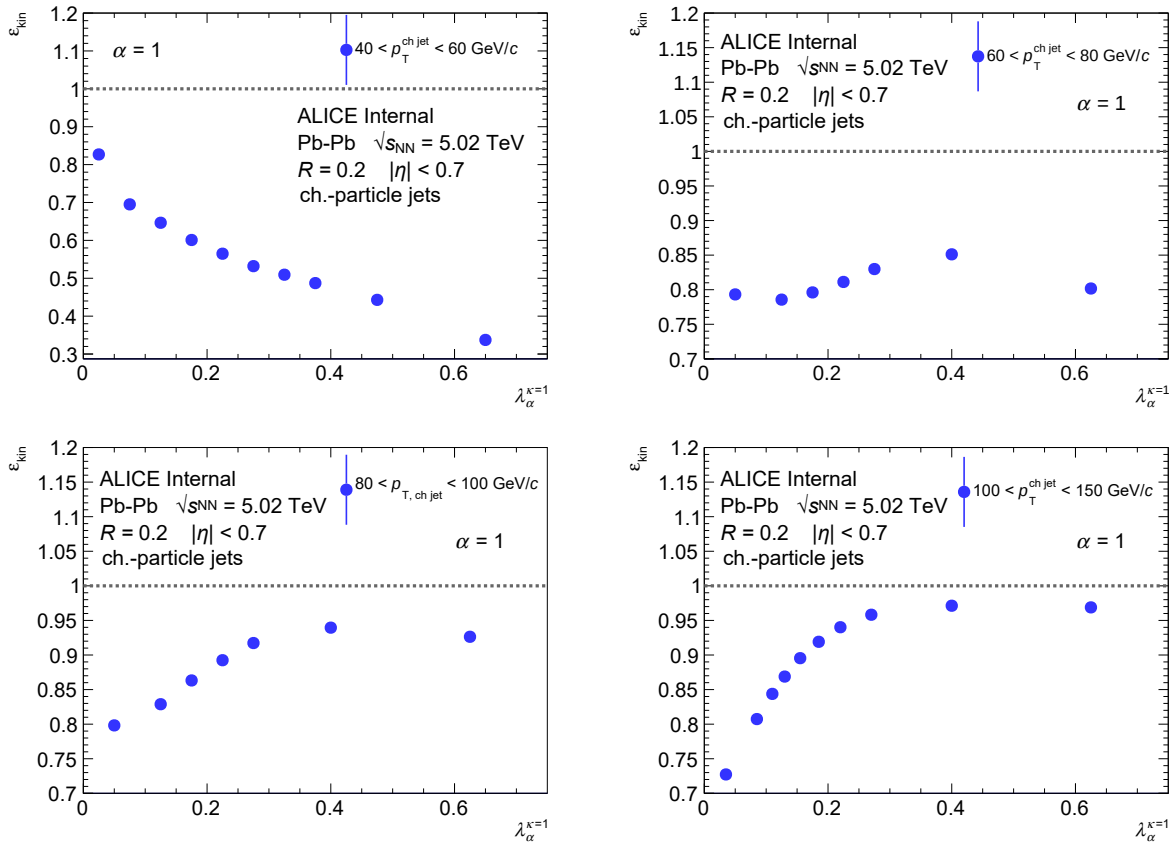


Figure 3.18: Example kinematic efficiency distributions of λ_1 for inclusive $R = 0.2$ charged jets in Pb–Pb collisions, using several $p_{\mathrm{T,truth}}^{\mathrm{ch\ jet}}$ selections, as evaluated using the anchored PYTHIA 8 + GEANT3 MC simulation [286]. The efficiency is worst for the lowest $40 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 60$ GeV/c bin, due to the requirement that $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}} > 40$ GeV/c to remove combinatorial jets from the unfolding procedure and thus ensure good thermal closure.

$p_{\mathrm{T,track}}$ bin (GeV/c)	0 - 0.5	0.5 - 1	1 - 2	2 - 3	3 - 4	4 - 5	. . .
LHC18q, 0-10% centrality (%)	3.6	3.6	6.1	6.4	5.1	4.2	. . .
LHC18r, 0-10% centrality (%)	3.7	3.7	6.0	6.4	5.1	4.2	. . .

. . .	5 - 6	6 - 7	7 - 8	8 - 9	9 - 10	10 - 11	11 - 12	12 - 13	13 - 14	14 - 999
. . .	2.4	3.3	3.3	3.3	3.3	2.4	2.4	1.6	2.4	2.4
. . .	2.3	3.3	2.4	3.4	2.5	2.5	2.5	1.8	1.8	1.8

Table 3.2: Estimated tracking efficiency uncertainty as calculated in PYTHIA 8 + GEANT3 for the relevant Pb–Pb data samples per $p_{\mathrm{T,track}}$ bin [297].

3.4 Systematic uncertainties

In the pp analyses, the systematic uncertainties in the unfolded results arise from uncertainties in the tracking efficiency and unfolding procedure, as well as the model-dependence of the response matrix, and the track mass assumption. The additional systematic uncertainty due to uncertainty on the track momentum resolution is taken to be negligible. For Pb–Pb data, an extra uncertainty is added due to the background subtraction method. The systematic variations for the analyses considered in this thesis are detailed below.

3.4.1 Tracking efficiency

In general, correcting for unmeasured tracks is a major effect of the unfolding procedure. It was originally estimated that, for hybrid tracks in pp collisions, the uncertainty on the tracking efficiency is approximately 4% [294, 295]; more recently, this number has been recalculated at about 3% [296, 297]. This number is attributed to two contributions: variation in the track selection parameters, and variation in the ITS-TPC matching requirements. In Pb–Pb collisions, a recent estimation was done based on $p_{\mathrm{T,track}}$, which shows that the effect is largest for low-p_{T} tracks and for the most central collisions [297]; a summary of specific values used for this analysis are given in Table 3.2.

Additionally, the tracking efficiency estimation is known to be slightly incorrect due to PYTHIA 8 over-predicting the production of strangeness with p_{T} in the range of a few GeV/c; however, strangeness in jets has been observed to be suppressed by a factor of 5-10 compared to inclusive event measurements [285]. This effect is therefore neglected in the present ALICE jet substructure analyses.

In order to assign a systematic uncertainty to the final result, an RM is constructed using the same techniques as for the nominal result, except that an additional percentage of tracks is rejected during jet finding. For the inclusive-jet pp analyses, λ_α, z_{g}, and θ_{g} have 4% of tracks randomly rejected; for m_{jet} analysis and the ALICE D^0-tagged jet analysis (see Appendix B), the value used is 3%, as the more-recent estimation had become available. For the Pb–Pb analysis, a $p_{\mathrm{T,track}}$-based rejection is applied using the above estimations. Since

the nominal response matrix which is used for the main result already randomly rejects 2% of all tracks to account for the worsened tracking efficiency in Pb–Pb data as compared to pp (see Sect. 3.1.2), this means that this varied response matrix in total randomly rejects approximately 2% more than the listed values. This response matrix is then used to unfold the same measured data as used in the final result. This variation is compared to the nominal result, with the differences in each bin taken as the tracking efficiency uncertainty. Since tracks can only be removed from the event and not added, the resulting uncertainties from this subtraction procedure are taken to be symmetric in the final distributions.

3.4.2 Unfolding uncertainties

In application of unfolding, the underlying systematic uncertainty is due to the regularization. In order to quantify the size of this uncertainty, several systematic variations are performed on the unfolding procedure, which assign a shape uncertainty arising from the regularization:

- **Variation of the regularization parameter** n_{iter} by ± 2 units (for λ_α, $\lambda_{\alpha,\text{g}}$, z_g, and θ_g) or ± 1 unit (for m_jet and $m_{\text{jet,g}}$). The average deviation of these two variations is taken as the systematic uncertainty.

- **Variation of the prior** by a reasonable shape, as described in Section 3.3.1. For m_jet, it is also mandated that $0.5 \leq N_\text{counts}^{\pm 0.1} \leq 1.5$, else it is set to the boundary values as to not over-modulate the tails of distributions. These variations are chosen since they vary the prior quite dramatically to demonstrate a broad range of independence on the prior, via an effect that would be reasonably expected in differing calculations (smoothing or sharpening of the distributions). The maximum of the $\pm$ case is chosen for each bin and used as the systematic uncertainty due to the prior in that bin.

- **Variation of the binning** of the observable. An alternate binning is constructed with slightly finer and/or coarser granularity than the main result, for both the data and RM. The unfolding is repeated, and the difference with respect to the nominal result is taken as the systematic uncertainty.

- **Variation of the truncation** of $p_{\text{T,det}}^{\text{ch jet}}$. The lower and upper $p_{\text{T,det}}^{\text{ch jet}}$ limits are truncated by 5 and 50 GeV/c, respectively, also taking into account the kinematic efficiency considerations described in Sect. 3.3.3. Similar to the binning variation, the unfolding is again repeated using the varied data and RM, and the difference with respect to the nominal result is taken as the systematic uncertainty.

Since each of these procedures comprise independent measurements of the same underlying systematic uncertainty in the regularization, the total unfolding systematic uncertainty is taken as the standard deviation of these variations, $\sqrt{\sum_{i=1}^{N} \sigma_i^2 / N}$, where σ_i is the systematic uncertainty due to a single variation. In this case, $N = 4$, as there are four unfolding

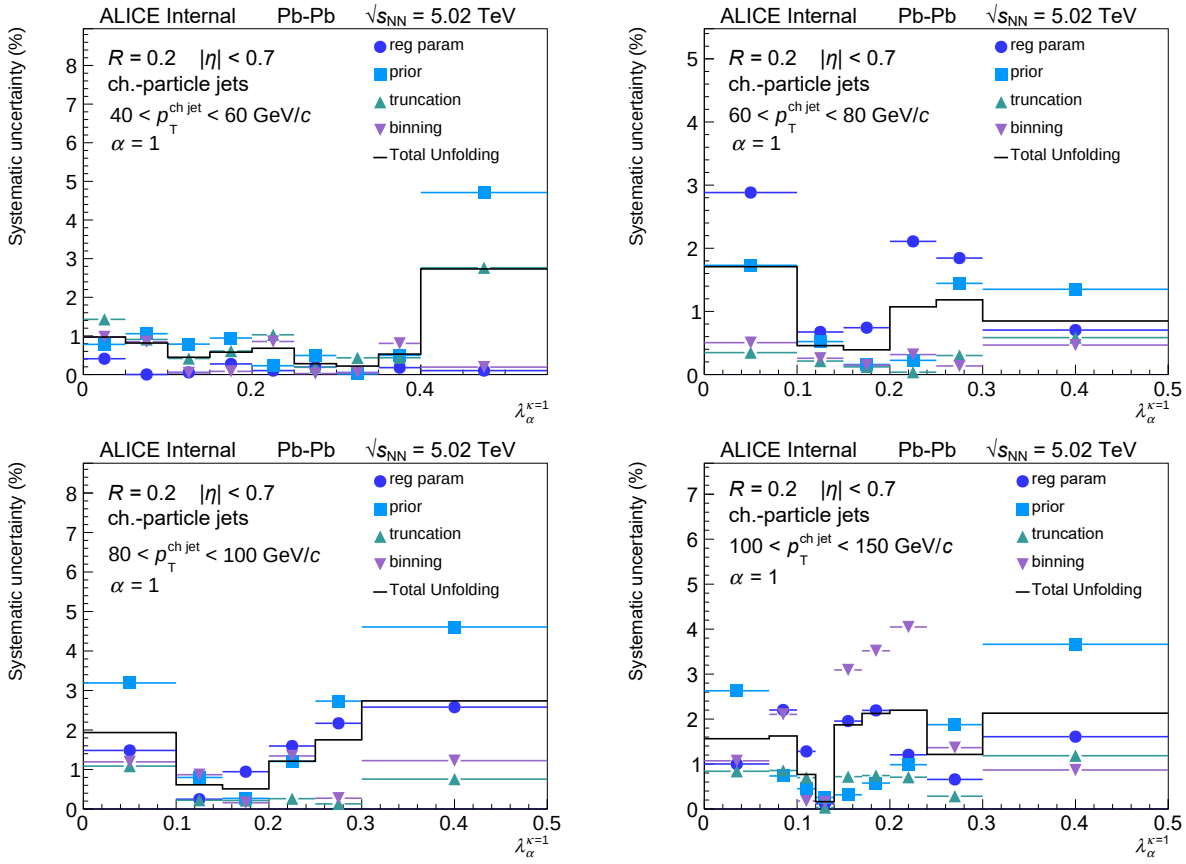


Figure 3.19: Example unfolding systematic uncertainties of λ_1 for $R = 0.2$ charged-particle jets in Pb–Pb collisions [286]. See text for details on each variation.

uncertainties calculated for these analyses. Some example distributions displaying the relative contributions from each source of unfolding systematic uncertainty are shown for λ_α in Fig. 3.19.

3.4.3 Generator model dependence

To quantify the model-dependence of using PYTHIA 8[38] to build the RM, alternate RMs are built using different MC generators. In particular, RMs are produced using Herwig 7 [42, 43] and JEWEL [192] (for Pb–Pb) using an identical binning and the same cuts as the nominal PYTHIA RM. To apply detector effects to the alternate MC events, a *fast simulation* is used, which applies tracking efficiency rejections and p_{T} smearing according to the GEANT3-estimated performance, both as a function of $p_{\mathrm{T,track}}$. This simulation was validated on truth-level PYTHIA by comparing to PYTHIA with the full GEANT3 simulation, with agreement to the few-percent-level. In order to make equitable comparisons,

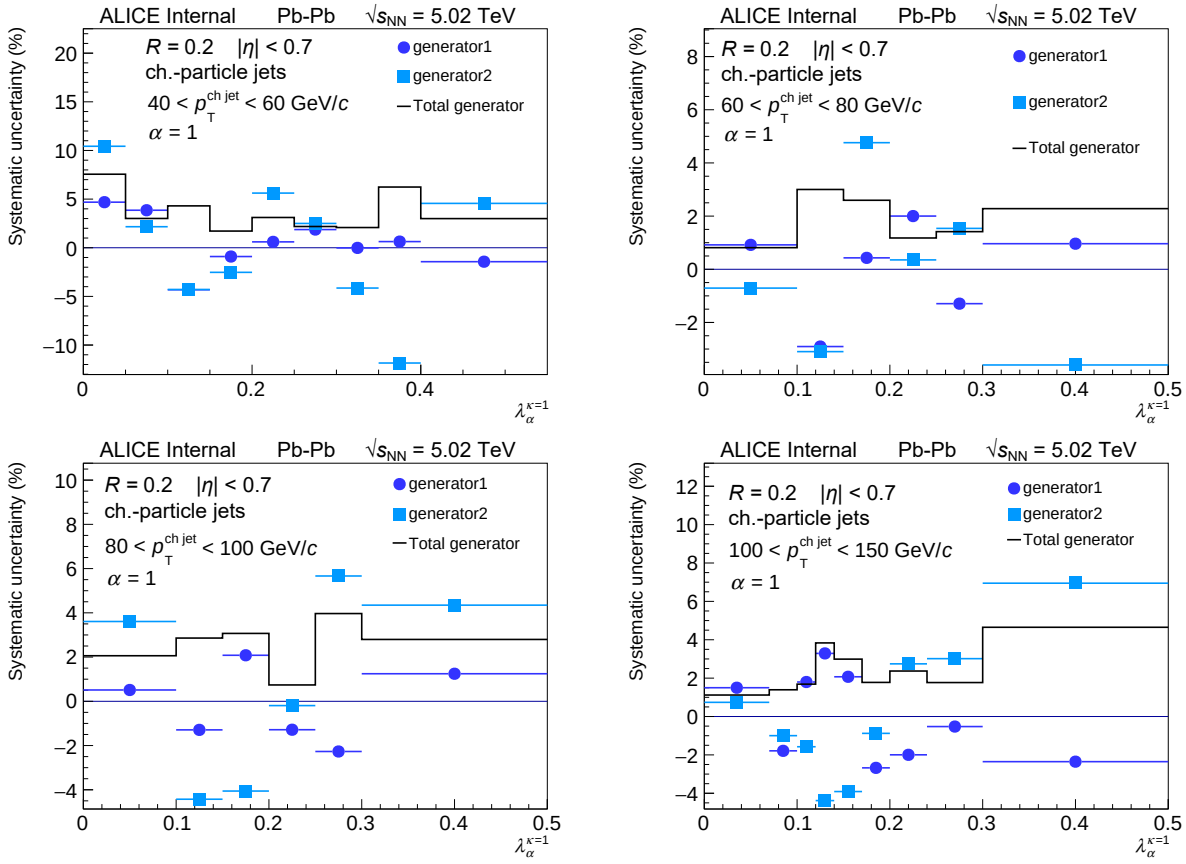


Figure 3.20: Generator systematic uncertainty of λ_1 for $R = 0.2$ charged-particle jets in Pb–Pb collisions [286]. Generator 1 is Herwig 7 default tune. Generator 2 is JEWEL Pb-Pb with recoils off. Each are taken as a ratio to PYTHIA 8 Monash 2013 through the same fast simulation.

both PYTHIA and the alternate MC events are run through this fast simulation, and the bin-by-bin differences are calculated in the final unfolded distributions, comparing the PYTHIA response to that from the alternate MCs. The average of the differences (for all variations) is taken as a systematic uncertainty. Example distributions of the calculated generator model dependence uncertainty for the alternate MCs is given for λ_α in Fig. 3.20.

Note that since the QA for the MC full-detector simulation generated for pp runs in this energy range (LHC18b8) [280] is reasonably similar to that for the Pb–Pb runs [282], and since only the difference between two fast simulations in considered, rather than a fast-to-full simulation comparison, the exact tuning of the fast simulation is not wholly necessary, and for convenience the same tracking efficiency and p_T smearing is used in the Pb–Pb analysis as the pp analysis [289], in addition to the extra 2% tracking efficiency cut described in Sect. 3.1.2 and embedding into Pb–Pb data to account for background effects.

3.4.4 Background subtraction

In Pb–Pb analyses, background subtraction via the CS approach introduces a bias in the observed distributions, since it implicitly makes a choice of how much to subtract the soft particles compared to hard particles, as well as their angular distributions. To estimate the size of the corresponding systematic uncertainty, the $R_{\max}$ parameter is varied generously from "undersubtraction" ($R_{\max} = 0.05$) to "oversubtraction" ($R_{\max} = 0.5$), around the nominal value of $R_{\max} = 0.1$ for $R = 0.2$ charged-particle jets. Figure 3.21 shows the δp_{T} distributions for these three cases. The maximum deviation of these two variations is taken as the systematic uncertainty.

3.4.5 Thermal non-closure

As mentioned in Sect. 3.3.2, any non-closure in the thermal closure test is addressed by the addition of a thermal non-closure systematic uncertainty. The thermal closure test for each of the distributions is converted into a percentage of non-closure, and the statistical uncertainty is 'subtracted' from the nominal ratio in order to minimize covariance with the statistical uncertainties introduced by the smearing procedure in the unfolding. Specifically, the non-closure uncertainty is approximated from these ratios as [286]

$$\sigma^2_{\text{total}} \approx \sigma^2_{\text{stat}} + \sigma^2_{\text{non-closure}}, \tag{3.11}$$

thereby requiring that $\sigma_{\text{total}} > \sigma_{\text{stat}}$, or else it is assumed that reasonable closure is observed for that bin, in which case the thermal non-closure uncertainty is set to 0. As before, this bin-by-bin uncertainty is taken as being symmetric on each individual reported data point.

3.4.6 Total systematic uncertainties

Each of the systematic uncertainties outlined in the above sections is assumed to be independent, and therefore they are summed in quadrature to obtain the total bin-by-bin systematic uncertainty. The assumption of independence is a very liberal approach, intended to capture any missing components of the systematic uncertainty that were not addressed in the above methodologies.[2] An example set of the total systematic uncertainties, showing the relative size of each contribution, are given for λ_α in Fig. 3.22. The dominant systematic uncertainties are typically due to uncertainty on the tracking efficiency and the dependence of unfolding on the model which is used to generate the RM.

[2] A more detailed procedure would involve reconstructing a covariance matrix for each of the variations, although this is nontrivial and is not typically done on ALICE for jet substructure analyses. In other words, the covariances between these systematic uncertainties are neglected.

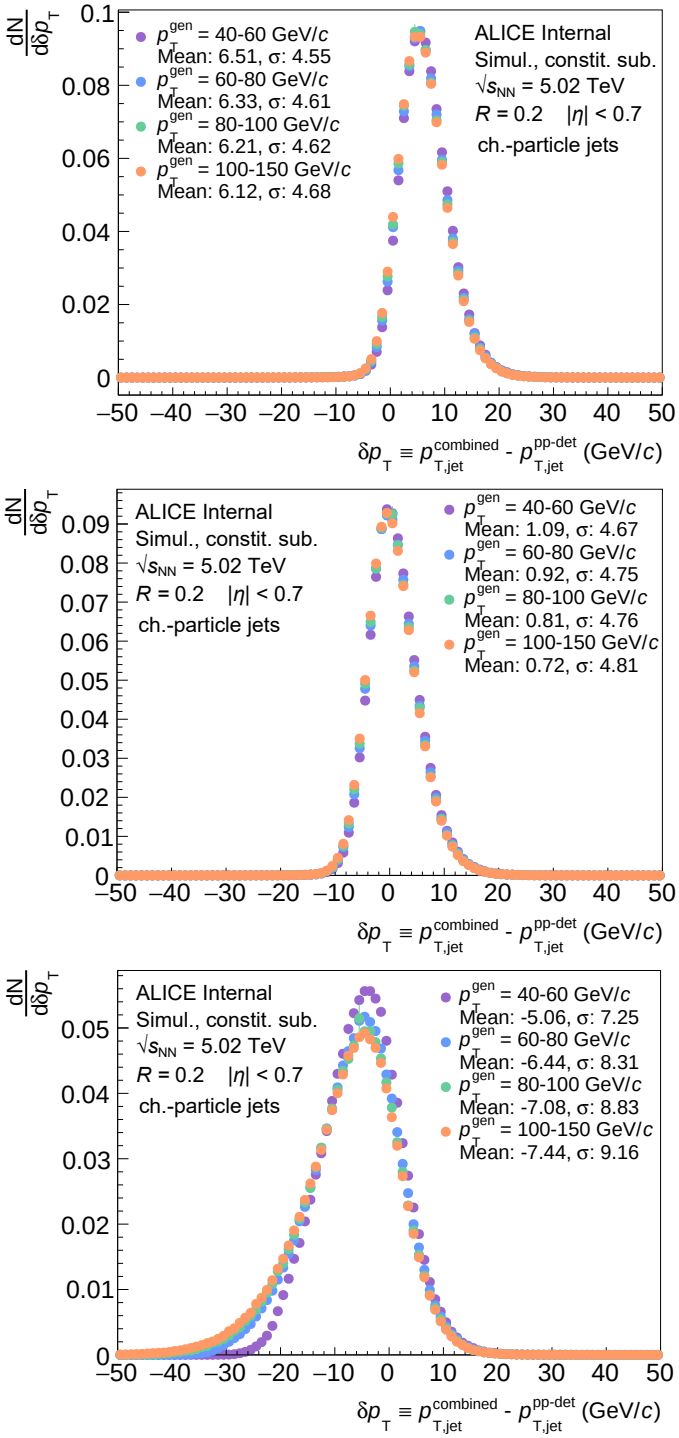


Figure 3.21: Distributions of δp_{T} in Pb–Pb collisions for $R = 0.2$ charged-particle jets [286]. The top plot is "undersubtracted", with $R_{\mathrm{max}} = 0.05$; the middle plot is the nominal case, with $R_{\mathrm{max}} = 0.1$; the bottom row is "oversubtracted" with $R_{\mathrm{max}} = 0.5$.

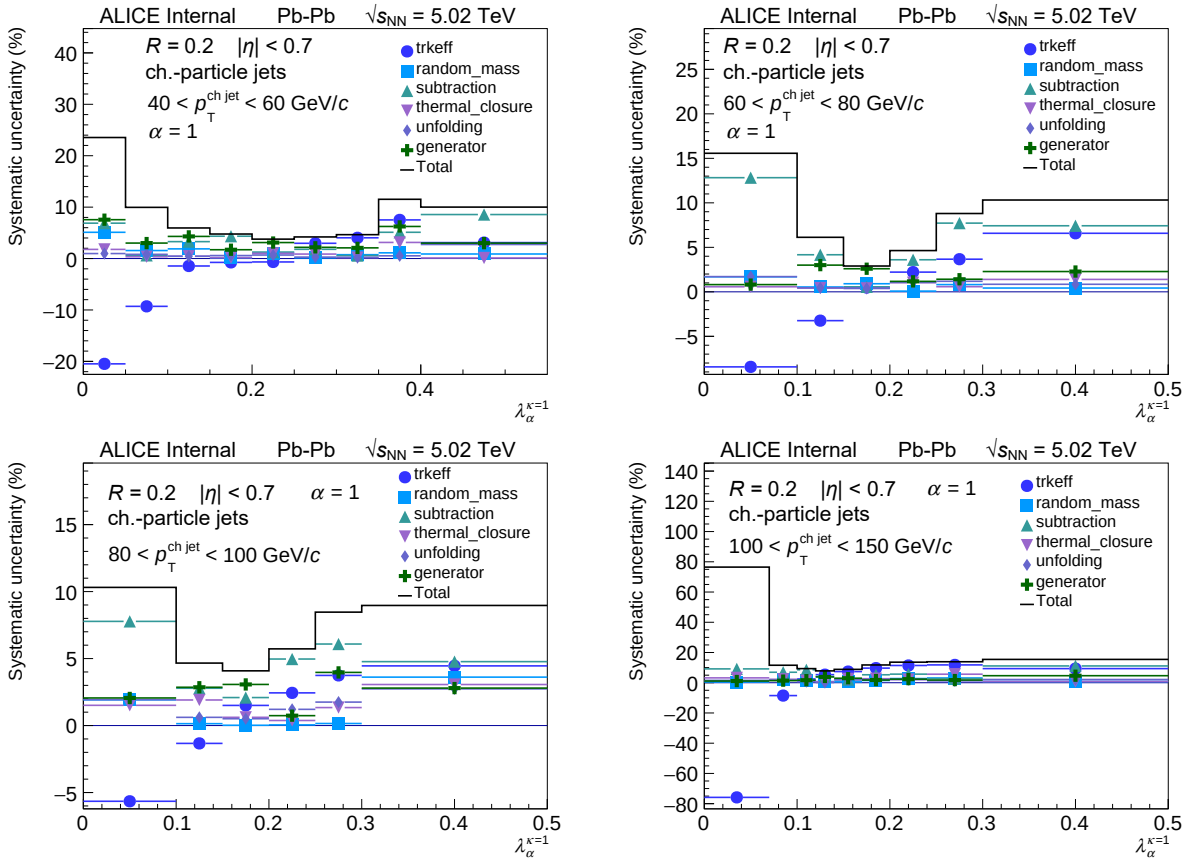


Figure 3.22: Total systematic uncertainty of λ_1 for $R = 0.2$ charged-particle jets in Pb–Pb collisions, along with individual constituents. The total uncertainty is taken as the quadrature sum of all contributions. [286]

Chapter 4

Theoretical corrections

This chapter overviews the produced pQCD predictions for λ_α, $\lambda_{\alpha,\mathrm{g}}$, R_g, and θ_g, and the corrections that were carried out on in order to compare them to pp data. In particular, the theoretical predictions, which are produced at parton level, are corrected the level of charged-particle (track) jets, which are reported alongside the experimental data. Details of the theoretical calculations that were carried out can be found in Refs. [99, 110, 231, 298]. No direct pQCD predictions were performed for the Pb–Pb analyses, though comparisons to theoretical models are reported in Ch. 5.

4.1 Parameters of obtained predictions

Theoretical predictions were obtained in the SCET formalism (see Sect. 1.3.1) at Next-to-Leading-Logarithmic (NLL′) accuracy for jets in the same kinematic regime as the ALICE detector acceptance and over a wide range of $p_\mathrm{T}^\mathrm{jet}$ bins, for both $R = 0.2$ and 0.4. The reported ALICE pp data covers $p_\mathrm{T}^{\mathrm{ch\ jet}} \in [20, 100]$ GeV/c, but the momentum of neutral constituents is lost in the reconstruction of charged-particle jets, and parton- to hadron-level smearing can be significant at such low $p_\mathrm{T}^\mathrm{jet}$ [289]. The parton-level pQCD predictions were therefore computed for $p_\mathrm{T}^\mathrm{jet} \in [10, 200]$ GeV/c, with a bin size of 5 GeV/c in the range [10, 50] GeV/c and a bin size of 10 GeV/c for the range [50, 200] GeV/c. The angularities naturally span the range $\lambda_\alpha \in [0, 1]$, as given by their definition in Eq. 1.27. The obtained predictions split this range into steps of varying size, with higher resolution at lower λ_α to account for the highly-peaked distributions, in particular for the groomed jet angularities $\lambda_{\alpha,\mathrm{g}}$ with larger values of α.

The SCET predictions are based on a factorization formalism that separates hard and soft physics at to certain momentum scales in order to calculate observables. This introduces a systematic uncertainty in the theoretical predictions, which are assessed by performing several scale variations [99]. For the jet angularities, a set of 15 scale variations was provided for each pair of R and exponent parameter α, including a central variation. Three scales variations are considered in calculation of the theoretical uncertainties,

- **the hard scale**, $\mu_{\mathrm{H}} \sim p_{\mathrm{T}}^{\mathrm{jet}}$,
- **the jet (hard-collinear) scale**, $\mu_{\mathrm{J}} \sim p_{\mathrm{T}}^{\mathrm{jet}} R$, and
- **the soft scale** $\mu_{\mathrm{S}}, \sim p_{\mathrm{T}}^{\mathrm{jet}} \lambda_{\alpha} R$.

These variations drive the uncertainty on the SCET calculations, and are propagated to the final charged-jet level. Each variation is read into individual 2D histograms of λ_{α} vs. $p_{\mathrm{T}}^{\mathrm{jet}}$, with the λ_{α} axis containing bins of varying size with edges [0, 0.001, ..., 0.01, 0.015, ..., 0.1, 0.11, ... 0.8], so that the central λ_{α} values of the bins correspond to the predicted values. Since the provided distributions are normalized by $1/\sigma_{\mathrm{jet}}$, it is necessary to include the $p_{\mathrm{T}}^{\mathrm{jet}}$ scaling in the correction procedure so that bin migration effects are handled correctly. Each distribution is scaled per $p_{\mathrm{T}}^{\mathrm{jet}}$ bin by the inclusive $p_{\mathrm{T}}^{\mathrm{jet}}$ cross-section, as calculated at NLO with NLL resummation of logarithms in the jet radius (sometimes called NLL_R) [299].

At finite perturbative accuracy, some regions of the observable distributions are understood better than others. The perturbatively limited regions are determined by the factorization scales listed above. To quantify the size of these scales and estimate the nonperturbative dependence, an average value $\langle p_{\mathrm{T}}^{\mathrm{jet}} \rangle$ is calculated for each $p_{\mathrm{T}}^{\mathrm{jet}}$ bin. The value of $\langle p_{\mathrm{T}}^{\mathrm{jet}} \rangle$ is calculated by assuming the $p_{\mathrm{T}}^{\mathrm{jet}}$ spectrum to follow a power law with exponent $\aleph$, and then performing a simple average value analysis for that spectrum. The result is

$$\langle p_{\mathrm{T}}^{\mathrm{jet}} \rangle = \frac{\aleph + 1}{\aleph + 2} \frac{(p_{\mathrm{T,jet}}^{\mathrm{min}})^{\aleph+2} - (p_{\mathrm{T,jet}}^{\mathrm{max}})^{\aleph+2}}{(p_{\mathrm{T,jet}}^{\mathrm{min}})^{\aleph+1} - (p_{\mathrm{T,jet}}^{\mathrm{max}})^{\aleph+1}} \tag{4.1}$$

where min and max denote the edges of the bin, and the power scaling is assumed to be $\aleph = -5.5$. The lowest $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$ bin reported in this analysis is [20, 40] GeV/c, for which $\langle p_{\mathrm{T}}^{\mathrm{jet}} \rangle \approx 25$ GeV/c. The hard scale $\mu_{\mathrm{H}} \sim p_{\mathrm{T}}^{\mathrm{jet}}$ is therefore minimized at this value, and remains a perturbative scale (that is, 25 GeV/c is sufficiently larger than Λ_{QCD}). The jet scale $\mu_{\mathrm{J}} \sim p_{\mathrm{T}}^{\mathrm{jet}} R$ is minimized with $R = 0.2$, giving a minimum of ≈ 5 GeV/c, which is also taken to be in a dominantly perturbative regime.

The size of the soft scale $\mu_{\mathrm{S}} \sim p_{\mathrm{T}}^{\mathrm{jet}} \lambda_{\alpha} R$ varies over the angularity distributions. At large values of $\lambda_{\alpha} \sim 1$, the soft scale will be sufficiently hard as to be perturbative. However, at small values of $\lambda_{\alpha} \sim 0$, the soft scale tends towards becoming infinitely soft. Therefore, a nonperturbatively-dominated region of the distributions is defined where the soft scale is comparable to some nonperturbative scale Λ [99],

$$\mu_{\mathrm{S}}^{\mathrm{NP}} \sim \lambda_{\alpha}^{\mathrm{NP}} \langle p_{\mathrm{T}}^{\mathrm{jet}} \rangle R \leq \Lambda \quad \rightarrow \quad \lambda_{\alpha}^{\mathrm{NP}} \leq \frac{\Lambda}{\langle p_{\mathrm{T}}^{\mathrm{jet}} \rangle R}, \tag{4.2}$$

where $\lambda_{\alpha}^{\mathrm{NP}}$ labels the nonperturbatively-dominated region. For the jet angularities groomed with SD using $\beta = 0$, this scale is replaced by [231]

$$\lambda_{\alpha,\mathrm{g}}^{\mathrm{NP}} \leq z_{\mathrm{cut}}^{1-\alpha} \left(\frac{\Lambda}{\langle p_{\mathrm{T}}^{\mathrm{jet}} \rangle R} \right)^{\alpha}. \tag{4.3}$$

This relation is used to separate the theoretical predictions into "non-perturbative" and "perturbative" regions, using the chosen value $\Lambda = 1$ GeV/c. The first bin above the one containing the value of $\lambda_\alpha^{\mathrm{NP}}$ is considered the leftmost bin of the "perturbative" region, with all bins below that called the "non-perturbative" region. The corresponding quantity for θ_g also comes from the collinear-soft scale $\mu_\mathrm{S} \sim z_\mathrm{cut}\,\theta_\mathrm{g}^{1+\beta}\,p_\mathrm{T}^\mathrm{jet}\,R$ [110],

$$\theta_\mathrm{g}^\mathrm{NP} \leq \left(\frac{\Lambda}{z_\mathrm{cut}\langle p_\mathrm{T}^\mathrm{jet}\rangle R}\right)^{\frac{1}{1+\beta}}, \tag{4.4}$$

which is similarly softer than μ_H or μ_J.

It should also be mentioned that the groomed jet angularity $\lambda_{\alpha,\mathrm{g}}$ predictions are made by matching and copying on the ungroomed jet angularity tail for $\lambda_\alpha > z_\mathrm{cut}$. For the case that the groomed $\lambda_{\alpha,\mathrm{g}}^\mathrm{NP} > z_\mathrm{cut}$, then the ungroomed definition of $\lambda_\alpha^\mathrm{NP}$ should be used; however, this does not occur for any of the reported distributions in this thesis.

4.1.1 Event generators

Both the PYTHIA 8 and Herwig 7 event generators are used to apply corrections to the parton-level predictions.

A sample of PYTHIA 8.244 events is generated using the Monash 2013 tune with process `HardQCD:all=on` and MPI turned off. In order to better match the perturbative predictions, the shower cutoff is set to `TimeShower:pTmin=0.2`, which is half of its default value. To generate adequate statistics across the entire $p_\mathrm{T}^{\mathrm{ch\ jet}}$ range, this simulation is run in batches for 20 $\hat{p}_\mathrm{T}$ bins, with the same edges as used in ALICE MC simulations (i.e., Eq. 3.1). The minimum and maximum $\hat{p}_\mathrm{T}$ ranges are initialized in PYTHIA using `PhaseSpace:pTHatMin` and `PhaseSpace:pTHatMax`, respectively, except for the last bin, for which only the minimum $\hat{p}_\mathrm{T}$ is set. Each $\hat{p}_\mathrm{T}$ bin is generated with $N_\mathrm{ev} = 2.1 \times 10^7$ events.

For each event in the simulation, the set of final-state partons (after the parton shower) is saved. The particles are then hadronized via `pythia.forceHadronLevel()`, and the set of full hadrons, as well as the subset of charged hadrons, are additionally saved for jet finding and matching. In the case that hadronization fails, the event is discarded and processing for a new event begins.

Similarly, Herwig 7.2.1 is run using the default tune with the QCD 2-to-2 scattering process via the `insert SubProcess:MatrixElements[0] MEQCD2to2` command. The phase space is set so that the pseudorapidity of $k_\mathrm{T}^\mathrm{jet}$ is generated within $|\eta_\mathrm{jet}| \leq 1.5$, considering the phase space requirements of ALICE, with an additional buffer to account for kinematic efficiency. Events are generated in $k_\mathrm{T}^\mathrm{jet}$ bins with edges

$$\begin{aligned}&[7,\ 9,\ 12,\ 16,\ 21,\ 28,\ 36,\ 45,\ 57,\ 70,\ 85,\ 99,\\ &\ 115,\ 132,\ 150,\ 169,\ 190,\ 212,\ 235,\ 260,\ \infty]\ \mathrm{GeV}/c,\end{aligned} \tag{4.5}$$

shifted one bin upward as compared to PYTHIA 8. This is due to divergence in the LO cross section when running Herwig 7 with $k_{\mathrm{T}}^{\mathrm{jet}} \in [5, 7]$ GeV/c, which generates an error for that bin. $N_{\mathrm{ev}} = 1.5 \times 10^7$ events are generated per $k_{\mathrm{T}}^{\mathrm{jet}}$ bin.

Since Herwig lacks integration with `Python`, events are parsed from the output log file generated with debugging level set to `-d2`. As before, final-state partons are saved, here by parsing the `final` subsection of the `ShowerHandler` section in the log file. Final-state hadrons (with minimum lifetime $\tau c \leq 10$ mm) are similarly saved by parsing the `final` subsection of the `DecayHandler` section of the log file. The set of charged hadrons is obtained by parsing the type of hadron and requiring that it has a $\pm$ charge.

Jets are reconstructed at parton, hadron, and charged-hadron levels using the same procedure and parameters as defined in Sect. 3.2. To match with the ALICE tracking requirements, jets are similarly required to be within psuedorapidity $|\eta_{\mathrm{jet}}| < 0.9 - R$. We also require that $p_{\mathrm{T}}^{\mathrm{jet}} > 5$ GeV/c at all levels, where model predictions are reasonably accurate. Before filling histograms, jets are matched at parton, hadron, and charged-hadron levels. The distance between matched jets is required to be $\Delta R < R/2$ for both charged-hadron to hadron level matching and hadron to parton level matching, and the match is required to be unique in both cases.

After filling all histograms, they are scaled by σ/N_{ev}. In PYTHIA, σ is obtained through the `pythia.info.sigmaGen()` method; in Herwig, σ is obtained by parsing the total integrated cross section from the end of the produced log file. All histograms from all the $\hat{p}_{\mathrm{T}}$ (or $k_{\mathrm{T}}^{\mathrm{jet}}$) bins are then merged into a single file.

This procedure is repeated with MPI turned on so that ratio histograms may be created for scaling, along with pure charged jet corrections, as described below.

4.2 Folding-based corrections

The parton-level predictions are corrected to charged-hadron-level via two different procedures. The first is a two-step procedure based purely on MC. In this approach, a two-dimensional forward-folding procedure is applied to account for hadronization and charged-jet effects. This is followed by a bin-by-bin scaling to account for MPI, and a final rescaling to reset the normalization of the distributions to 1.

4.2.1 Theory response matrices

The two-dimensional folding procedure is performed by first building a four-dimensional RM for each observable O, similar to those produced for unfolding the data in Sect. 3.3,

$$\mathrm{R}^{\mathrm{p}\to\mathrm{ch}}\left(p_{\mathrm{T}}^{\mathrm{p\ jet}},\ p_{\mathrm{T}}^{\mathrm{ch\ jet}},\ O^{\mathrm{p\ jet}},\ O^{\mathrm{ch\ jet}}\right), \tag{4.6}$$

where the superscript p designates a parton-level quantity, while ch designates a charged-hadron-level one. The parton-level dimensions are set so that they match the same parameters as given in Sect. 4.1, while the charged-hadron-level dimensions are set the same as the

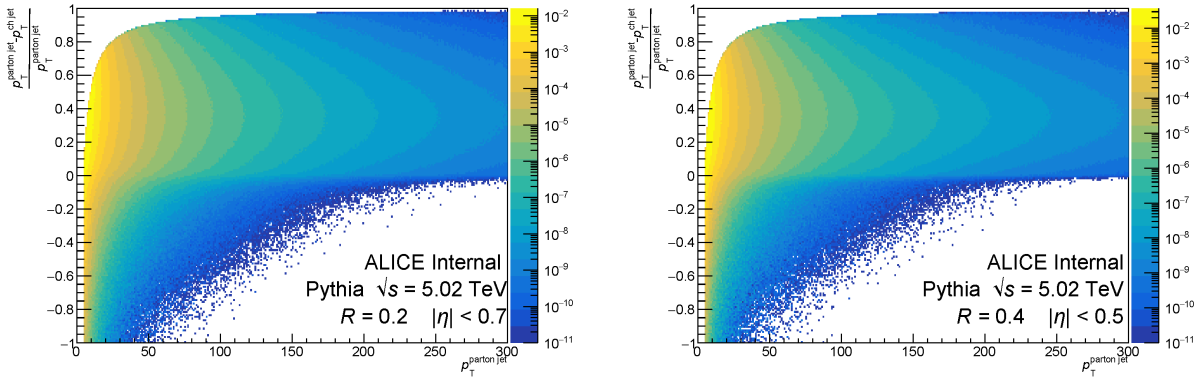


Figure 4.1: $p_{\mathrm{T}}^{\mathrm{jet}}$ residuals between parton and charged-hadron level for $R = 0.2$ (left) and $R = 0.4$ (right) charged-particle jets in PYTHIA 8 [289].

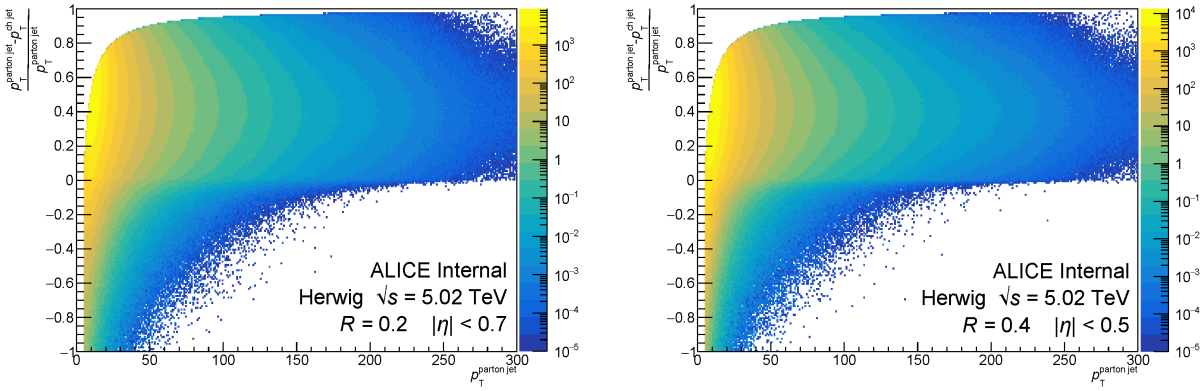


Figure 4.2: $p_{\mathrm{T}}^{\mathrm{jet}}$ residuals between parton and charged-hadron level for $R = 0.2$ (left) and $R = 0.4$ (right) charged-particle jets in Herwig 7 [289].

parton-level ones on the $p_{\mathrm{T}}^{\mathrm{jet}}$ axis, and to a comparable binning for the measured observable on the O axis. For the purpose of checking hadronization and charged effects separately, a second response matrix is also generated for full jets,

$$\mathrm{R}^{\mathrm{p}\to\mathrm{h}}\left(p_{\mathrm{T}}^{\mathrm{p\ jet}},\ p_{\mathrm{T}}^{\mathrm{h\ jet}},\ O^{\mathrm{p\ jet}},\ O^{\mathrm{h\ jet}}\right) \tag{4.7}$$

and uses the same dimensions as the response matrix including charged jet corrections.

Residuals of $p_{\mathrm{T}}^{\mathrm{jet}}$ between parton and charged-hadron level for PYTHIA and Herwig are shown in Figs. 4.1 and 4.2, respectively. Two-dimensional projections of the RMs onto the $p_{\mathrm{T}}^{\mathrm{jet}}$ axes are shown for PYTHIA and Herwig in Figs. 4.3 and 4.4, respectively. Note that the peak in parton-level (vertical) slices of these figures suggests that for the average jet

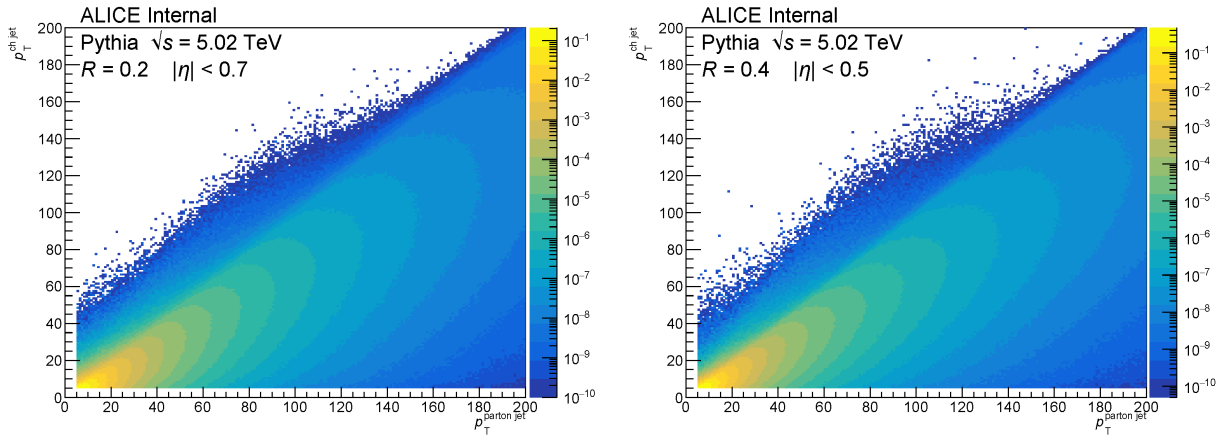


Figure 4.3: $p_{\mathrm{T}}^{\mathrm{jet}}$ response matrix projection for $R = 0.2$ (left) and $R = 0.4$ (right) charged-particle jets in PYTHIA 8 between parton and charged-hadron levels [289].

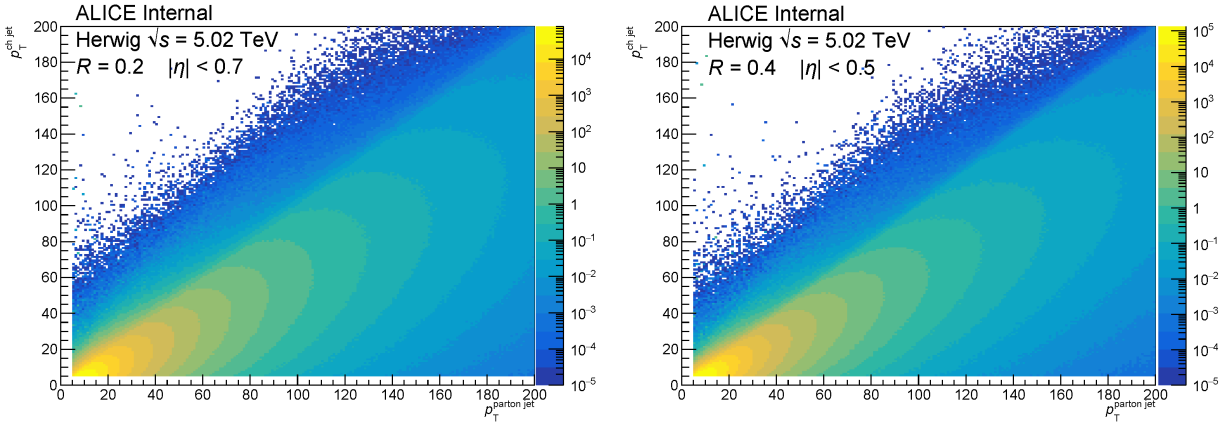


Figure 4.4: $p_{\mathrm{T}}^{\mathrm{jet}}$ response matrix projection for $R = 0.2$ (left) and $R = 0.4$ (right) charged-particle jets in Herwig 7 between parton and charged-hadron levels [289].

roughly a third of the $p_{\mathrm{T,jet}}$ is lost at charged-hadron level, as one would expect in the approximation that one third of particles in a given event are electrically neutral. The jets which appear above the diagonal on the RM projections correspond to the rare case that a jet picks up a fragmented hadron from another process in the event and actually gains p_{T} at charged-hadron level as compared to parton level.

Large off-diagonal components are observed in the RMs, especially at lower $p_{\mathrm{T}}^{\mathrm{jet}}$ and smaller R [289]. These large off-diagonal components introduce a model dependence bias into the folding procedure, which weakens the sensitivity to the original parton-level predictions at low $p_{\mathrm{T}}^{\mathrm{jet}}$ and small R. We account for this model dependence by following the procedure

with two different hadronization models: the Lund string model in PYTHIA 8, and cluster hadronization in Herwig 7, which are both tuned to data (see Sect. 1.1.3).

A large peak also occurs at $\lambda_\alpha^{\mathrm{p\,jet}} = 0$, due to single-parton 'jets' for which $\Delta R_i = 0 \rightarrow \lambda_\alpha = 0$, exactly. Projections of this feature can be seen in Sect. 4.2.2. During hadronization these single-particle jets fragment into jets with more than one constituent, spreading across the $\lambda_\alpha^{\mathrm{ch}}$ spectrum according to the hadronization model. This peak decreases with higher $p_{\mathrm{T}}^{\mathrm{jet}}$ and larger R, but constitutes a direct model bias in the form of an efficiency correction across the entire measured range.

4.2.2 Parton-level matching

Using these tuned MC hadronization models to correct NLL′ predictions is an approximation which may not be valid at low-$p_{\mathrm{T}}^{\mathrm{jet}}$, since these models are typically matched to a LO matrix element and parton shower which has no clear perturbative accuracy. In order for the folding procedure to both properly account for nonperturbative effects and to work without significant statistical fluctuations, the parton-level distributions from these event generators must reasonably match the shape and range of the theory predictions. If there is a large mismatch, then the RM will not give the correct behavior nor an accurate result.

Example λ_α distributions for parton-level jets are shown in Figures 4.5 through 4.6 (for $R = 0.2$) for $10 < p_{\mathrm{T}}^{\mathrm{p\,jet}} < 200$ GeV/c. There is a mismatch between all generators and the NLL′ predictions at low values of $p_{\mathrm{T}}^{\mathrm{p\,jet}}$. For this reason, the distributions sensitive to this region are not reported, since the folding procedure itself breaks down; specifically, the folding approach is only used for $p_{\mathrm{T}}^{\mathrm{ch\,jet}} \geq 60$ GeV/c; the contribution from parton-level jets with $p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 60$ GeV/c are still included in the folding procedure, but their influence is reduced.

4.2.3 Applying corrections

The four-dimensonal RMs (defined by Eq. 4.6 and 4.7) are used to build `RooUnfoldResponse` objects using the RooUnfold framework for ROOT [291]. These objects are then multiplied against the parton-level two-dimensonal ($O^{\mathrm{p\,jet}}$ vs. $p_{\mathrm{T}}^{\mathrm{p\,jet}}$) histograms via `ApplyToTruth()`, and the resulting two-dimensional histograms represent the distributions at charged-hadron level ($O^{\mathrm{ch\,jet}}$ vs. $p_{\mathrm{T}}^{\mathrm{ch\,jet}}$). This is the so-called "forward-folding" procedure.

MPI cannot be accounted for in a folding-based approach, because attempting to match between two iterations of PYTHIA – one with MPI on and the other with MPI off – is non-trivial even when starting with identical random seeds. The MPI setting adds additional (low-momentum) perturbative scatterings to events, which causes the generator to loop through interactions faster than with MPI off. Therefore, to account for corrections due to MPI, a simple bin-by-bin correction is used, whereby a scaling histogram is created for each individual result in order to correct for MPI. Two-dimensional histograms of O versus $p_{\mathrm{T}}^{\mathrm{jet}}$, created with both MPI off and on, are projected onto the O-axis for the proper $p_{\mathrm{T}}^{\mathrm{jet}}$ range, and a ratio of these histograms is taken (MPI-on divided by MPI-off). This one-dimensional

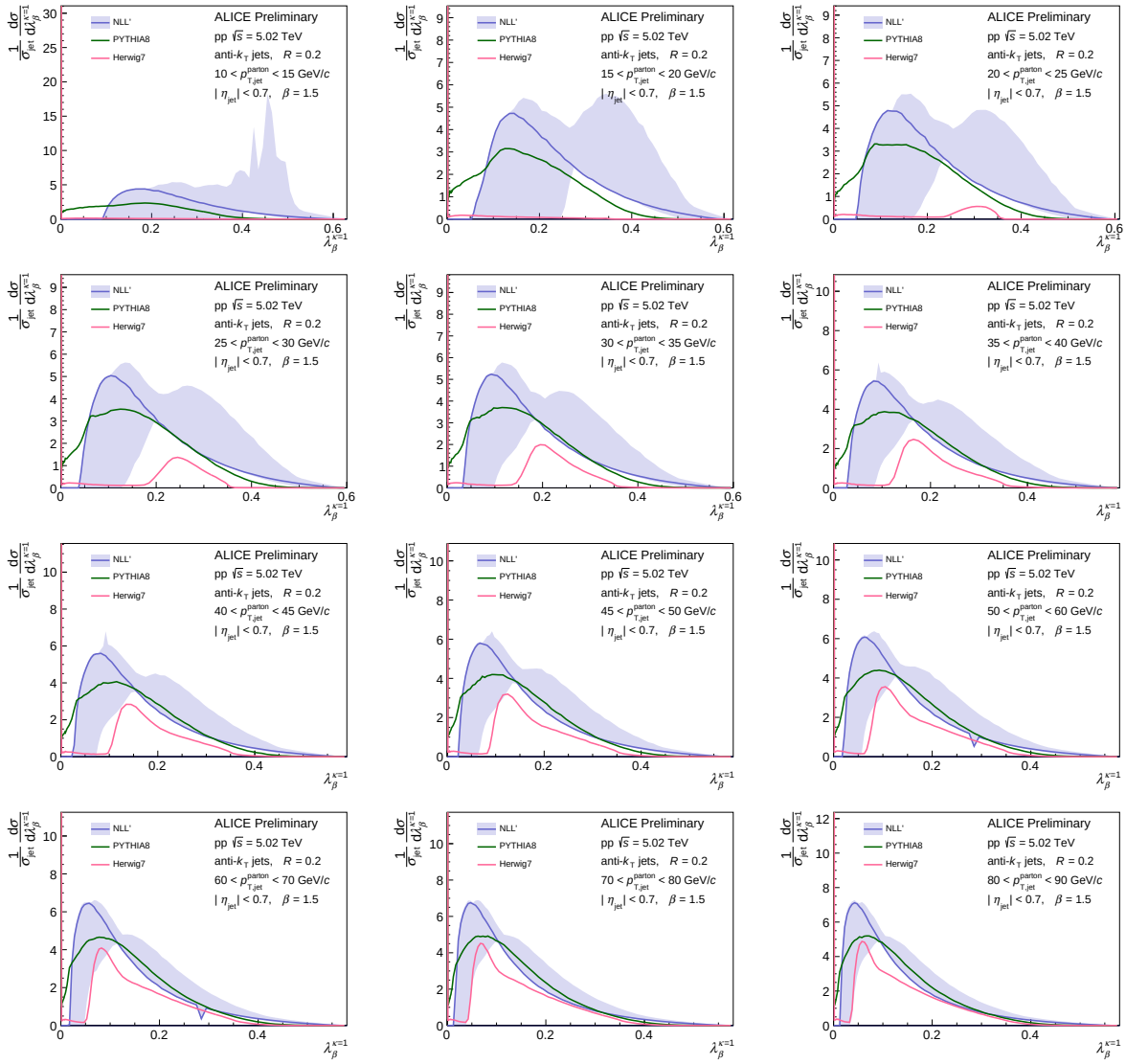


Figure 4.5: Comparison of the $\lambda_{1.5}$ distributions according to parton-level theory predictions with the selected event generators for $R = 0.2$ jets the 10-90 GeV/c range [289].

histogram is then multiplied by each of the 15 scale variations for each observable. Since this procedure does not preserve the normalization to 1 required for a self-normalized cross section, the histograms are rescaled by the integral, here also incorporating the proper bin width correction [using `h.Scale(1/h.Integral("width"))`].

Example ratios of the histograms at each level of the procedure are plotted in Fig. 4.7 to explore the size of each individual correction (hadronization, charge selection, and MPI scaling). These checks are the only place where the full hadron distributions are used, forward-folded using the RMs given by Eq. 4.7. The parton-to-hadron-level corrections are

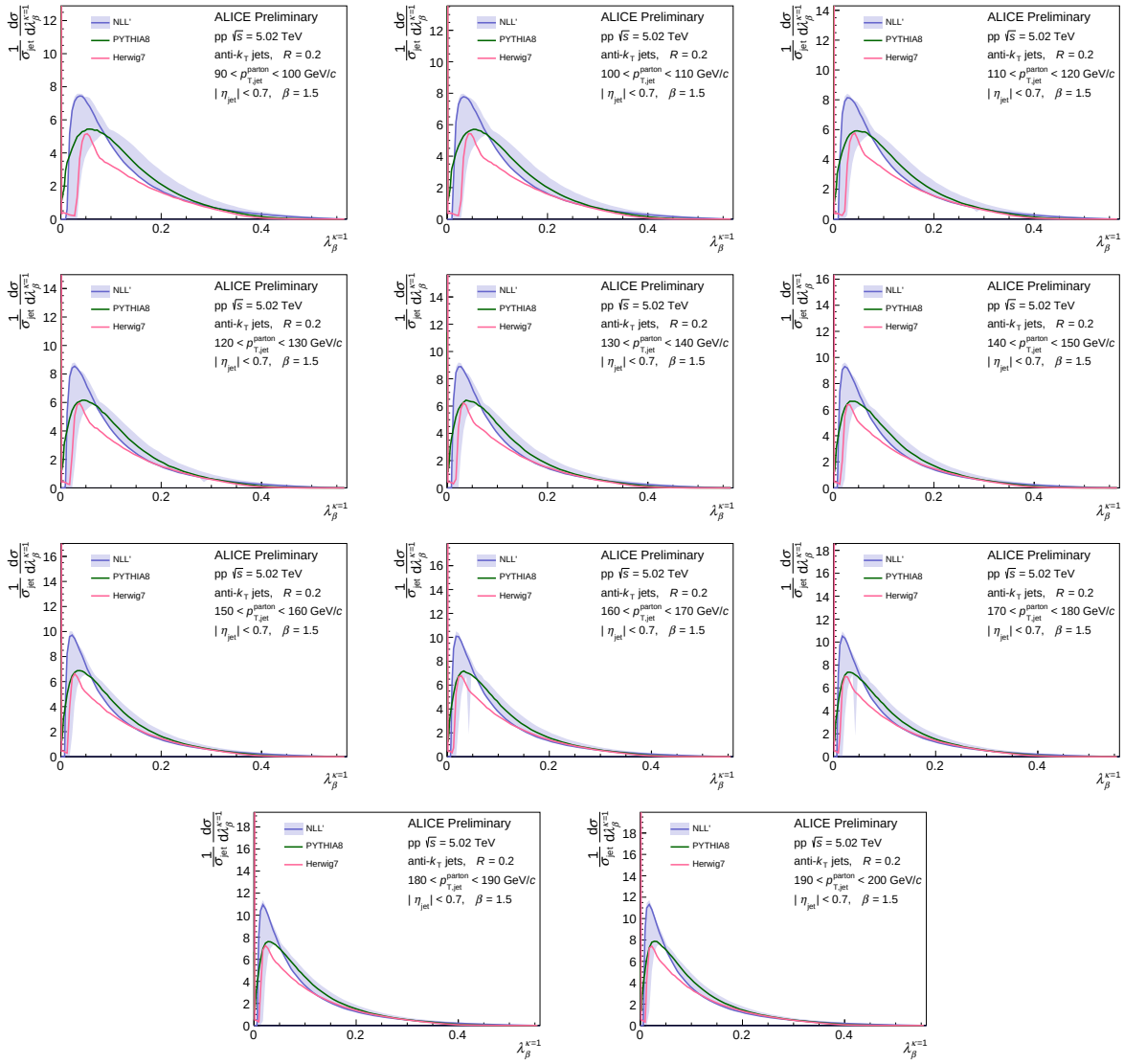


Figure 4.6: Comparison of the $\lambda_{1.5}$ distributions according to parton-level theory predictions with the selected event generators for $R = 0.2$ jets the 90-200 GeV/c range [289].

typically large (factor 2 or more) and result in a broadening of the jet, meaning λ_α becomes larger. The mapping of hadron-to-charged level is also typically large, and results in a corresponding narrowing of the jet. Together, these corrections mostly cancel to the $\sim 50\%$ level, though a large MC bias is still introduced from the two different approaches. The MPI correction is observed to be considerably smaller, rarely exceeding 5 to 10%.

The systematic uncertainty due to this model bias is displayed in the final results by reporting the folded theory distributions with both PYTHIA and Herwig corrections. These two MC approaches do not span the space of possible models, but at least give an estimate

on the size of the differences that are introduced when two reasonable models are used for the forward-folding procedure.

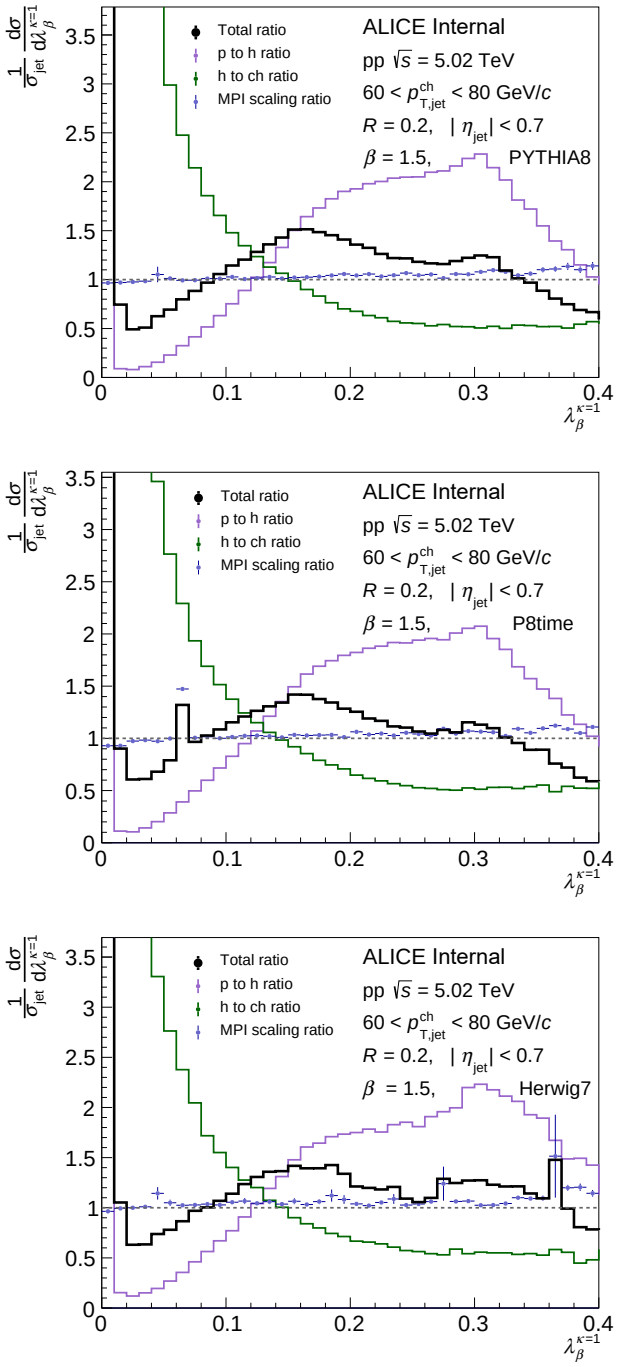


Figure 4.7: Example ratios depicting the size of $\lambda_{1.5}$ correction between parton, hadron, charged-hadron, and charged-hadron with MPI levels, for default PYTHIA 8 (top), PYTHIA 8 with lower `TimeShower` cutoff (middle), and Herwig 7 (bottom). The matched jets use radius parameter $R = 0.2$ and are projected for $p_{\mathrm{T}}^{\mathrm{ch\ jet}} \in [60, 80]$ GeV/c.

4.3 Nonperturbative shape function

The second correction procedure – applied only for the jet angularities – is a multi-step process. First, a nonperturbative shape function $F(k)$ [234] is used to account for hadronization and underlying event effects. Secondly, a two-dimensional folding (with four-dimensional RM) is applied to the corrected distributions to account for charged jet effects, followed by a similar scaling to reset the normalization of the distributions to 1, as required by the definition of the vertical axis. The shape function is defined as [235]

$$F(k) = \frac{4k}{\Omega_\alpha^2} \exp\left(-\frac{2k}{\Omega_\alpha}\right), \tag{4.8}$$

where Ω_α is described by a single parameter $\Omega = \mathcal{O}(1\ \mathrm{GeV}/c)$ obeying the scaling relation

$$\Omega_\alpha = \Omega/(\alpha - 1) \tag{4.9}$$

and is expected to hold universally for hadronization corrections (but not for underlying event corrections). To correct the parton level calculations to hadron level, this shape function is convolved with the perturbative (parton-level) jet angularity distribution via numerical integration over argument k [235],

$$\frac{\mathrm{d}\sigma}{\mathrm{d}p_\mathrm{T}^\mathrm{jet}\mathrm{d}\lambda_\alpha} = \int F(k)\ \frac{\mathrm{d}\sigma_\mathrm{pert}}{\mathrm{d}p_\mathrm{T}^\mathrm{jet}\mathrm{d}\lambda_\alpha}\left[\lambda_\alpha - \lambda_\alpha^\mathrm{shift}(k)\right]\ \mathrm{d}k, \tag{4.10}$$

where the shift term $\lambda_\alpha^\mathrm{shift}(k)$ is either [231, 235]

$$\begin{aligned}\lambda_\alpha^\mathrm{shift}(k) &= \frac{k}{p_\mathrm{T}^\mathrm{jet} R}\ \text{(ungroomed)}, \quad \text{or} \\ &= z_\mathrm{cut}^{1-\alpha}\left(\frac{k}{p_\mathrm{T}^\mathrm{jet} R}\right)^\alpha\ \text{(SD groomed, with } \beta = 0).\end{aligned} \tag{4.11}$$

Since the nonperturbative parameter Ω is not fundamentally known, four values (0.2, 0.4, 0.8, and 2) are chosen to observe the different shifting effects. These distributions are then corrected once more using a similar PYTHIA 8 folding procedure as described in Sect. 4.2 above to account for the effects of only reconstructing charged jets, which is dominated by a shift in $p_\mathrm{T}^\mathrm{jet}$. For this, an RM is produced similar to the ones described in Eq. 4.6 and 4.7,

$$\mathrm{R}^{\mathrm{h}\to\mathrm{ch}}\left(p_\mathrm{T}^{\mathrm{h\ jet}},\ p_\mathrm{T}^{\mathrm{ch\ jet}},\ \lambda_\alpha{}^{\mathrm{h\ jet}},\ \lambda_\alpha^{\mathrm{ch\ jet}}\right). \tag{4.12}$$

The shape function approach, specifically the scaling given in Eq. 4.9, is not fully justified in the groomed case, due to a change in leading power corrections and β-dependence of the shape function [91, 92]. However, reasonable agreement has been seen at larger R and higher $p_\mathrm{T}^\mathrm{jet}$, so this comparison provides a means to test its accuracy with smaller R and lower $p_\mathrm{T}^\mathrm{jet}$.

Figures 4.8 and 4.9 show example comparisons between final-state, full hadron jets in PYTHIA 8 (using the same tune settings as described above) and the NLL′ predictions convolved with $F(k)$ for these four values of Ω. Since this convolution procedure is done with discrete bins, by Eq. 4.11 above, the first bin must always be zero in the convolved theory. The lowest bin is therefore split in the theoretical predictions ([0, 0.001]) into much finer bins, [0, 0.0001, 0.0002, ..., 0.0009, 0.001], which are then combined before folding with the RM, such that the loss of the first bin is insignificant; otherwise, statistical fluctuations result in the final folded distributions due to the missing phase space at $F(k)$ level. This happens in both the groomed and ungroomed distributions, but the shape of the distributions in the groomed case (and especially at high $p_{\mathrm{T}}^{\mathrm{jet}}$ and large α) is much more strongly peaked towards 0, where the effect is significant, whereas most λ_α predictions for the ungroomed distributions are already 0 at these low values of λ_α.

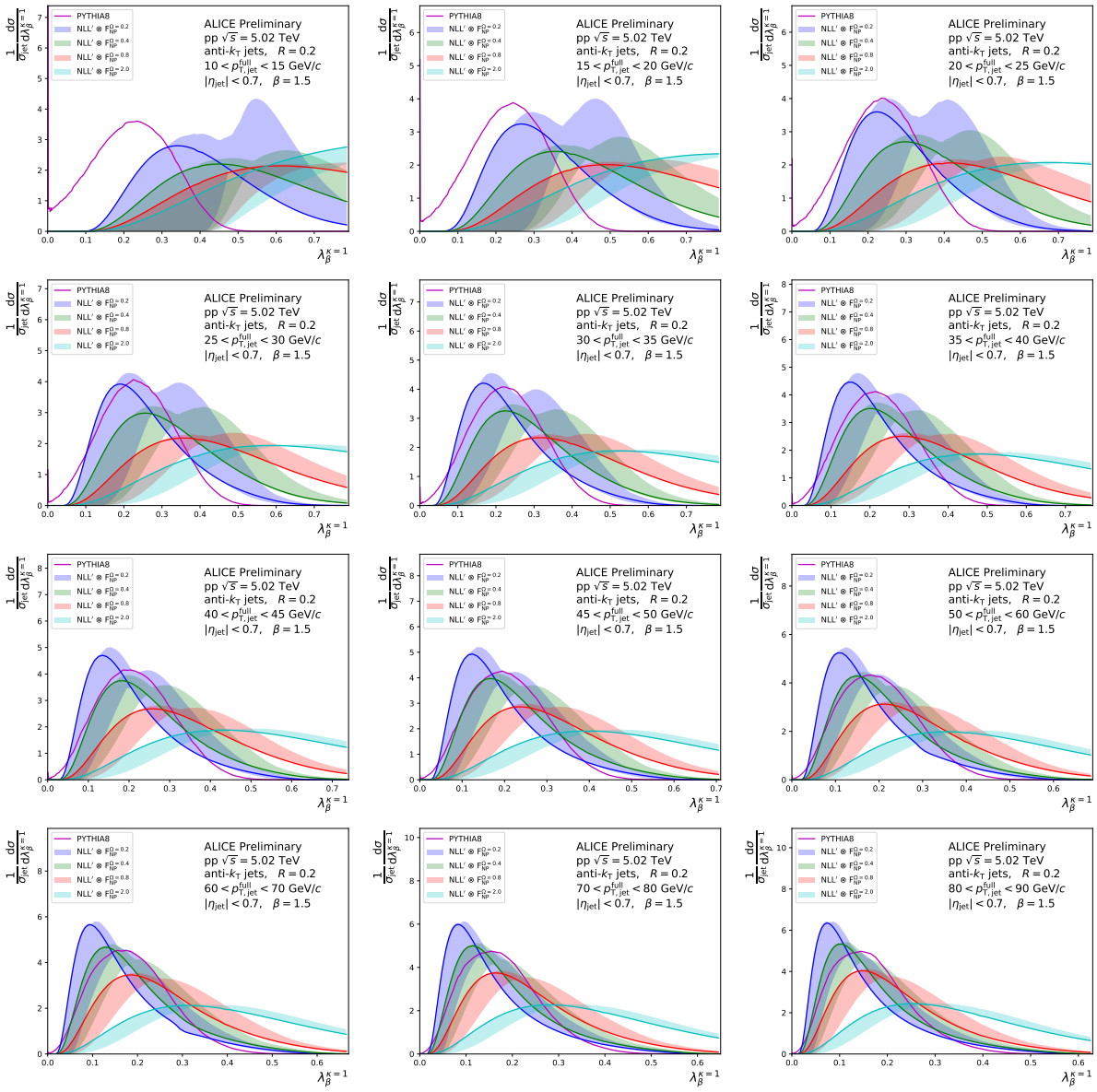


Figure 4.8: Comparison of the parton-level theory predictions for $\lambda_{1.5}$ with nonperturbative shape function convolution versus PYTHIA 8 for $R = 0.2$ jets in the 10-90 GeV/c range.

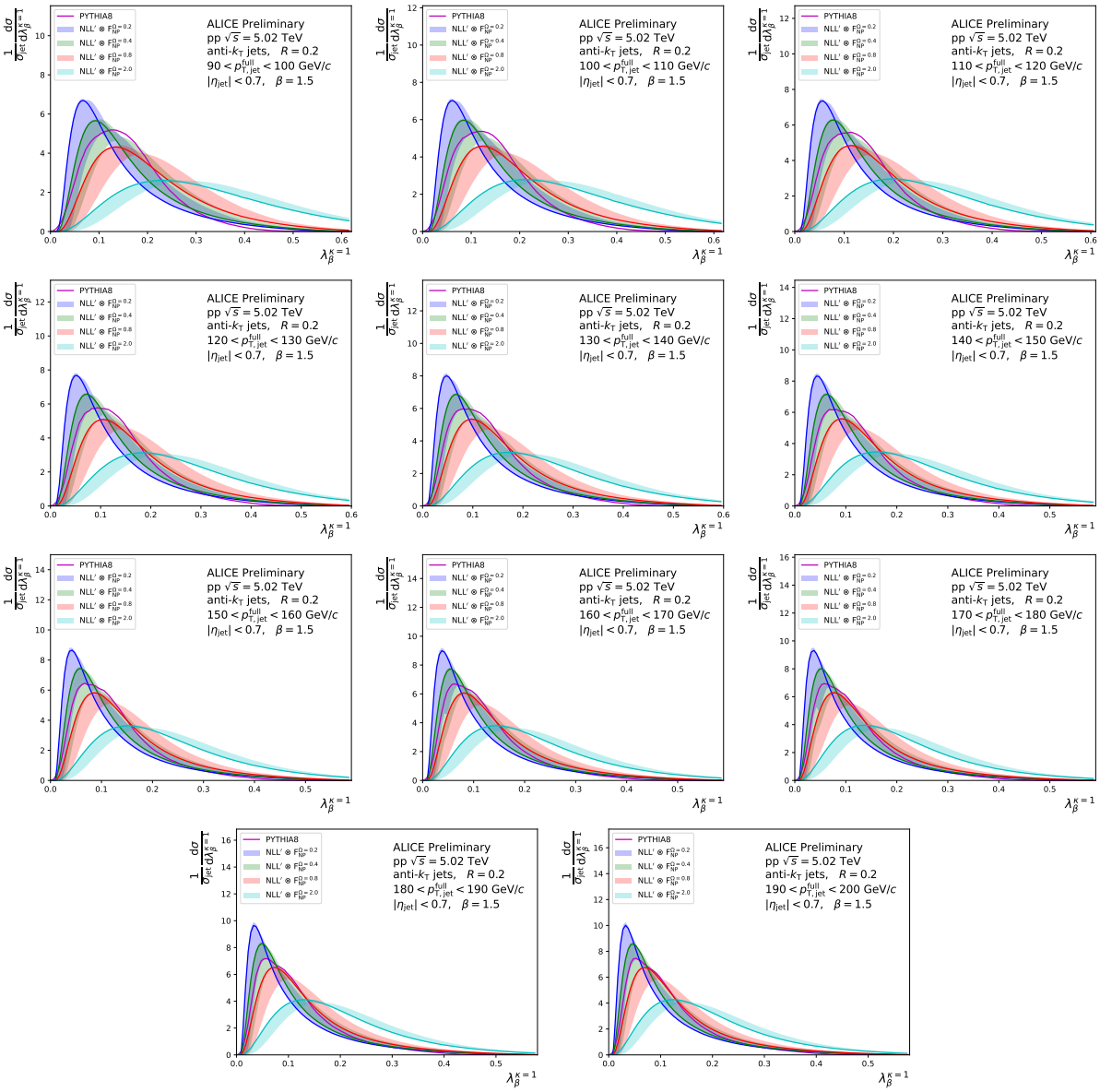


Figure 4.9: Comparison of the parton-level theory predictions for $\lambda_{1.5}$ with nonperturbative shape function convolution versus PYTHIA 8 for $R = 0.2$ jets in the 90-200 GeV/c range.

Chapter 5

Results

Following the analysis procedure outlined in Ch. 3 and the correction of theory predictions as discussed in Ch. 4, this chapter presents results on the jet substructure observables motivated in Ch. 1, and specifically stated in Sect. 1.5. The results include jet angularity and mass, and other observables measured by ALICE which contribute to the physics conclusions of this thesis. Measurements of the groomed jet radius R_{g} (or θ_{g}) and momentum splitting fraction z_{g} which have been analyzed by ALICE colleagues are first presented in Sect. 5.1, including comparisons to theoretical predictions as part of this work. This is followed by the groomed and ungroomed jet angularities in pp collisions in Sect. 5.2, with theoretical predictions also prepared following Ch. 4. The jet angularities are reported for D^0-tagged jets in Sect. 5.2.1. Finally, the jet angularities and jet mass are compared between pp and Pb–Pb data in Sect. 5.3. Results are briefly discussed, with a more detailed discussion of the results and closing remarks delivered in Ch. 6.

The observable O distributions are reported as self-normalized differential cross sections,

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}O} \equiv \frac{1}{N_{\mathrm{jets}}}\frac{\mathrm{d}N_{\mathrm{jets}}}{\mathrm{d}O} \text{ (ungroomed)}, \quad \text{or} \quad \frac{1}{\sigma_{\mathrm{inc}}}\frac{\mathrm{d}\sigma}{\mathrm{d}O_{\mathrm{g}}} \equiv \frac{1}{N_{\text{inc jets}}}\frac{\mathrm{d}N_{\text{gr jets}}}{\mathrm{d}O_{\mathrm{g}}} \text{ (groomed)}, \tag{5.1}$$

where N_{jets} is the number of jets within a given $p_{\mathrm{T}}^{\text{ch jet}}$ range and σ is the corresponding cross section. For the groomed case, some jets are removed by the grooming procedure ("untagged"), and therefore two different quantities are defined: $N_{\text{gr jets}}$, the number of jets which have at least one splitting satisfying the SD condition, and $N_{\text{inc jets}}$, the total number of inclusive jets, with both $N_{\text{gr jets}}$ and $N_{\text{inc jets}}$ being within the given $p_{\mathrm{T}}^{\text{ch jet}}$ range. $N_{\text{inc jets}}$ is corrected from the raw data by including the number of untagged jets as an extra bin in the unfolding procedure. σ_{inc} is the cross section corresponding to the latter inclusive quantity. For the ungroomed case, $N_{\text{inc jets}} = N_{\mathrm{jets}}$ and $\sigma = \sigma_{\mathrm{inc}}$, so the redundant labels are dropped. It is useful to normalize the groomed differential cross section by the number of inclusive jets since the groomed jet angularities are a property of the inclusively-measured jet population and are thus typically normalized as such in theoretical calculations [231].

5.1 Groomed jet splitting observables in vacuum

This section reports ALICE results of the groomed jet radius R_{g} (or θ_{g}) and momentum splitting fraction z_{g} using SD and dynamical grooming (see Sect. 1.3.3) in pp collisions at $\sqrt{s} = 5.02$ TeV, which were recently published in Ref. [300]. These results build off the first fully-unfolded measurements of these observables in Pb–Pb collisions at the same center-of-mass energy, which were published in Ref. [103].

Results are reported for inclusive charged-particle jets with transverse momentum $60 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 80$ GeV/c using two different grooming algorithms: SD grooming and dynamical grooming, as defined in Sect. 1.3.3. Variations of the grooming parameters probe the impact of collinear radiation on jet substructure. These results are compared to MC predictions as well as perturbative calculations that include resummation of large logarithms at all orders in the strong coupling constant [115, 117].

The SD z_{g} and θ_{g} distributions are shown with comparisons to PYTHIA 8 in Fig. 5.1, while θ_{g} is compared to theoretical predictions in Fig. 5.2. Theoretical comparisons for z_{g} are not available, as pQCD calculations require a minimum $\theta_{\mathrm{g}} > \theta_{\mathrm{g}}^{\mathrm{cut}}$ to remain sufficiently perturbative [298]; z_{g} was not measured differentially in θ_{g} for the analysis reported here. Measured distributions of z_{g} and θ_{g} with dynamical grooming are shown in Fig. 5.3, and theoretical comparisons are given in Fig. 5.4. Good agreement is found between the theoretical predictions and the data for all grooming settings considered. Results are discussed further in Sect. 6.1.

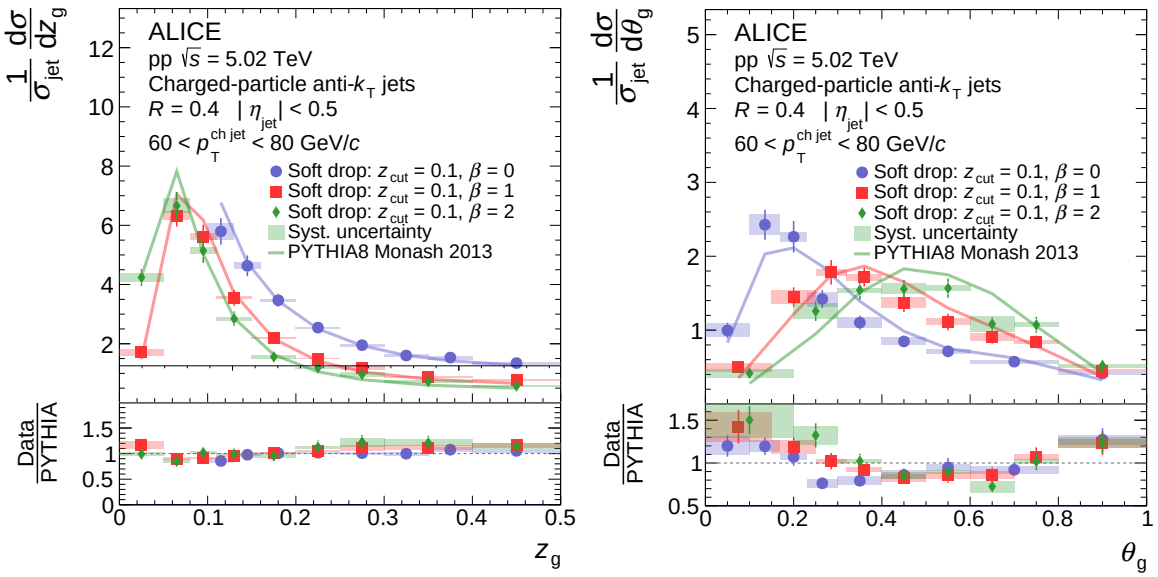


Figure 5.1: ALICE measurements of z_{g} (left) and θ_{g} (right) distributions in pp collisions at $\sqrt{s} = 5.02$ TeV with SD for three values of the grooming parameter β, compared with PYTHIA 8 Monash 2013 [37, 38] calculations [300].

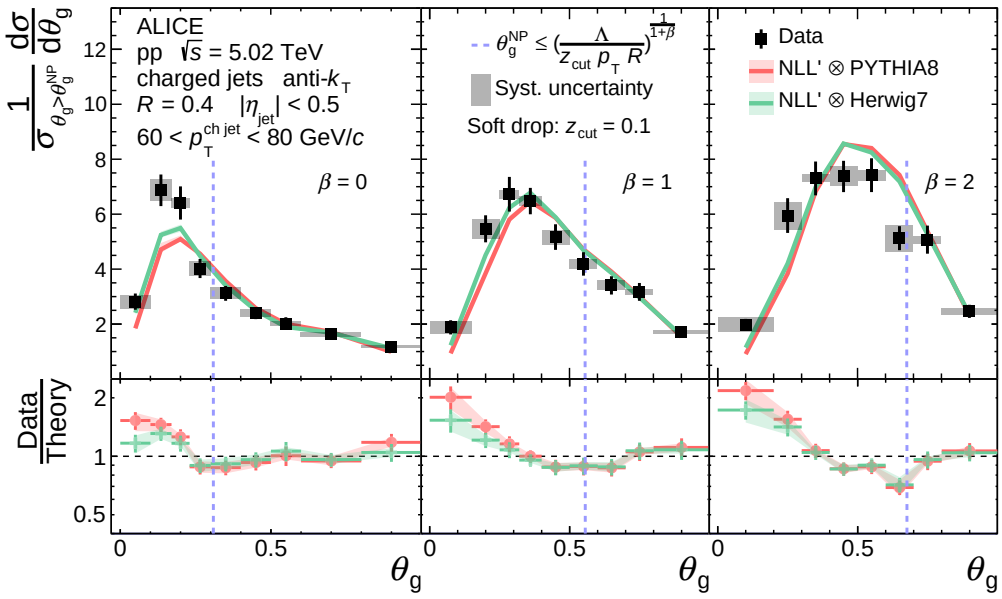


Figure 5.2: ALICE measurements of θ_g distributions with SD, compared with pQCD predictions including nonperturbative corrections. The integral of the perturbative region, $\theta_g > \theta_g^{NP}$ (to the right of the dashed vertical blue line; see Sect. 4.1), is set at unity. The nonperturbative scale is taken to be $\Lambda = 1$ GeV/c [300].

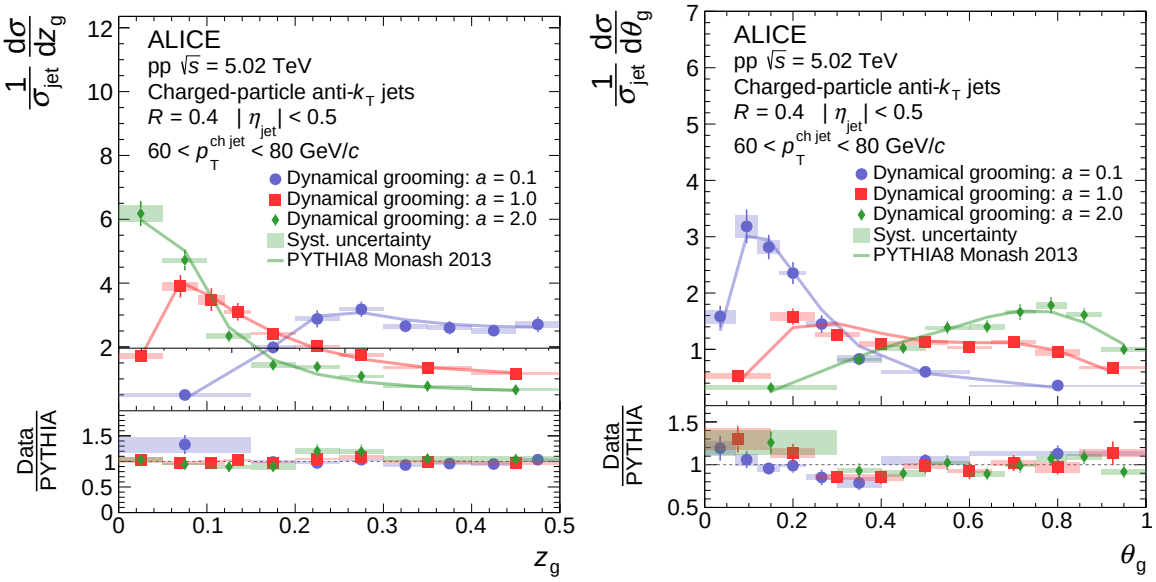


Figure 5.3: ALICE measurements of z_g (left) and θ_g (right) distributions in pp collisions at $\sqrt{s} = 5.02$ TeV with dynamical grooming for three values of the grooming parameter a, compared with PYTHIA 8 Monash 2013 [37, 38] calculations [300].

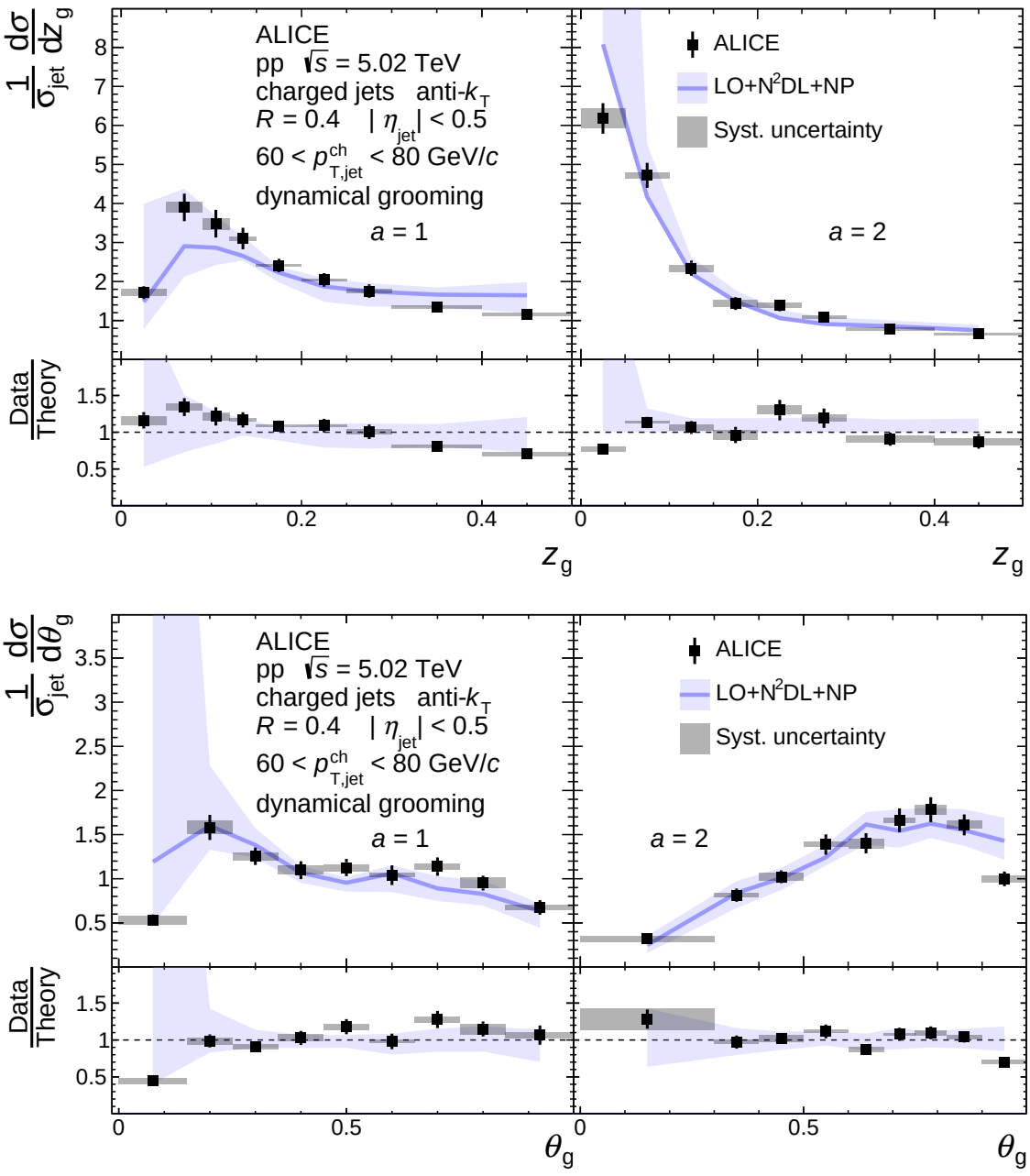


Figure 5.4: ALICE measurements of z_g (top) and θ_g (bottom) distributions in pp collisions at $\sqrt{s} = 5.02$ TeV with dynamical grooming for two values of the grooming parameter a as compared with pQCD calculations [300].

5.2 Jet angularities in vacuum

This section reports ALICE results of the groomed and ungroomed jet angularities λ_α (see Sect. 1.3.2) in pp collisions at $\sqrt{s} = 5.02$ TeV, which were recently published in Ref. [301]. SD grooming is used with $z_{\rm cut} = 0.2$ and $\beta = 0$ in order to minimize mistagged splittings when comparing to Pb–Pb collisions at the same center-of-mass energy [214], the results of which are shown in Sect. 5.3.

The ungroomed jet angularity distributions are shown in Fig. 5.5 and Fig. 5.6 for $R = 0.4$ and $R = 0.2$, respectively. As α increases, the distributions skew towards small λ_α, since θ_i is smaller than unity. For larger R, the distributions are narrower than for smaller R, as expected due to the collinear nature of jet fragmentation. For small R and low $p_{\rm T}^{\rm ch\ jet}$ there is a visible peak at $\lambda_\alpha = 0$, which is due to single particle jets. These distributions are compared to PYTHIA 8 Monash 2013 [37, 38] and Herwig 7 (default tune) [42, 43] from truth-level projections of the respective response matrices, with jet reconstruction assigning tracks the $\pi^\pm$ meson mass as in the measured data. These comparisons show deviations up to approximately $+50\%(-30\%)$. The largest deviations are for small values of λ_α, where nonperturbative physics becomes significant (see Ch. 4 for discussion).

The groomed jet angularity distributions for $z_{\rm cut} = 0.2$ and $\beta = 0$ are shown in Fig. 5.7 for $R = 0.4$ and Fig. 5.8 for $R = 0.2$. Note that these distributions are shown on a logarithmic scale due to the distributions being more strongly peaked and falling faster with λ_α as compared to the ungroomed distributions. The groomed jet angularities have significantly smaller values than the ungroomed jet angularities, due to the removal of soft wide-angle radiation. The fraction of "untagged" jets, those that do not contain a splitting which passes the SD condition, ranges from 10 to 12%. Unlike the ungroomed jet angularities, which are normalized to unity, the groomed jet angularities are normalized to the SD tagging fraction. Since the tagging rate is fairly large, the measured distributions are therefore normalized close to unity. PYTHIA and Herwig describe the groomed jet angularities slightly better than the ungroomed jet angularities, with most deviations seen in the ungroomed distributions improving by 10–20% in the groomed case. This is largely due to the reduction of nonperturbative effects in groomed jet distributions. Comparing to the two MC generators, the data are in slightly better agreement with Herwig 7 than with PYTHIA 8, especially for $R = 0.4$.

The data cover a wide range of α and multiple R down to low $p_{\rm T}$, and therefore are subject to varying influence from nonperturbative effects. Accordingly, these data can be used to study nonpertubative effects. The level and location of the disagreements with PYTHIA and Herwig provide further constraints on nonperturbative effects in MC event generators. Moreover, the comparison of the groomed and the ungroomed jet angularities with MC event generators allows direct sensitivity to radiation that was groomed away, which is highly nonperturbative.

Figure 5.9 shows comparisons of the measured ungroomed jet angularities to the folded theoretical predictions for $60 < p_{\rm T}^{\rm ch\ jet} < 80$ GeV/c, for both $R = 0.2$ (top) and $R = 0.4$ (bottom) and for $\alpha = 1.5$ (left), 2 (middle), and 3 (right). Figure 5.10 shows the corresponding

comparisons for the groomed jet angularities. The same comparisons for $80 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 100$ GeV/c are shown in Figs. 5.11 and 5.12. Predictions for the $\alpha = 1$ distributions are not currently available due to enhanced sensitivity to soft quark recoil, which requires a different factorization [99, 302].

Similar to Fig. 5.2, a dashed vertical line is drawn as a rough estimate for the division of perturbative- and nonperturbative-dominated regions, via Eq. 4.2 or Eq. 4.3 with nonperturbative scale $\Lambda = 1$ GeV/c and the mean $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ for each interval; see Sect. 4.3 for details on these theory predictions. Note that the transition from values of λ_α which are dominated by perturbative versus nonperturbative physics is actually smooth, and this vertical line is merely intended as a visual guide. The nonperturbative-dominated region of the jet angularities is denoted as $\lambda_\alpha^{\mathrm{NP}}$.

Since the integral for all of the distributions in Fig. 5.5 through Fig. 5.8 is fixed at unity by construction, disagreement in the nonperturbative-dominated region induces disagreement in the perturbative-dominated region. Discrepancy in the nonperturbative region is expected, since as $\lambda_\alpha \to 0$, higher-order terms in the perturbative expansion become significant. Therefore, for these theoretical comparisons, the distributions are normalized such that the integral above $\lambda_\alpha^{\mathrm{NP}}$ is unity.

The theoretical predictions using the nonperturbative shape function approach are also compared to the ungroomed and groomed predictions. The comparisons for $40 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 60$ GeV/c are shown in Figs. 5.13 and 5.14, for $60 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 80$ GeV/c in Figs. 5.15 and 5.16, and for $80 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 100$ GeV/c in Figs. 5.17 and 5.18. As discussed in Sect. 4.3, the shape function approach is not fully justified in the groomed case [91, 92]; nevertheless, reasonable agreement is observed. Since this shape convolution does not require matching to MC models at the parton level, the comparisons are extended to the $40 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 60$ GeV/c interval, but below this the perturbative accuracy of the parton-level predictions is insufficient for rigorous comparisons.

Despite the pQCD predictions having an arguably higher perturbative accuracy than that of the MC simulations, large deviations are seen in many of the NLL′ distributions at low λ_α, which are not observed in MC. This is expected, and is due to the breakdown of the perturbative prediction within the nonperturbative region, which the correction procedures of Ch. 4 cannot account for. The MC predictions are able to perform better since they are tuned to experimental data in both the perturbative and nonperturbative regions, while higher-order perturbative calculations would be needed to reduce the size of the nonperturbative region in SCET calculations.

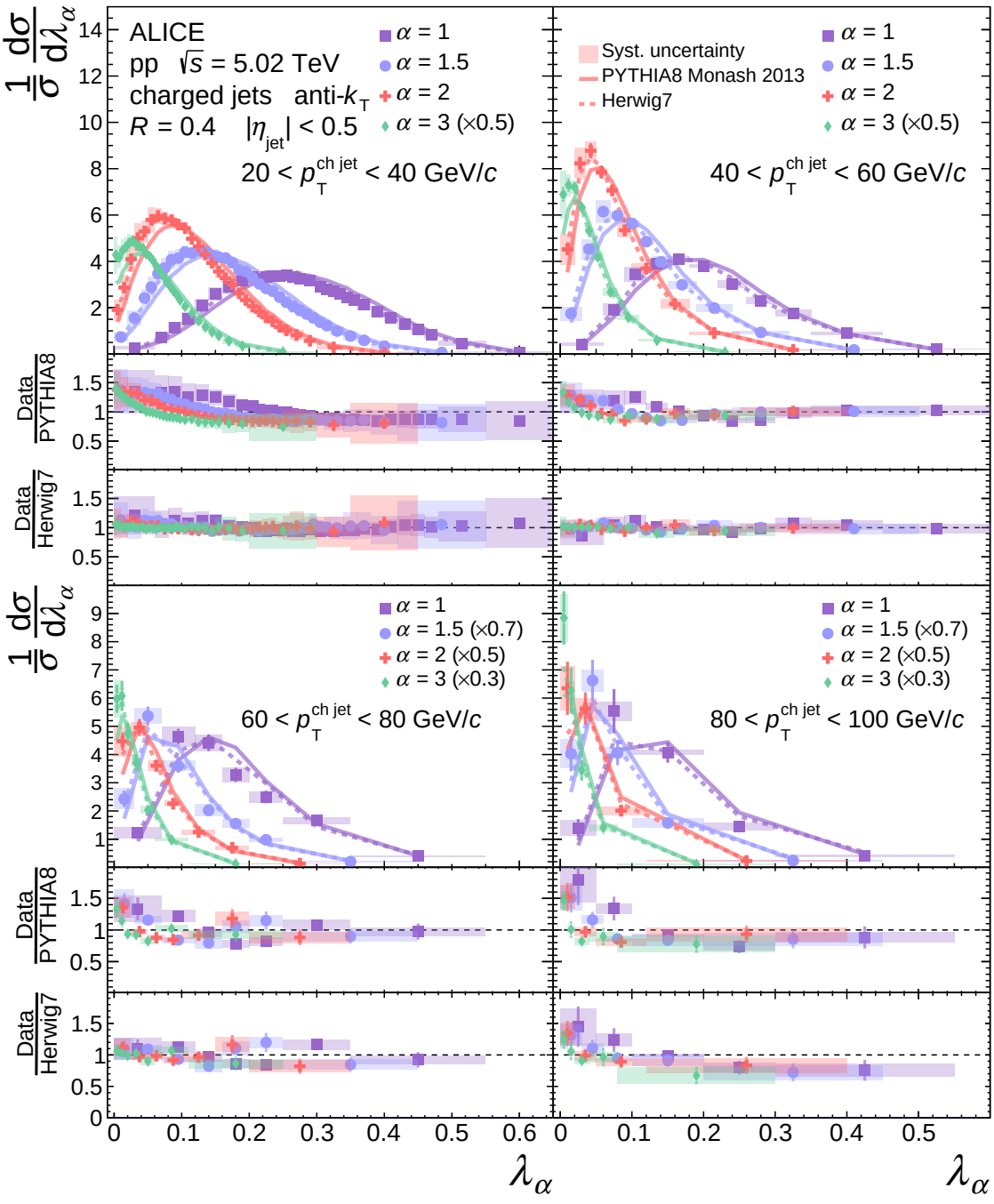


Figure 5.5: Comparison of ungroomed jet angularities λ_α in pp collisions for $R = 0.4$ to MC predictions using PYTHIA 8 and Herwig 7, as described in the text. Four equally-sized $p_{\mathrm{T}}^{\text{ch jet}}$ intervals are shown, with edges ranging between 20 and 100 GeV/c. The distributions are normalized to unity [301].

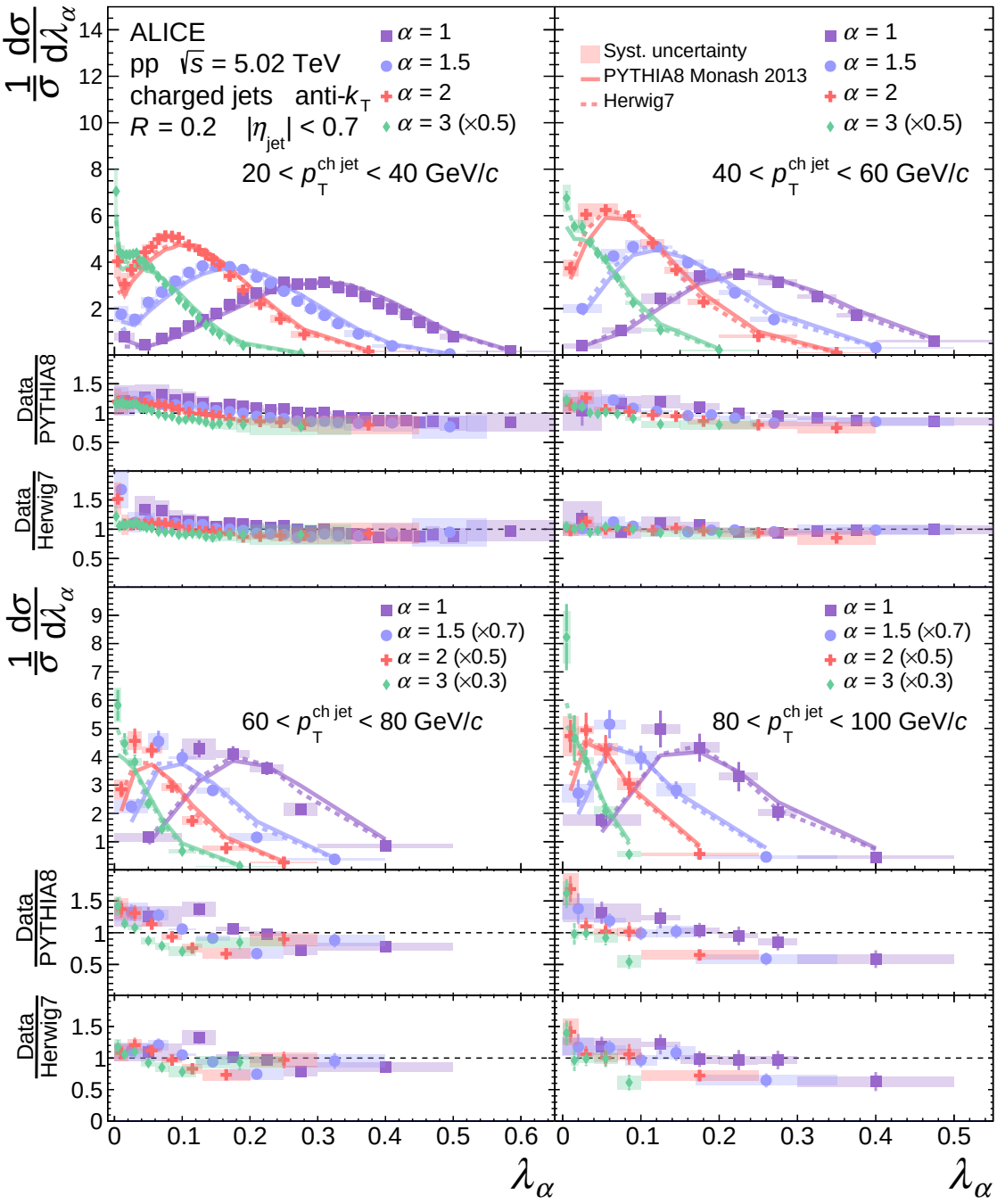


Figure 5.6: Comparison of ungroomed jet angularities λ_α in pp collisions for $R = 0.2$ to MC predictions using PYTHIA 8 and Herwig 7, as described in the text. Four equally-sized $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ intervals are shown, with edges ranging between 20 and 100 GeV/c. The distributions are normalized to unity [301].

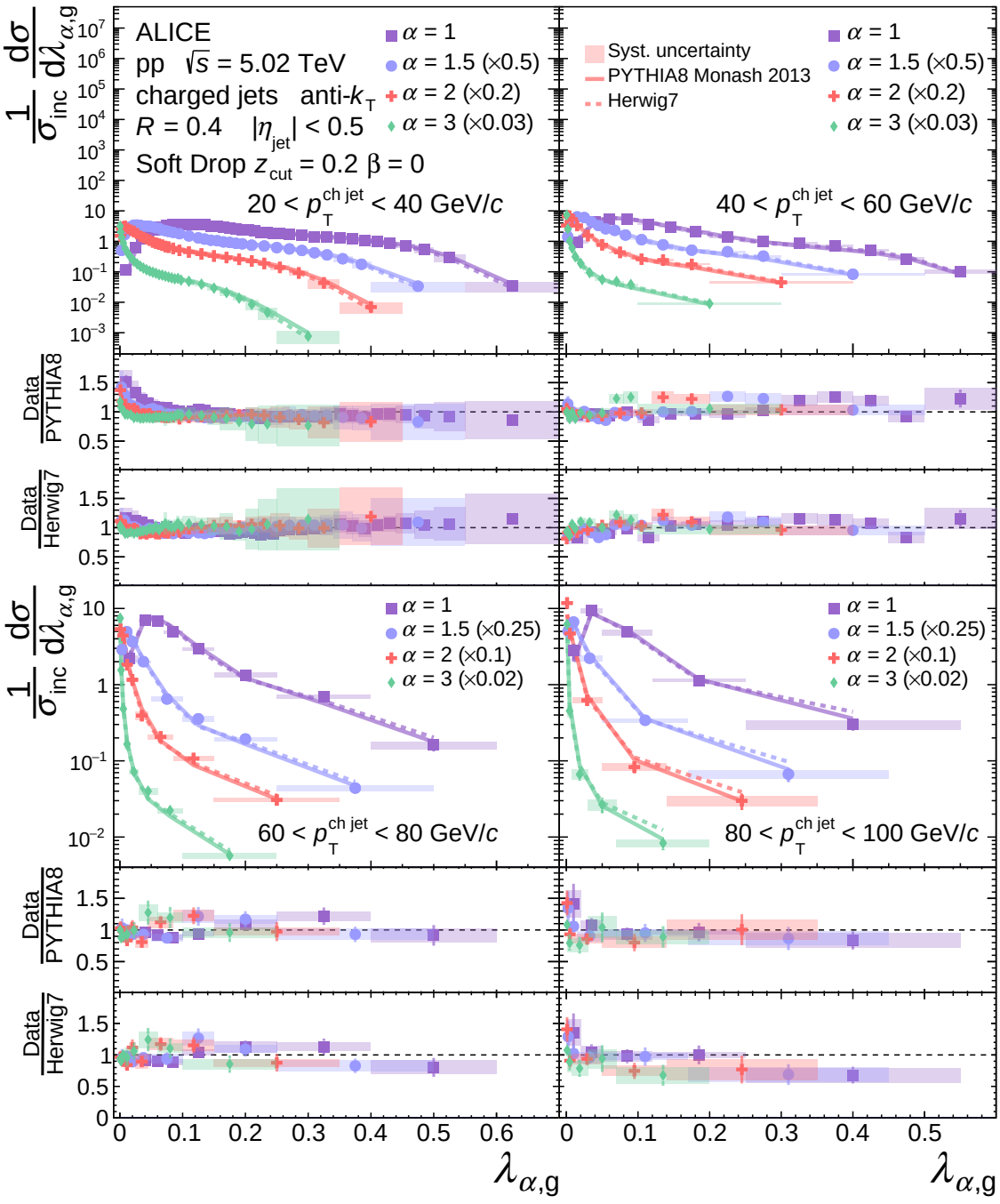


Figure 5.7: Comparison of groomed jet angularities $\lambda_{\alpha,g}$ in pp collisions for $R = 0.4$ to MC predictions using PYTHIA 8 and Herwig 7, as described in the text. Four equally-sized $p_T^{ch\ jet}$ intervals are shown between 20 and 100 GeV/c. The distributions are normalized to the groomed jet tagging fraction [301].

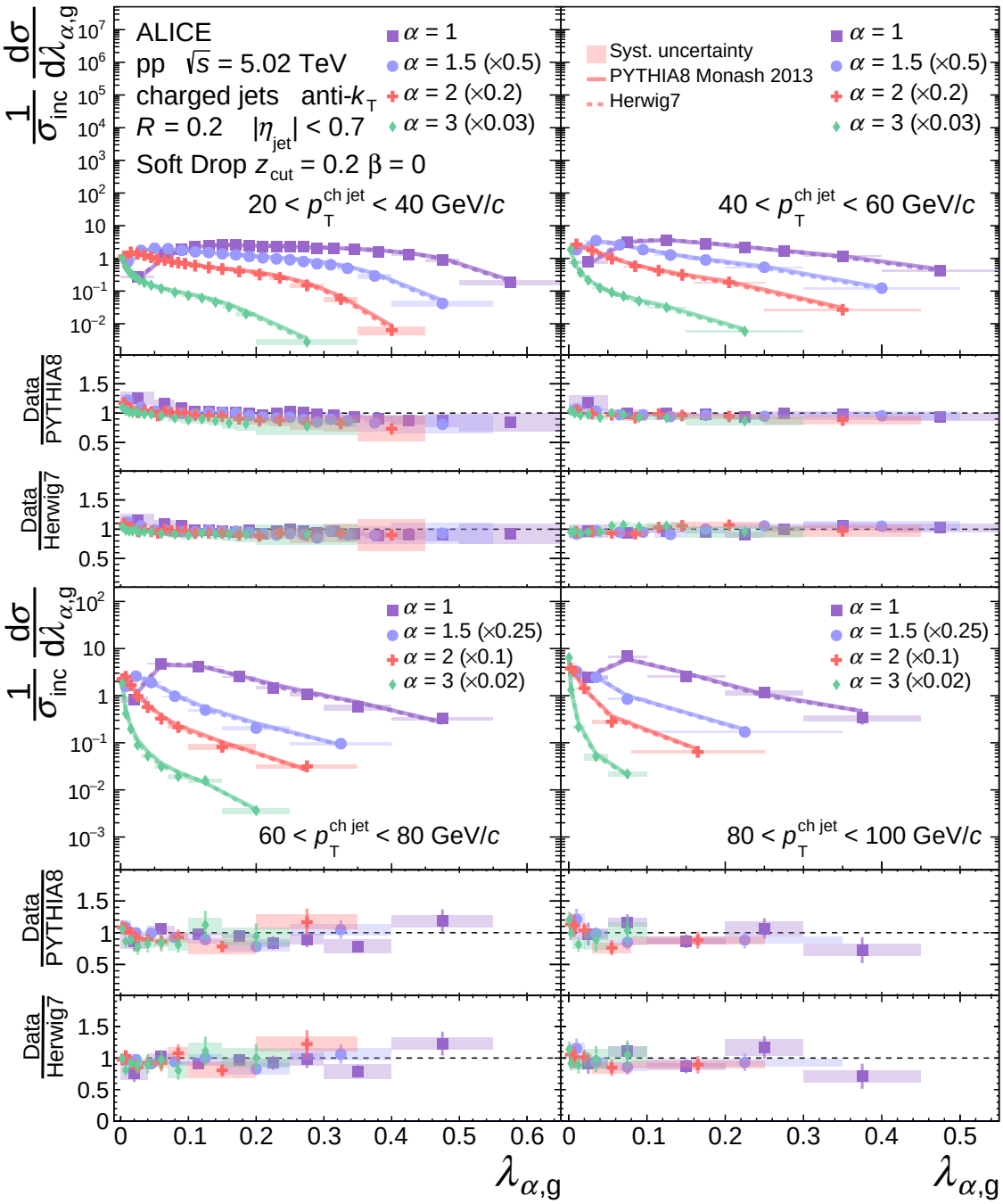


Figure 5.8: Comparison of groomed jet angularities $\lambda_{\alpha,g}$ in pp collisions for $R = 0.2$ to MC predictions using PYTHIA 8 and Herwig 7, as described in the text. Four equally-sized $p_T^{ch\ jet}$ intervals are shown between 20 and 100 GeV/c. The distributions are normalized to the groomed jet tagging fraction [301].

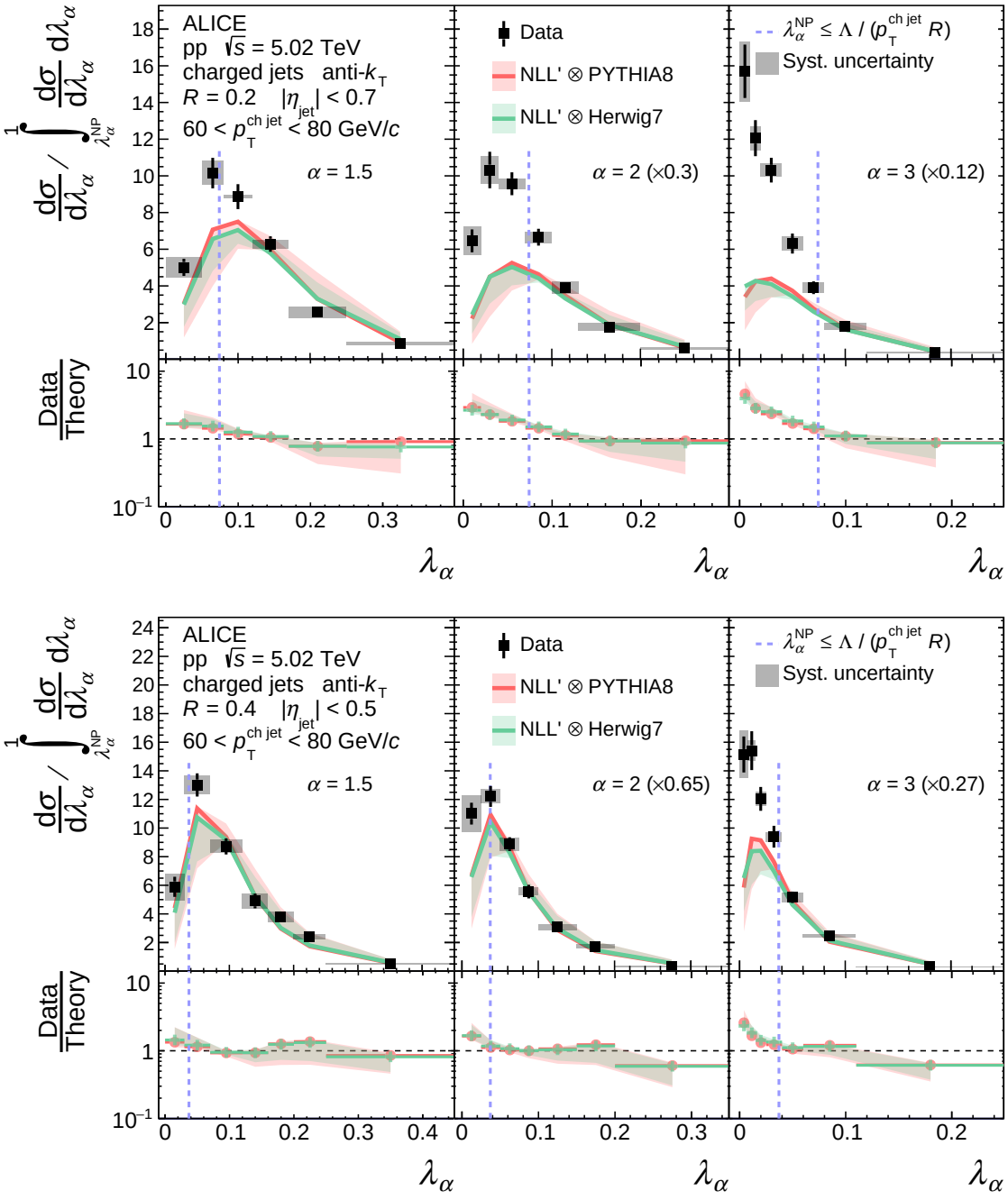


Figure 5.9: Comparison of ungroomed jet angularities λ_α in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions with MC hadronization corrections in the range $60 < p_{\mathrm{T}}^{\text{ch jet}} < 80$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_\alpha > \lambda_\alpha^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region.

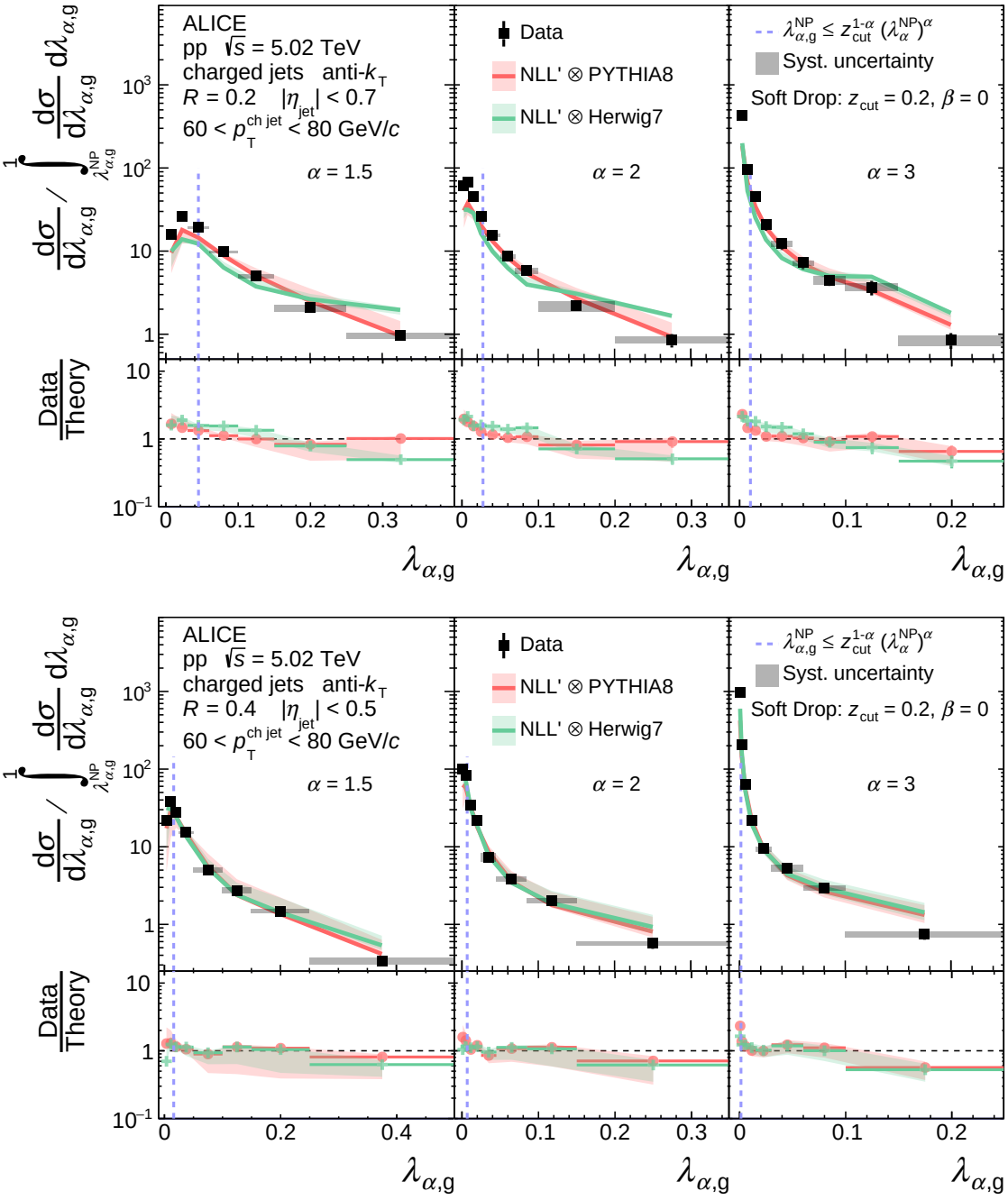


Figure 5.10: Comparison of groomed jet angularities $\lambda_{\alpha,\mathrm{g}}$ in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions with MC hadronization corrections in the range $60 < p_{\mathrm{T}}^{\text{ch jet}} < 80$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_{\alpha,\mathrm{g}} > \lambda_{\alpha,\mathrm{g}}^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region.

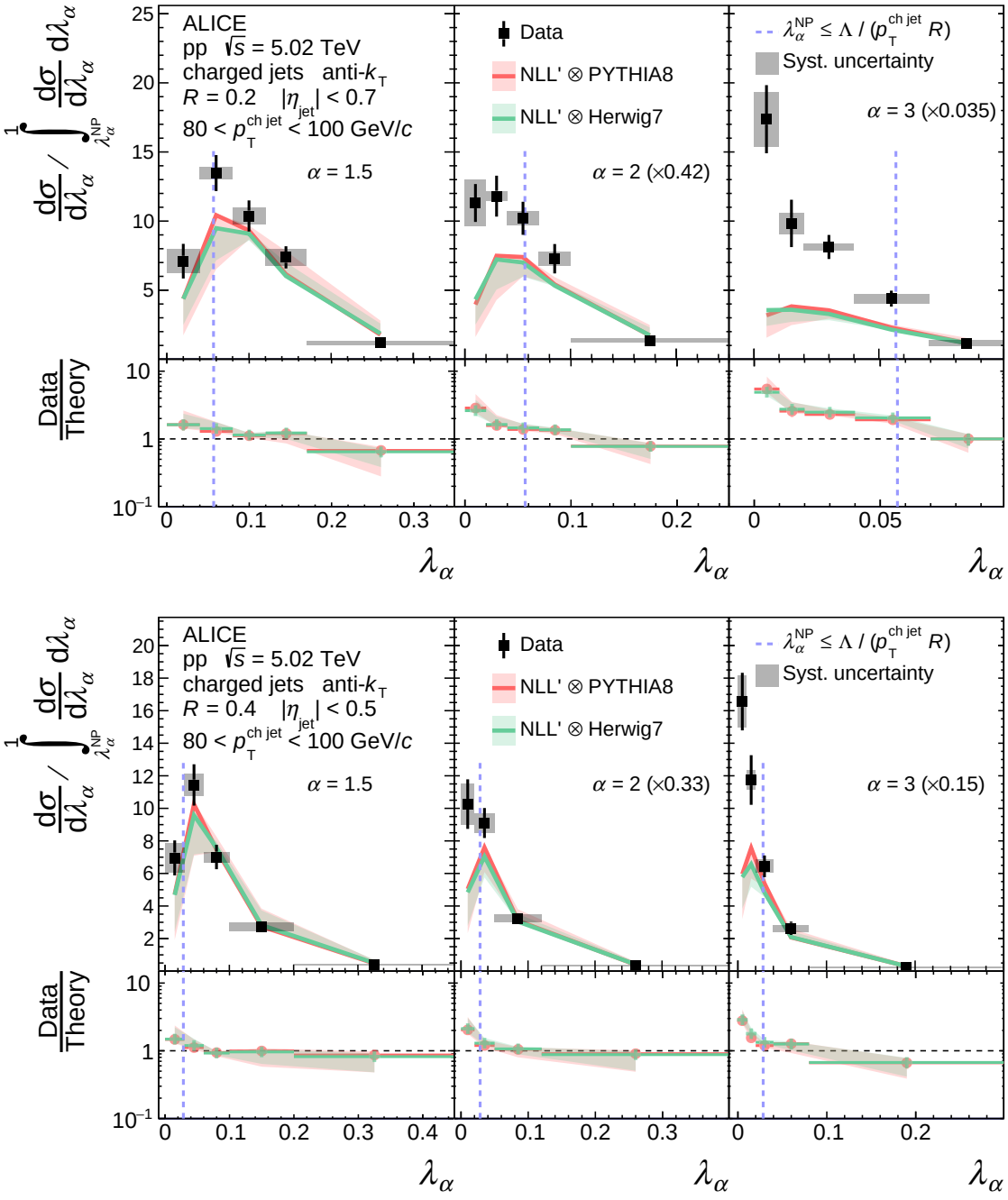


Figure 5.11: Comparison of ungroomed jet angularities λ_α in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions with MC hadronization corrections in the range $80 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 100$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_\alpha > \lambda_\alpha^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region.

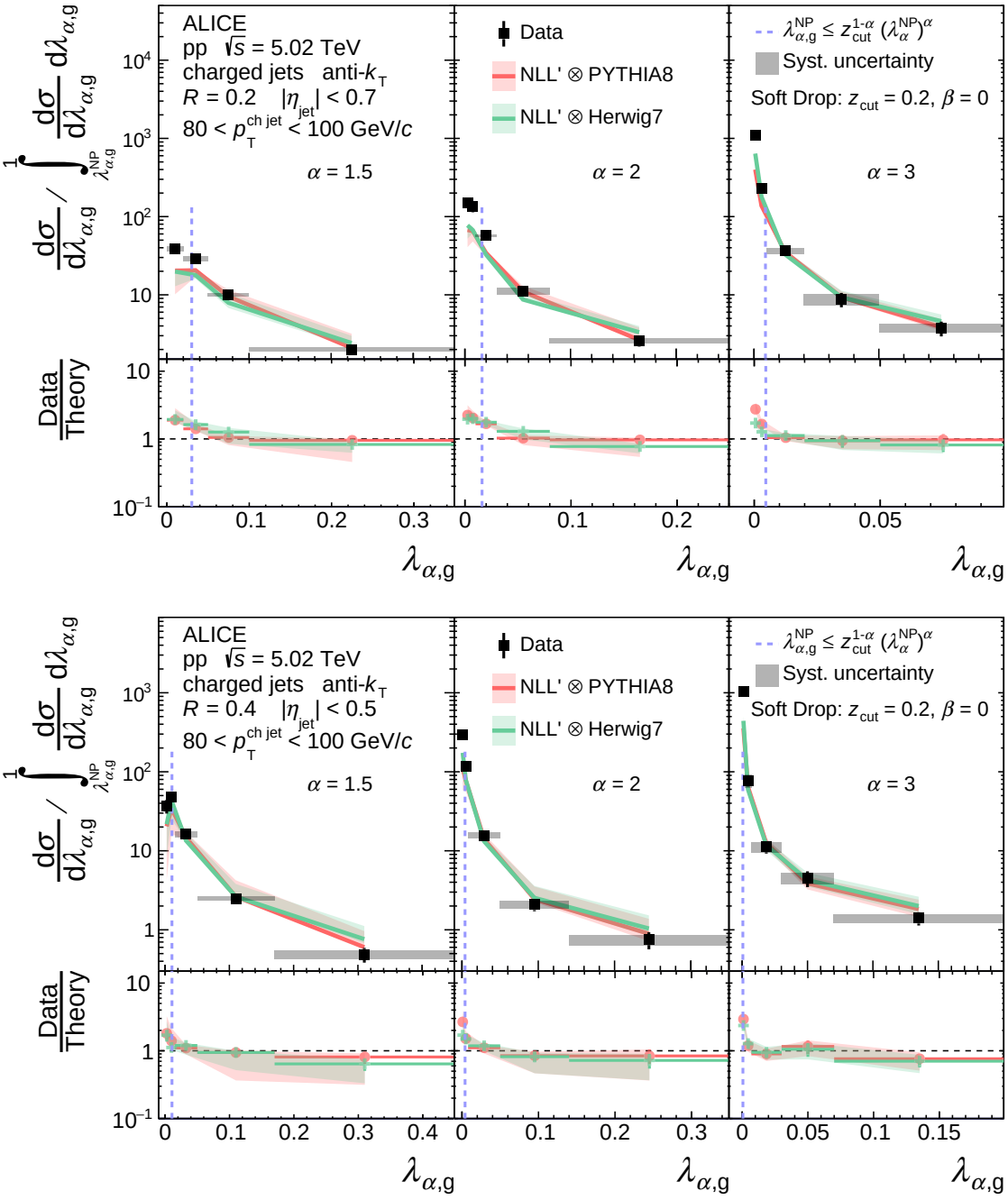


Figure 5.12: Comparison of groomed jet angularities $\lambda_{\alpha,\mathrm{g}}$ in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions with MC hadronization corrections in the range $80 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 100$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_{\alpha,\mathrm{g}} > \lambda_{\alpha,\mathrm{g}}^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region.

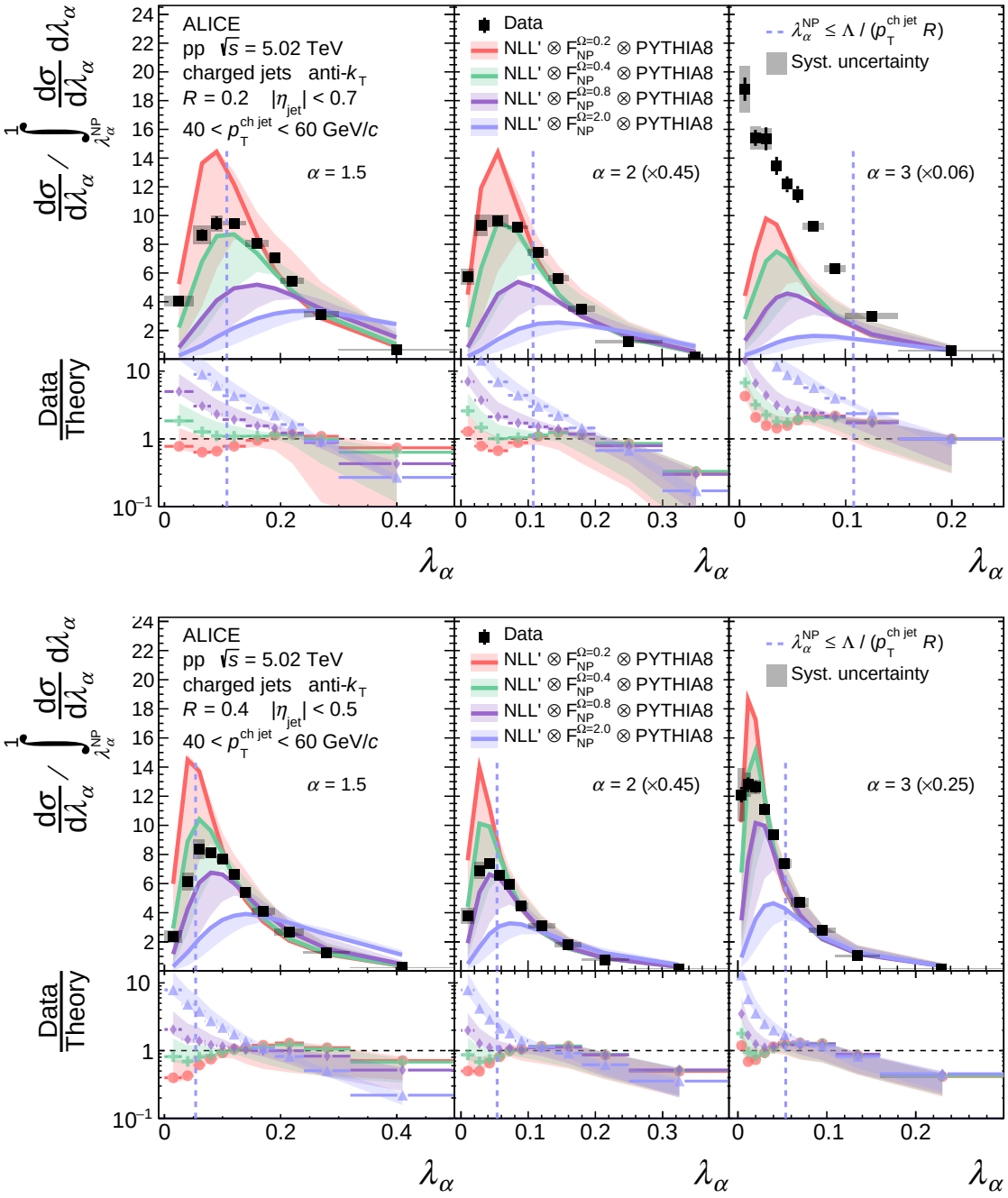


Figure 5.13: Comparison of ungroomed jet angularities λ_α in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions using $F(k)$ convolution in the range $40 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 60$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_\alpha > \lambda_\alpha^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region.

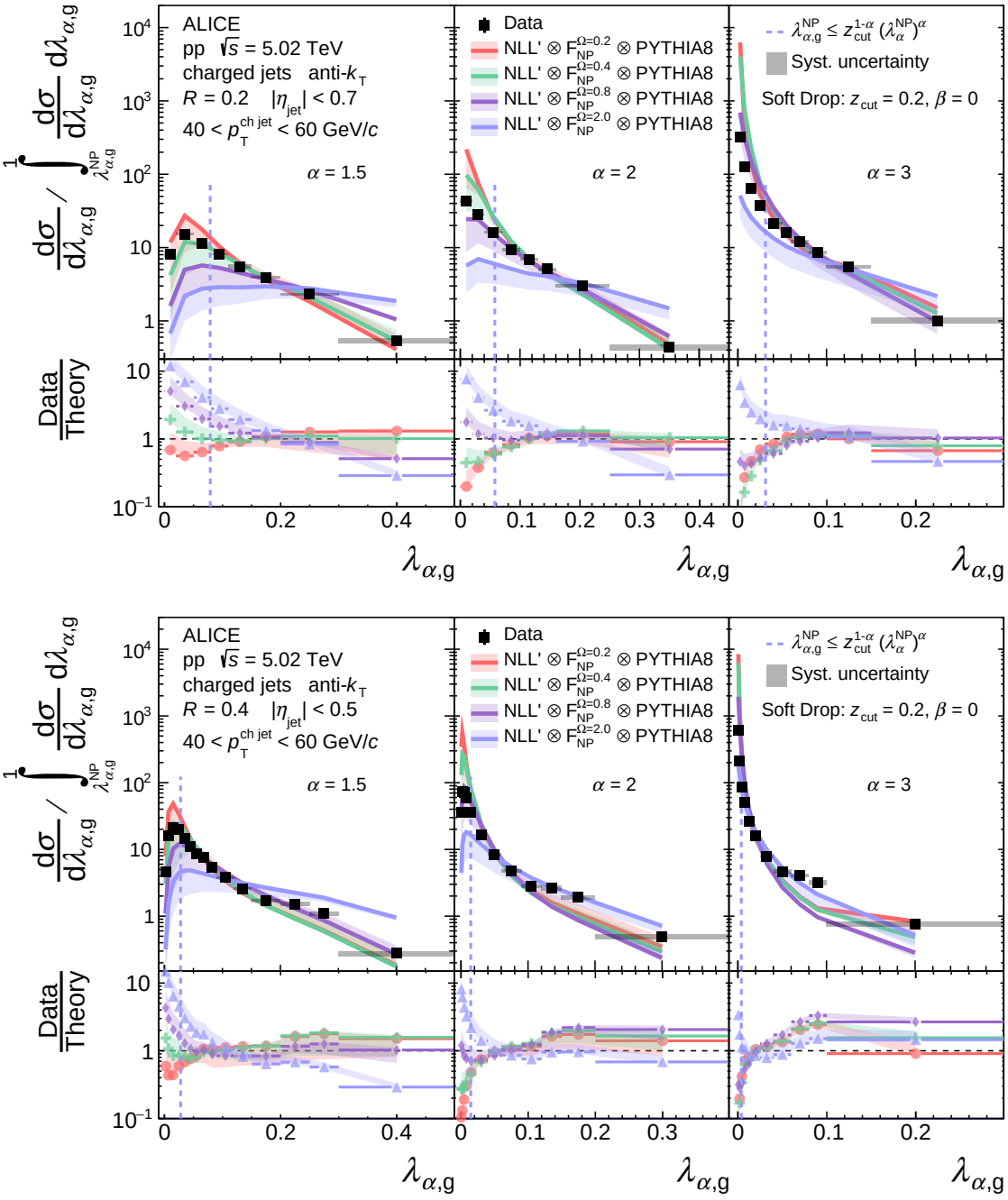


Figure 5.14: Comparison of groomed jet angularities $\lambda_{\alpha,\mathrm{g}}$ in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions using $F(k)$ convolution in the range $40 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 60$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_{\alpha,\mathrm{g}} > \lambda_{\alpha,\mathrm{g}}^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region.

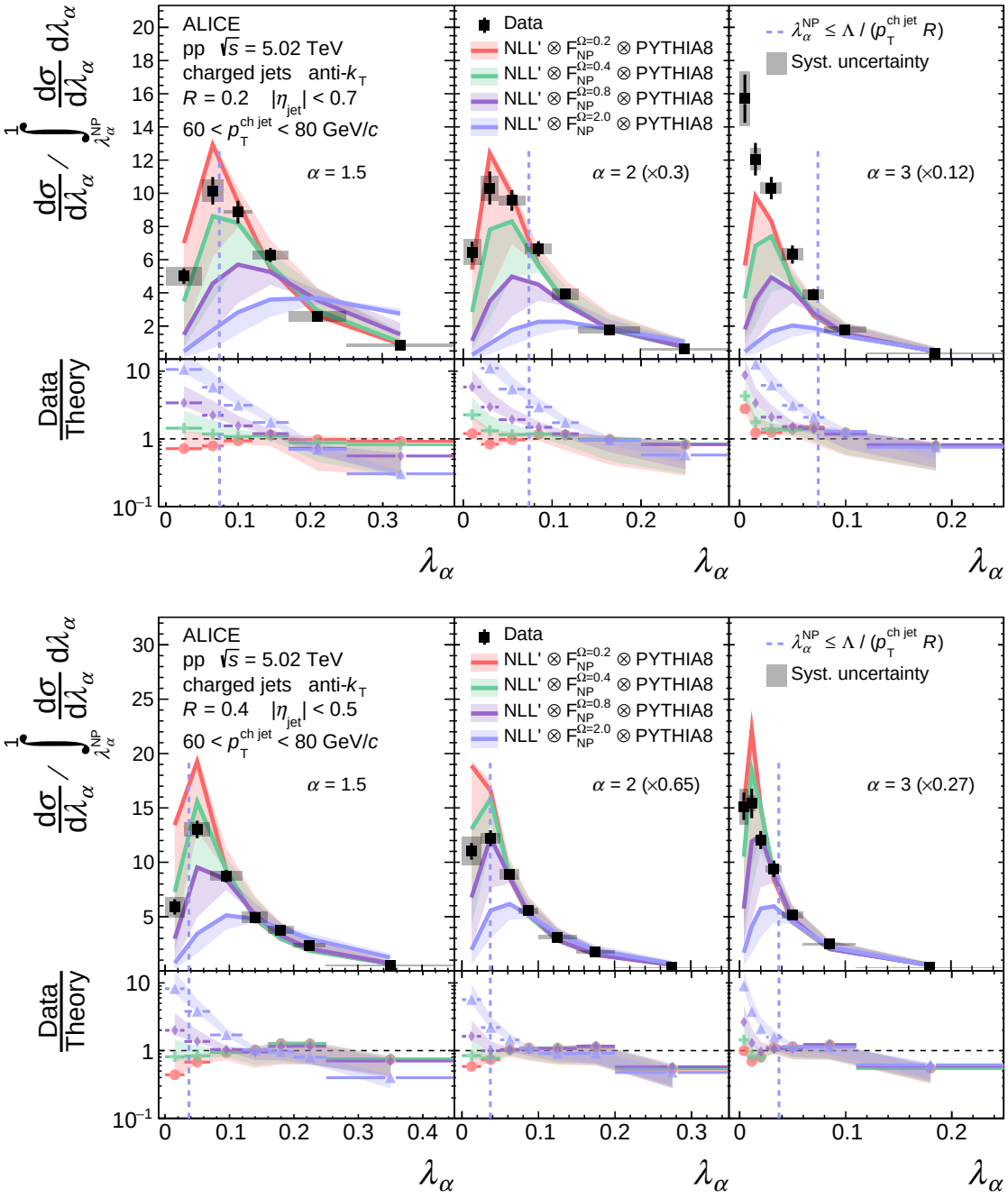


Figure 5.15: Comparison of ungroomed jet angularities λ_α in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions using $F(k)$ convolution in the range $60 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 80$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_\alpha > \lambda_\alpha^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region.

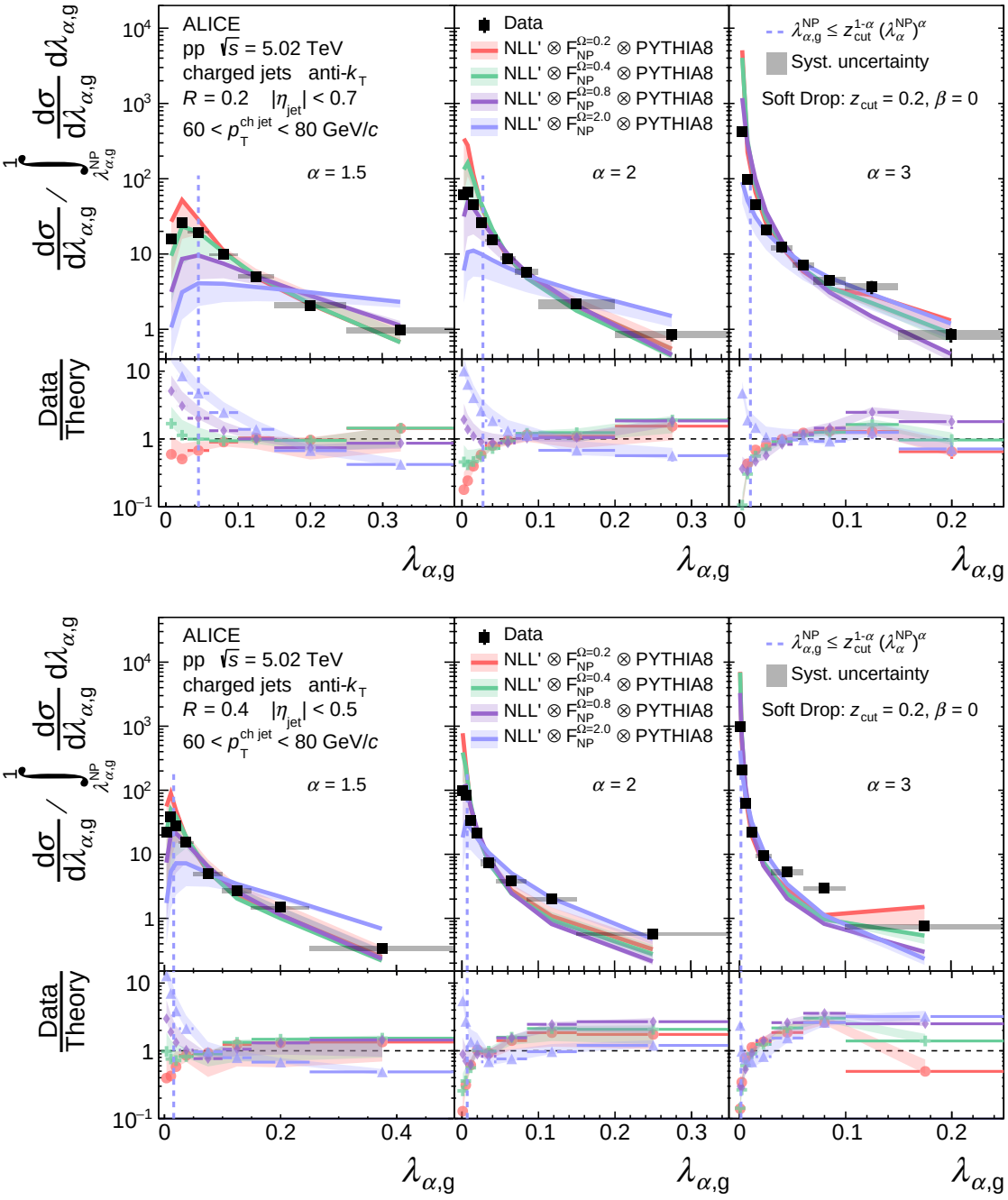


Figure 5.16: Comparison of groomed jet angularities $\lambda_{\alpha,\mathrm{g}}$ in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions using $F(k)$ convolution in the range $60 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 80$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_{\alpha,\mathrm{g}} > \lambda_{\alpha,\mathrm{g}}^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region.

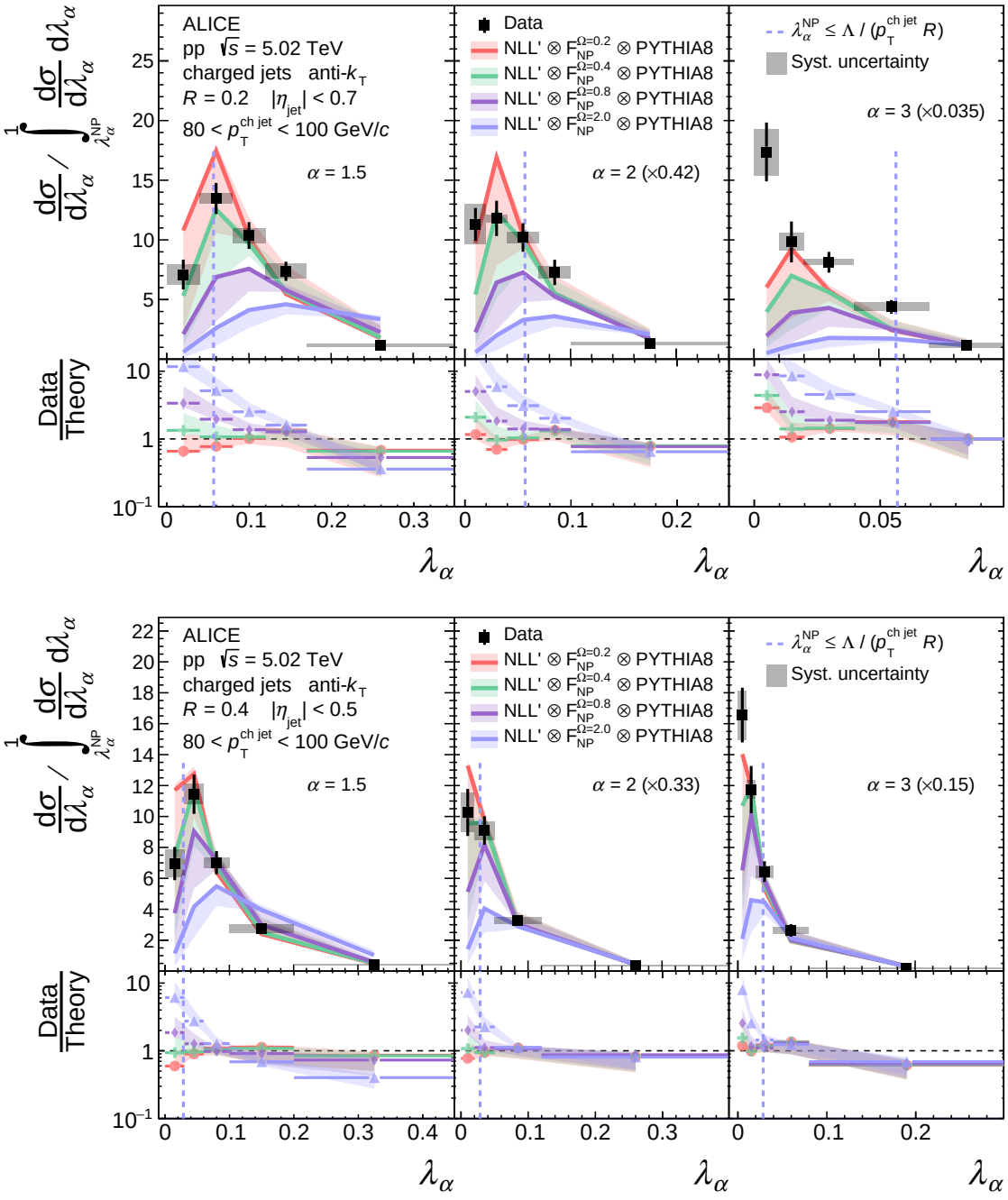


Figure 5.17: Comparison of ungroomed jet angularities λ_α in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions using $F(k)$ convolution in the range $80 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 100$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_\alpha > \lambda_\alpha^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region.

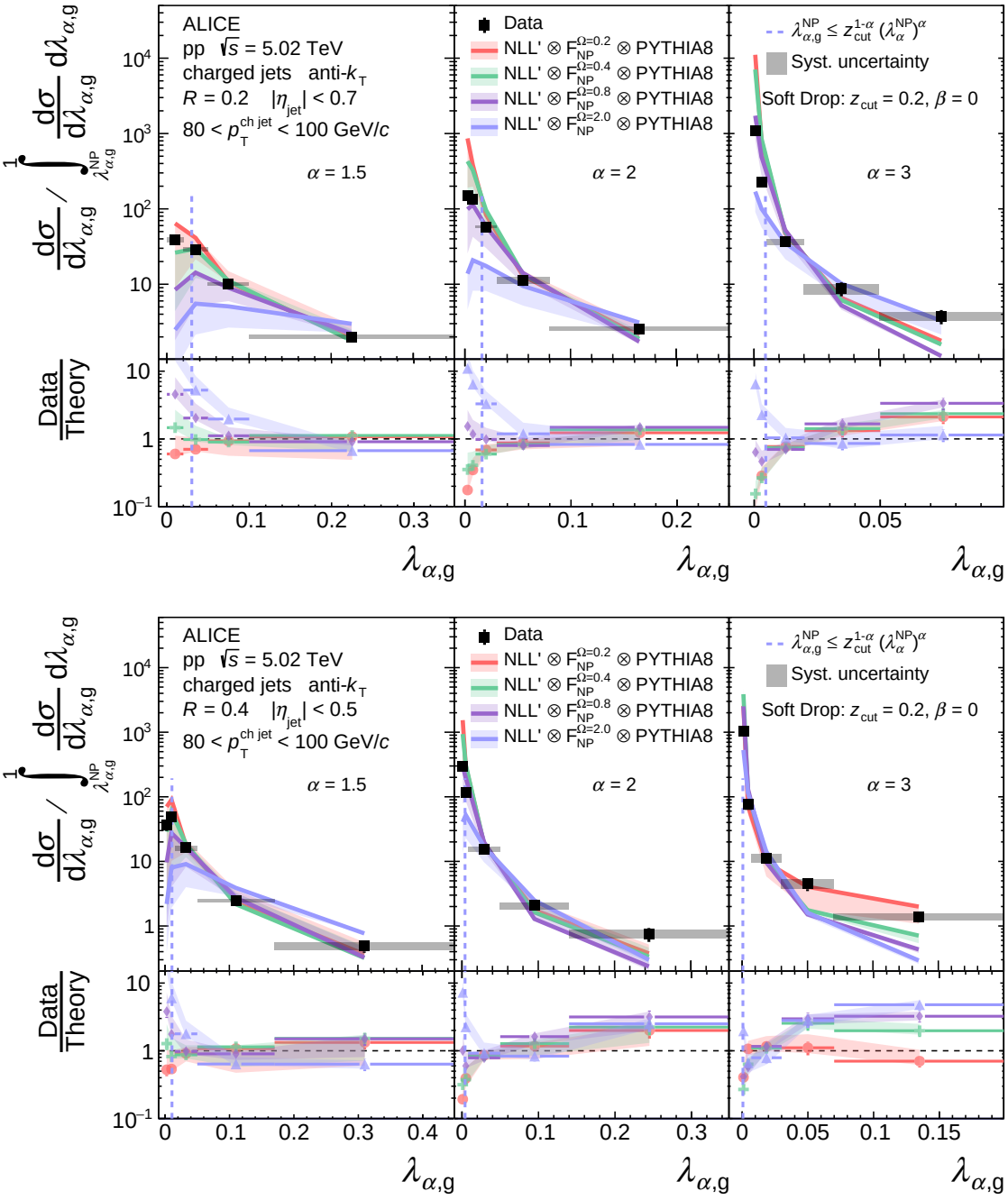


Figure 5.18: Comparison of groomed jet angularities $\lambda_{\alpha,\mathrm{g}}$ in pp collisions for $R = 0.2$ (top) and $R = 0.4$ (bottom) to analytical NLL′ predictions using $F(k)$ convolution in the range $80 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 100$ GeV/c. The distributions are normalized such that the integral of the perturbative region defined by $\lambda_{\alpha,\mathrm{g}} > \lambda_{\alpha,\mathrm{g}}^{\mathrm{NP}}$ (to the right of the dashed vertical line) is unity. Divided bins are placed into the left (NP) region.

5.2.1 D^0-tagged jet angularities

This section includes ALICE results of the ungroomed jet angularities λ_α for D^0-tagged jets in pp collisions at $\sqrt{s} = 5.02$ TeV [303]. The D^0-tagged angularities are compared to a semi-inclusive baseline with a cut on the p_T of the leading track inside the jets. The selection criteria is based on the transverse mass m_T of the D^0,

$$m_T^2 = m^2 + p_T^2, \tag{5.2}$$

and requiring that the lowest transverse mass of reconstructed D^0 mesons inside jets is equal to that of a $\pi^\pm$ meson in the semi-inclusive sample, $m_T^{D^0} = m_T^{\pi^\pm}$. Thus for a minimum $p_T^{D^0} \geq 5$ GeV/c, the semi-inclusive measurement requires $p_T^{\text{leading track}} \geq 5.336$ GeV/c in order to induce a similar fragmentation bias [119]. The semi-inclusive results are unfolded in a similar way to the inclusive results, using their own RM and satisfying all unfolding checks and systematic criteria.

The D^0-tagged jet angularity distributions for D^0-tagged jets are given in Fig. 5.19 for $\alpha = 1$ and 1.5 and in Fig. 5.20 for $\alpha = 2$ and 3. A 10-30% discrepancy is observed for both the individual heavy-flavor and semi-inclusive results as compared to PYTHIA (bottom left two panels), though the fragmentation modification, as described by the ratio of the heavy-flavor to semi-inclusive distributions, agrees with the data within the given experimental uncertainties (bottom right panel). D^0-tagged jets are slightly narrower than the inclusive sample due to their larger mass reducing emissions (QCD dead cone). As the exponent α is increased, the difference between teh HF and inclusive jet angularities decreases, implying that the modification is mainly within the jet core rather than its edge.

To better understand the behavior of the distributions, a sample of charged-particle jets was generated with PYTHIA 8 and tagged as being initiated by a quark or gluon. To obtain a clean sample, jets were required to have a unique geometrical match to one of the two hard-scattered partons (PYTHIA particle number 5 and 6), else they were discarded. Comparisons of these distributions to those containing a D^0 meson are shown in Fig. 5.21. Gluon distributions are observed to be the broadest, corresponding to the larger gluon color factor enhancing perturbative emissions. As also observed in the experimental data, D^0-tagged jets are narrower than the inclusive quark-sample. The D^0-tagged jets begin to converge on the inclusive quark result at larger values of α, where the large-angle weighting reduces the effect of small-angle dead cone suppression. This simultaneously increases the sensitivity to Casimir color effects [303].

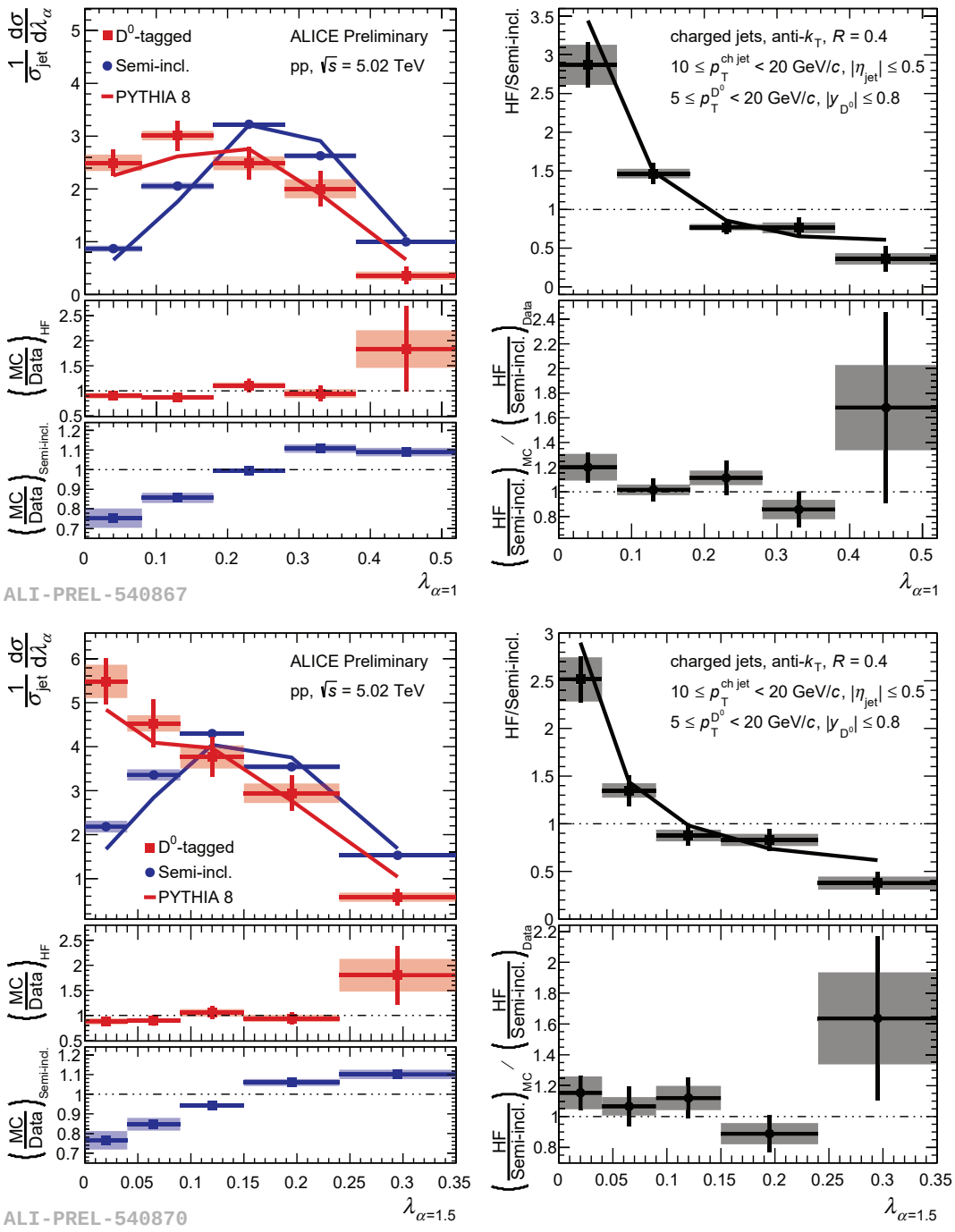


Figure 5.19: Comparison of D^0-tagged jet angularities λ_α to semi-inclusive baseline in pp collisions for $R = 0.4$ charged-particle jets and $\alpha = 1$ (top) or 1.5 (bottom). Results are compared to PYTHIA MC predictions for both results. Jets are reported in $10 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 20$ GeV/c with $p_{\mathrm{T}}^{\mathrm{D}^0} > 5$ GeV/c, and the distributions are normalized to unity [303].

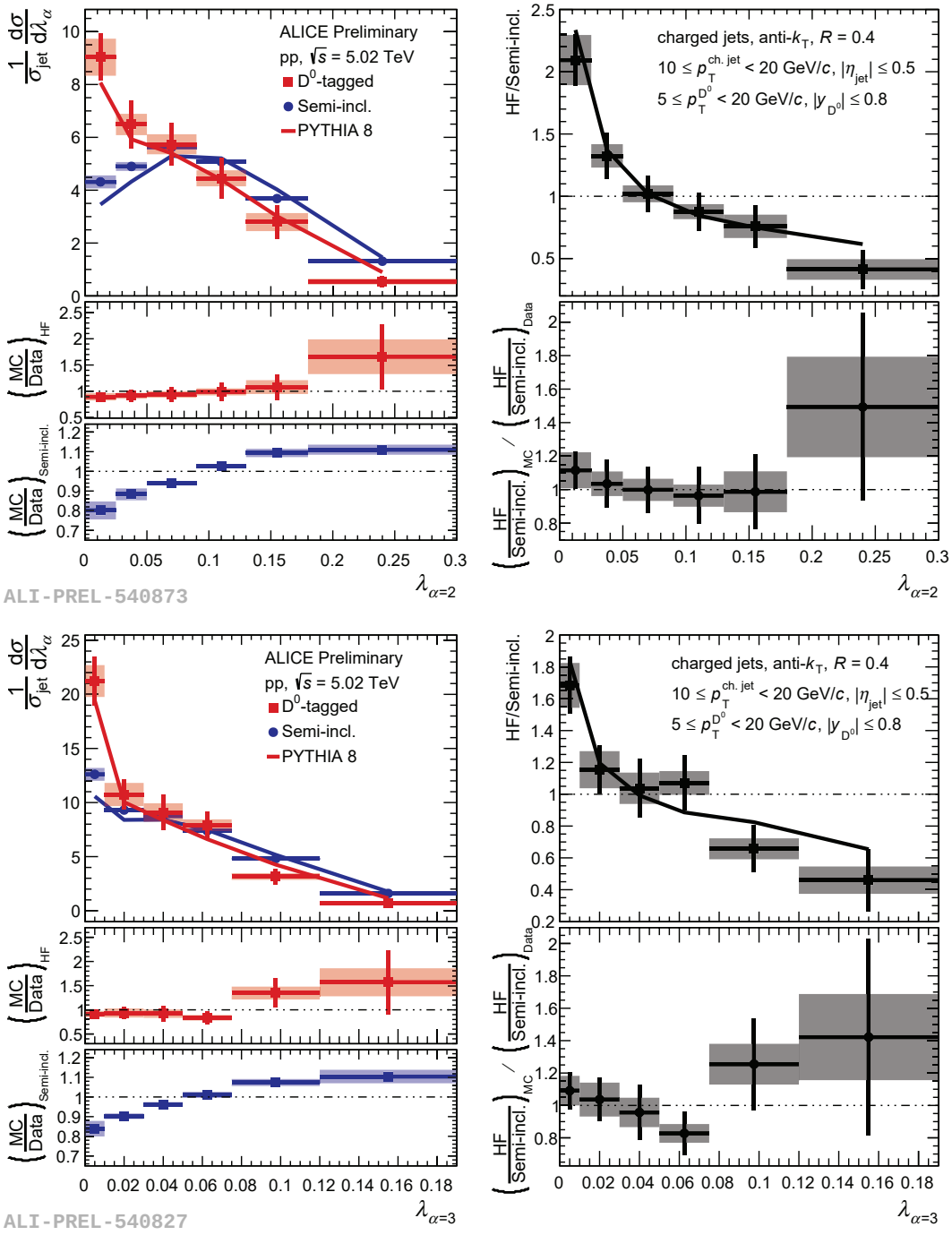


Figure 5.20: Comparison of D^0-tagged jet angularities λ_α to semi-inclusive baseline in pp collisions for $R = 0.4$ charged-particle jets and $\alpha = 2$ (top) or 3 (bottom). Results are compared to PYTHIA MC predictions for both results. Jets are reported in $10 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 20$ GeV/c with $p_{\mathrm{T}}^{\mathrm{D}^0} > 5$ GeV/c, and the distributions are normalized to unity [303].

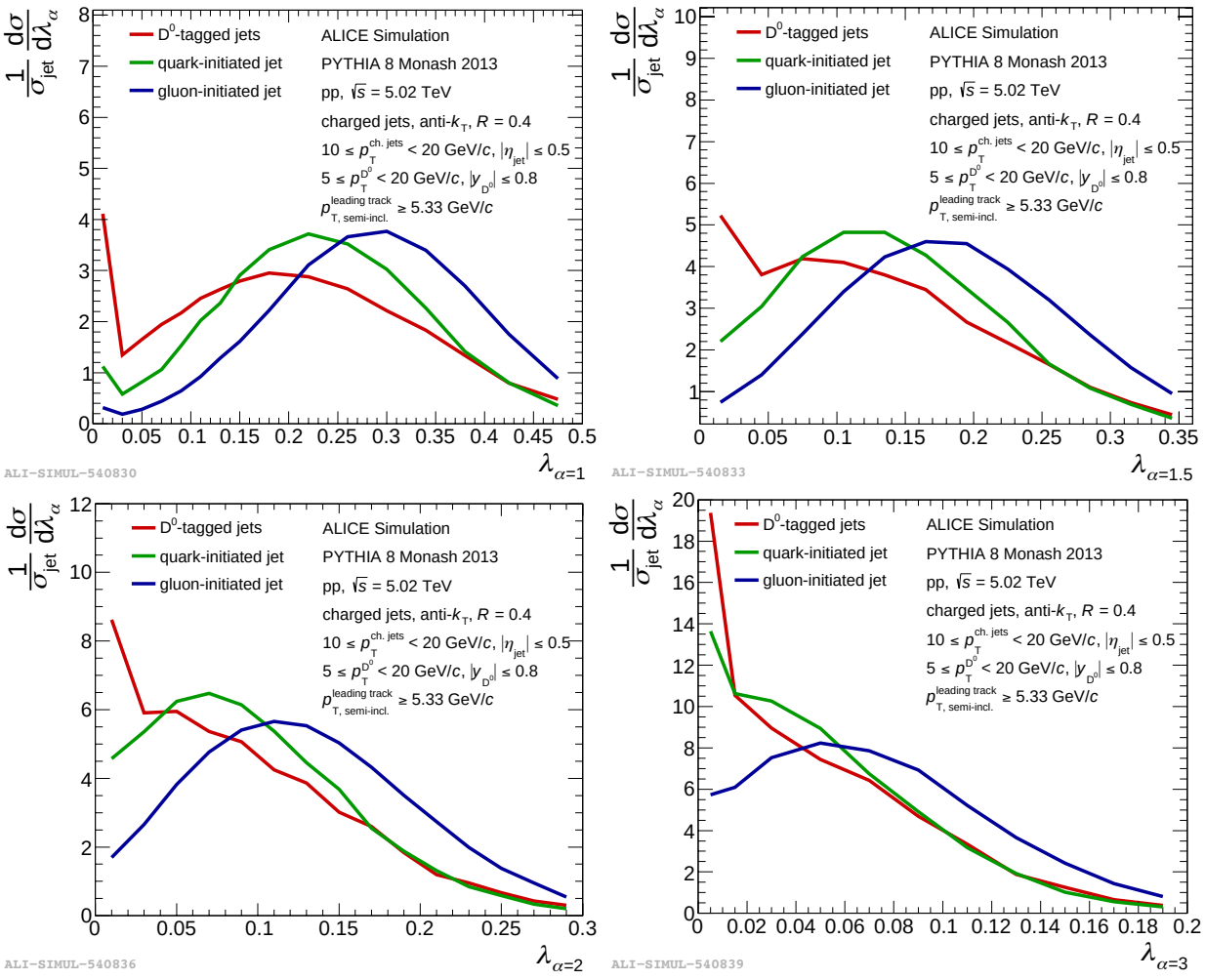


Figure 5.21: Comparisons of the inclusive quark-, gluon-, and D^0-initiated charged-particle jet angularities in PYTHIA 8 at $\sqrt{s} = 5.02$ TeV. Generated distributions require a unique angular matching between a hard-scattered parton and the charged-particle jet [303].

5.3 Quenched jet mass and angularities

This section reports ALICE results of the ungroomed and SD-groomed jet angularities λ_α and jet mass $m_{\rm jet}$ for inclusive jets in Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV [286]. Results are compared to inclusive pp measurements at $\sqrt{s} = 5.02$ TeV, reported for λ_α in Sect. 5.2 (the vacuum $m_{\rm jet}$ results are still preliminary and are included as part of this analysis).

Results are compared to theoretical predictions from several different models, with implementations of the phenomenological approaches described in Sect. 1.4.6:

- **JEWEL** [192] with medium recoils both on and off [194]. In the recoils on case, recoils are reconstructed inside jets using negative four-momentum, and their substructure is correspondingly corrected.[1] The Pb–Pb results are compared to a JEWEL pp baseline (based on a uniquely tuned PYTHIA), generated using the nominal settings.

- **Higher-Twist formalism** for energy loss of the hard-scattered parton traversing the QCD medium [191, 305]. These predictions are calculated via MC simulations, using POWHEG matrix elements (NLO) matched to PYTHIA 6 parton shower and final-state hadronization [44, 306] as a baseline.

- **Hybrid model** [195], both with and without elastic Molière scattering [197], and with the medium wake effect enabled [198]. Results are computed in a MC approach using a tuned PYTHIA baseline, labelled as "Hybrid model vacuum" for clarity.

- **JETSCAPE** [199], using **MATTER** [200] and **LBT** [201] shower modifications and a full hydrodynamic simulation. The MC also uses tuned PYTHIA as a baseline, labelled as "JETSCAPE pp" for clarity.

Results for the ungroomed and SD-groomed jet angularities λ_α with $\alpha = 1$ (girth) are shown in Figs. 5.22-5.25, for $\alpha = 1.5$ in Figs. 5.26-5.29, for $\alpha = 2$ (thrust) in Figs. 5.30-5.33, and for $\alpha = 3$ in Figs. 5.34-5.37. The same are shown for the groomed and ungroomed jet mass $m_{\rm jet}$ in Figs. 5.38-5.41. Models are compared to the Pb–Pb result (center panels) while their baselines are compared to the pp result (left panels). The Pb–Pb / pp ratio, which quantifies the quenching modification, is also shown in the right panels. The highest $100 < p_{\rm T}^{\rm ch\ jet} < 150$ GeV/c bin does not have a baseline result from pp, so only the Pb–Pb result is reported.

In comparisons of Pb–Pb and pp distributions, a significant narrowing effect is observed at low α which decreases at larger values of α, corresponding to a strongly quenched jet core. This conclusion is supported by a significant enhancement in the narrowing for SD groomed jets, which remove soft radiation at wide angles. The strength of this quenching effect appears to be consistent across the reported $p_{\rm T}^{\rm ch\ jet}$ range within the reported uncertainties, though the behavior is most obvious at low-$p_{\rm T}^{\rm ch\ jet}$ where statistical uncertainties are the smallest.

[1] Recently a new method has been proposed to account for thermal recoils by applying a CS procedure. This has shown improvement in data comparisons for the jet mass; for details, see Ref. [304].

JEWEL with recoils on typically approximates the data more closely than with recoils off for λ_α, while the recoils off case is a better approximation of Pb–Pb data for m_{jet}. JEWEL pp exhibits the strongest tension of all models with the ALICE pp data, which correspondingly skews the agreement in both recoils on and off Pb–Pb / pp ratios. In other words, despite some reasonable agreement in the Pb–Pb predictions alone, the jet substructure modification due to quenching as quantified by Pb–Pb / pp exhibits strong tension with the data, up to a factor of 2 or larger in the tails of the distributions.

These results are discussed in more detail in Ch. 6.

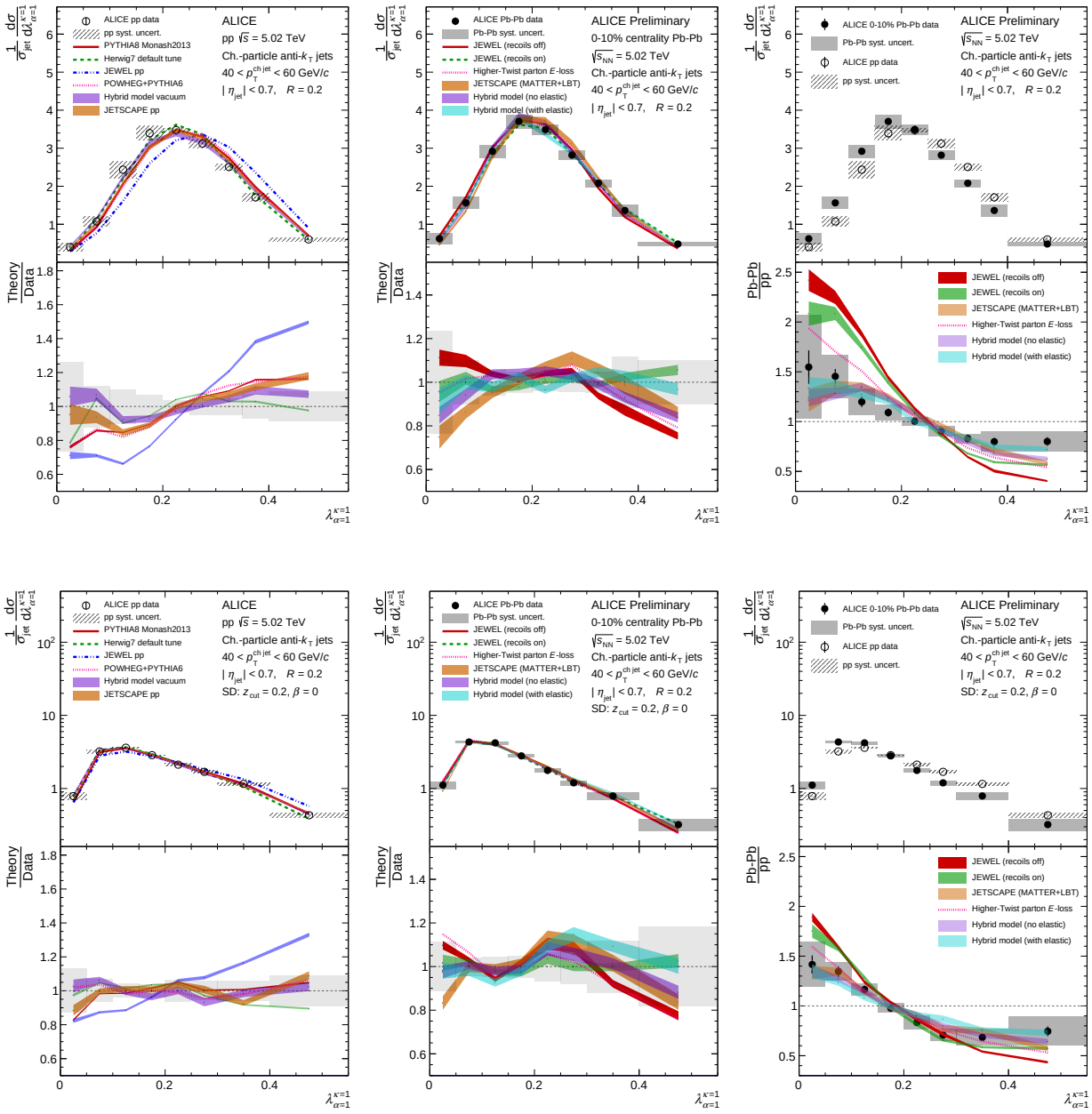


Figure 5.22: ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_1 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV for $40 < p_{\rm T}^{\rm ch\ jet} < 60$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

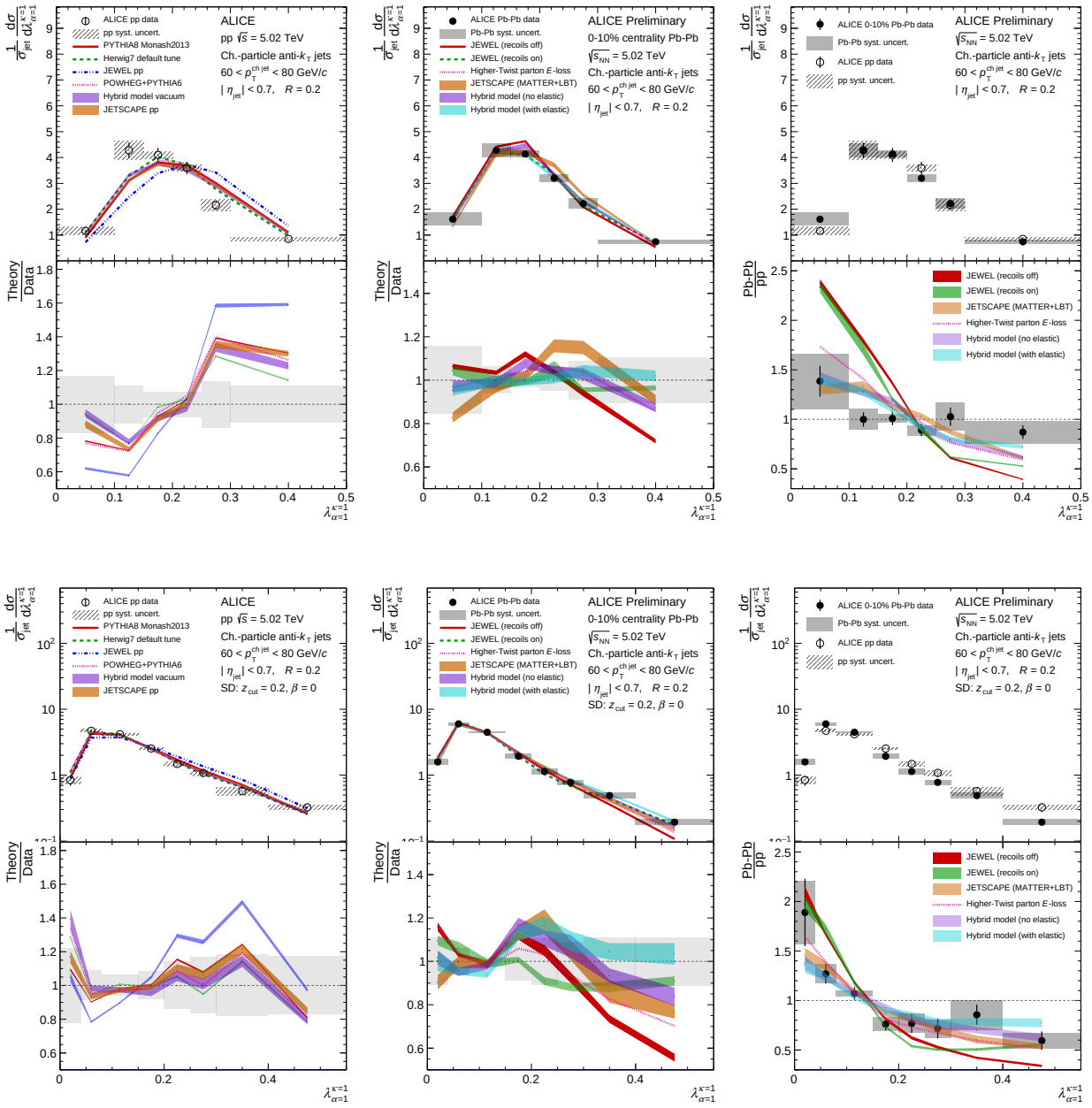


Figure 5.23: ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_1 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV for $60 < p_{\rm T}^{\rm ch\ jet} < 80$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

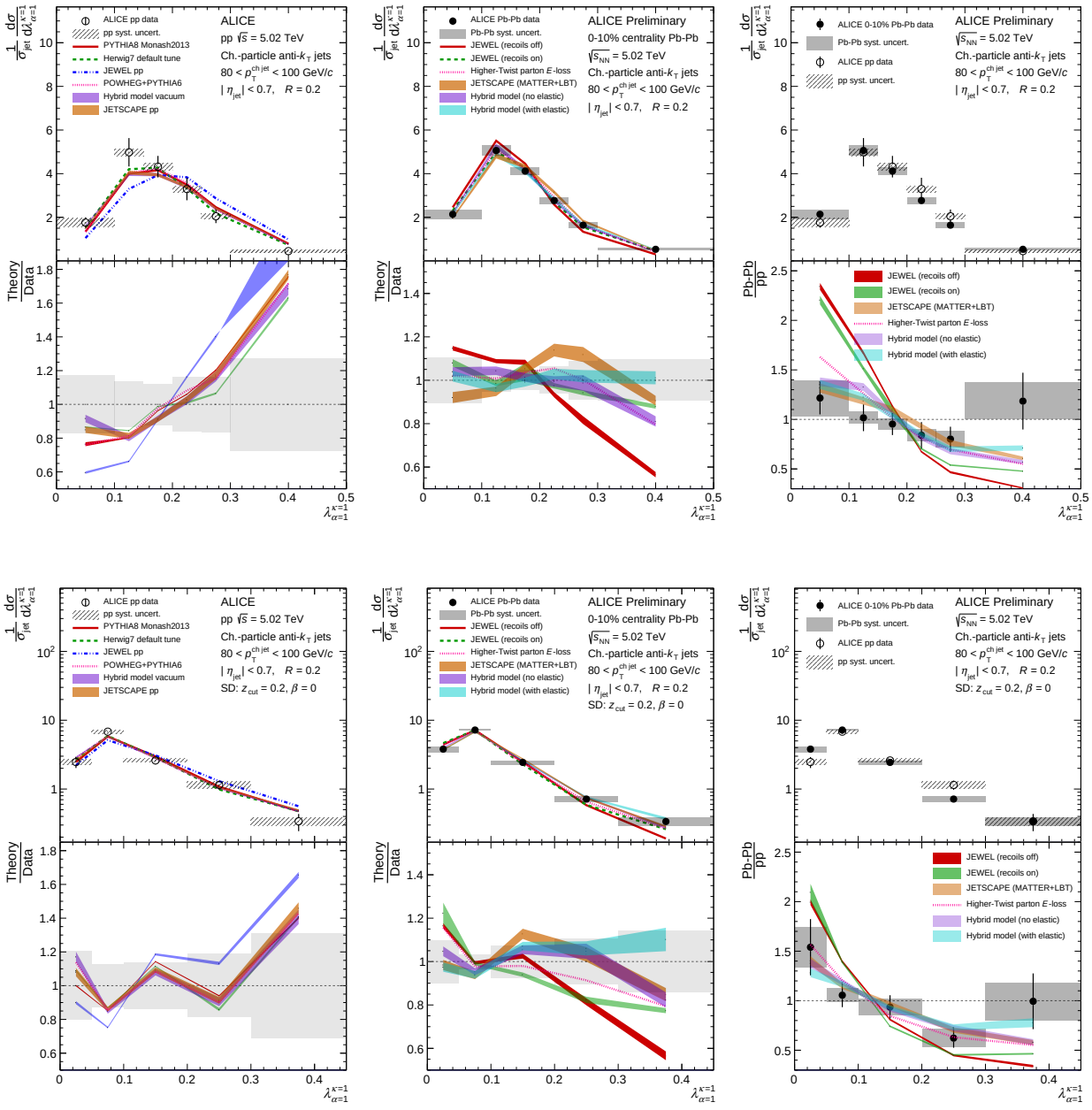


Figure 5.24: ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_1 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV for $80 < p_{\rm T}^{\rm ch\ jet} < 100$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

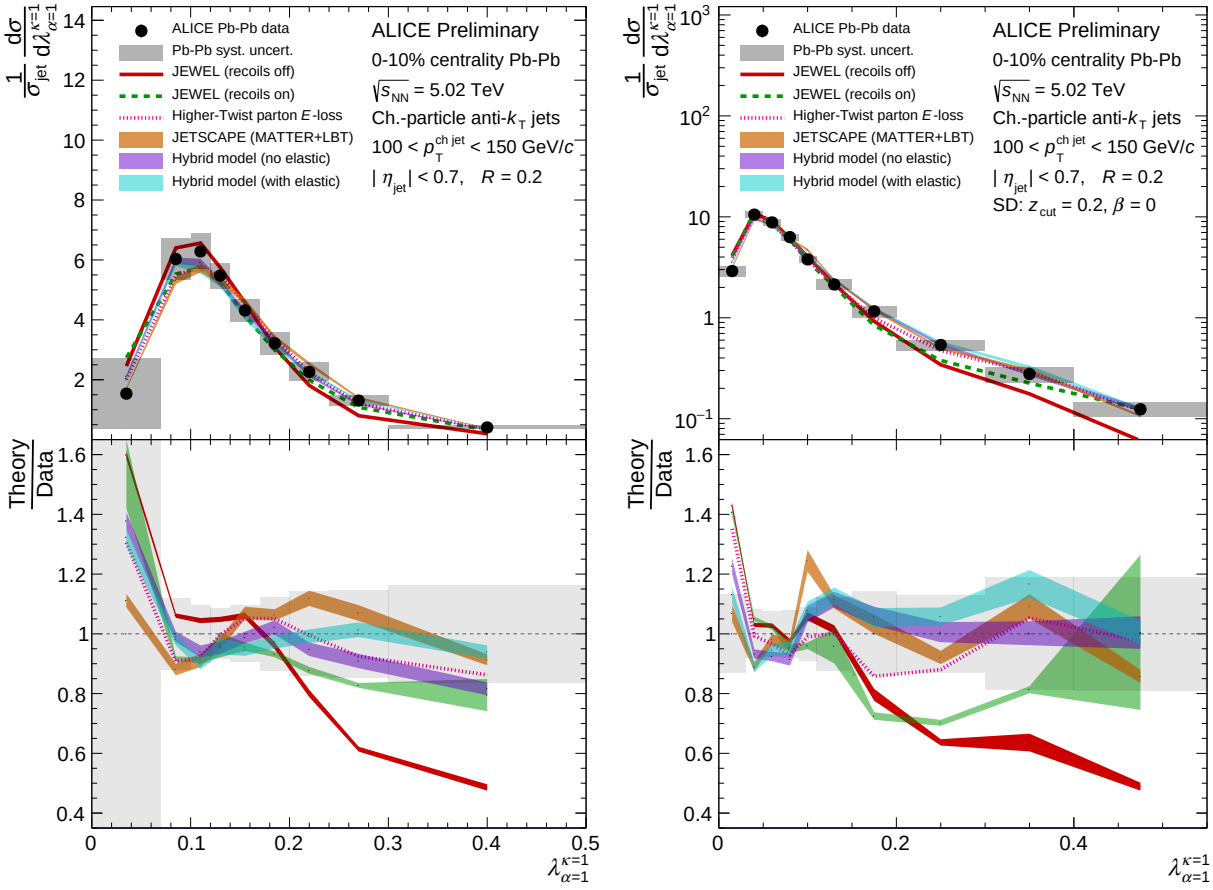


Figure 5.25: ALICE measurement of ungroomed (left) and SD groomed (right) λ_1 for $R = 0.2$ charged-particle jets in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $100 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 150$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

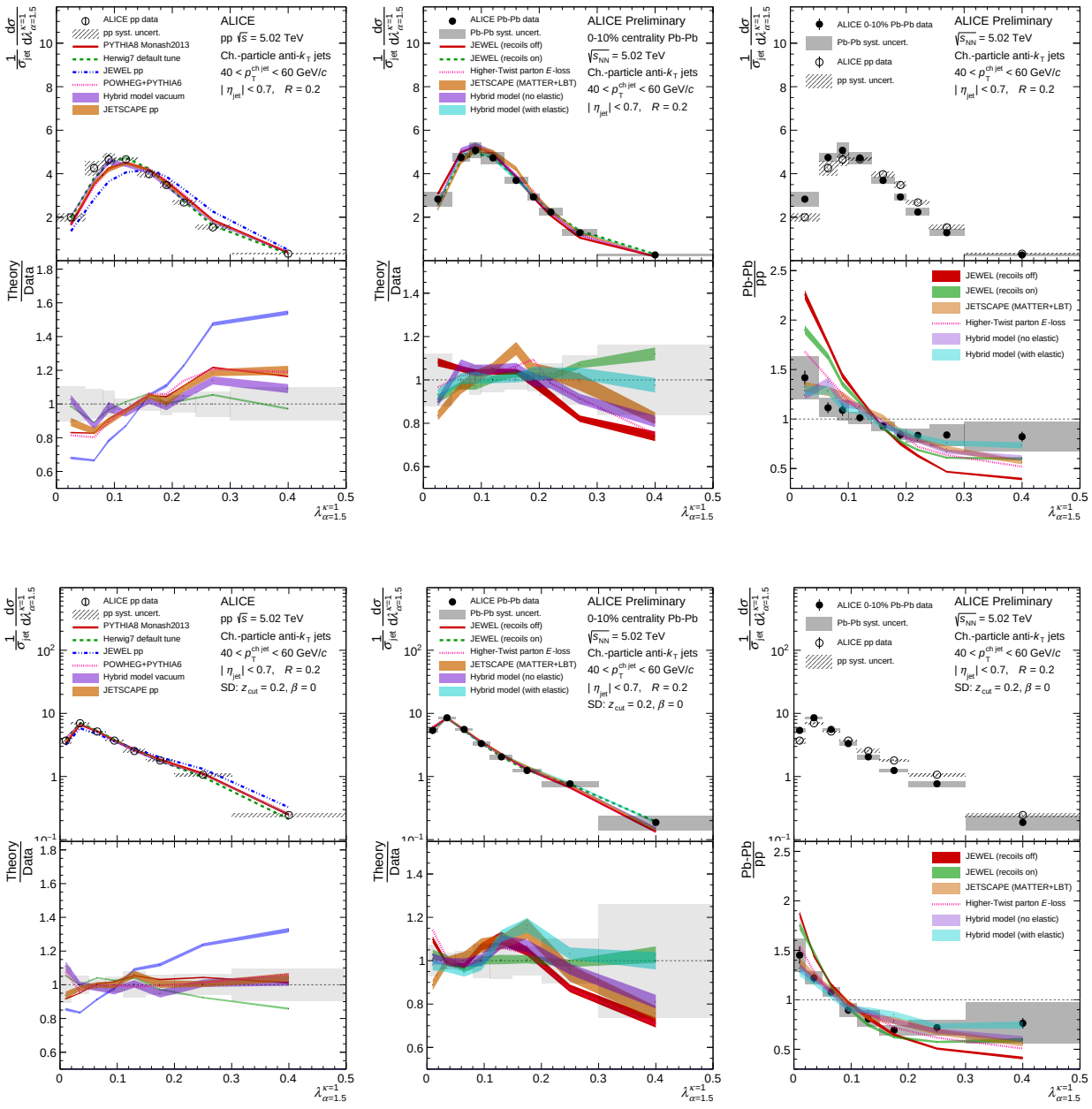


Figure 5.26: ALICE measurement of ungroomed (top) and SD groomed (bottom) $\lambda_{1.5}$ for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV for $40 < p_{\rm T}^{\rm ch\ jet} < 60$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

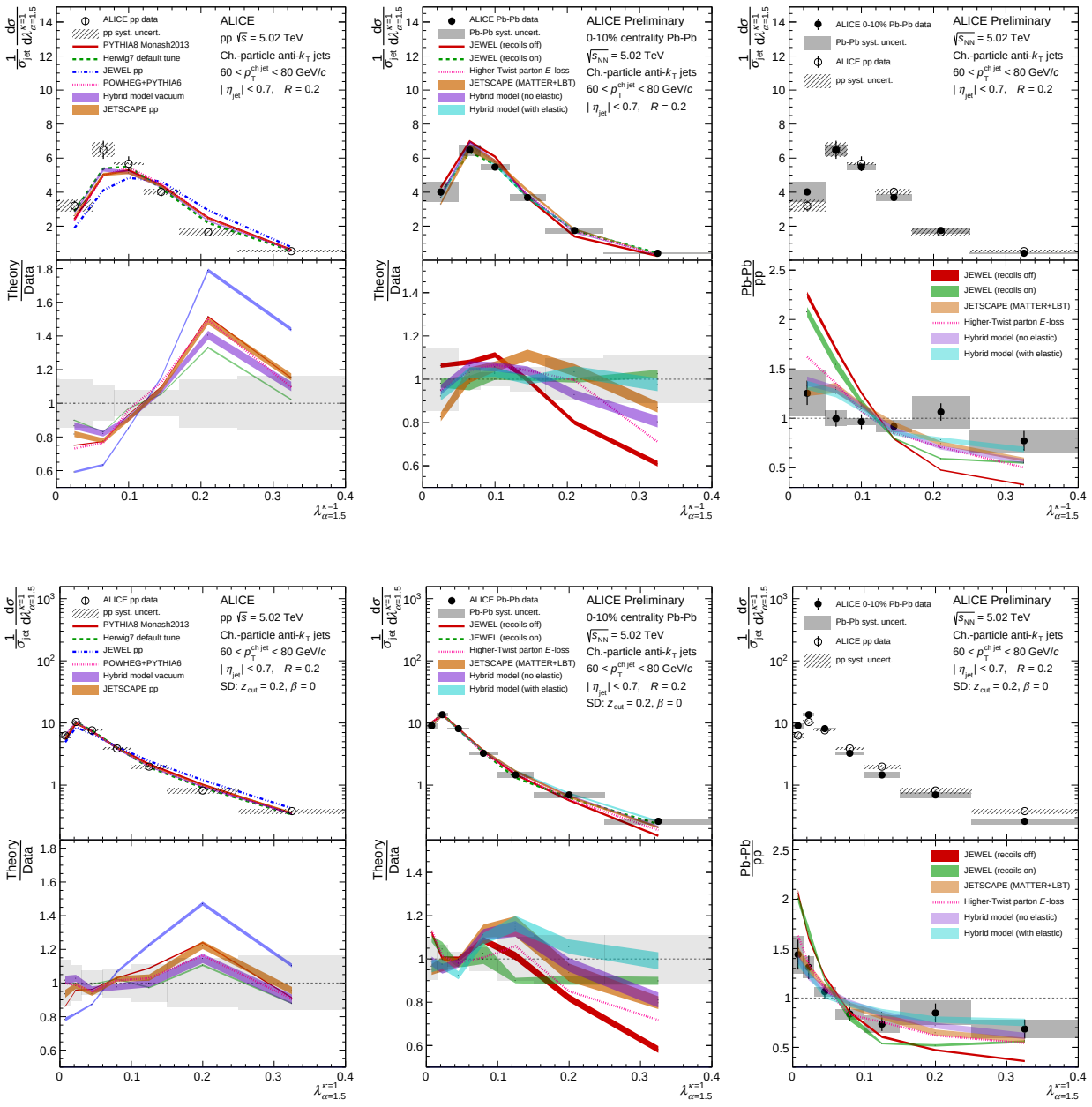


Figure 5.27: ALICE measurement of ungroomed (top) and SD groomed (bottom) $\lambda_{1.5}$ for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV for $60 < p_{\rm T}^{\rm ch\ jet} < 80$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

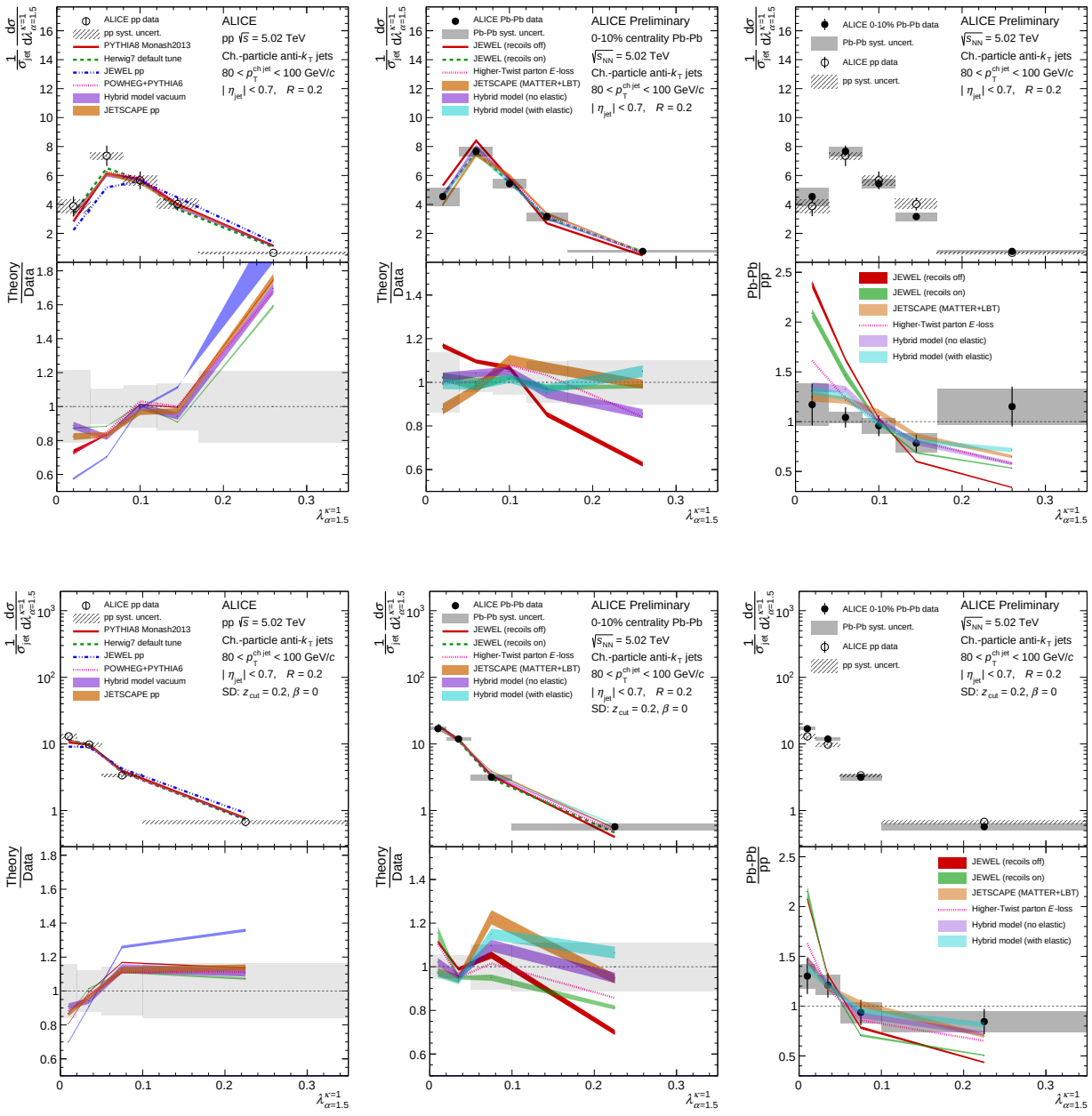


Figure 5.28: ALICE measurement of ungroomed (top) and SD groomed (bottom) $\lambda_{1.5}$ for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $80 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 100$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

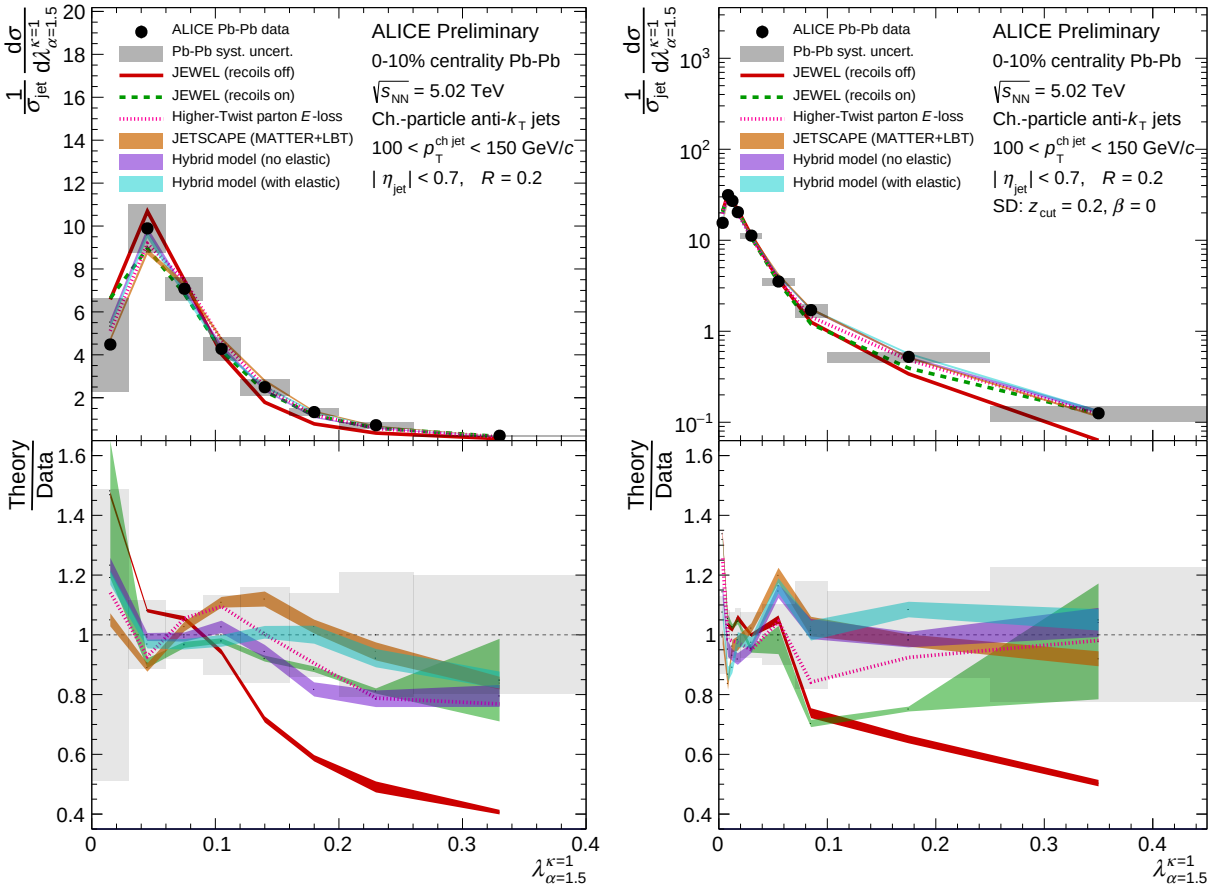


Figure 5.29: ALICE measurement of ungroomed (left) and SD groomed (right) λ_1 for $R = 0.2$ charged-particle jets in Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV for $100 < p_{\rm T}^{\rm ch\ jet} < 150$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

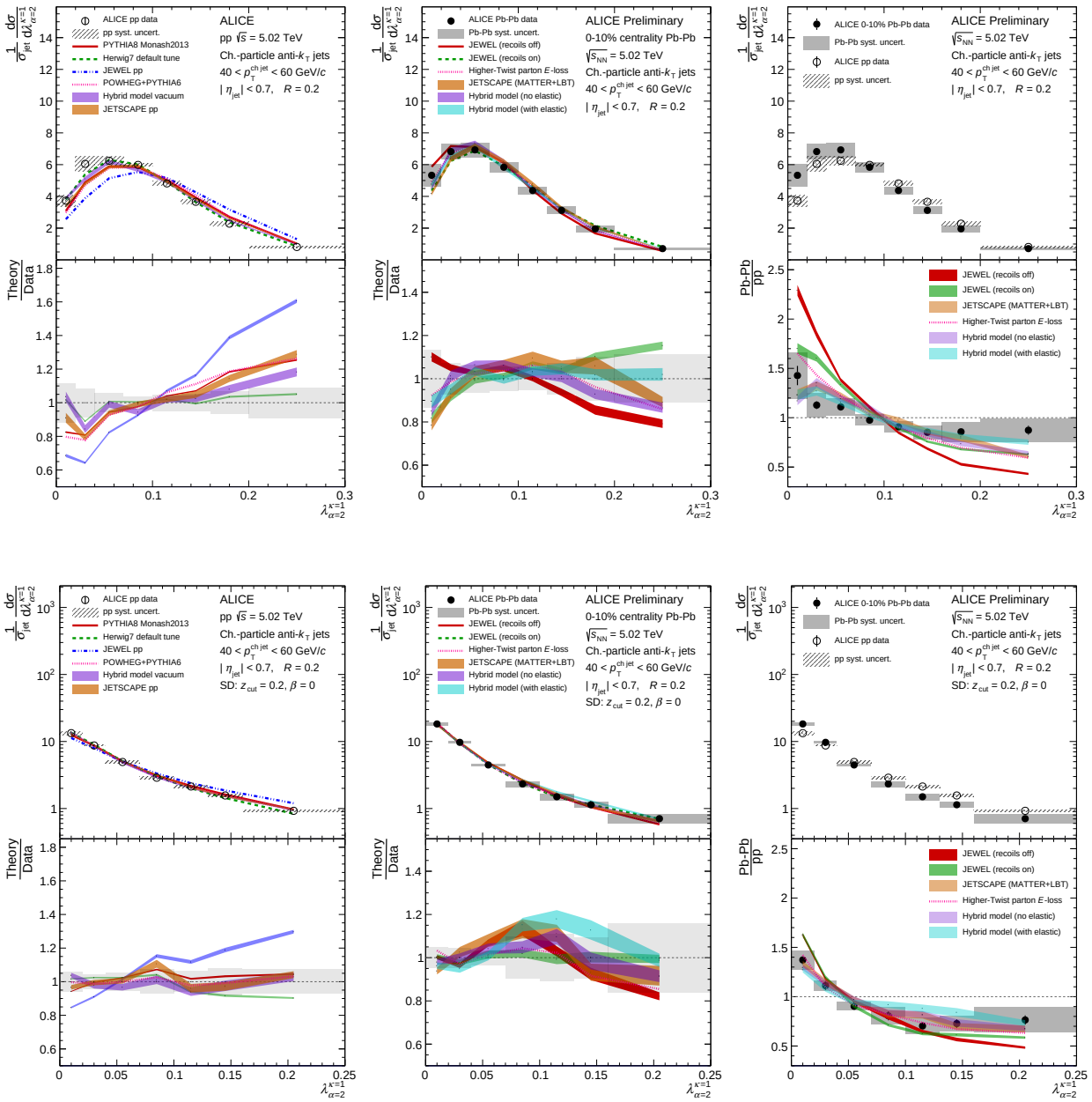


Figure 5.30: ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_2 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV for $40 < p_{\rm T}^{\rm ch\ jet} < 60$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

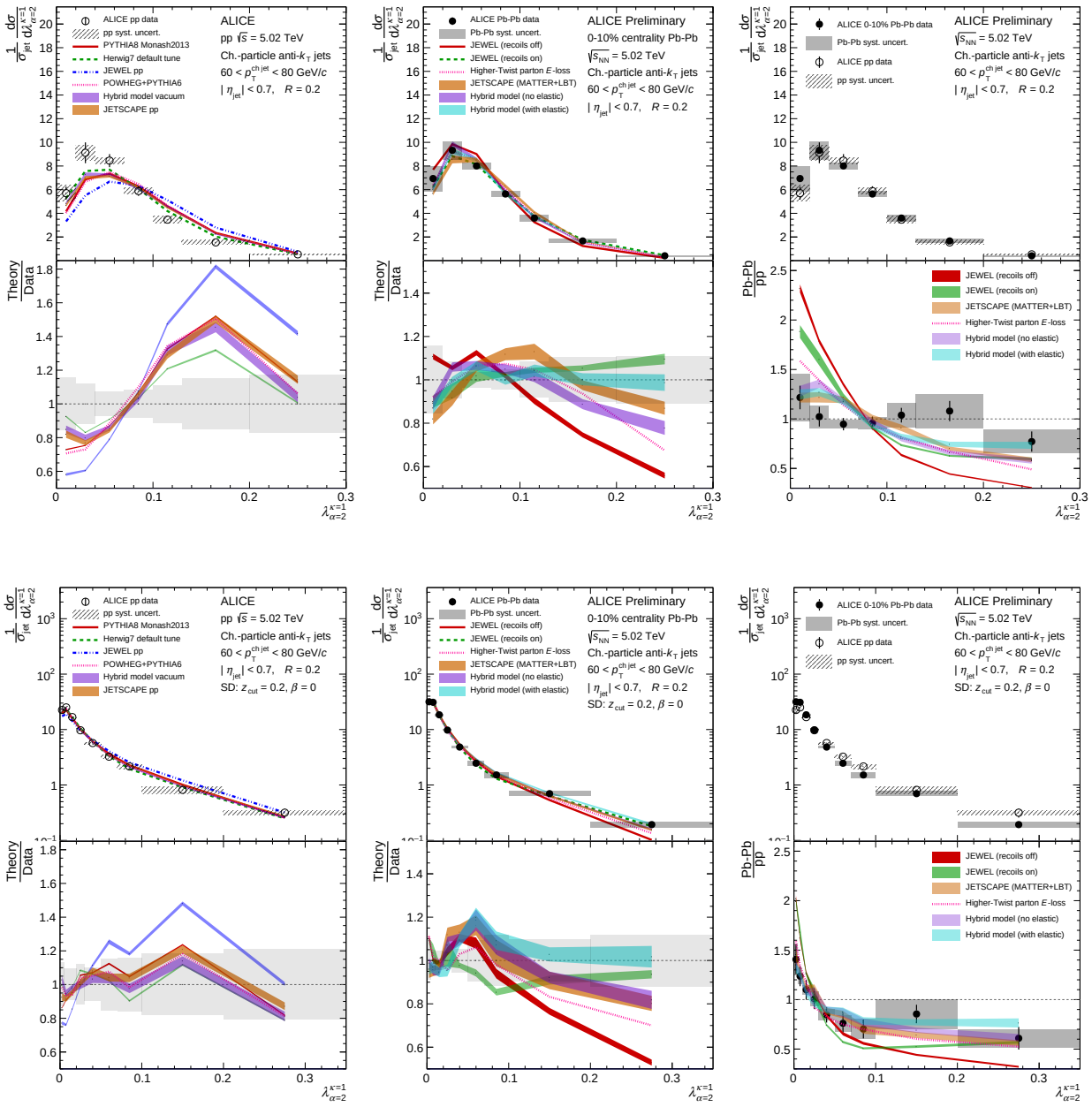


Figure 5.31: ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_2 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV for $60 < p_{\rm T}^{\rm ch\ jet} < 80$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

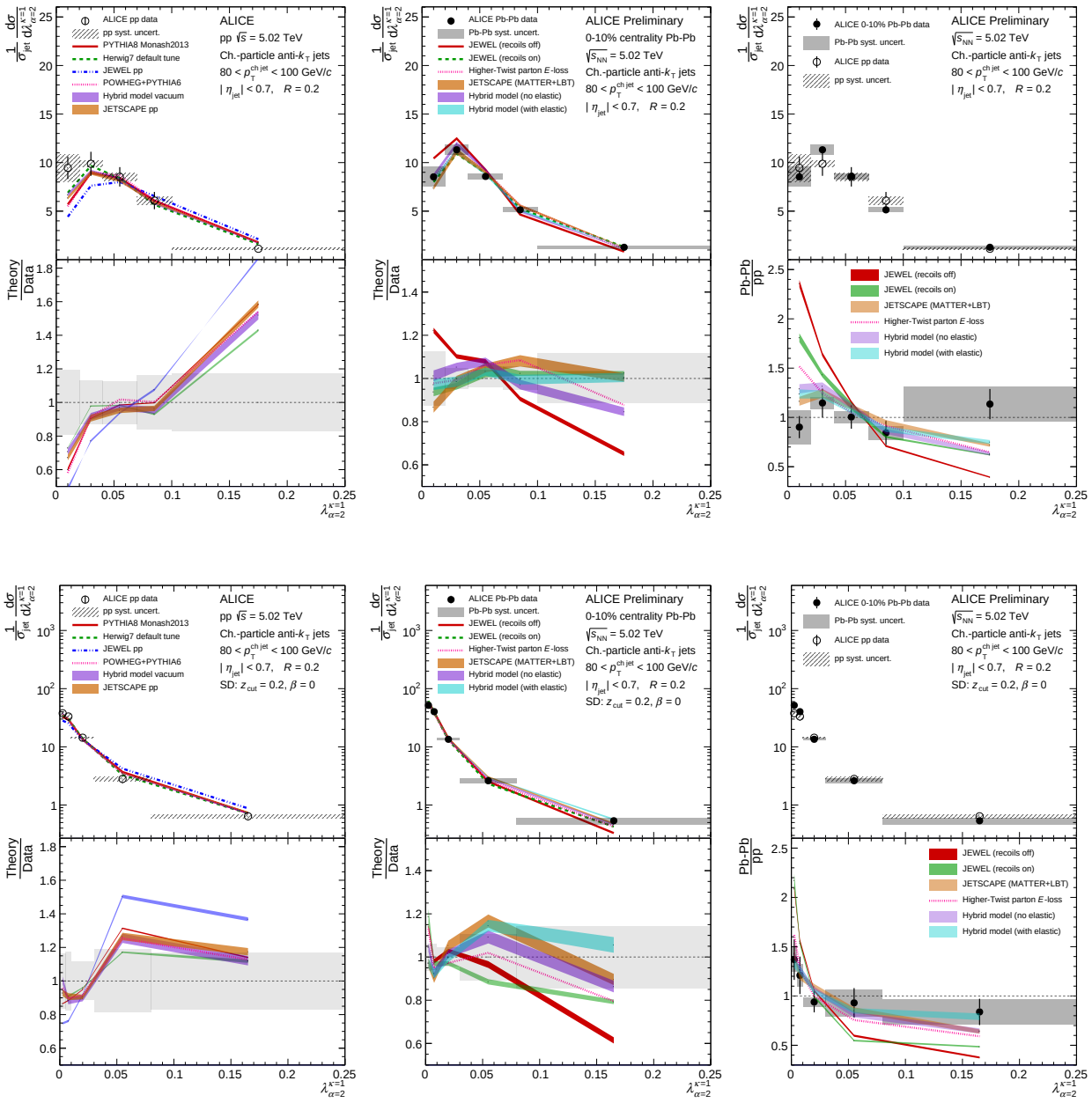


Figure 5.32: ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_2 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV for $80 < p_{\rm T}^{\rm ch\ jet} < 100$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

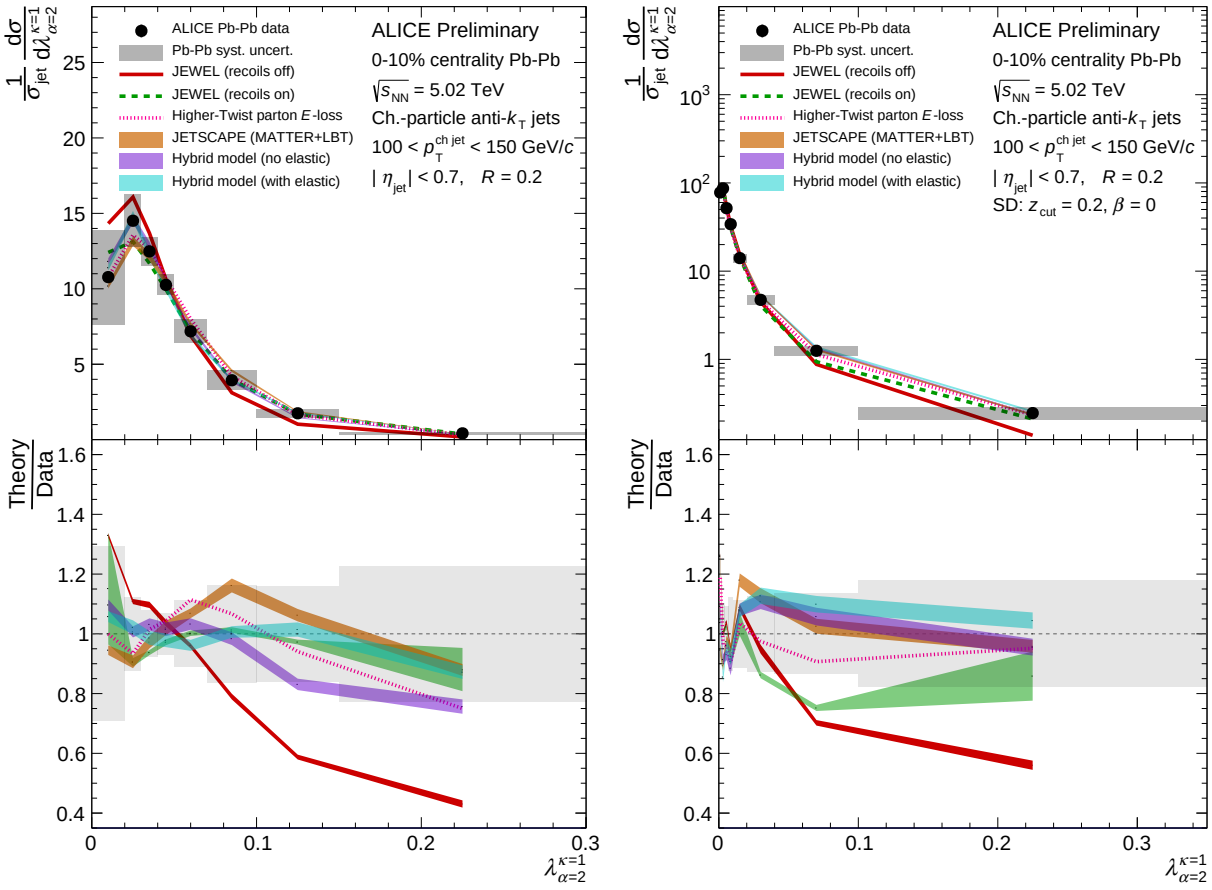


Figure 5.33: ALICE measurement of ungroomed (left) and SD groomed (right) λ_2 for $R = 0.2$ charged-particle jets in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $100 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 150$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

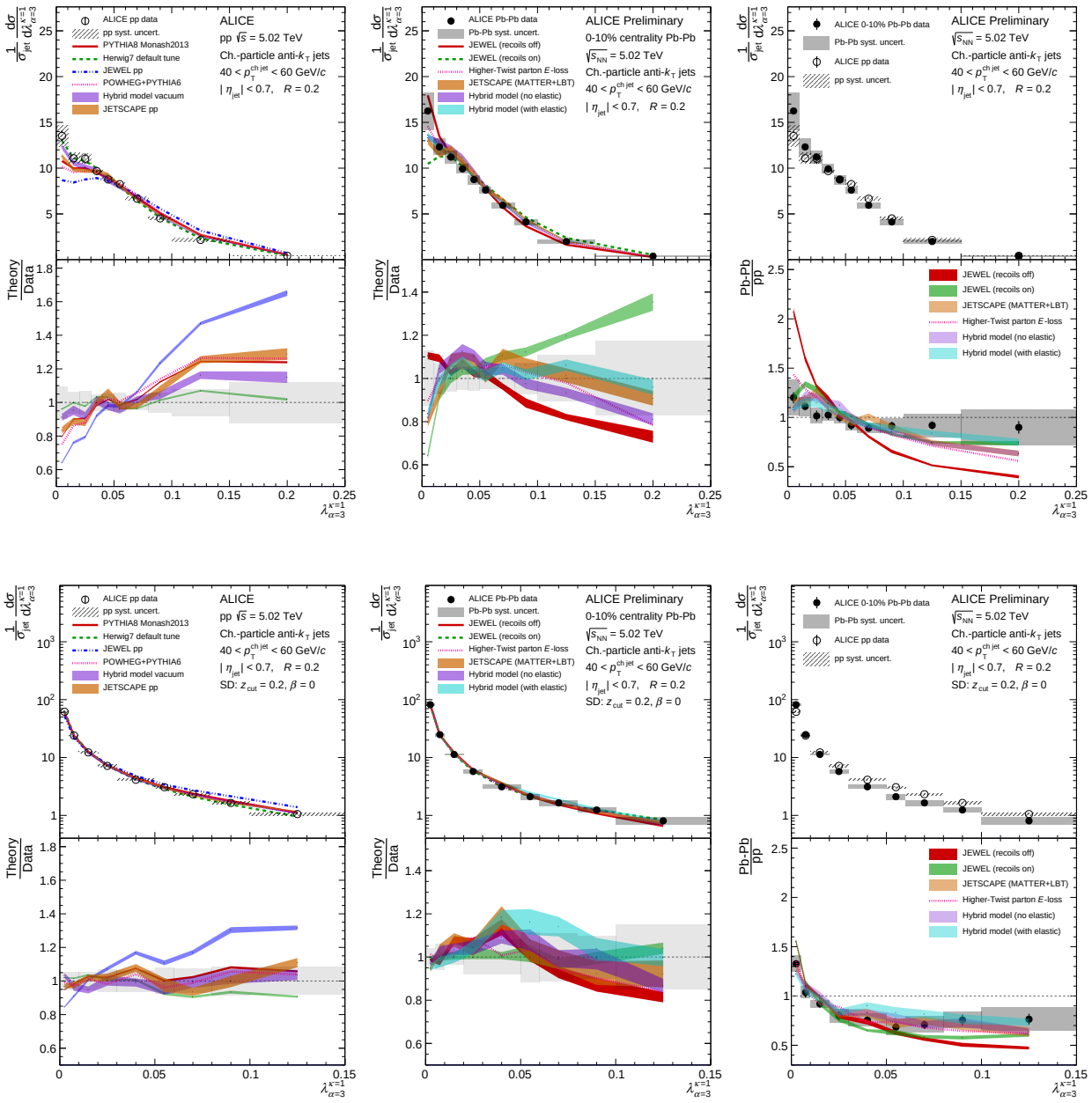


Figure 5.34: ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_3 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $40 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 60$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

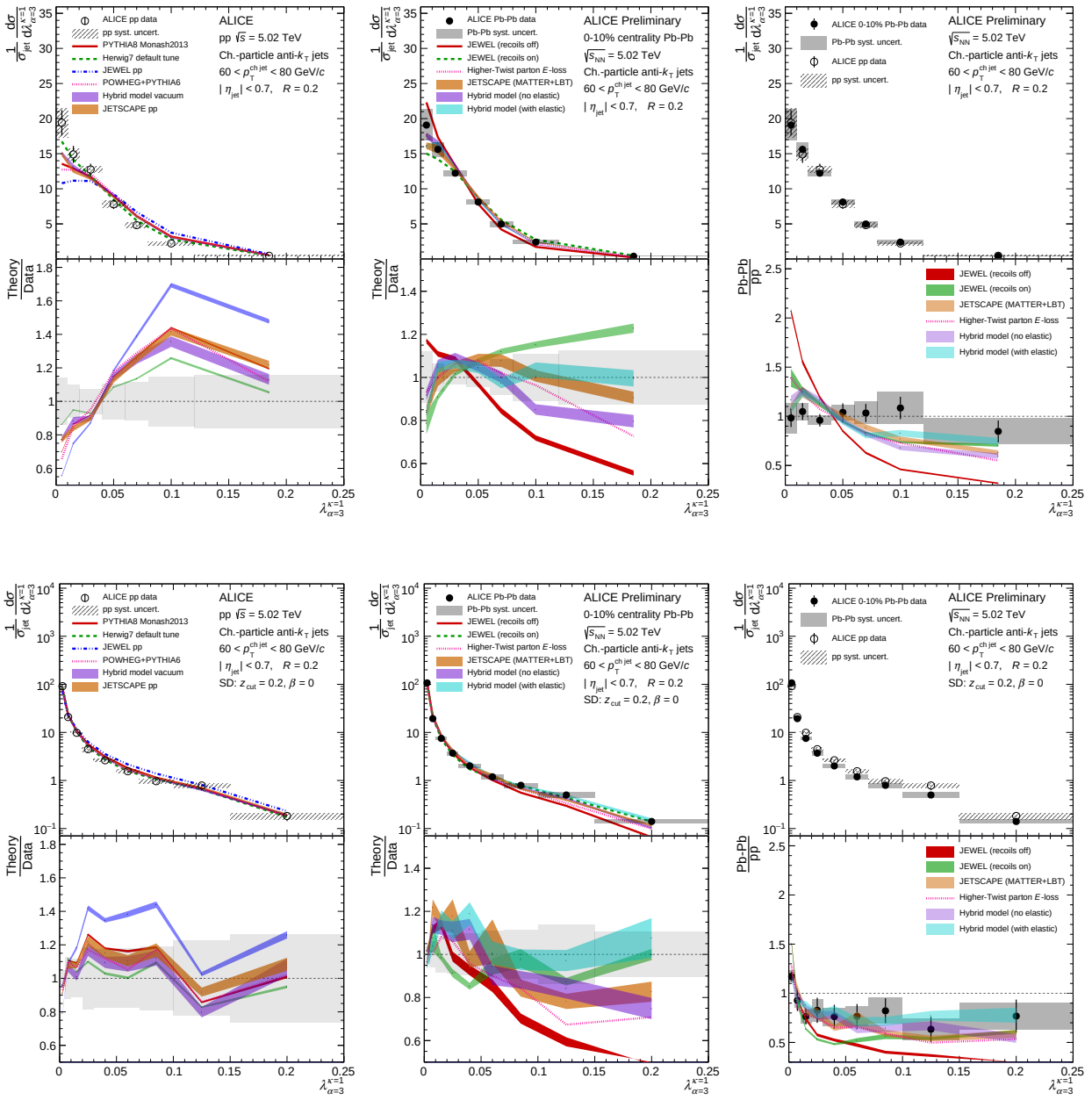


Figure 5.35: ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_3 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV for $60 < p_{\rm T}^{\rm ch\ jet} < 80$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

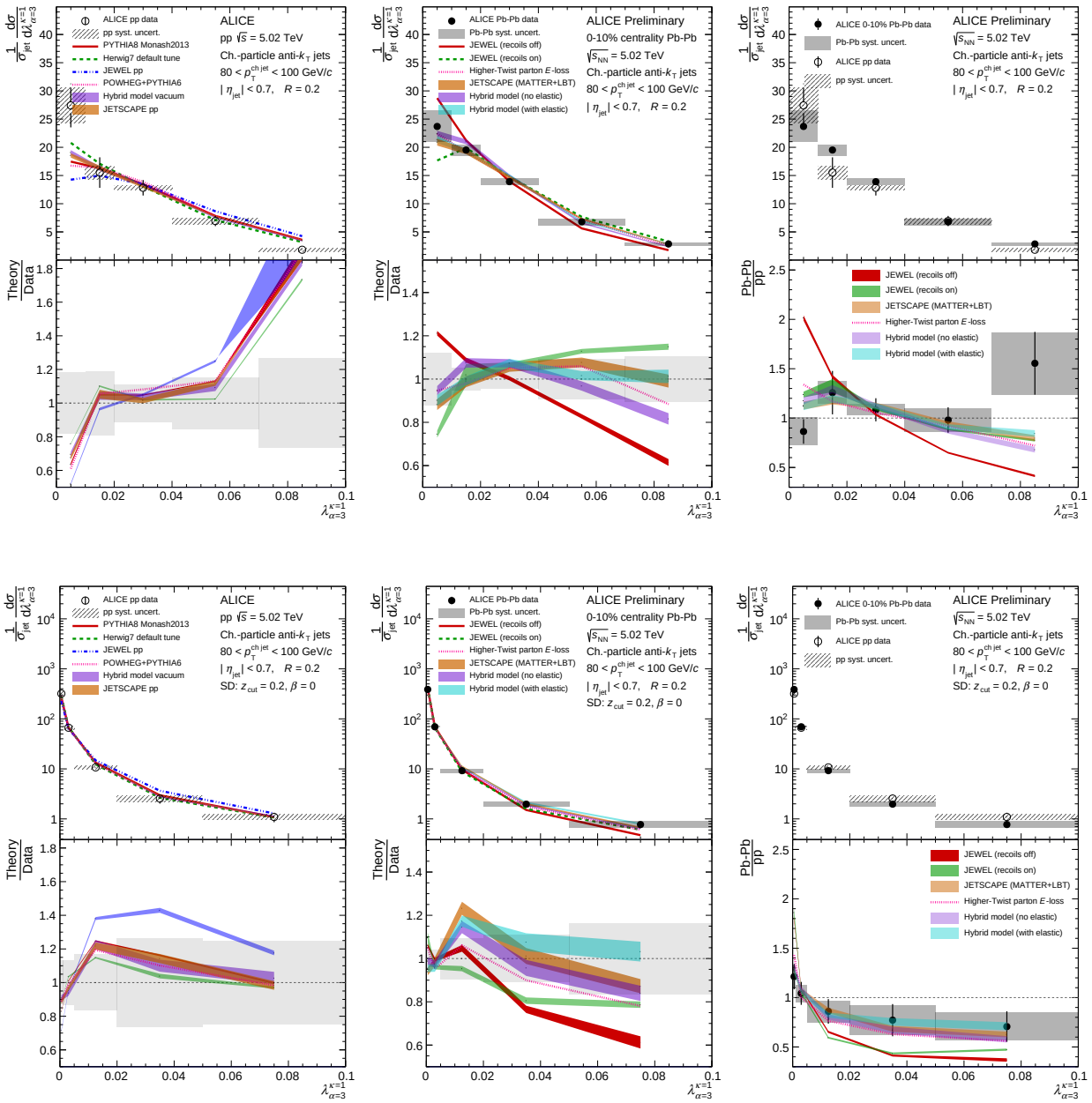


Figure 5.36: ALICE measurement of ungroomed (top) and SD groomed (bottom) λ_3 for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $80 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 100$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

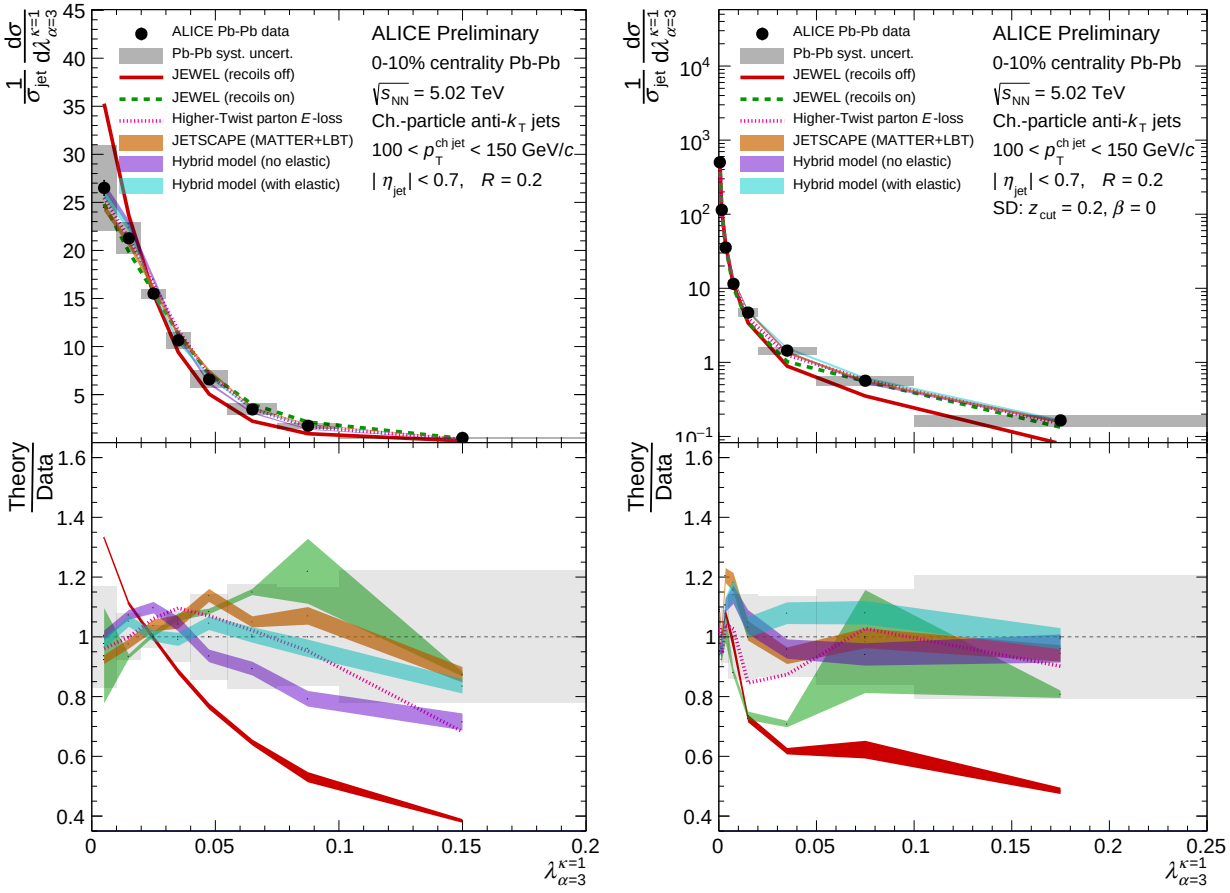


Figure 5.37: ALICE measurement of ungroomed (left) and SD groomed (right) λ_3 for $R = 0.2$ charged-particle jets in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $100 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 150$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

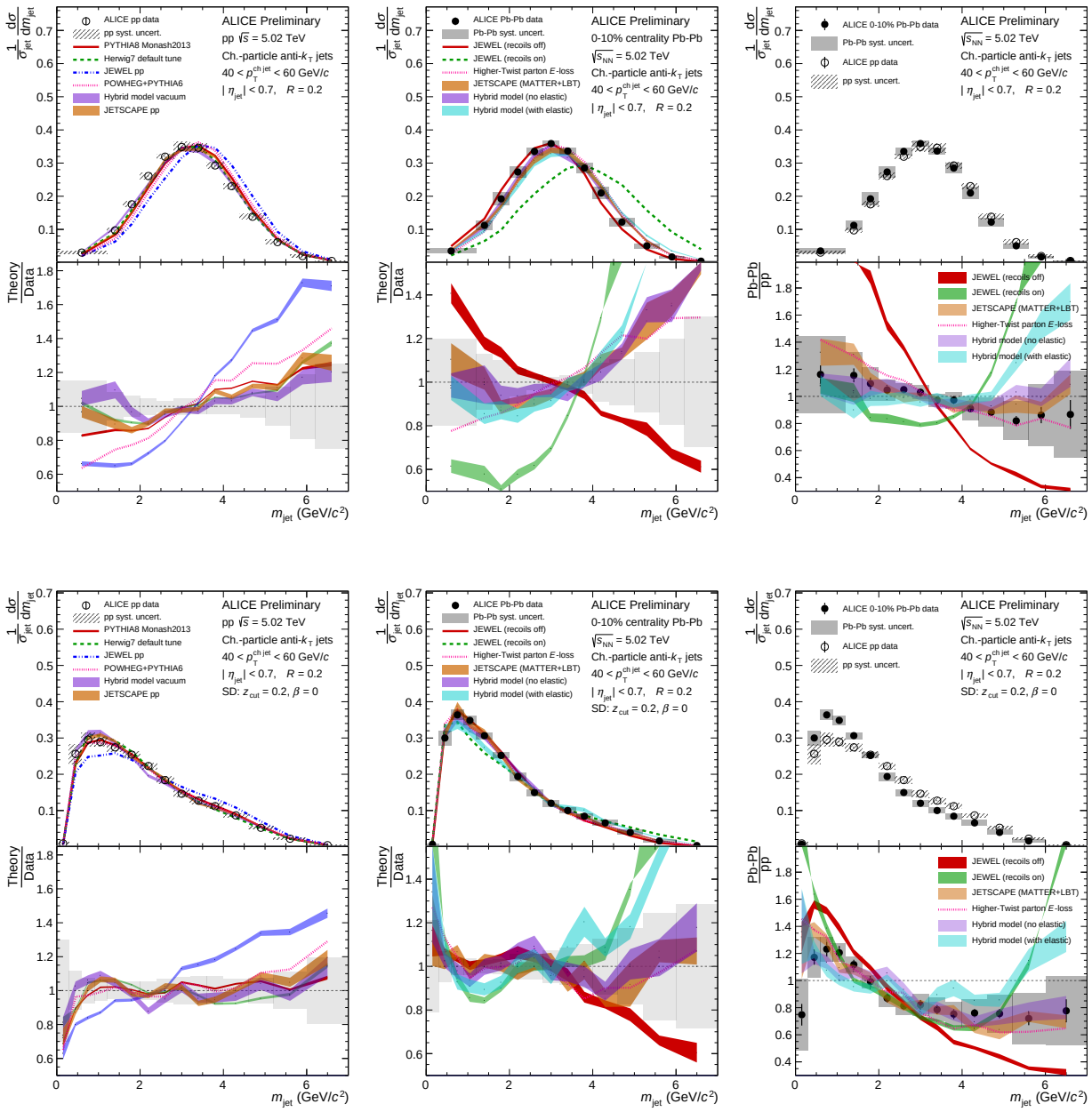


Figure 5.38: ALICE measurement of ungroomed (top) and SD groomed (bottom) m_{jet} for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $40 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 60$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

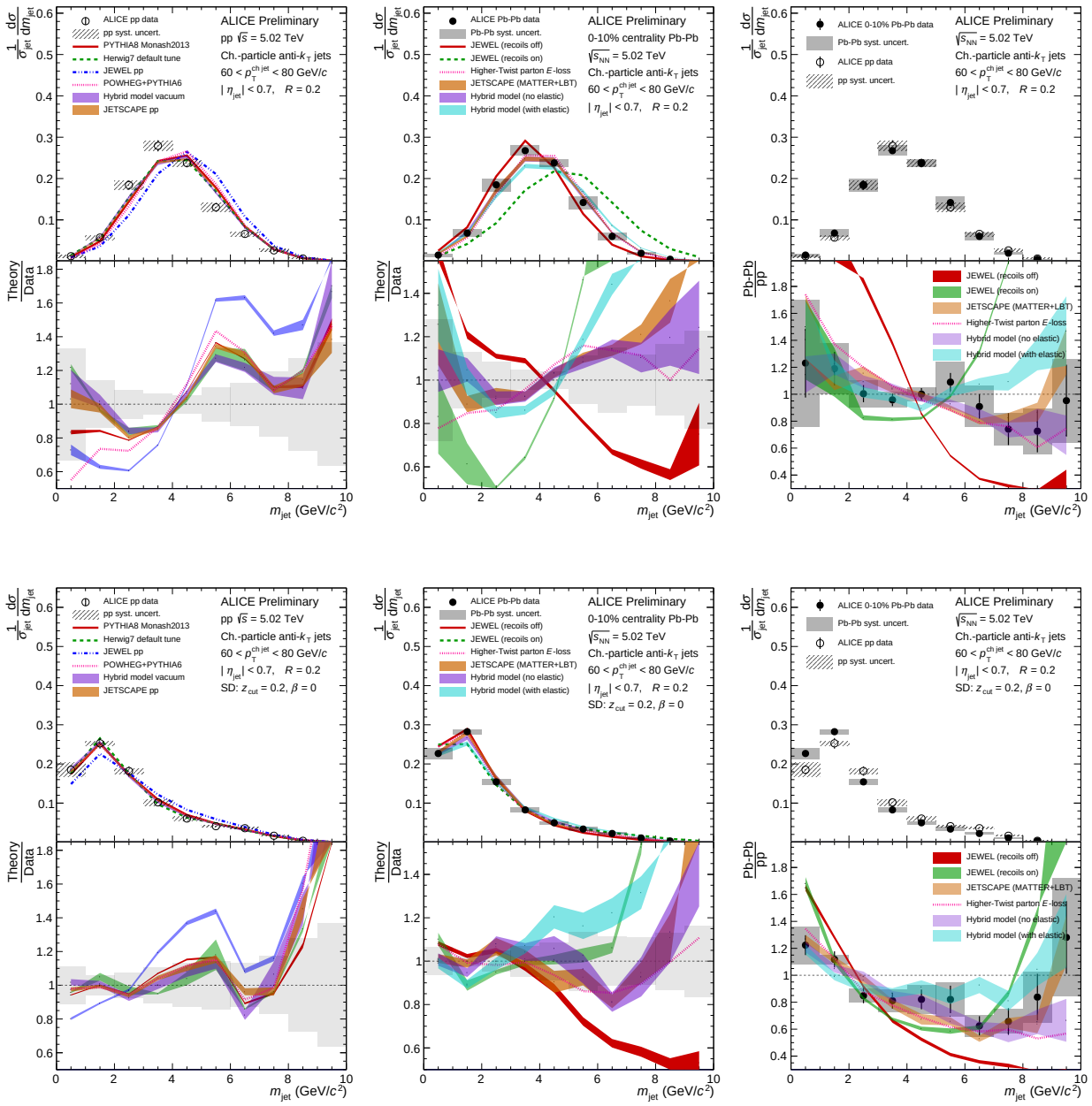


Figure 5.39: ALICE measurement of ungroomed (top) and SD groomed (bottom) m_{jet} for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $60 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 80$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

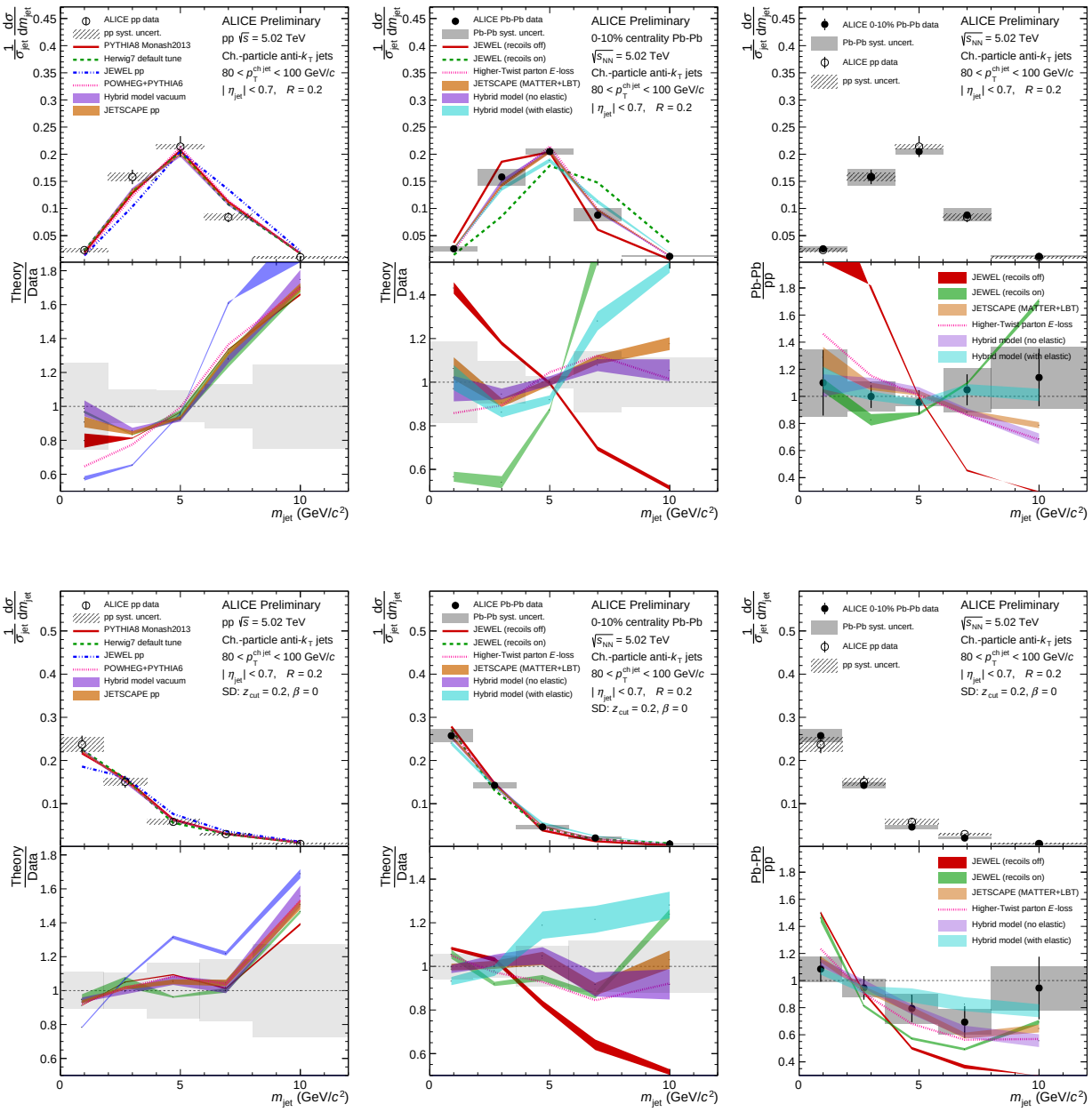


Figure 5.40: ALICE measurement of ungroomed (top) and SD groomed (bottom) m_{jet} for $R = 0.2$ charged-particle jets in pp (left) and Pb–Pb (middle) collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $80 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 100$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

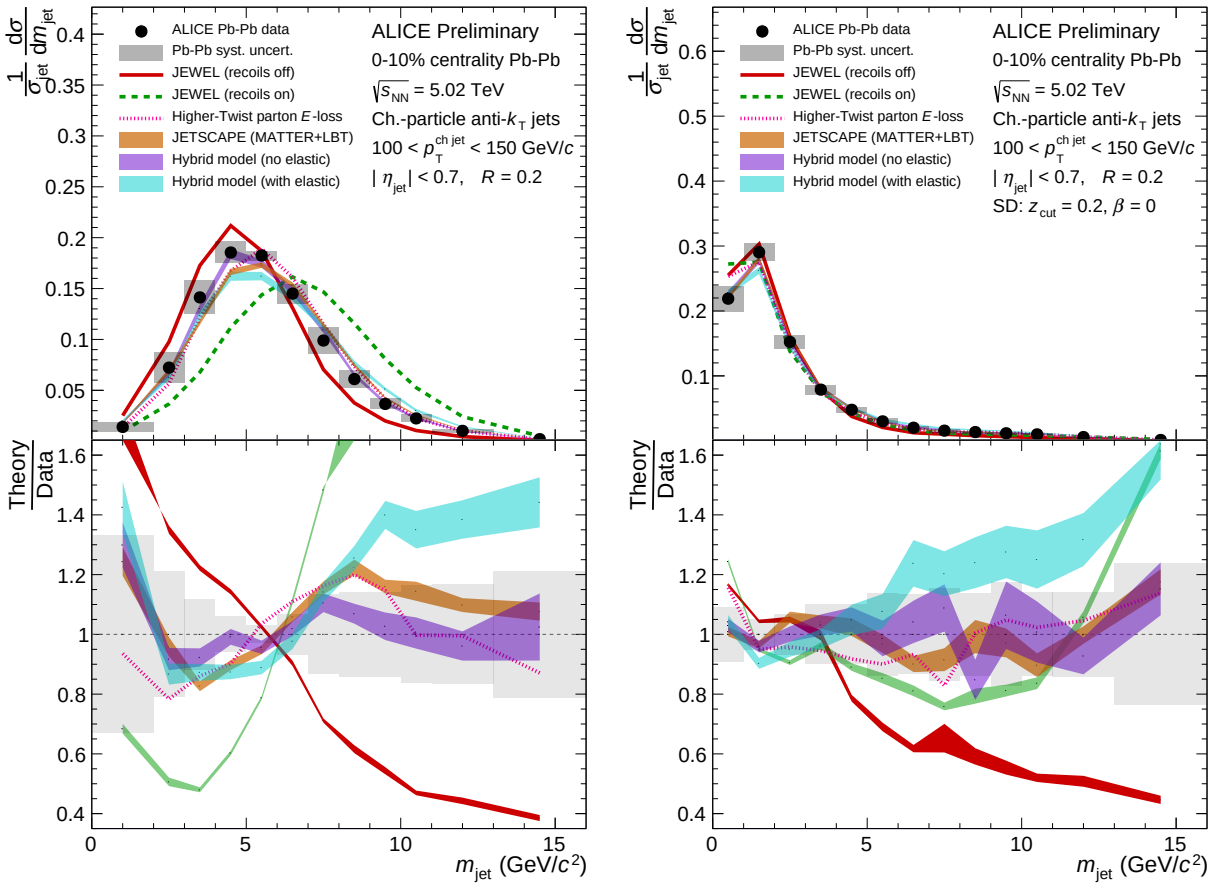


Figure 5.41: ALICE measurement of ungroomed (left) and SD groomed (right) m_{jet} for $R = 0.2$ charged-particle jets in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV for $100 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 150$ GeV/c as compared to models. The ratio of Pb–Pb to pp is also shown (right), which quantifies the substructure modifications from quenching.

Chapter 6

Discussion and conclusions

The selection of experimental studies presented here offer several insights into QCD in vacuum as well as collective QCD effects at high temperature and energy density. In particular, these studies of jet substructure offer insight into jet fragmentation, the "parton shower" of QCD, as well as nonperturbative effects such as hadronization. When compared to data from heavy-ion collisions, these results extend insight into the dynamics of a strongly-coupled QGP.

6.1 Discussion

Jet grooming can be used to identify specific splittings in the jet fragmentation. The momentum fraction z_{g} and angle θ_{g} of the tagged splittings can then be calculated, as presented in Sect. 5.1 (see also Ref. [300]). The SD distributions show that soft jet fragmentation is typically observed, with a peak in the z_{g} distribution at $\lesssim 0.1$ when using grooming settings $z_{\mathrm{cut}} = 0.1$ and $\beta \geq 0$. The splitting angle θ_{g} is typically smaller with stricter grooming settings ($\beta = 0$) than with looser grooming setting ($\beta > 0$), which is consistent with the jet having soft radiation at broad angles.

The results with dynamical grooming offer additional insights. The TimeDrop case ($a = 2$), which selects the splitting with the shortest formation time, confirms that early emissions peak at small momentum with large splitting angles. Splittings with the largest relative k_{T}, as selected by $a = 1$, also seem to carry z_{g} but occur at more moderate angles. When the most symmetric z_{g} is demanded ($a = 0.1$), the angle of the splitting becomes much more collinear. Understanding the covariance between momentum and angle for emissions produced in the parton shower provides guidance for interpreting their involvement in observables where they are mixed, such as λ_α and m_{jet}.

Reasonable agreement is also observed between these experimental data and predictions from pQCD, which include resummation of large logarithms at all orders in α_s. This conclusion holds for both algorithms and for all grooming settings considered. Predictions for θ_{g} agree with data within the perturbatively-dominated region; however, some deviations

are seen in comparisons to both pQCD predictions and MC event generators at small θ_g, in the predominantly-nonperturbative region where higher-order perturbative effects as well as nonperturbative influences such as hadronization could be more significant. The deviations grow larger for increased grooming parameter β, which signifies an increased dominance of these effects for jets containing more collinear radiation.

The groomed and ungroomed jet angularities λ_α shown in Sect. 5.2 (see also Ref. [289]) provide constraints on models and capture the interplay between perturbative and nonperturbative effects in QCD. Systematic variations of the contributions from collinear and soft radiation of the shower, captured within a given R, are provided by measuring the jet angularities for a selection of α parameters. These results consequently provide tests of pQCD calculations.

The theoretical predictions at NLL$'$ in SCET show an overall agreement with the data for jets with values of λ_α in the perturbatively-dominated regime ($\lambda_\alpha > \lambda_\alpha^{\mathrm{NP}}$) delimited by a collinear-soft momentum scale of about 1 GeV/c in the factorization framework. The calculations, after accounting for nonperturbative effects by two different methods, are compatible within about 20% or better with the data in the perturbative region for all explored values of R and α. However, larger deviations of up to about 50% are observed in the tails of some distributions, which may be due to missing perturbative corrections which are expected to become significant at large-λ_α [99].

By making comparisons solely in the perturbatively-dominated regime, consistency is seen with a predicted universal scaling of the nonperturbative shape function parameter Ω_α with value $\Omega < 1$. A clear breakdown of the agreement is observed for small λ_α, where the perturbative calculation is expected to fail. Such nonperturbative effects include soft splittings and hadronization, and these effects dominate over significant regions of the phase space of moderate and low-energy jets. This is corroborated by the comparison of the measured groomed jet angularities to the equivalent theoretical calculations, which demonstrate a wider range of agreement with the perturbative calculations.

These perturbative tests are extended by comparing these results to the angularities for jets containing a D^0 meson, as detailed in Appendix B and reported in Sect. 5.2.1. The QCD dead cone modifies jet fragmentation through the relative suppression of gluon radiation, which is observed in substructure observables such as λ_α. Dead cone modification is extracted via comparison of the D^0-tagged jet distributions to a semi-inclusive baseline, which MC predictions reproduce within experimental uncertainties. Further MC studies using jets tagged with an initiating quark or gluon show that dead cone effects are increasingly suppressed at large α, where Casimir color effects play a more important role. These studies therefore test both the angular dependence of the QCD dead cone as well as the differing fragmentation of quark- versus gluon-initiated jets.

All these results provide critical guidance for measurements in high-energy heavy-ion collisions where the internal structure of jets undergo modifications via scatterings of jet fragments with the hot and dense QCD medium. Measurements of the groomed and ungroomed λ_α and m_{jet}, reported in Sect. 5.3 (also see Ref. [286, 307]), probe the angular dependence of jet quenching and seek to clarify previous measurements where a vacuum

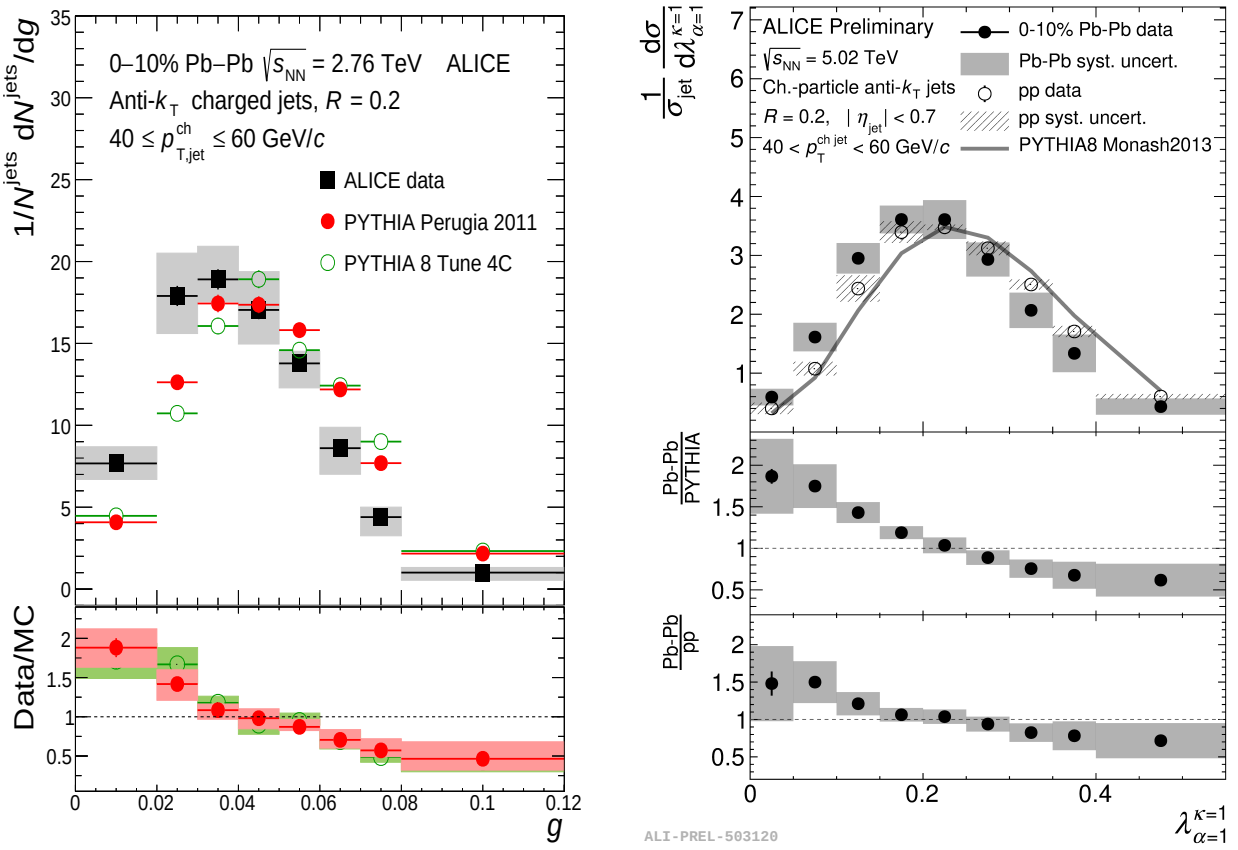


Figure 6.1: Measurement of the jet girth $g = \lambda_1^1 R$ in Pb–Pb data at $\sqrt{s_{\rm NN}} = 2.76$ TeV from LHC Run 1 (left) and at $\sqrt{s_{\rm NN}} = 5.02$ TeV from Run 2 (right). On the left figure, the bottom ratio panel compares data to PYTHIA baselines, as does the central ratio panel on the right figure. However, the comparison to a proper pp baseline in the bottom ratio panel significantly pushes the ratio closer to unity.

baseline was unavailable. Comparisons of heavy-ion data to pp MC simulations show strong modification in both the earlier LHC Run 1 [223] and current Run 2 [286] data sets, with quenched jets exhibiting a "narrowing" behavior via an enhancement at small values of angularity (girth), and a corresponding suppression at large values, with both tails modified by an approximate factor of 2. The results from both of these data samples are juxtaposed in Fig. 6.1. However, the new Run 2 result is also compared to the baseline of pp data taken at equivalent center-of-mass energy, significantly reducing the narrowing effect. The MC model-skewed ratio arbitrarily enhances perceived quenching of this observable. A proper pp baseline is therefore essential for correctly interpreting measurements of jet quenching in an unbiased way. This conclusion has far-reaching implications for future runs at the LHC: heavy-ion data must pair with statistically consistent jet samples in pp, where smaller collision systems result in fewer jets.

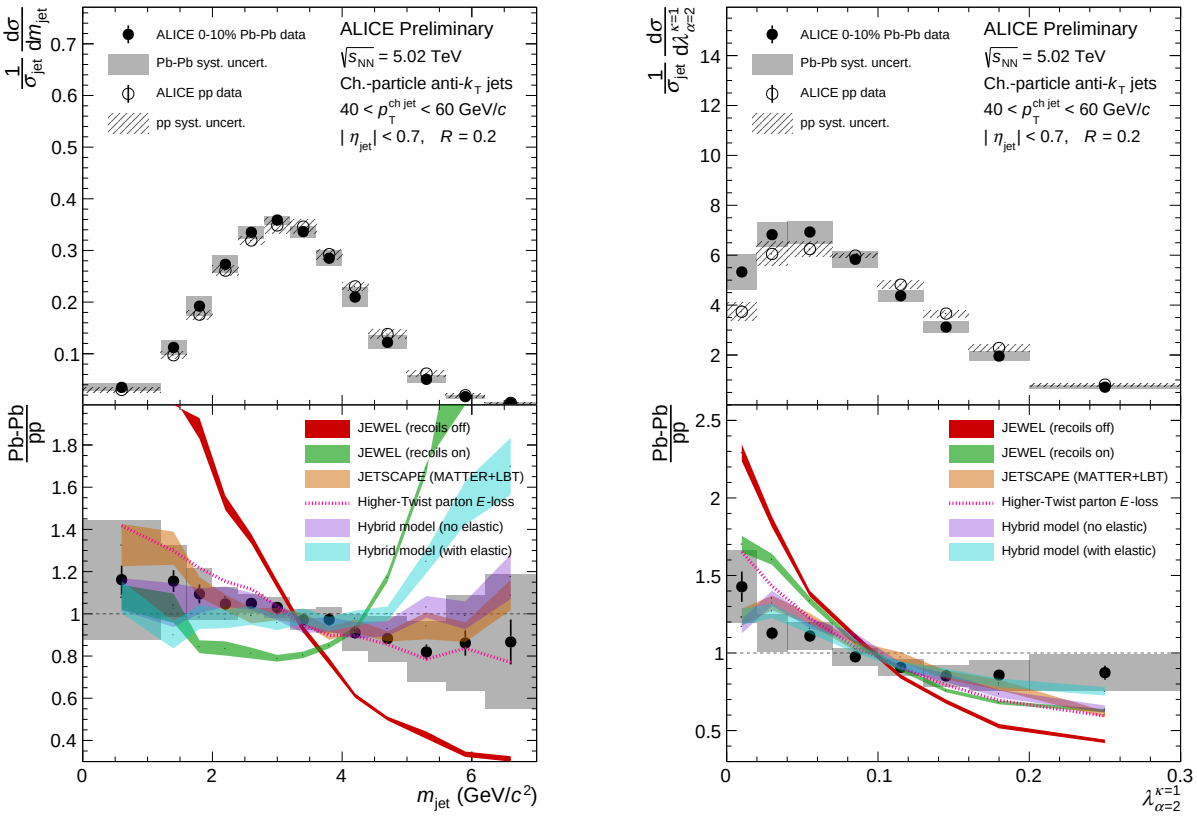


Figure 6.2: Measurement of the jet mass m_{jet} (left) and thrust λ_2^1 (right) in Pb–Pb and pp data at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV, as presented in Sect. 5.3. Systematic uncertainties are assumed to be totally uncorrelated between pp and Pb–Pb. Comparisons are made to several models [192, 195, 197–201, 305], which show varying behavior with respect to measured data.

In order to study the girth-mass puzzle in light of Eq. 1.28, ALICE has performed new measurements of m_{jet} and λ_2^1 using the same jet sample for the first time. Figure 6.2 compares the two distributions using identical $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ intervals. While Eq. 1.28 relates m_{jet} and λ_2 directly to one another, model comparisons show differing behavior. JEWEL with recoils on, for example, overestimates enhancement at large values of m_{jet}, while it underestimates the yield at large λ_2. Since the distributions are positive definite and obey square proportionality following Eq. 1.28, large corrections to Eq. 1.28 must apply at these values of $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$. This could include nonperturbative effects such as hadronization as well as higher-order correction terms $\mathcal{O}[(\lambda_2)^2]$. Despite their mathematical similarity, underlying physical differences between the two observables exist: the jet mass is sensitive to quark masses, whereas the IRC-safe jet angularities are sensitive to fragmentation and quark- versus gluon-initiated jet differences. Identifying the variations in the measured distributions as these physical differences of the observables explains the girth-mass difference.

This observation highlights the importance of making broad measurements of quenched jet substructure, as closely-related observables can provide significantly different probes of underlying physical phenomena. Studies of quenched jets using N-subjettiness variables as a basis suggest that dozens of such observables may be required to optimally characterize quenched jet behavior [308].

These measurements are not able to significantly differentiate between JETSCAPE, the Higher-Twist formalism, and the Hybrid model without elastic scattering. The addition of elastic Molière scattering to the Hybrid approach enhances wide-angle constituents and correspondingly boosts the quenching modification at large m_{jet} and λ_α; however, this effect reduces agreement with experimental data. This suggests that such scattering may not occur or may not have a significant effect in the QGP. This conclusion also qualitatively agrees with recent measurements of the groomed jet k_{T} [309], which show similar evidence using a different approach based on grooming to tag individual splittings. The jet narrowing observed in Pb–Pb data as compared to pp, which is strongest at smaller values of α and with SD grooming as discussed in Sect. 5.3, is also consistent with other recent measurements of jet quenching, such as the jet axis differences [310].

As discussed above, jet grooming enhances the perturbative calculability of jet substructure observables, while its use in heavy-ion collisions additionally reduces contamination from the thermal background. Compared to their ungroomed counterparts, the groomed λ_α and m_{jet} distributions display reduced systematic uncertainties and an enhanced narrowing effect, consistent with a strongly quenched jet core. Increased similarity between models also suggests perturbative agreement and a continuing need for probing nonperturbative effects. For the jet angularities, several model predictions converge with grooming, limiting the differentiating power between them – despite different theoretical approaches. Scrutinizing jet quenching models therefore requires consideration of both the significant nonperturbative and perturbative effects.

6.2 Outlook

Despite these studies successfully yielding several lessons outlined above, a general understanding of QCD at all momentum scales is still incomplete. While lattice calculations continue to describe increasingly large systems [311–313], pQCD remains the dominant means for making predictions. This approach can successfully describe a wide range of observables, but high-precision data from the LHC across broad transverse momenta have made perturbative limitations equally apparent. Nonperturbative effects exhibit significant influence on jet observables at low $p_{\text{T}}^{\text{ch jet}}$ and at small R, where the existing phenomenological models do not fully account for discrepancies. Theoretical work is expanding these horizons: since the publication of the λ_α measurements presented here, a more careful shape function approach (using three $\mathcal{O}(\Lambda_{\text{QCD}})$ parameters instead of one) has been developed for hadronization corrections to the groomed jet mass [314]. This enables a future experimental study of hadronization power corrections in a completely model-independent way.

Similarly, the apparent hadronization universality breaking for jets containing heavy-flavor hadrons [125] remains unexplained; see Sect. 1.3.4. The results presented here show that heavy-flavor jet fragmentation is well-described by PYTHIA 8 MC simulations, which are generated from pQCD at LO and coupled to a phenomenological parton shower and tuned hadronization model. However, this same MC model under-predicts the yield for charm baryons as compared to mesons, which is tuned from e^+e^- collisions [125]. Future studies of jet substructure must search for the origin of this discrepancy to identify the momentum and angular scale at which deviations in the charm meson versus baryon clustering occur.

These studies of jet substructure in heavy-ion collisions have been successful at explaining the girth-mass difference as well as resolving differences in model comparisons. In the Hybrid model approach, the data favors a QCD medium which produces a strong wake behind hard-scattered partons, but without elastic Molière scattering. Other substructure studies have been able to identify that medium-induced energy loss is generally incoherent, with data suggesting a resolution length $L_{\text{res}} = 0$ [310], and to distinguish momentum broadening as a subleading effect to jet narrowing [103]. An understanding of how these behaviors arise from first-principles QCD, however, is yet to be ascertained, and a formal study of factorization in the presence of a QGP is yet to be performed.

As the LHC continues data taking into Runs 3 and 4, and as theoretical precision continues to improve, the experimental collaborations at the LHC have an excellent opportunity to develop a more complete understanding of strong interactions. An increased instantaneous luminosity, higher collision energies, and various detector upgrades will enable the LHC experiments to probe QCD and search for new physics at scales previously unreachable. The construction of the Electron-Ion Collider (EIC) [315] at the current RHIC facility in New York state will offer opportunities to broaden lessons from the LHC by studying polarized and unpolarized nucleon and nuclear structure in previously inaccessible regimes. The implications of lessons learned here and in the next several years will be paramount for successful execution of the EIC physics program in completing a generalized understanding of the many wonderful physical marvels which arise from Eq. 1.1.

Bibliography

[1] Michael Peskin and Daniel Schroeder. *An Introduction To Quantum Field Theory*. Westview Press, 1995. ISBN: 978-0-367-32056-0.

[2] R. K. Ellis, W. J. Stirling, and B. R. Webber. *QCD and Collider Physics*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 1996. DOI: 10.1017/CBO9780511628788.

[3] **Particle Data Group**: Ron Workman et al. *Review of Particle Physics*. Prog. Theor. Exp. Phys. (2022) 083C01. DOI: 10.1093/ptep/ptac097.

[4] Walter Greiner, Stefan Schramm, and Eckart Stein. *Quantum Chromodynamics*. New York: Springer, 2007. ISBN: 978-3-540-48534-6. DOI: 10.1007/978-3-540-48535-3.

[5] Matthew Schwartz. *Quantum Field Theory and the Standard Model*. New York: Cambridge University Press, 2013. ISBN: 978-1-107-03473-0.

[6] Mark Thomson. *Modern Particle Physics*. New York: Cambridge University Press, 2013. ISBN: 978-1-107-03426-6. DOI: 10.1017/CBO9781139525367.

[7] David J. Gross and Frank Wilczek. "Ultraviolet Behavior of Non-Abelian Gauge Theories". In: *Phys. Rev. Lett.* 30.26 (1973), pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343.

[8] H. David Politzer. "Reliable Perturbative Results for Strong Interactions?" In: *Phys. Rev. Lett.* 30.26 (1973), pp. 1346–1349. DOI: 10.1103/PhysRevLett.30.1346.

[9] Kenneth G. Wilson. "Confinement of quarks". In: *Phys. Rev. D* 10.8 (1974), pp. 2445–2459. DOI: 10.1103/PhysRevD.10.2445.

[10] R. Vogt. "Physics of the nucleon sea quark distributions". In: *Prog. Part. Nucl. Phys.* 45 (2000), S105–S169. ISSN: 0146-6410. DOI: 10.1016/S0146-6410(00)90012-7. arXiv: hep-ph/0011298.

[11] D. F. Geesaman and P. E. Reimer. "The sea of quarks and antiquarks in the nucleon". In: *Rep. Prog. Phys.* 82.4 (2019), p. 046301. DOI: 10.1088/1361-6633/ab05a7. arXiv: 1812.10372 [nucl-ex].

[12] Xiang Gao et al. "Unpolarized proton PDF at NNLO from lattice QCD with physical quark masses". In: *Phys. Rev. D* 107.7 (2023), p. 074509. DOI: 10.1103/PhysRevD.107.074509.

[13] Rajan Gupta. *Introduction to Lattice QCD*. 1998. arXiv: hep-lat/9807028.

[14] C. T. H. Davies et al. "High-Precision Lattice QCD Confronts Experiment". In: *Phys. Rev. Lett.* 92.2 (2004), p. 022001. DOI: 10.1103/PhysRevLett.92.022001.

[15] Tim Byrnes and Yoshihisa Yamamoto. "Simulating lattice gauge theories on a quantum computer". In: *Phys. Rev. A* 73.2 (2006), p. 022328. DOI: 10.1103/PhysRevA.73.022328.

[16] Steven Weinberg. "Phenomenological Lagrangians". In: *Physica A* 96.1 (1979), pp. 327–340. ISSN: 0378-4371. DOI: 10.1016/0378-4371(79)90223-1.

[17] Steven Weinberg. "Nuclear forces from chiral lagrangians". In: *Phys. Lett. B* 251.2 (1990), pp. 288–292. ISSN: 0370-2693. DOI: 10.1016/0370-2693(90)90938-3.

[18] B. Andersson et al. "Parton fragmentation and string dynamics". In: *Phys. Rep.* 97.2 (1983), pp. 31–145. ISSN: 0370-1573. DOI: 10.1016/0370-1573(83)90080-7.

[19] Gerhard A. Schuler and Torbjörn Sjöstrand. "Hadronic diffractive cross sections and the rise of the total cross section". In: *Phys. Rev. D* 49.5 (1994), pp. 2257–267. DOI: 10.1103/PhysRevD.49.2257.

[20] Bryan R. Webber. "A QCD model for jet fragmentation including soft gluon interference". In: *Nucl. Phys. B* 238.3 (1984), pp. 492–528. ISSN: 0550-3213. DOI: 10.1016/0550-3213(84)90333-X.

[21] Giuseppe Marchesini and Bryan R. Webber. "Monte Carlo simulation of general hard processes with coherent QCD radiation". In: *Nucl. Phys. B* 310.3 (1988), pp. 461–526. ISSN: 0550-3213. DOI: 10.1016/0550-3213(88)90089-2.

[22] Giuseppe Marchesini et al. "HERWIG 5.1 - a Monte Carlo event generator for simulating hadron emission reactions with interfering gluons". In: *Comput. Phys. Commun.* 67.3 (1992), pp. 465–508. ISSN: 0010-4655. DOI: 10.1016/0010-4655(92)90055-4.

[23] Francesco Becattini. *An introduction to the Statistical Hadronization Model*. 2009. arXiv: 0901.3643 [hep-ph].

[24] Enrico Fermi. "High Energy Nuclear Events". In: *Progress of Theoretical Physics* 5.4 (1950), pp. 570–583. ISSN: 0033-068X. DOI: 10.1143/PTP/5.4.570.

[25] Rolf Hagedorn. "Statistical thermodynamics of strong interactions at high energies". In: *Nuovo Cimento, Suppl.* 3 (1965), pp. 147–186. URL: https://cds.cern.ch/record/346206.

[26] V. Greco and C.M. Ko. "Hadronization via Coalescence". In: *Acta Phys. Hung. A* 24.1-4 (2005), pp. 235–240. DOI: 10.1556/APH.24.2005.1-4.32. arXiv: nucl-th/0405040.

[27] Johann Rafelski. "Strange anti-baryons from quark-gluon plasma". In: *Phys. Lett. B* 262.2 (1991), pp. 333–340. ISSN: 0370-2693. DOI: 10.1016/0370-2693(91)91576-H.

[28] S. M. Berman, J. D. Bjorken, and J. B. Kogut. "Inclusive Processes at High Transverse Momentum". In: *Phys. Rev. D* 4.11 (1971), pp. 3388–3418. DOI: 10.1103/PhysRevD.4.3388.

[29] A. Metz and A. Vossen. "Parton fragmentation functions". In: *J. PPNP* 91 (2016), pp. 136–202. DOI: 10.1016/j.ppnp.2016.08.003.

[30] S. Bailey et al. "Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs". In: *Eur. Phys. J. C* 81.4 (2021). DOI: 10.1140/EPJC/s10052-021-09057-0.

[31] Yu. L. Dokshitser. "Calculation of structure functions of deep-inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics". In: *Sov. Phys. JETP* 46.4 (1977), pp. 641–653. URL: http://www.jetp.ras.ru/cgi-bin/dn/e_046_04_0641.pdf.

[32] V. N. Gribov and L. N. Lipatov. "Deep inelastic ep scattering in perturbation theory". In: *Sov. J. Nucl. Phys.* 15.4 (1972), pp. 438–450.

[33] G. Altarelli and G. Parisi. "Asymptotic freedom in parton language". In: *Nucl. Phys. B* 126.2 (1977), pp. 298–318. ISSN: 0550-3213. DOI: 10.1016/0550-3213(77)90384-4.

[34] John C. Collins, Davison E. Soper, and George Sterman. "Factorization of Hard Processes in QCD". In: *Adv. Ser. Direct. High Energy Phys.* 5 (1989), pp. 1–91. DOI: 10.1142/9789814503266_0001. arXiv: hep-ph/0409313.

[35] Dennis Sivers. "Single-spin production asymmetries from the hard scattering of point-like constituents". In: *Phys. Rev. D* 41.1 (1990), pp. 83–90. DOI: 10.1103/PhysRevD.41.83.

[36] D. Boer and P. J. Mulders. "Time-reversal odd distribution functions in leptoproduction". In: *Phys. Rev. D* 57.9 (1998), pp. 5780–5786. DOI: 10.1103/PhysRevD.57.5780. arXiv: hep-ph/9711485.

[37] P. Skands, S. Carrazza, and J. Rojo. "Tuning PYTHIA 8.1: the Monash 2013 tune". In: *Eur. Phys. J. C* 74.8 (2014). DOI: 10.1140/EPJC/s10052-014-3024-y. arXiv: 1404.5630 [hep-ph].

[38] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/J.CPC.2015.01.024. arXiv: 1410.3012 [hep-ph].

[39] Christian Bierlich et al. *A comprehensive guide to the physics and usage of PYTHIA 8.3*. 2022. arXiv: 2203.11601 [hep-ph].

[40] R. Brun et al. *GEANT 3: user's guide Geant 3.10, Geant 3.11; rev. version*. Geneva: CERN, 1987. URL: https://cds.cern.ch/record/1119728.

[41] *PYTHIA Manual: Hard QCD Processes*. URL: https://pythia.org/latest-manual/QCDHardProcesses.html.

[42] Manuel Bähr et al. "Herwig++ physics and manual". In: *Eur. Phys. J. C* 58.4 (2008), pp. 639–707. DOI: 10.1140/EPJC/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].

[43] Johannes Bellm et al. "Herwig 7.0/Herwig++ 3.0 release note". In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/EPJC/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[44] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with parton shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[45] Michael H. Seymour and Marilyn Marx. "Monte Carlo Event Generators". In: (2013). arXiv: 1304.6677 [hep-ph].

[46] G. Hanson et al. "Evidence for Jet Structure in Hadron Production by e^+e^- Annihilation". In: *Phys. Rev. Lett.* 35.24 (1975), pp. 1609–1612. DOI: 10.1103/PhysRevLett.35.1609.

[47] Gavin P. Salam. "Towards jetography". In: *Eur. Phys. J. C* 67 (2010), pp. 637–686. DOI: 10.1140/EPJC/s10052-010-1314-6. arXiv: 0906.1833 [hep-ph].

[48] **TASSO** Collaboration: R. Brandelik et al. "Evidence for planar events in e^+e^- annihilation at high energies". In: *Phys. Lett. B* 86.2 (1979), pp. 243–249. ISSN: 0370-2693. DOI: 10.1016/0370-2693(79)90830-X.

[49] Carola F. Berger and George Sterman. "Scaling rule for nonperturbative radiation in a class of event shapes". In: *JHEP* 09 (2003), p. 058. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2003/09/058. arXiv: hep-ph/0307394.

[50] George Sterman and Steven Weinberg. "Jets from Quantum Chromodynamics". In: *Phys. Rev. Lett.* 39.23 (1977), pp. 1436–1439. DOI: 10.1103/PhysRevLett.39.1436.

[51] Christian W. Bauer, Sean Fleming, and Michael Luke. "Summing Sudakov logarithms in $B \to X_s$ + gamma in effective field theory". In: *Phys. Rev. D* 63.1 (2000), p. 014006. DOI: 10.1103/PhysRevD.63.014006. arXiv: hep-ph/0005275.

[52] Christian W. Bauer et al. "An effective field theory for collinear and soft gluons: Heavy to light decays". In: *Phys. Rev. D* 63.11 (2001), p. 114020. DOI: 10.1103/PhysRevD.63.114020. arXiv: hep-ph/0011336.

[53] Christian W. Bauer and Iain W. Stewart. "Invariant operators in collinear effective theory". In: *Phys. Lett. B* 516 (2001), pp. 134–142. DOI: 10.1016/s0370-2693(01)00902-9. arXiv: hep-ph/0107001.

[54] Christian W. Bauer, Dan Pirjol, and Iain W. Stewart. "Soft-collinear factorization in effective field theory". In: *Phys. Rev. D* 65.5 (2002), p. 054022. DOI: 10.1103/PhysRevD.65.054022. arXiv: hep-ph/0109045.

[55] Christian W. Bauer et al. "Hard scattering factorization from effective field theory". In: *Phys. Rev. D* 66.1 (2002), p. 014017. DOI: 10.1103/PhysRevD.66.014017. arXiv: hep-ph/0202088.

[56] Andrew J. Larkoski and Jesse Thaler. "Unsafe but calculable: ratios of angularities in perturbative QCD". In: *JHEP* 2013.09 (2013), p. 137. DOI: 10.1007/JHEP09(2013)137. arXiv: 1307.1699 [hep-ph].

[57] Andrew J. Larkoski, Simone Marzani, and Jesse Thaler. "Sudakov safety in perturbative QCD". In: *Phys. Rev. D* 91.11 (2015), p. 111501. DOI: 10.1103/PhysRevD.91.111501. arXiv: 1502.01719 [hep-ph].

[58] Gavin P. Salam and Grégory Soyez. "A practical seedless infrared-safe cone jet algorithm". In: *JHEP* 05 (2007), p. 086. DOI: 10.1088/1126-6708/2007/05/086. arXiv: 0704.0292 [hep-ph].

[59] **JADE** Collaboration: W. Bartel et al. "Experimental Studies on Multi-Jet Production in e^+e^- Annihilation at PETRA Energies". In: *Z. Phys. C* 33 (1986), pp. 23–31. DOI: 10.1007/BF01410449.

[60] **JADE** Collaboration: S. Bethke et al. "Experimental investigation of the energy dependence of the strong coupling strength". In: *Phys. Lett. B* 213.2 (1988), pp. 235–241. DOI: 10.1016/0370-2693(88)91032-5.

[61] Stefano Catani et al. "New clustering algorithm for multi-jet cross-sections in e^+e^- annihilation". In: *Phys. Lett. B* 269 (1991), pp. 432–438. DOI: 10.1016/0370-2693(91)90196-W.

[62] Stefano Catani et al. "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions". In: *Nucl. Phys. B* 406 (1993), pp. 187–224. DOI: 10.1016/0550-3213(93)90166-M.

[63] Stephen D. Ellis and Davison E. Soper. "Successive combination jet algorithm for hadron collisions". In: *Phys. Rev. D* 48.7 (1993), pp. 3160–3166. DOI: 10.1103/PhysRevD.48.3160. arXiv: hep-ph/9305266.

[64] E. W. N. Glover and David A. Kosower. "Recombination methods for jets in $p\bar{p}$ collisions". In: *Phys. Lett. B* 367.1-4 (1996), pp. 369–376. DOI: 10.1016/0370-2693(95)01496-9. arXiv: hep-ph/9510420.

[65] Yu. L. Dokshitzer et al. "Better jet clustering algorithms". In: *JHEP* 08 (1997), p. 001. DOI: 10.1088/1126-6708/1997/08/001. arXiv: hep-ph/9707323.

[66] Matteo Cacciari, Gavin P. Salam, and Grégory Soyez. "The anti-k_{T} jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[67] Matteo Cacciari and Gavin P. Salam. "Dispelling the N^3 myth for the k_t jet-finder". In: *Phys. Lett. B* 641 (2006), pp. 57–61. DOI: 10.1016/J.PhysLetB.2006.08.037. arXiv: hep-ph/0512210.

[68] Matteo Cacciari, Gavin P. Salam, and Grégory Soyez. "FastJet user manual". In: *Eur. Phys. J. C* 72.1896 (2012). DOI: 10.1140/EPJC/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].

[69] Carola F. Berger, Tibor Kúcs, and George Sterman. "Event shape/energy flow correlations". In: *Phys. Rev. D* 68.1 (2003), p. 014012. DOI: 10.1103/PhysRevD.68.014012. arXiv: hep-ph/0303051.

[70] Stephen D. Ellis, Zoltan Kunszt, and Davison E. Soper. "Jets at hadron colliders at order α_s^3: A look inside". In: *Phys. Rev. Lett.* 69.25 (1992), pp. 3615–3618. DOI: 10.1103/PhysRevLett.69.3615.

[71] Jonathan M. Butterworth et al. "Jet Substructure as a New Higgs-Search Channel at the Large Hadron Collider". In: *Phys. Rev. Lett.* 100.24 (2008), p. 242001. DOI: 10.1103/PhysRevLett.100.242001. arXiv: 0802.2470 [hep-ph].

[72] Edward Farhi. "Quantum Chromodynamics Test for Jets". In: *Phys. Rev. Lett.* 39.25 (1977), pp. 1587–1588. DOI: 10.1103/PhysRevLett.39.1587.

[73] **ALICE** Collaboration: S. Acharya et al. "First measurement of jet mass in Pb–Pb and p–Pb collisions at the LHC". In: *Phys. Lett. B* 776 (2018), pp. 249–264. DOI: 10.1016/J.PhysLetB.2017.11.044. arXiv: 1702.00804 [nucl-ex].

[74] **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". In: *JHEP* 11 (2018), p. 113. DOI: 10.1007/JHEP11(2018)113. arXiv: 1807.05974 [hep-ex].

[75] **ATLAS** Collaboration: G. Aad et al. "A measurement of soft-drop jet observables in pp collisions with the ATLAS detector at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. D* 101.5 (2020), p. 052007. DOI: 10.1103/PhysRevD.101.052007. arXiv: 1912.09837 [hep-ex].

[76] **ATLAS** Collaboration: G. Aad et al. "Measurement of the jet mass in high transverse momentum $Z(\to b\bar{b})\gamma$ production at $\sqrt{s} = 13$ TeV using the ATLAS detector". In: *Phys. Lett. B* 812 (2021), p. 135991. DOI: 10.1016/J.PhysLetB.2020.135991. arXiv: 1907.07093 [hep-ex].

[77] **CMS** Collaboration: A. M. Sirunyan et al. "Measurement of the jet mass distribution and top quark mass in hadronic decays of boosted top quarks in pp collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. Lett.* 124.20 (2020), p. 202001. DOI: 10.1103/PhysRevLett.124.202001. arXiv: 1911.03800 [hep-ex].

[78] **CMS** Collaboration. "Measurement of the jet mass distribution and top quark mass in hadronic decays of boosted top quarks in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: (2022). DOI: 10.48550/arXiv.2211.01456. arXiv: 2211.01456 [hep-ex].

[79] **ATLAS** Collaboration. "Measurement of $R = 0.4$ jet mass in Pb+Pb and pp collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV with the ATLAS detector". ATLAS-CONF-2018-014. CERN, Geneva, CH, 2018. URL: https://cds.cern.ch/record/2319867.

[80] **CDF** Collaboration: T. Aaltonen et al. "Study of Substructure of High Transverse Momentum Jets Produced in Proton-Antiproton Collisions at $\sqrt{s} = 1.96$ TeV". In: *Phys. Rev. D* 85 (2012), p. 091101. DOI: 10.1103/PhysRevD.85.091101. arXiv: 1106.5952 [hep-ex].

[81] **ATLAS** Collaboration: G. Aad et al. "Jet mass and substructure of inclusive jets in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS experiment". In: *JHEP* 05 (2012), p. 128. DOI: 10.1007/JHEP05(2012)128. arXiv: 1203.4606 [hep-ex].

[82] **ATLAS** Collaboration: G. Aad et al. "ATLAS measurements of the properties of jets for boosted particle searches". In: *Phys. Rev. D* 86 (2012), p. 072006. ISSN: 1550-2368. DOI: 10.1103/PhysRevD.86.072006. arXiv: 1206.5369 [hep-ex].

[83] **ATLAS** Collaboration: M. Aaboud et al. "Properties of $g \to b\bar{b}$ at small opening angles in pp collisions with the ATLAS detector at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. D* 99.5 (2019), p. 052004. DOI: 10.1103/PhysRevD.99.052004. arXiv: 1812.09283 [hep-ex].

[84] **CMS** Collaboration: S. Chatrchyan et al. "Studies of jet mass in dijet and W/Z + jet events". In: *JHEP* 05 (2013), p. 090. DOI: 10.1007/JHEP05(2013)090. arXiv: 1303.4811 [hep-ex].

[85] **CMS** Collaboration: A. M. Sirunyan et al. "Measurement of the jet mass in highly boosted $\mathrm{t\bar{t}}$ events from pp collisions at $\sqrt{s} = 8$ TeV". In: *Eur. Phys. J. C* 77 (2017), p. 467. DOI: 10.1140/EPJC/s10052-017-5030-3. arXiv: 1703.06330 [hep-ex].

[86] **ATLAS** Collaboration: M. Aaboud et al. "A measurement of the soft-drop jet mass in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. Lett.* 121.9 (2018), p. 092001. DOI: 10.1103/PhysRevLett.121.092001. arXiv: 1711.08341 [hep-ex].

[87] **CMS** Collaboration: A. M. Sirunyan et al. "Measurement of the groomed jet mass in PbPb and pp collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV". In: *JHEP* 10 (2018), p. 161. DOI: 10.1007/JHEP10(2018)161. arXiv: 1805.05145 [hep-ex].

[88] Mrinal Dasgupta et al. "Towards an understanding of jet substructure". In: *JHEP* 09 (2013), p. 029. DOI: 10.1007/JHEP09(2013)029. arXiv: 1307.0007 [hep-ph].

[89] Simone Marzani, Lais Schunk, and Grégory Soyez. "A study of jet mass distributions with grooming". In: *JHEP* 07 (2017), p. 132. DOI: 10.1007/JHEP07(2017)132. arXiv: 1704.02210 [hep-ph].

[90] Simone Marzani, Lais Schunk, and Grégory Soyez. "The jet mass distribution after Soft Drop". In: *Eur. Phys. J. C* 78.2 (2018), p. 96. DOI: 10.1140/EPJC/s10052-018-5579-5. arXiv: 1712.05105 [hep-ph].

[91] André H. Hoang et al. "Nonperturbative Corrections to Soft Drop Jet Mass". In: *JHEP* 12 (2019), p. 002. DOI: 10.1007/JHEP12(2019)002. arXiv: 1906.11843 [hep-ph].

[92] Aditya Pathak et al. "EFT for Soft Drop Double Differential Cross Section". In: (2020). arXiv: 2012.15568 [hep-ph].

[93] Adam Kardos, Andrew J. Larkoski, and Zoltán Trócsányi. "Groomed jet mass at high precision". In: *Phys. Lett. B* 809 (2020), p. 135704. ISSN: 0370-2693. DOI: 10.1016/J.PhysLetB.2020.135704. arXiv: 2002.00942 [hep-ph].

[94] Carola F. Berger, Tibor Kúcs, and George F. Sterman. "Interjet Energy Flow/Event Shape Correlations". In: *Int. J. Mod. Phys. A* 18.23 (2003), pp. 4159–4168. DOI: 10.1142/s0217751x03016161. arXiv: hep-ph/0212343.

[95] Carola F. Berger and Lorenzo Magnea. "Scaling of power corrections for angularities from dressed gluon exponentiation". In: *Phys. Rev. D* 70.9 (2004), p. 094010. DOI: 10.1103/PhysRevD.70.094010. arXiv: hep-ph/0407024.

[96] Leandro G. Almeida et al. "Substructure of high-p_{T} jets at the LHC". In: *Phys. Rev. D* 79.7 (2009), p. 074017. DOI: 10.1103/PhysRevD.79.074017. arXiv: 0807.0234 [hep-ph].

[97] Andrew J. Larkoski, Jesse Thaler, and Wouter J. Waalewijn. "Gaining (mutual) information about quark/gluon discrimination". In: *JHEP* 11 (2014), p. 129. ISSN: 1029-8479. DOI: 10.1007/JHEP11(2014)129. arXiv: 1408.3122 [hep-ph].

[98] S. Catani, G. Turnock, and B.R. Webber. "Jet broadening measures in e^+e^- annihilation". In: *Phys. Lett. B* 295.3 (1992), pp. 269–276. ISSN: 0370-2693. DOI: 10.1016/0370-2693(92)91565-Q.

[99] Zhong-Bo Kang, Kyle Lee, and Felix Ringer. "Jet angularity measurements for single inclusive jet production". In: *JHEP* 04 (2018), p. 110. ISSN: 1029-8479. DOI: 10.1007/JHEP04(2018)110. arXiv: 1801.00790 [hep-ph].

[100] Simone Marzani, Grégory Soyez, and Michael Spannowsky. *Looking Inside Jets: An Introduction to Jet Substructure and Boosted-object Phenomenology*. Lecture Notes in Physics. Springer International Publishing, 2019. DOI: 10.1007/978-3-030-15709-8. arXiv: 1901.10342 [hep-ph].

[101] Roman Kogler et al. "Jet substructure at the Large Hadron Collider: Experimental Review". In: *Rev. Mod. Phys.* 91.4 (2019), p. 45003. DOI: 10.1103/RevModPhys.91.045003. arXiv: 1803.06991 [hep-ex].

[102] **ALICE** Collaboration. *The ALICE experiment – A journey through QCD*. 2022. DOI: 10.48550/arXiv.2211.04384. arXiv: 2211.04384 [nucl-ex].

[103] **ALICE** Collaboration: S. Acharya et al. "Measurement of the groomed jet radius and momentum splitting fraction in pp and Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV". In: *Phys. Rev. Lett.* 128.10 (2022), p. 102001. DOI: 10.1103/PhysRevLett.128.102001. arXiv: 2107.12984 [nucl-ex].

[104] Andrew J. Larkoski et al. "Soft Drop". In: *JHEP* 05 (2014), p. 146. DOI: 10.1007/JHEP05(2014)146. arXiv: 1402.2657 [hep-ph].

[105] David Krohn, Jesse Thaler, and Lian-Tao Wang. "Jet trimming". In: *JHEP* 02 (2010), p. 084. DOI: 10.1007/JHEP02(2010)084. arXiv: 0912.1342 [hep-ph].

[106] Stephen D. Ellis, Christopher K. Vermilion, and Jonathan R. Walsh. "Techniques for improved heavy particle searches with jet substructure". In: *Phys. Rev. D* 80.5 (2009), p. 051501. DOI: 10.1103/PhysRevD.80.051501. arXiv: 0903.5081 [hep-ph].

[107] Stephen D. Ellis, Christopher K. Vermilion, and Jonathan R. Walsh. "Recombination algorithms and jet substructure: Pruning as a tool for heavy particle searches". In: *Phys. Rev. D* 81.9 (2010), p. 094023. DOI: 10.1103/PhysRevD.81.094023. arXiv: 0912.0033 [hep-ph].

[108] V. V. Sudakov. "Vertex Parts at Very High Energies in Quantum Electrodynamics". In: *Sov. Phys. JETP* 3.1 (1956), pp. 65–71. URL: http://jetp.ras.ru/cgi-bin/dn/e_003_01_0065.pdf.

[109] M. Dasgupta and G. P. Salam. "Resummation of non-global QCD observables". In: *Phys. Lett. B* 512 (2001), pp. 323–330. DOI: 10.1016/s0370-2693(01)00725-0. arXiv: hep-ph/0104277.

[110] Zhong-Bo Kang et al. "The soft drop groomed jet radius at NLL". In: *JHEP* 02 (2020), p. 054. DOI: 10.1007/JHEP02(2020)054. arXiv: 1908.01783 [hep-ph].

[111] Aashish Tripathee et al. "Jet substructure studies with CMS open data". In: *Phys. Rev. D* 96.7 (2017), p. 074003. DOI: 10.1103/PhysRevD.96.074003. arXiv: 1704.05842 [hep-ph].

[112] **CMS** Collaboration: A. M. Sirunyan et al. "Measurement of jet substructure observables in $t\bar{t}$ events from proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. D* 98.9 (2018), p. 092014. DOI: 10.1103/PhysRevD.98.092014. arXiv: 1808.07340 [hep-ex].

[113] **ALICE** Collaboration: S. Acharya et al. "Exploration of jet substructure using iterative declustering in pp and Pb–Pb collisions at LHC energies". In: *Phys. Lett. B* 802 (2020), p. 135227. DOI: 10.1016/J.PhysLetB.2020.135227. arXiv: 1905.02512 [hep-ph].

[114] **STAR** Collabroation: J. Adam et al. "Measurement of Groomed Jet Substructure Observables in pp Collisions at $\sqrt{s} = 200$ GeV with STAR". In: *Phys. Lett. B* 811 (2020), p. 135846. DOI: 10.1016/J.PhysLetB.2020.135846. arXiv: 2003.02114 [hep-ex].

[115] Yacine Mehtar-Tani, Alba Soto-Ontoso, and Konrad Tywoniuk. "Dynamical grooming of QCD jets". In: *Phys. Rev. D* 101.3 (2020), p. 034004. DOI: 10.1103/physrevd.101.034004. arXiv: 1911.00375 [hep-ph].

[116] Yacine Mehtar-Tani, Alba Soto-Ontoso, and Konrad Tywoniuk. "Tagging boosted hadronic objects with dynamical grooming". In: *Phys. Rev. D* 102.11 (2020), p. 114013. DOI: 10.1103/PhysRevD.102.114013. arXiv: 2005.07584 [hep-ph].

[117] Paul Caucal, Alba Soto-Ontoso, and Adam Takacs. "Dynamical Grooming meets LHC data". In: *JHEP* 07 (2021), p. 020. DOI: 10.1007/JHEP07(2021)020. arXiv: 2103.06566 [hep-ph].

[118] Paul Caucal, Alba Soto-Ontoso, and Adam Takacs. "Dynamically groomed jet radius in heavy-ion collisions". In: *Phys. Rev. D* 105.11 (2022), p. 114046. DOI: 10.1103/PhysRevD.105.114046. arXiv: 2111.14768 [hep-ph].

[119] **ALICE** Collaboration: S. Acharya et al. "Direct observation of the dead-cone effect in quantum chromodynamics". In: *Nature* 605.7910 (2022), pp. 440–446. DOI: 10.1038/s41586-022-04572-w. arXiv: 2106.05713 [hep-ph].

[120] G. Marchesini and B. R. Webber. "Simulation of QCD coherence in heavy quark production and decay". In: *Nucl. Phys. B* 330.1 (1990), pp. 261–283. ISSN: 0550-3213. DOI: 10.1016/0550-3213(90)90310-A.

[121] Yu. L. Dokshitzer, V. A. Khoze, and S. I. Troyan. "On specific QCD properties of heavy quark fragmentation ('dead cone')". In: *J. Phys. G: Nucl. Part. Phys.* 17.10 (1991), p. 1602. DOI: 10.1088/0954-3899/17/10/023.

[122] Yu. L. Dokshitzer and D. E. Kharzeev. "Heavy-quark colorimetry of QCD matter". In: *Phys. Lett. B* 519.3 (2001), pp. 199–206. ISSN: 0370-2693. DOI: 10.1016/S0370-2693(01)01130-3.

[123] **ATLAS** Collaboration: G. Aad et al. "Measurement of jet shapes in top pair events at $\sqrt{s} = 7$ TeV using the ATLAS detector". In: *Eur. Phys. J. C* 73.12 (2013). DOI: 10.1140/EPJC/s10052-013-2676-3. arXiv: 1307.5749 [hep-ph].

[124] Javier Llorente and Josu Cantero. "Determination of the b-quark mass m_b from the angular screening effects in the ATLAS b-jet shape data". In: *Nucl. Phys. B* 889 (2014), pp. 401–418. DOI: 10.1016/J.NuclPhysB.2014.10.021. arXiv: 1407.8001 [hep-ph].

[125] **ALICE** Collaboration: S. Acharya et al. "Measurement of prompt D^0, Λ_c^+, and $\Sigma_c^{0,++}(2455)$ production in pp collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. Lett.* 128.1 (2022), p. 012001. DOI: 10.1103/PhysRevLett.128.012001. arXiv: 2106.08278 [hep-ex].

[126] Roderick V. Reid. "Local phenomenological nucleon-nucleon potentials". In: *Annals of Physics* 50.3 (1968), pp. 411–448. ISSN: 0003-4916. DOI: 10.1016/0003-4916(68)90126-7.

[127] Michele Arneodo. "Nuclear effects in structure functions". In: *Phys. Rep.* 240.5 (1994), pp. 301–393. ISSN: 0370-1573. DOI: 10.1016/0370-1573(94)90048-5.

[128] **European Muon Collaboration (EMC)**: J.J. Aubert et al. "The ratio of the nucleon structure functions F_2^N for iron and deuterium". In: *Phys. Lett. B* 123.3 (1983), pp. 275–278. ISSN: 0370-2693. DOI: 10.1016/0370-2693(83)90437-9.

[129] Or Hen et al. "Nucleon-nucleon correlations, short-lived excitations, and the quarks within". In: *Rev. Mod. Phys.* 89.4 (2017), p. 045002. DOI: 10.1103/RevModPhys.89.045002.

[130] P. R. Norton. "The EMC effect". In: *Rep. Prog. Phys.* 66.8 (2003), p. 1253. DOI: 10.1088/0034-4885/66/8/201.

[131] J. Gomez et al. "Measurement of the A dependence of deep-inelastic electron scattering". In: *Phys. Rev. D* 49.9 (1994), pp. 4348–4372. DOI: 10.1103/PhysRevD.49.4348.

[132] Néstor Armesto. "Nuclear shadowing". In: *J. Phys. G* 32.11 (2006), R367–R393. DOI: 10.1088/0954-3899/32/11/r01. arXiv: hep-ph/0604108.

[133] Wei Zhu, Jianhong Ruan, and Fengyao Hou. "From EMC- and Cronin-effects to signals of quark-gluon plasma". In: *Int. J. Mod. Phys. E* 22.03 (2013), p. 1350013. DOI: 10.1142/s0218301313500134. arXiv: 1012.4224 [hep-ph].

[134] J. W. Cronin et al. "Production of hadrons at large transverse momentum at 200, 300, and 400 GeV". In: *Phys. Rev. D* 11 (1975), pp. 3105–3123. DOI: 10.1103/PhysRevD.11.3105.

[135] B. Z. Kopeliovich et al. "Cronin Effect in Hadron Production Off Nuclei". In: *Phys. Rev. Lett.* 88.23 (2002), p. 232303. DOI: 10.1103/PhysRevLett.88.232303. arXiv: hep-ph/0201010.

[136] **PHENIX** Collaboration: S. S. Adler et al. "Jet Structure from Dihadron Correlations in d+Au collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV". In: *Phys. Rev. C* 73.5 (2006), p. 054903. DOI: 10.1103/PhysRevC.73.054903. arXiv: nucl-ex/0510021.

[137] **ALICE** Collaboration: S. Acharya et al. "Jet fragmentation transverse momentum measurements from di-hadron correlations in $\sqrt{s} = 7$ TeV pp and $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV p–Pb collisions". In: *JHEP* 03 (2019), p. 169. DOI: 10.1007/JHEP03(2019)169. arXiv: 1811.09742 [nucl-ex].

[138] **ALICE** Collaboration: J. Adam et al. "Measurement of dijet k_{T} in p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV". In: *Phys. Lett. B* 746 (2015), pp. 385–395. DOI: 10.1016/J.PhysLetB.2015.05.033. arXiv: 1503.03050 [nucl-ex].

[139] **PHENIX** Collaboration: A. Adare et al. "Suppression of back-to-back hadron pairs at forward rapidity in d+Au Collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV". In: *Phys. Rev. Lett.* 107.17 (2011), p. 172301. DOI: 10.1103/PhysRevLett.107.172301. arXiv: 1105.5112 [nucl-ex].

[140] **PHENIX** Collaboration: C. Aidala et al. "Nonperturbative transverse momentum broadening in dihadron angular correlations in $\sqrt{s_{\mathrm{NN}}} = 200$ GeV proton-nucleus collisions". In: *Phys. Rev. C* 99.4 (2019), p. 044912. DOI: 10.1103/PhysRevC.99.044912. arXiv: 1809.09045 [hep-ex].

[141] **ATLAS** Collaboration: M. Aaboud et al. "Measurement of jet fragmentation in 5.02 TeV proton–lead and proton–proton collisions with the ATLAS detector". In: *Nucl. Phys. A* 978 (2018), pp. 65–106. DOI: 10.1016/J.NuclPhysA.2018.07.006. arXiv: 1706.02859 [hep-ex].

[142] Edward V. Shuryak. "Quantum chromodynamics and the theory of superdense matter". In: *Phys. Rep.* 61.2 (1980), pp. 71–158. ISSN: 0370-1573. DOI: 10.1016/0370-1573(80)90105-2.

[143] J. D. Bjorken. "Highly relativistic nucleus-nucleus collisions: The central rapidity region". In: *Phys. Rev. D* 27.1 (1983), pp. 140–151. DOI: 10.1103/PhysRevD.27.140.

[144] John W. Harris and Berndt Müller. "The search for the quark-gluon plasma". In: *Annu. Rev. Nucl. Part. Sci.* 46.1 (1996), pp. 71–107. DOI: 10.1146/AnnuRev.Nucl.46.1.71.

[145] Patrick Steinbrecher. "The QCD crossover at zero and non-zero baryon densities from Lattice QCD". In: *Nucl. Phys. A* 982 (2019), pp. 847–850. DOI: 10.1016/J.NuclPhysA.2018.08.025. arXiv: 1807.05607 [hep-lat].

[146] Ulrich Heinz and Maurice Jacob. *Evidence for a New State of Matter: An Assessment of the Results from the CERN Lead Beam Programme*. 2000. arXiv: nucl-th/0002042.

[147] Eemeli Annala et al. "Evidence for quark-matter cores in massive neutron stars". In: *Nature Phys.* 16.9 (2020), pp. 907–910. DOI: 10.1038/s41567-020-0914-9. arXiv: 1903.09121 [astro-ph.HE].

[148] Wit Busza, Krishna Rajagopal, and Wilke van der Schee. "Heavy Ion Collisions: The Big Picture and the Big Questions". In: *Annu. Rev. Nucl. Part. Sci.* 68.1 (2018), pp. 339–376. DOI: 10.1146/AnnuRev-Nucl-101917-020852. arXiv: 1802.04801 [hep-ph].

[149] Rustamov, Anar. "Energy dependence of strangeness production and event-by-event fluctuations". In: *EPJ Web Conf.* 171 (2018), p. 02003. DOI: 10.1051/EPJConf/201817102003.

[150] **WA97** Collaboration: E. Andersen et al. "Enhancement of central Λ, Ξ and Ω yields in Pb–Pb collisions at 158 A GeV/c". In: *Phys. Lett. B* 433.1 (1998), pp. 209–216. ISSN: 0370-2693. DOI: 10.1016/S0370-2693(98)00689-3.

[151] **NA49** Collaboration: H. Appelshäuser et al. "Ξ and $\overline{\Xi}$ production in 158 GeV/nucleon Pb+Pb collisions". In: *Phys. Lett. B* 444.3 (1998), pp. 523–530. ISSN: 0370-2693. DOI: 10.1016/S0370-2693(98)01399-9.

[152] **WA97** Collaboration: F. Antinori et al. "Production of strange and multistrange hadrons in nucleus–nucleus collisions at the SPS". In: *Nucl. Phys. A* 661.1 (1999), pp. 130–139. ISSN: 0375-9474. DOI: 10.1016/S0375-9474(99)85015-5.

[153] **WA97** Collaboration: F. Antinori et al. "Enhancement of strange and multi-strange baryons in central Pb–Pb interactions at 158 GeV/c per nucleon". In: *Nucl. Phys. A* 663-664 (2000), pp. 717c–720c. ISSN: 0375-9474. DOI: 10.1016/S0375-9474(99)00700-9.

[154] **NA49** Collaboration: H. Appelshäuser et al. "Hadronic expansion dynamics in central Pb+Pb collisions at 158 GeV per nucleon". In: *Eur. Phys. J. C* 2.4 (1998), pp. 661–670. DOI: 10.1007/s100529800866.

[155] Ulrich Heinz and Barbara V. Jacak. "Two-Particle Correlations in Relativisitic Heavy-Ion Collisions". In: *Annu. Rev. Nucl. Part. Sci.* 49.1 (1999), pp. 529–579. DOI: 10.1146/AnnuRev.Nucl.49.1.529. arXiv: nucl-th/9902020.

[156] R. Hanbury Brown and R. Q. Twiss. "A Test of a New Type of Stellar Interferometer on Sirius". In: *Nature* 178 (1956), pp. 1046–1048. DOI: 10.1038/1781046a0.

[157] Szabolcs Borsányi et al. "The QCD equation of state with dynamical quarks". In: *JHEP* 11 (2010), p. 077. DOI: 10.1007/JHEP11(2010)077. arXiv: 1007.2580 [hep-lat].

[158] **HotQCD** Collaboration: A. Bazavov et al. "Equation of state in (2+1)-flavor QCD". In: *Phys. Rev. D* 90 (9 2014), p. 094503. DOI: 10.1103/PhysRevD.90.094503.

[159] Peter Braun-Munzinger, Krzysztof Redlich, and Johanna Stachel. "Particle Production in Heavy Ion Collisions". In: *Quark–Gluon Plasma 3*. World Scientific Pub Co Pte Lt, 2004, pp. 491–599. DOI: 10.1142/9789812795533_0008. arXiv: nucl-th/0304013.

[160] Sebastian N. White. "Inelastic Diffraction at Heavy Ion Colliders". In: *Nucl. Phys. Proc. Suppl.* 146 (2005), pp. 48–52. DOI: 10.1016/J.NuclPhysBPS.2005.02.059. arXiv: nucl-ex/0501004.

[161] E. Gourgoulhon. "An introduction to relativistic hydrodynamics". In: *EAS Publications Series* 21 (2006), pp. 43–79. DOI: 10.1051/eas:2006106. arXiv: gr-qc/0603009.

[162] Norbert Herrmann, Johannes P. Wessels, and Thomas Wienold. "Collective Flow in Heavy-Ion Collisions". In: *Annu. Rev. Nucl. Part. Sci.* 49.1 (1999), pp. 581–632. DOI: 10.1146/AnnuRev.Nucl.49.1.581.

[163] **STAR** Collaboration: K. H. Ackermann et al. "Elliptic flow in Au+Au collisions at $\sqrt{s_{\mathrm{NN}}} = 130$ GeV". In: *Phys. Rev. Lett.* 86.3 (2001), pp. 402–407. DOI: 10.1103/PhysRevLett.86.402. arXiv: nucl-ex/0009011.

[164] **ALICE** Collaboration: S. Acharya et al. "Centrality determination in heavy ion collisions". In: (2018).

[165] **ALICE** Collaboration: B. Abelev et al. "Centrality determination of Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV with ALICE". In: *Phys. Rev. C* 88.4 (2013), p. 044909. DOI: 10.1103/PhysRevC.88.044909.

[166] Michael L. Miller et al. "Glauber Modeling in High-Energy Nuclear Collisions". In: *Annu. Rev. Nucl. Part. Sci.* 57.1 (2007), pp. 205–243. DOI: 10.1146/AnnuRev.Nucl.57.090506.123020. arXiv: nucl-ex/0701025.

[167] **ALICE** Collaboration: B. Abelev et al. "Elliptic flow of identified hadrons in Pb-Pb collisions at $\sqrt{s_{\mathrm{NN}}}$ = 2.76 TeV". In: *JHEP* 06 (2015), p. 190. DOI: 10.1007/JHEP06(2015)190. arXiv: 1405.4632 [nucl-ex].

[168] G. Policastro, D. T. Son, and A. O. Starinets. "Shear viscosity of strongly coupled N=4 supersymmetric Yang-Mills plasma". In: *Phys. Rev. Lett.* 87.8 (2001). DOI: 10.1103/PhysRevLett.87.081601. arXiv: hep-th/0104066.

[169] Jorge Casalderrey-Solana et al. *Gauge/String Duality, Hot QCD and Heavy Ion Collisions*. 2012. arXiv: 1101.0618 [hep-th].

[170] Veronika E. Hubeny, Shiraz Minwalla, and Mukund Rangamani. *The fluid/gravity correspondence*. 2011. arXiv: 1107.5780 [hep-th].

[171] C. Aidala et al. *A new era of discover: The 2023 long range plan for nuclear science*. Version 1.1. 2023. URL: https://science.osti.gov/-/media/np/nsac/pdf/202310/October-4-LRP-Report.pdf.

[172] M. A. Stephanov. "QCD phase diagram and the critical point". In: *Int. J. Mod. Phys. A* 20.19 (2005), pp. 4387–4392. DOI: 10.1142/s0217751x05027965. arXiv: hep-ph/0402115.

[173] Philippe de Forcrand. *Simulating QCD at finite density*. 2010. arXiv: 1005.0539 [hep-lat].

[174] Mark G. Alford et al. "Color superconductivity in dense quark matter". In: *Rev. Mod. Phys.* 80.4 (2008), pp. 1455–1515. DOI: 10.1103/RevModPhys.80.1455. arXiv: 0709.4635 [hep-ph].

[175] Krishna Rajagopal. "Mapping the QCD phase diagram". In: *Nucl. Phys. A* 661 (1999), pp. 150–161. DOI: 10.1016/s0375-9474(99)85017-9. arXiv: hep-ph/9908360.

[176] Grazyna Odyniec. "The RHIC Beam Energy Scan program in STAR and what's next ..." In: *J. Phys.: Conf. Ser.* 455.1 (2013), p. 012037. DOI: 10.1088/1742-6596/455/1/012037.

[177] Jana N. Guenther. *Overview of the QCD phase diagram – Recent progress from the lattice*. 2021. arXiv: 2010.15503 [hep-lat].

[178] **PHENIX** Collaboration: K. Adcox et al. "Suppression of Hadrons with Large Transverse Momentum in Central Au+Au Collisions at $\sqrt{s_{\mathrm{NN}}}$ = 130 GeV". In: *Phys. Rev. Lett.* 88.2 (2001), p. 022301. DOI: 10.1103/PhysRevLett.88.022301. arXiv: nucl-ex/0109003.

[179] **STAR** Collaboration: C. Adler et al. "Centrality Dependence of High-p_{T} Hadron Suppression in Au+Au Collisions at $\sqrt{s_{\mathrm{NN}}}$ = 130 GeV". In: *Phys. Rev. Lett.* 89.20 (2002), p. 202301. DOI: 10.1103/PhysRevLett.89.202301. arXiv: nucl-ex/0206011.

[180] **ATLAS** Collaboration: M. Aaboud et al. "Measurement of the nuclear modification factor for inclusive jets in Pb+Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV with the ATLAS detector". In: *Phys. Lett. B* 790 (2019), pp. 108–128. ISSN: 0370-2693. DOI: 10.1016/J.PhysLetB.2018.10.076. arXiv: 1805.05635 [nucl-ex].

[181] John W. Harris. "Introduction to Hard Scattering Processes and Recent Results from Hard Probes at RHIC and LHC". In: *J. Phys.: Conf. Ser.* 630.1 (2015), p. 012052. DOI: 10.1088/1742-6596/630/1/012052.

[182] **ATLAS** Collaboration: G. Aad et al. "Observation of a Centrality-Dependent Dijet Asymmetry in Lead-Lead Collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV with the ATLAS Detector at the LHC". In: *Phys. Rev. Lett.* 105.25 (2010), p. 252303. DOI: 10.1103/PhysRevLett.105.252303. arXiv: 1011.6182 [hep-ex].

[183] **CMS** Collaboration: S. Chatrchyan et al. "Observation and studies of jet quenching in PbPb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV". In: *Phys. Rev. C* 84.2 (2011), p. 024906. DOI: 10.1103/PhysRevC.84.024906. arXiv: 1102.1957 [nucl-ex].

[184] Yacine Mehtar-Tani. "Gluon bremsstrahlung in finite media beyond multiple soft scattering approximation". In: *JHEP* 07 (2019), p. 057. DOI: 10.1007/JHEP07(2019)057. arXiv: 1903.00506 [hep-ph].

[185] Sangyong Jeon and Guy D. Moore. "Energy loss of leading partons in a thermal QCD medium". In: *Phys. Rev. C* 71.3 (2005), p. 034901. DOI: 10.1103/PhysRevC.71.034901. arXiv: hep-ph/0309332.

[186] Jean-Paul Blaizot et al. "Probabilistic picture for medium-induced jet evolution". In: *JHEP* 06 (2014), p. 075. DOI: 10.1007/JHEP06(2014)075. arXiv: 1311.5823 [hep-ph].

[187] J.-P. Blaizot, E. Iancu, and Y. Mehtar-Tani. "Medium-Induced QCD Cascade: Democratic Branching and Wave Turbulence". In: *Phys. Rev. Lett.* 111.5 (2013), p. 052001. DOI: 10.1103/PhysRevLett.111.052001. arXiv: 1301.6102 [hep-ph].

[188] M. Gyulassy, P. Levai, and I. Vitev. "Reaction operator approach to non-abelian energy loss". In: *Nucl. Phys. B* 594 (2001), pp. 371–419. DOI: 10.1016/s0550-3213(00)00652-0. arXiv: nucl-th/0006010.

[189] Urs Achim Wiedemann. "Gluon radiation off hard quarks in a nuclear environment: opacity expansion". In: *Nucl. Phys. B* 588 (2000), pp. 303–344. DOI: 10.1016/s0550-3213(00)00457-0. arXiv: hep-ph/0005129.

[190] Xin-Nian Wang and Xiaofeng Guo. "Multiple parton scattering in nuclei: parton energy loss". In: *Nucl. Phys. A* 696 (2001), pp. 788–832. DOI: 10.1016/s0375-9474(01)01130-7. arXiv: hep-ph/0102230.

[191] Yuan-Yuan Zhang, Guang-You Qin, and Xin-Nian Wang. "Parton Energy Loss in the Generalized High-Twist Approach". In: *Nucl. Phys. A* 1005 (2021). The 28th International Conference on Ultra-relativistic Nucleus-Nucleus Collisions: Quark Matter 2019, p. 122003. ISSN: 0375-9474. DOI: 10.1016/J.NuclPhysA.2020.122003.

[192] Korinna Zapp. "JEWEL 2.0.0: Directions for Use". In: *Eur. Phys. J. C* 74.2 (2014), p. 2762. ISSN: 1434-6052. DOI: 10.1140/EPJC/s10052-014-2762-1. arXiv: 1311.0048 [hep-ph].

[193] Andy Buckley et al. "LHAPDF6: parton density access in the LHC precision era". In: *Eur. Phys. J. C* 75.3 (2015). DOI: 10.1140/EPJC/s10052-015-3318-8. arXiv: 1412.7420 [hep-ph].

[194] Raghav K. Elayavalli and Korinna C. Zapp. "Medium Recoils and background subtraction in JEWEL". In: (2016). arXiv: 1612.05116 [nucl-th].

[195] Jorge Casalderrey-Solana et al. "A hybrid strong/weak coupling approach to jet quenching". In: *JHEP* 10 (2014), p. 019. DOI: 10.1007/JHEP10(2014)019. arXiv: 1405.3864 [nucl-th].

[196] Eric D'Hoker and Daniel Z. Freedman. *Supersymmetric Gauge Theories and the AdS/CFT Correspondence*. 2002. arXiv: hep-th/0201253.

[197] Francesco D'Eramo, Krishna Rajagopal, and Yi Yin. "Molière scattering in quark-gluon plasma: finding point-like scatterers in a liquid". In: *JHEP* 01 (2019), p. 172. DOI: 10.1007/JHEP01(2019)172. arXiv: 1808.03250 [hep-ph].

[198] Jorge Casalderrey-Solana et al. "Jet wake from linearized hydrodynamics". In: *JHEP* 05 (2021), p. 230. DOI: 10.1007/JHEP05(2021)230. arXiv: 2010.01140 [hep-ph].

[199] J. H. Putschke et al. "The JETSCAPE framework". In: (2019). arXiv: 1903.07706 [nucl-th].

[200] Abhijit Majumder. "Incorporating space-time within medium-modified jet-event generators". In: *Phys. Rev. C* 88.1 (2013), p. 014909. DOI: 10.1103/PhysRevC.88.014909. arXiv: 1301.5323 [nucl-th].

[201] Yayun He et al. "Linear Boltzmann transport for jet propagation in the quark-gluon plasma: Elastic processes and medium recoil". In: *Phys. Rev. C* 91.5 (2015), p. 054908. DOI: 10.1103/PhysRevC.91.054908. arXiv: 1503.03313 [nucl-th].

[202] Y. Hama, T. Kodama, and O. Socolowski Jr. "Topics on hydrodynamic model of nucleus-nucleus collisions". In: *Brazilian Journal of Physics* 35.1 (2005), pp. 24–51. ISSN: 0103-9733. DOI: 10.1590/S0103-97332005000100003. arXiv: hep-ph/0407264.

[203] Björn Schenke, Sangyong Jeon, and Charles Gale. "(3+1)D hydrodynamic simulation of relativistic heavy-ion collisions". In: *Phys. Rev. C* 82.1 (2010), p. 014903. DOI: 10.1103/PhysRevC.82.014903. arXiv: 1004.1408 [hep-ph].

[204] Björn Schenke, Sangyong Jeon, and Charles Gale. "Elliptic and Triangular Flow in Event-by-Event D = 3+1 Viscous Hydrodynamics". In: *Phys. Rev. Lett.* 106.4 (2011), p. 042301. DOI: 10.1103/PhysRevLett.106.042301. arXiv: 1009.3244 [hep-ph].

[205] Long-Gang Pang, Hannah Petersen, and Xin-Nian Wang. "Pseudorapidity distribution and decorrelation of anisotropic flow within the open-computing-language implementation CLVisc hydrodynamics". In: *Phys. Rev. C* 97.6 (2018), p. 064918. DOI: 10.1103/PhysRevC.97.064918. arXiv: 1802.04449 [nucl-th].

[206] Longgang Pang, Qun Wang, and Xin-Nian Wang. "Effects of initial flow velocity fluctuation in event-by-event (3+1)D hydrodynamics". In: *Phys. Rev. C* 86.2 (2012), p. 024911. DOI: 10.1103/PhysRevC.86.024911. arXiv: 1205.5019 [nucl-th].

[207] Derek Soeder et al. *Bayesian parameter estimation with a new three-dimensional initial-conditions model for ultrarelativistic heavy-ion collisions*. 2023.

[208] Dmytro Oliinychenko et al. *smash-transport/smash: SMASH-3.0*. DOI: 10.5281/zenodo.7870822.

[209] *JETSCAPE*. GitHub repository. URL: https://github.com/JETSCAPE/JETSCAPE.

[210] Andrew J. Larkoski, Ian Moult, and Benjamin Nachman. "Jet substructure at the Large Hadron Collider: A review of recent advances in theory and machine learning". In: *Phys. Rept.* 841 (2020), pp. 1–63. ISSN: 0370-1573. DOI: 10.1016/J.PhysRep.2019.11.001. arXiv: 1709.04464 [hep-ph].

[211] Zachary Hulcher, Daniel Pablos, and Krishna Rajagopal. "Resolution effects in the hybrid strong/weak coupling model". In: *JHEP* 03 (2018), p. 010. DOI: 10.1007/JHEP03(2018)010. arXiv: 1707.05245 [hep-ph].

[212] Yacine Mehtar-Tani, Soeren Schlichting, and Ismail Soudi. "Jet thermalization in QCD kinetic theory". In: *JHEP* 05 (2023), p. 091. DOI: 10.1007/JHEP05(2023)091. arXiv: 2209.10569 [hep-ph].

[213] Yang-Ting Chien and Ivan Vitev. "Probing the Hardest Branching within Jets in Heavy-Ion Collisions". In: *Phys. Rev. Lett.* 119.11 (2017), p. 112301. DOI: 10.1103/PhysRevLett.119.112301. arXiv: 1608.07283 [hep-ph].

[214] James Mulligan and Mateusz Płoskoń. "Identifying groomed jet splittings in heavy-ion collisions". In: *Phys. Rev. C* 102.4 (2020). DOI: 10.1103/PhysRevC.102.044913. arXiv: 2006.01812 [hep-ph].

[215] Yacine Mehtar-Tani and Konrad Tywoniuk. "Groomed jets in heavy-ion collisions: sensitivity to medium-induced bremsstrahlung". In: *JHEP* 04 (2017), p. 125. DOI: 10.1007/JHEP04(2017)125. arXiv: 1610.08930 [hep-ph].

[216] Raghav Kunnawalkam Elayavalli and Korinna Christine Zapp. "Medium response in JEWEL and its impact on jet shape observables in heavy ion collisions". In: *JHEP* 07 (2017), p. 141. DOI: 10.1007/JHEP07(2017)141. arXiv: 1707.01539 [hep-ph].

[217] Harry Arthur Andrews et al. "Novel tools and observables for jet physics in heavy-ion collisions". In: *J. Phys. G* 47 (2020), p. 065102. DOI: 10.1088/1361-6471/ab7cbc. arXiv: 1808.03689 [hep-ph].

[218] Ning-Bo Chang, Shanshan Cao, and Guang-You Qin. "Probing medium-induced jet splitting and energy loss in heavy-ion collisions". In: *Phys. Lett. B* 781 (2018), pp. 423–432. DOI: 10.1016/J.PhysLetB.2018.04.019. arXiv: 1707.03767 [hep-ph].

[219] Guilherme Milhano, Urs Achim Wiedemann, and Korinna Christine Zapp. "Sensitivity of jet substructure to jet-induced medium response". In: *Phys. Lett. B* 779 (2018), pp. 409–413. DOI: 10.1016/J.PhysLetB.2018.01.029. arXiv: 1707.04142 [hep-ph].

[220] P. Caucal, E. Iancu, and G. Soyez. "Deciphering the z_g distribution in ultrarelativistic heavy ion collisions". In: *JHEP* 10 (2019), p. 273. DOI: 10.1007/JHEP10(2019)273. arXiv: 1907.04866 [hep-ph].

[221] J. Casalderrey-Solana et al. "Modification of jet substructure in heavy ion collisions as a probe of the resolution length of quark-gluon plasma". In: *JHEP* 01 (2020), p. 044. DOI: 10.1007/JHEP01(2020)044. arXiv: 1907.11248 [hep-ph].

[222] Felix Ringer, Bo-Wen Xiao, and Feng Yuan. "Can we observe jet P_T-broadening in heavy-ion collisions at the LHC?" In: *Phys. Lett. B* 808 (2020), p. 135634. DOI: 10.1016/J.PhysLetB.2020.135634. arXiv: 1907.12541 [hep-ph].

[223] **ALICE** Collaboration: S. Acharya et al. "Medium modification of the shape of small-radius jets in central Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV". In: *JHEP* 10 (2018), p. 139. ISSN: 1029-8479. DOI: 10.1007/JHEP10(2018)139. arXiv: 1807.06854 [nucl-ex].

[224] **CMS** Collaboration: A. M. Sirunyan et al. "Measurement of the splitting function in pp and PbPb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV". In: *Phys. Rev. Lett.* 120.14 (2018). DOI: 10.1103/PhysRevLett.120.142302. arXiv: 1708.09429 [nucl-ex].

[225] **STAR** Collabroation: M. S. Abdallah et al. "Differential measurements of jet substructure and partonic energy loss in Au + Au collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV". In: *Phys. Rev. C* 105.4 (2022), p. 044906. DOI: 10.1103/PhysRevC.105.044906. arXiv: 2109.09793 [nucl-ex].

[226] Yi Chen et al. "Jet energy spectrum and substructure in e^+e^- collisions at 91.2 GeV with ALEPH Archived Data". In: *JHEP* 06 (2022), p. 008. DOI: 10.1007/JHEP06(2022)008. arXiv: 2111.09914 [hep-ex].

[227] Andrew J. Larkoski, Ian Moult, and Duff Neill. "Toward Multi-Differential Cross Sections: Measuring Two Angularities on a Single Jet". In: *JHEP* 09 (2014), p. 046. DOI: 10.1007/JHEP09(2014)046. arXiv: 1401.4458 [hep-ph].

[228] Massimiliano Procura, Wouter J. Waalewijn, and Lisa Zeune. "Joint resummation of two angularities at next-to-next-to-leading logarithmic order". In: *JHEP* 10 (2018), p. 098. DOI: 10.1007/JHEP10(2018)098. arXiv: 1806.10622 [hep-ph].

[229] Iain W. Stewart, Frank J. Tackmann, and Wouter J. Waalewijn. "Dissecting Soft Radiation with Factorization". In: *Phys. Rev. Lett.* 114.9 (2015), p. 092001. DOI: 10.1103/PhysRevLett.114.092001. arXiv: 1405.6722 [hep-ph].

[230] Christopher Lee and George Sterman. "Momentum flow correlations from event shapes: Factorized soft gluons and soft-collinear effective theory". In: *Phys. Rev. D* 75.1 (2007). ISSN: 1550-2368. DOI: 10.1103/PhysRevD.75.014022. arXiv: hep-ph/0611061.

[231] Zhong-Bo Kang et al. "Soft drop groomed jet angularities at the LHC". In: *Phys. Lett. B* 793 (2019), pp. 41–47. ISSN: 0370-2693. DOI: 10.1016/J.PhysLetB.2019.4.018. arXiv: 1811.06983 [hep-ph].

[232] Simone Caletti et al. "Jet Angularities in Z+jet production at the LHC". In: (2021). arXiv: 2104.06920 [hep-ph].

[233] Leandro G. Almeida et al. "Comparing and counting logs in direct and effective methods of QCD resummation". In: *JHEP* 04 (2014), p. 174. DOI: 10.1007/JHEP04(2014)174. arXiv: 1401.4460 [hep-ph].

[234] Gregory P. Korchemsky and George Sterman. "Power corrections to event shapes and factorization". In: *Nucl. Phys. B* 555.1-2 (1999), pp. 335–351. ISSN: 0550-3213. DOI: 10.1016/s0550-3213(99)00308-9. arXiv: hep-ph/9902341.

[235] Elke-Caroline Aschenauer et al. "Jet angularities in photoproduction at the Electron-Ion Collider". In: *Phys. Rev. D* 101.5 (2020). ISSN: 2470-0029. DOI: 10.1103/PhysRevD.101.054028. arXiv: 1910.11460 [hep-ph].

[236] Hsi-Ming Chang et al. "Calculating Track-Based Observables for the LHC". In: *Phys. Rev. Lett.* 111 (2013), p. 102002. DOI: 10.1103/PhysRevLett.111.102002. arXiv: 1303.6637 [hep-ph].

[237] Néstor Armesto, Carlos Salgado, and Urs Wiedemann. "Medium-induced gluon radiation off massive quarks fills the dead cone". In: *Phys. Rev. D* 69.11 (2004), p. 114003. DOI: 10.1103/PhysRevD.69.114003. arXiv: hep-ph/0312106.

[238] Preeti Dhankher. "First measurement of jet angularities with D^0-meson tagged jets with ALICE". Proceedings of Science: 11th International Conference on Hard and Electromagnetic Probes of High-Energy Nuclear Collisions. 2023.

[239] **LHC Study Group**: P. Lefèvre et al. *The Large Hadron Collider : conceptual design (LHC Yellow Book)*. CERN, 1995. URL: https://cds.cern.ch/record/291782/.

[240] Oliver Sim Brüning et al. *LHC Design Report*. CERN, 2004. ISBN: 9789290832249. DOI: 10.5170/CERN-2004-003-V-1.

[241] Lyndon Evans and Philip Bryant. "LHC Machine". In: *J. Inst.* 3.08 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[242] *Ernest Lawrence - Biographical*. The Nobel Prize in Physics 1939. URL: https://www.nobelprize.org/prizes/physics/1939/lawrence/biographical/.

[243] Vladimir Veksler. "A new method of accelerating relativistic particles". In: *Compt. Rend. Acad. Sci. URSS* 43.8 (1944), pp. 346–348. URL: lhe.jinr.ru/rus/veksler/wv0/publikacii/1944Veksler.pdf.

[244] Edwin M. McMillan. "The Synchrotorn - A Proposed High Energy Particle Accelerator". In: *Phys. Rev.* 68.5-6 (1945), pp. 143–144. DOI: 10.1103/PhysRev.68.143.

[245] Esma Mobs. *The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018*. General Photo. 2018. URL: https://cds.cern.ch/record/2636343.

[246] Gerhard John Krige. *History of CERN*. Amsterdam: North-Holland, 1996. ISBN: 0444896554. URL: https://cds.cern.ch/record/318964.

[247] The LEP Injector Study Group. *LEP design report*. Report. Geneva: CERN, 1983. URL: https://cds.cern.ch/record/98881.

[248] Luigi Di Lella and Carlo Rubbia. "The Discovery of the W and Z Particles". In: *Adv. Ser. Dir. High Energy Phys.* 23 (2015), pp. 137–163. DOI: 10.1142/9789814644150_0006. URL: https://cds.cern.ch/record/2103277.

[249] **ATLAS** Collaboration. *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. 2022. arXiv: 2212.09379 [hep-ex].

[250] **CMS** Collaboration. *CMS Luminosity – Public Results*. 2023. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[251] **ATLAS** Collaboration: M. Aaboud et al. "Jet reconstruction and performance using particle flow with the ATLAS Detector". In: *EPJC* 77.7 (2017). DOI: 10.1140/epjc/s10052-017-5031-2. arXiv: 1703.10485 [hep-ex].

[252] Daniele Bertolini et al. "Pileup per particle identification". In: *JHEP* 10 (2014), p. 059. DOI: 10.1007/JHEP10(2014)059. arXiv: 1407.6013 [hep-ph].

[253] Patrick T. Komiske et al. "Pileup Mitigation with Machine Learning (PUMML)". In: *JHEP* 12 (2017), p. 051. DOI: 10.1007/JHEP12(2017)051. arXiv: 1707.08600 [hep-ph].

[254] *ALICE: Technical proposal for a Large Ion collider Experiment at the CERN LHC*. LHC technical proposal. Geneva: CERN, 1995. ISBN: 92-9083-077-8. URL: https://cds.cern.ch/record/293391.

[255] Arturo Tauro. "ALICE Schematics". General Photo. 2017. URL: https://cds.cern.ch/record/2263642.

[256] **ALICE** Collaboration: K. Aamodt et al. "The ALICE experiment at the CERN LHC". In: *J. Inst.* 3.S08002 (2008). DOI: 10.1088/1748-0221/3/08/S08002.

[257] **ALICE** Collaboration: B. Abelev et al. "Performance of the ALICE experiment at the CERN LHC". In: *Int. J. Mod. Phys. A* 29.24 (2014), p. 1430044. DOI: 10.1142/s0217751x14300440. arXiv: 1402.4476 [nucl-ex].

[258] **ALICE** Collaboration: K. Aamodt et al. "Alignment of the ALICE Inner Tracking System with cosmic-ray tracks". In: *J. Inst.* 5.03 (2010), P03003. DOI: 10.1088/1748-0221/5/03/P03003. arXiv: 1001.0502 [physics.ins-det].

[259] J. Alme et al. "The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events". In: *NIMA* 622.1 (2010), pp. 316–367. ISSN: 0168-9002. DOI: 10.1016/J.NIMA.2010.04.042. arXiv: 1001.1950 [physics.ins-det].

[260] D. R. Nygren. "The Time Projection Chamber: A New 4 pi Detector for Charged Particles". In: *eConf* C740805 (1974). Ed. by J. Kadyk et al., p. 58. URL: https://lss.fnal.gov/conf/C740805/p58.pdf.

[261] Steffen Georg Weber and Anton Andronic. "ALICE event display of a Pb-Pb collision at 5.02A TeV". General Photo. 2016. URL: https://cds.cern.ch/record/2202730.

[262] R. Veenhof. "Choosing a gas mixture for the ALICE TPC". ALICE-INT-2003-29. 2003.

[263] **ALICE** Collaboration: S. Acharya et al. "ALICE 2016-2017-2018 luminosity determination for pp collisions at $\sqrt{s} = 13$ TeV". In: (2021). ALICE-PUBLIC-2021-005. URL: https://cds.cern.ch/record/2776672.

[264] **ALICE** Collaboration: S. Acharya et al. "ALICE 2017 luminosity determination for pp collisions at $\sqrt{s} = 5$ TeV". In: (2018). URL: https://cds.cern.ch/record/2648933.

[265] **ALICE** Collaboration: S. Acharya et al. "ALICE luminosity determination for Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV". In: (2021). ALICE-PUBLIC-2021-001. URL: https://cds.cern.ch/record/2749127.

[266] F. Reidt. "Upgrade of the ALICE ITS detector". In: *NIMA* 1032 (2022), p. 166632. DOI: 10.1016/J.NIMA.2022.166632. arXiv: 2111.08301 [physics.ins-det].

[267] **ALICE** Collaboration: S. Acharya et al. "The ALICE Transition Radiation Detector: Construction, operation, and performance". In: *NIMA* 881 (2018), pp. 88–127. ISSN: 0168-9002. DOI: 10.1016/J.NIMA.2017.09.028. arXiv: 1709.02743 [physics.ins-det].

[268] **ALICE TOF** Group: A. Akindinov et al. "Particle identification with the ALICE TOF detector at very high particle multiplicity". In: *Eur. Phys. J. CS* 32S1 (2003), pp. 165–177. DOI: 10.1140/EPJCd/s2003-01-0013-5. URL: https://cds.cern.ch/record/2633870.

[269] **ALICE** Collaboration: E. Abbas et al. "Performance of the ALICE VZERO system". In: *JINST* 8 (2013), P10016. DOI: 10.1088/1748-0221/8/10/P10016. arXiv: 1306.3130 [nucl-ex].

[270] B. Rapp. "Production de dimuons en réactions pp et Pb–Pb avec ALICE: le détecteur V0 et les résonances de basse masse". CERN-THESIS-2004-034. PhD thesis. Lyon University, Lyon, France, 2004. URL: https://cds.cern.ch/record/1065744.

[271] **ALICE** Collaboration. *Performance of the ALICE Electromagnetic Calorimeter*. 2022. arXiv: 2209.04216 [physics.ins-det].

[272] **ALICE** Collaboration: P. Cortese et al. *ALICE Electromagnetic Calorimeter Technical Design Report*. Tech. rep. 2008. URL: https://cds.cern.ch/record/1121574.

[273] R. Frühwirth. "Application of Kalman filtering to track and vertex fitting". In: *NIMA* 262.2 (1987), pp. 444–450. ISSN: 0168-9002. DOI: 10.1016/0168-9002(87)90887-4.

[274] Rongrong Ma. *Jet Analysis QA: Charged Tracks*. URL: https://twiki.cern.ch/twiki/bin/view/Main/QA:ChargedTracks#Hybrid_tracks.

[275] **ALICE** Collaboration: B. Abelev et al. "Suppression of high transverse momentum D mesons in central Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV". In: *JHEP* 09 (2012), p. 112. DOI: 10.1007/JHEP09(2012)112. arXiv: 1203.2160 [nucl-ex].

[276] **ALICE** Data Preparation Group. *Run 2 Data Sets*. URL: https://twiki.cern.ch/twiki/bin/viewauth/ALICE/AliDPGRun2DataSets.

[277] **ALICE** Collaboration: G. Dellacasa and others. *ALICE Zero-Degree Calorimeter (ZDC): Technical Design Report*. Technical design report. ALICE. Geneva: CERN, 1999. URL: https://cds.cern.ch/record/381433.

[278] **ALICE DPG**: Francesco Prino et al. "Tools for pileup tagging, removal, mitigation". 2016. URL: https://twiki.cern.ch/twiki/bin/view/ALICE/AliDPGtoolsPileup#Tools_for_pileup_tagging_removal.

[279] *PYTHIA Manual: Phase Space Cuts*. URL: https://www.pythia.org/latest-manual/PhaseSpaceCuts.html.

[280] Eliane Epple. "PWG-JE/EMCal presentation: full jets at 5.02 TeV pp". URL: https://indico.cern.ch/event/750113/.

[281] James Mulligan. "AliPhysics: Jet QA documentation". URL: http://alidoc.cern.ch/AliPhysics/master/README jetQA.html.

[282] Laura B. Havener. "PWG-JE presentation: jet QA for 5.02 TeV Pb–Pb". URL: https://indico.cern.ch/event/825841/.

[283] Marta Verweij. "Measurement of jet spectra with charged particles in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV with the ALICE detector at the LHC". ALICE Analysis Note ANA-34. 2012. URL: https://alice-notes.web.cern.ch/node/34.

[284] Miklos Gyulassy and Xin-Nian Wang. "HIJING 1.0: A Monte Carlo program for parton and particle production in high energy hadronic and nuclear collisions". In: *Comput. Phys. Commun.* 83.2-3 (1994), pp. 307–331. DOI: 10.1016/0010-4655(94)90057-4. arXiv: nucl-th/9502021.

[285] James Mulligan. "Inclusive full jet measurements in Pb–Pb collisions at 5.02 TeV with ALICE". ALICE Analysis Note ANA-814. 2019. URL: https://alice-notes.web.cern.ch/node/814.

[286] Ezra D. Lesser, James Mulligan, and Mateusz Płoskoń. "Measurements of the groomed and ungroomed jet angularities and mass in central Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV". ALICE Analysis Note ANA-1306. 2022. URL: https://alice-notes.web.cern.ch/node/1306.

[287] Peter Berta et al. "Particle-level pileup subtraction for jets and jet shapes". In: *JHEP* 06 (2014), p. 092. DOI: 10.1007/JHEP06(2014)092. arXiv: 1403.3108 [hep-ex].

[288] M. Cacciari, G. Salam, and G. Soyez. *fastjet::GridMedianBackgroundEstimator Class Reference*. URL: https://fastjet.fr/repo/doxygen-3.3.4/classfastjet_1_1GridMedianBackgroundEstimator.html.

[289] Ezra D. Lesser, James Mulligan, and Mateusz Płoskoń. "Measurements of generalized jet angularities in pp collisions at $\sqrt{s} = 5.02$ TeV with ALICE". ALICE Analysis Note ANA-1068. URL: https://alice-notes.web.cern.ch/node/1068.

[290] G. D'Agostini. "A multidimensional unfolding method based on Bayes' theorem". In: *NIMA* 362.2 (1995), pp. 487–498. DOI: 10.1016/0168-9002(95)00274-X. URL: https://cds.cern.ch/record/265717.

[291] RooUnfold development team. *RooUnfold: ROOT Unfolding Framework*. URL: https://gitlab.cern.ch/RooUnfold/RooUnfold.

[292] James Mulligan and Mateusz Płoskoń. "Measurements of the groomed jet radius and groomed momentum fraction in pp and central Pb–Pb collisions at 5.02 TeV". ALICE Analysis Note ANA-1069. 2021. URL: https://alice-notes.web.cern.ch/node/1069.

[293] Auro P. Mohanty. "D^0-tagged jets in pp collisions at 5.02 TeV". ALICE Analysis Note ANA-840. 2018. URL: https://alice-notes.web.cern.ch/node/840.

[294] Ritsuya Hosokawa. "Measurement of charged jet cross section in pp collisions at $\sqrt{s} = 5.02$ TeV". ALICE Analysis Note ANA-534. 2016. URL: https://alice-notes.web.cern.ch/node/534.

[295] Tatsuya Chujo. "Measurement of charged jet spectra in Pb–Pb collisions at $\sqrt{s} = 5.02$ TeV with ALICE at LHC (update including high interaction Pb–Pb runs)". ALICE Analysis Note ANA-818. 2018. URL: https://alice-notes.web.cern.ch/node/818.

[296] ALICE PWG-JE meeting on tracking efficiency uncertainty. 2022. URL: https://indico.cern.ch/event/1172063/.

[297] Jaime Norman. TrackEfficiencyConfiguration.yaml. 2022. URL: https://github.com/alisw/AliPhysics/blob/541e459b8591a5b8b1854cf758a89c1566472f9c/PWGJE/EMCALJetTasks/macros/TrackEfficiencyConfiguration.yaml.

[298] Pedro Cal et al. "The soft drop momentum sharing fraction z_{g} beyond leading-logarithmic accuracy". In: *Phys. Lett. B* 833 (2022), p. 137390. DOI: 10.1016/J.PhysLetB.2022.137390. arXiv: 2106.04589 [hep-ph].

[299] Zhong-Bo Kang, Felix Ringer, and Ivan Vitev. "The semi-inclusive jet function in SCET and small radius resummation for inclusive jet production". In: *JHEP* 10 (2016), p. 125. ISSN: 1029-8479. DOI: 10.1007/JHEP10(2016)125. arXiv: 1606.06732 [hep-ph].

[300] **ALICE** Collaboration: S. Acharya et al. "Measurements of the groomed jet radius and momentum splitting fraction with the soft drop and dynamical grooming algorithms in pp collisions at $\sqrt{s} = 5.02$ TeV". In: *JHEP* 05 (2023), p. 244. DOI: 10.1007/JHEP05(2023)244. arXiv: 2204.10246 [nucl-ex].

[301] **ALICE** Collaboration: S. Acharya et al. "Measurements of the groomed and ungroomed jet angularities in pp collisions at $\sqrt{s} = 5.02$ TeV". In: *JHEP* 05 (2022), p. 061. DOI: 10.1007/JHEP05(2022)061. arXiv: 2107.11303 [nucl-ex].

[302] Yuri L. Dokshitzer et al. "On the QCD analysis of jet broadening". In: *JHEP* 01 (1998), p. 011. DOI: 10.1088/1126-6708/1998/01/011. arXiv: hep-ph/9801324.

[303] Preeti Dhankher, Ezra D. Lesser, and Mateusz Płoskoń. "D^0 meson-tagged jet substructure measurement in pp collisions". ALICE Analysis Note ANA-1373. 2023. URL: https://alice-notes.web.cern.ch/node/1373.

[304] José Guilherme Milhano and Korinna Zapp. *Improved background subtraction and a fresh look at jet sub-structure in JEWEL*. 2022. arXiv: 2207.14814 [hep-ph].

[305] Jun Yan et al. "Medium modifications of girth distributions for inclusive jets and Z^0 + jet in relativistic heavy-ion collisions at the LHC". In: *Chin. Phys. C* 45.2 (2021), p. 024102. DOI: 10.1088/1674-1137/abca2b. arXiv: 2005.01093 [hep-ph].

[306] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "PYTHIA 6.4 Physics and Manual". In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175.

[307] Ezra D. Lesser. "Measurement of the jet mass and angularities in Pb–Pb collisions at 5.02 TeV with ALICE". Proceedings of Science: 11th International Conference on Hard and Electromagnetic Probes of High-Energy Nuclear Collisions. 2023.

[308] Yue Shi Lai et al. "The information content of jet quenching and machine learning assisted observable design". In: *JHEP* 06 (2022), p. 011. DOI: 10.1007/JHEP10(2022)011. arXiv: 2111.14589 [hep-ph].

[309] Raymond J. Ehlers. "Hardest k_{T} Jet Splittings in pp and Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV". ALICE Analysis Note ANA-1071. 2020. URL: https://alice-notes.web.cern.ch/node/1071.

[310] **ALICE** Collaboration. *Measurement of the angle between jet axes in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV*. 2023. arXiv: 2303.13347 [nucl-ex].

[311] **NPLQCD** Collaboration: S. R. Beane et al. "Magnetic Moments of Light Nuclei from Lattice Quantum Chromodynamics". In: *Phys. Rev. Lett.* 113.25 (2014), p. 252001. DOI: 10.1103/PhysRevLett.113.252001. arXiv: 1409.3556 [hep-lat].

[312] Zohreh Davoudi. "Lattice QCD input for nuclear structure and reactions". In: *EPJ Web of Conferences* 175 (2018). Ed. by M. Della Morte et al., p. 01022. DOI: 10.1051/EPJConf/201817501022. arXiv: 1711.02020 [hep-lat].

[313] Zohreh Davoudi. "Light Nuclei from Lattice QCD: Spectrum, Structure and Reactions". In: *Recent Progress in Few-Body Physics*. Ed. by N. A. Orr et al. Springer International Publishing, 2020. DOI: 10.1007/978-3-030-32357-8. arXiv: 1902.04959 [hep-lat].

[314] Anna Ferdinand, Kyle Lee, and Aditya Pathak. *Field-Theoretic Analysis of Hadronization Using Soft Drop Jet Mass*. 2023. arXiv: 2301.03605 [hep-ph].

[315] R. Abdul Khalek et al. "Science Requirements and Detector Concepts for the Electron-Ion Collider: EIC Yellow Report". In: *Nucl. Phys. A* 1026 (2022), p. 122447. DOI: 10.1016/J.NuclPhysA.2022.122447. arXiv: 2103.05419 [physics.ins-det].

[316] AliRDHFCutsD0toKpi.cxx – AliPhysics software repository. URL: https://github.com/alisw/AliPhysics/blob/master/PWGHF/vertexingHF/AliRDHFCutsD0toKpi.cxx#L845-L859.

[317] **ALICE** Collaboration: S. Acharya et al. "Measurement of the production of charm jets tagged with D^0 mesons in pp collisions at $\sqrt{s} = 5.02$ and 13 TeV". In: *JHEP* 06 (2023), p. 133. DOI: 10.1007/JHEP06(2023)133. arXiv: 2204.10167 [nucl-ex].

Appendix A

Additional kinematic efficiency checks

Given the large MC correction involved in the kinematic efficiency for the bins which are directly adjacent to the $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}}$ cuts, it is worthwhile to investigate what the effects are of these corrections. This appendix looks at various tests for this bin to ensure that the MC corrections being applied to the result are reasonable.

A.1 Generator resilience

One way to test whether or not the MC corrections being applied in these bins are reasonable is to examine the effects of several different generators. If different generators give reasonably similar results, then it can be inferred that the bias induced is smaller than the size of the actual correction. Note that the figures do not have to be identical, since a systematic uncertainty is assigned to the overall choice of generator, which is designed to account for effects such as this. The generator and total systematic uncertainties can be seen in the relevant figures in Sect. 3.4.

Here four additional calculations are shown of the kinematic efficiency, using four different MC generators:

1. PYTHIA 8 Monash 2013 (taken as default case in the generator systematic);
2. Herwig7 default tune;
3. JEWEL Pb–Pb, recoils off;
4. JEWEL Pb–Pb, recoils on, without recoil subtraction.

The four different simulations are passed through a fast simulation which accounts for the ALICE tracking efficiency and track p_{T} resolution. This means that the PYTHIA 8 efficiency shown here is not identical to the one used for the final measurement; however, they are observed to be very similar, as expected. JEWEL is used both with recoils on and off to probe the size of the perturbation that model-dependent effects can have. While there

are different prescriptions to apply a "recoil subtraction" procedure, the case of no recoil subtraction is investigated as a maximal deviation one could expect from this effect.

Example kinematic efficiencies ε_{kin} for the ungroomed λ_α distributions are given in Fig. A.1. Note that by default the statistical uncertainties are set to 0. In general, good agreement is seen with the shape of the corrections. The last bin in these plots is used for capturing outliers in the unfolding and is not reported; for the actual binning that is reported, see Ch. 5. Also recall that, as stated in Sect. 3.3.3, since final distributions are all normalized to 1, ε_{kin} enters as a correction to the shape rather than the overall number of jets.

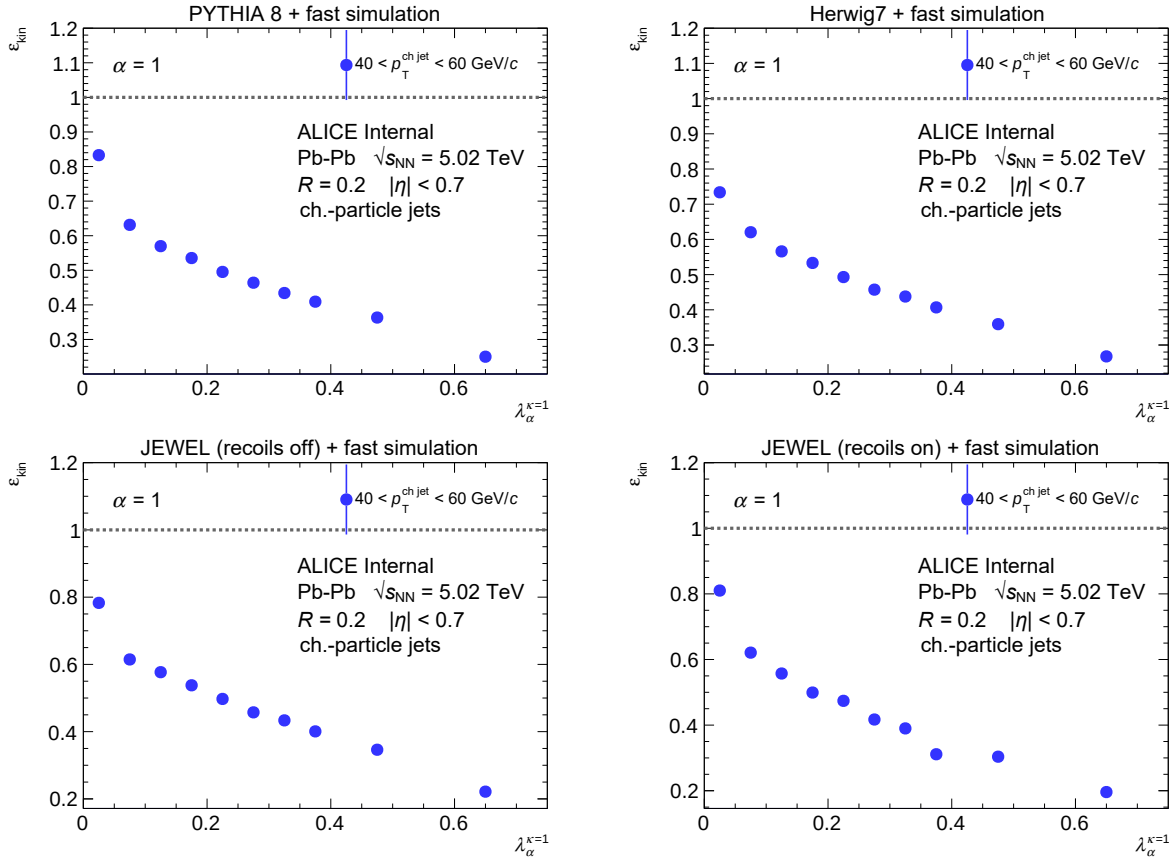


Figure A.1: Kinematic efficiency of λ_1 for $R = 0.2$ jets in Pb–Pb collisions for $40 < p_{\text{T,truth}}^{\text{ch jet}} < 60$ GeV/c with four different generator selections, each combined with the same fast simulation [286]. Top left: PYTHIA8 Monash 2013. Top right: Herwig7 default tune. Bottom left: JEWEL Pb–Pb, recoils off. Bottom right: JEWEL Pb–Pb, recoils on, without recoil subtraction. The final bin is truncated from the reported distributions.

A.2 Test at higher jet transverse momentum

Another test is purposefully increasing the applied cuts on $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}}$ to be adjacent to the second-lowest reported $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$ bin, and then see how this larger cut affects the result in that bin as compared to the nominal result (using a less liberal cut). This of course verifies mostly the behavior of the distributions at higher $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$, but seeing reasonable agreement within uncertainties there would lends some confidence on the distributions at lower $p_{\mathrm{T}}^{\mathrm{ch\ jet}}$.

To implement this test, the nominal detector-level cut is changed to be $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}} > 60$ GeV/c for $R = 0.2$ jets. The effect of this aggressive cut is then investigated for the $60 < p_{\mathrm{T}}^{\mathrm{ch\ jet}} < 80$ GeV/c bin in the unfolded data.

Example kinematic efficiencies for this larger cut are given in Fig. A.2. The drop in kinematic efficiency is observed to be similar as in the nominal $40 < p_{\mathrm{T}}^{\mathrm{jet}} < 60$ GeV/c range, with significant MC scaling needed for all distributions. To verify that the unfolding is still behaving as expected, the unfolding tests are required to converge given this stricter cut. Example closure tests are given in Fig. A.3, which can be compared to the nominal closure tests in Sect. 3.3.1.

Finally, the unfolded results with this variation are compared to the nominal results presented in Ch. 5. The distributions are found to be in complete agreement in all cases. There is no significant effect on the the physics conclusions from this stricter $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}}$ cut, and the much larger unfolding uncertainty accounts for the differences.

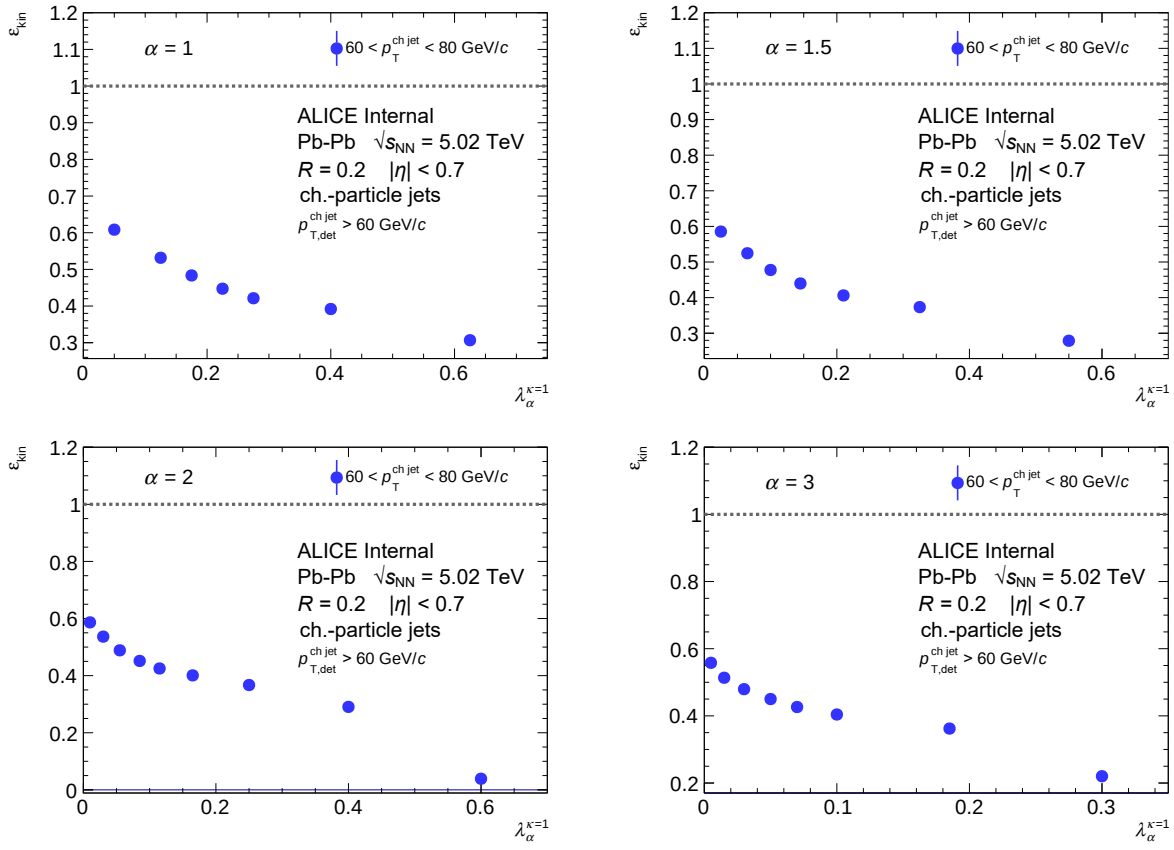


Figure A.2: Kinematic efficiency of λ_α for $R = 0.2$ charged-particle jets for $60 < p_{\mathrm{T,truth}}^{\mathrm{ch\,jet}} < 80$ GeV/c with a stricter $p_{\mathrm{T,det}}^{\mathrm{ch\,jet}}$ cut, as simulated using the anchored PYTHIA 8 + GEANT3 production embedded in Pb–Pb data and background-subtracted with CS [286]. The final bin is truncated from the reported distributions. All four plots are computed using the same jets, with $\alpha = 1$, 1.5, 2, and 3.

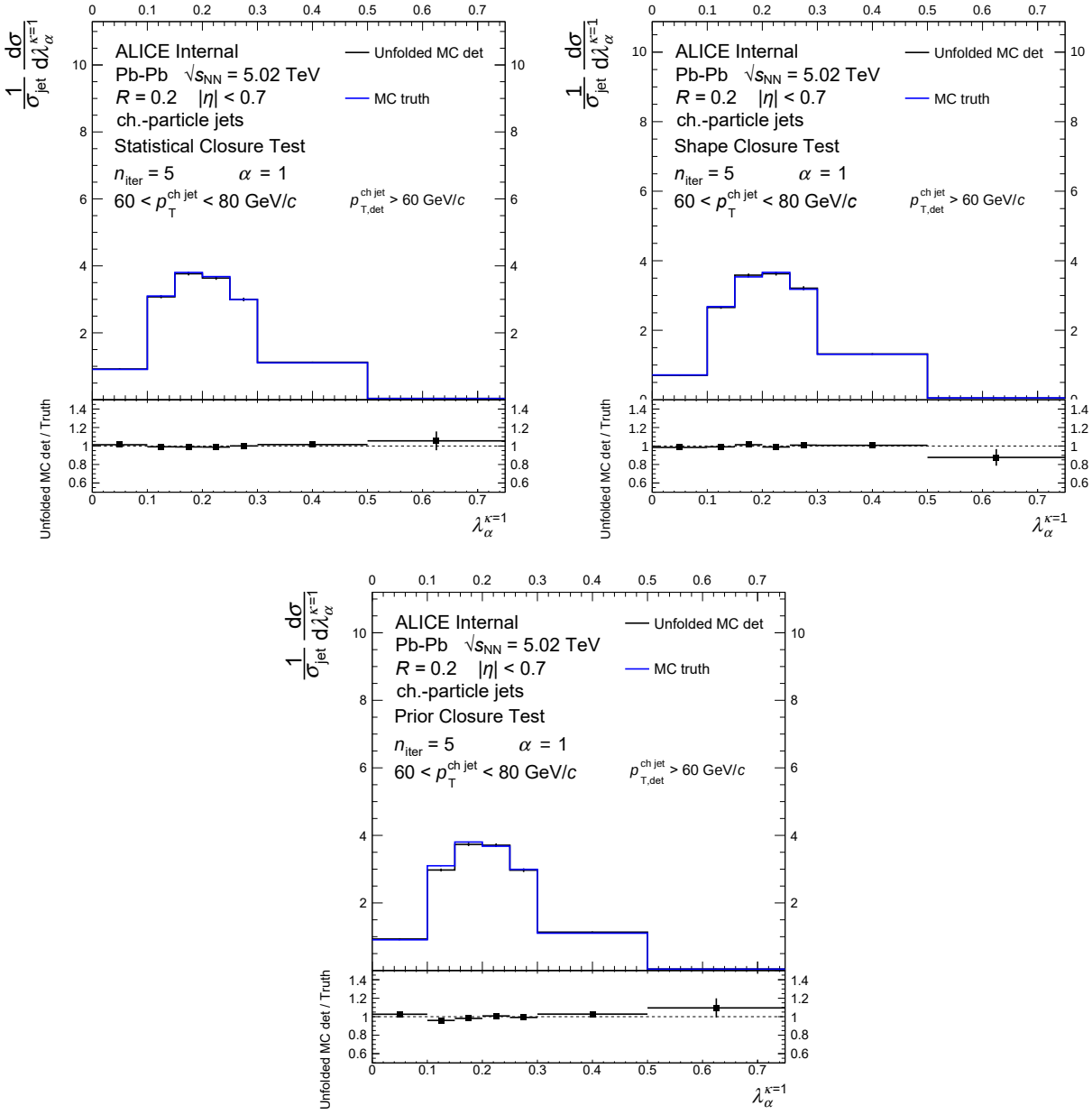


Figure A.3: Example closure tests for λ_1 with $R = 0.2$ charged-particle jets using a stricter $p_{\mathrm{T,det}}^{\mathrm{ch\ jet}}$ cut, as simulated using the anchored PYTHIA 8 + GEANT3 production embedded in Pb–Pb data and background-subtracted with CS [286]. The final bin is truncated from the reported distributions. The test is considered successful in each case.

Appendix B

Heavy-flavor tagged jet analysis

This appendix details the ALICE heavy-flavor (HF) tagged jet analysis, which is compared to the results of this thesis (see Sect. 5.2.1). Further details are also given in Ref. [238, 303].

B.1 Data sets

For the HF-jet analysis, the same MB data is used (see Sect. 3.1) with a separate MC production. The inclusive production LHC18b8 does not extend to low-p_{T} (< 5 GeV/c), and is statistically limited for HF processes. To account for this, an alternate MC was simulated using the PYTHIA 6 event generator [306] with the `Pythia6_Perugia2011_HF001:hv0` tune, and similarly transported through GEANT3. The charm content in this production (named LHC18a4a2) is enhanced by requesting a $\mathrm{c\bar{c}}$ pair in 50% of the events, and a $\mathrm{b\bar{b}}$ pair in the remaining half. Furthermore, all D mesons are forced to decay hadronically. This MC production is used to compute the D^0-meson efficiency with jets, acceptance corrections, a response matrix of D^0-tagged jets, and to correct for the effects of prompt D^0 ($\mathrm{c} \rightarrow \mathrm{D}^0$) versus non-prompt D^0 ($\mathrm{b} \rightarrow \mathrm{c} + \mathrm{W}^-$, with the "non-prompt" c then $\rightarrow \mathrm{D}^0$). The "good" runs are again the same as in data:

- **LHC18a4a2** (44 runs): 282343, 282342, 282341, 282340, 282314, 282313, 282312, 282309, 282307, 282306, 282305, 282304, 282303, 282302, 282247, 282230, 282229, 282227, 282224, 282206, 282189, 282147, 282146, 282127, 282126, 282125, 282123, 282122, 282120, 282119, 282118, 282099, 282098, 282078, 282051, 282050, 282031, 282025, 282021, 282016, 282008, 282367, 282366, 282365

with approximately 6.7M events generated.

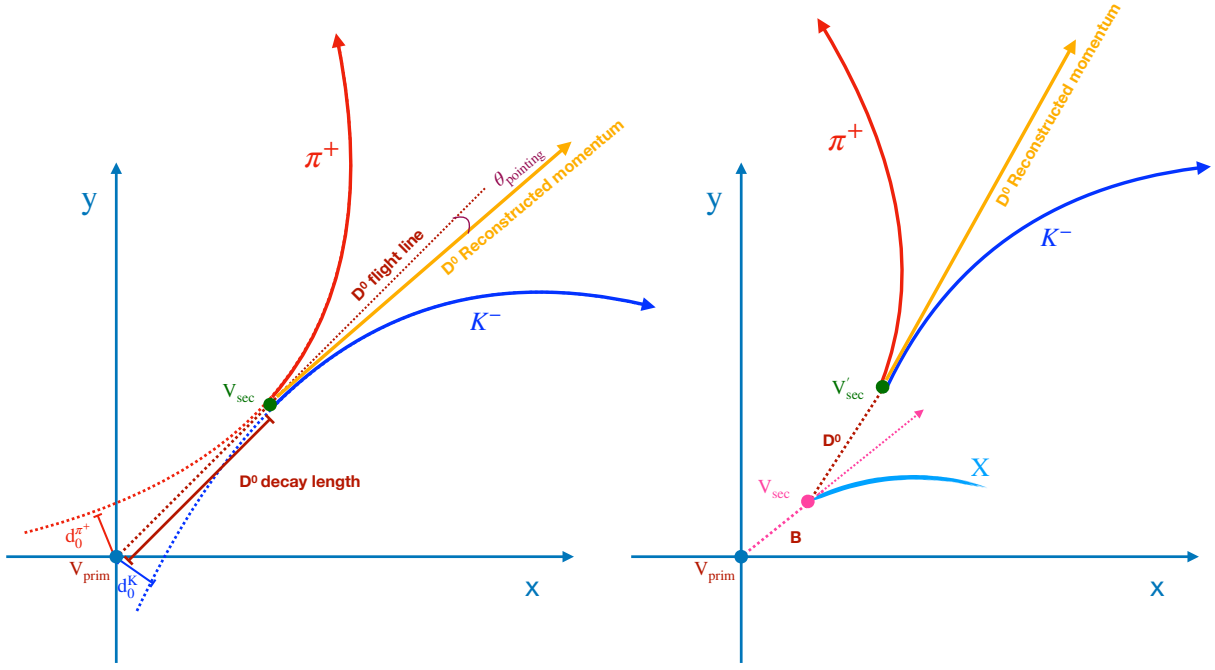


Figure B.1: Two diagrams of a $D^0 \to K^- \pi^+$ decay [303]. *Left:* a "prompt" D^0 meson, originating from the hard scattering, decays at a secondary vertex V_{sec}, nominally 122 μm away from the IP in its rest frame. *Right:* a "non-prompt" D^0 meson, originating from the "feed-down" decay of a B meson, decays with slightly different kinematics.

B.2 HF meson and jet reconstruction

Details on the reconstruction of D^0-tagged jets is given in Ref. [238]. For D^0-meson tagged jets, the D^0 candidate and its charge conjugate[1] are first reconstructed, prior to jet reconstruction, via the hadronic decay mode $D^0 \to K^-\pi^+$ (or $\overline{D}^0 \to K^+\pi^-$), which carries a branching ratio of 3.947(30)% [3]. See Fig. B.1 for a decay diagram.

D^0 mesons are reconstructed by applying topological selections, described in Table B.1, and by PID (see Sect. 2.2 for PID capabilities). Each candidate also passes a fiducial acceptance cut in rapidity: as the D^0 decay angle becomes large at low $p_T^{D^0}$, the acceptance rapidly falls due to edge effects of the TPC acceptance. The cut values are therefore $p_T^{D^0}$-dependent, and grow from $y < 0.5$ to 0.8 (according to a second-order polynomial function) for $p_T^{D^0} \leq 5$ GeV/c, and remain at $y < 0.8$ for $p_T^{D^0} > 5$ GeV/c [316]. These requirements remove a significant amount of combinatorial background from uncorrelated tracks. Decay daughters of each D^0 candidate are then replaced by their four-momentum vector sum (i.e., the D^0). This mitigates cases where the angle between decay daughters is larger than the jet radius, which would degrade the jet energy resolution. D^0 mesons are reconstructed in

[1]This analysis does not discriminate between D^0 and $\overline{D}^0$, so for simplification, both are referred to a D^0.

$p_{\mathrm{T}}^{\mathrm{D}^0}$ (GeV/c)		2 – 3	3 – 4	4 – 5	5 – 6	6 – 7	7 – 8	8 – 10	10 – 12	12 – 16	16 – 24	24 – 36
ΔM_{D^0} (GeV/c^2)	<	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
DCA (cm)	<	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
$\lvert\cos(\theta^*)\rvert$	<	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	1	1	1
$p_{\mathrm{T,K}}$ (GeV/c)	>	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
$p_{\mathrm{T},\pi}$ (GeV/c)	>	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
$\lvert d_0^K\rvert$ (cm)	<	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
$\lvert d_0^\pi\rvert$ (cm)	<	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
$d_0^K \cdot d_0^\pi$ (10^{-4}cm^2)	<	-3	-3	-1.5	-1	-0.8	-0.8	-0.5	-0.5	1	1	1
$\cos(\theta_{\mathrm{point}})$	>	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.90	0.90
$L_\perp/\sigma_{L_\perp}$(cm)	>	5	5	5	4	4	4	3	3	3	3	3

Table B.1: List of the topological cuts used to reconstruct D^0 mesons across various $p_{\mathrm{T}}^{\mathrm{D}^0}$ intervals in pp collisions at $\sqrt{s} = 5.02$ TeV [303].

the transverse momentum interval $2 \leq p_{\mathrm{T}}^{\mathrm{D}^0} \leq 36$ GeV/c.

D^0 candidates are reconstructed using the measured value for m_{D^0} (at detector level) or using the accepted value $m_{\mathrm{D}^0} = 1.864$ GeV/c^2 (at generator level). All other tracks are still reconstructed using the mass of the $\pi^\pm$ meson, as described in Sect. 3.2.2.

D^0-meson tagged jets are reconstructed in pp data using the same approach as the inclusive jets, with radius parameter $R = 0.4$. It is possible for more than one D^0 candidate to pass the topological selections per event, and these candidates can sometimes share the same daughter. For this reason, jet reconstruction is repeated independently for each candidate D^0 meson in the event, meaning that each candidate is treated as if it were the only one in the event. After jet reconstruction, only those jets which contain the D^0 candidate are considered. Since these jets are from the same MB data as the inclusive jet sample, they are statistically limited by the charm production cross section, and are reconstructed with $5 < p_{\mathrm{T}}^{\mathrm{ch\,jet}} < 50$ GeV/c.

B.3 Jet corrections

Despite suppressing much of the combinatorial background using the selection criteria detailed in Sect. B.2 above, a significant number of combinatorial "fake" D^0 candidates remain. In order to correct for this, a data-driven technique utilizing the invariant mass spectrum of D^0 candidates is employed, wherein the invariant mass peak is fit with a Gaussian function while the background is fit with an exponential decay. From this invariant mass spectrum, two different regions are defined: a "signal plus background" region, within $\pm 2\sigma$ of the Gaussian mean, and a "background" region, estimated by taking the 4–9σ region on either side of the Gaussian peak. The side-bands are selected far enough away from the peak that they are background-dominated, but are truncated before statistical fluctuations become too significant. Using the jets which contain the D^0 candidates in either region, the jet substructure observable distribution is computed, and the "background" distribution is then scaled by the fractional area and subtracted from the "signal plus background" distribution to ob-

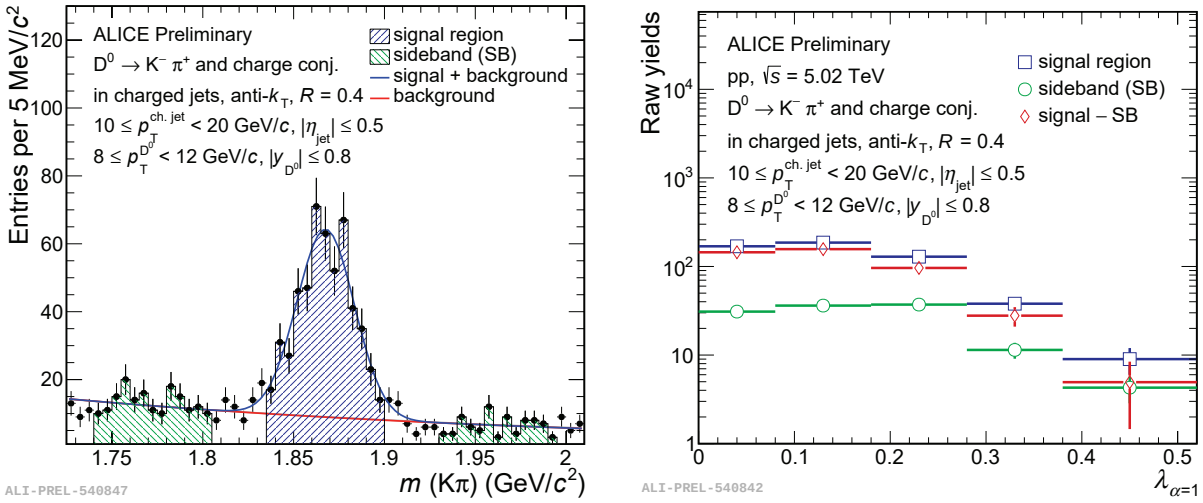


Figure B.2: *Left:* example D^0 invariant mass spectra with signal and background fits, using $8 \leq p_{\mathrm{T}}^{\mathrm{D}^0} < 13$ GeV/c candidates inside of jets with $10 \leq p_{\mathrm{T}}^{\text{ch jet}} < 20$ GeV/c. *Right:* the corresponding correction of the jet angularity λ_1 (girth) raw yield.

tain a background-subtracted "signal" distribution. These distributions are finally scaled by $(0.9545)^{-1}$ to account for the limited 2σ width of the "signal plus background" region. An example of this fit and the resulting subtraction is given in Fig. B.2.

Sometimes neither of the D^0 daughters can be definitively identified as a $\pi^{\pm}$ or $K^{\mp}$; for example, around $p_{\text{track}} \approx 1$ GeV/c, π and K separation fails due to crossing in the $\mathrm{d}E/\mathrm{d}x$ spectra (see right panel of Fig. 2.3). In such cases, both π and K assignment hypotheses are taken, and the pair is accepted as both a D^0 and a $\overline{\mathrm{D}}^0$ candidate, with the two respective invariant mass values being used. The duplicate D^0 (or $\overline{\mathrm{D}}^0$) candidate, using the wrong invariant mass, is called a *reflection signal*. The number of reflected candidates is $p_{\mathrm{T}}^{\mathrm{D}^0}$dependent, and is strongest at low $p_{\mathrm{T}}^{\mathrm{D}^0}$ [303]. The reflection templates are obtained from the PYTHIA 8 simulations described in Sect. 3.1 and parameterized as a sum of two Gaussians with the means, widths, and the D^0 signal-over-reflection ratio fixed to values obtained in the simulations [317]. This component is included in the invariant mass fitting procedure described above, and thus subtracted from the "signal plus background" distributions.

The background-subtracted distributions for D^0-tagged jets require additional corrections. First, an efficiency correction is needed to account for the topological cuts and PID selection. This is evaluated using the anchored MC production (LHC18a4a2). The ratio is taken of the matched, detector-level D^0-tagged jets passing all selection criteria (see Table B.1) versus all truth-level D^0-tagged jets within the detector acceptance. This is evaluated for each detector-level $p_{\mathrm{T}}^{\mathrm{D}^0}$ bin,

$$\varepsilon^{\mathrm{c}\to\mathrm{D}^0}(p_{\mathrm{T,det}}^{\mathrm{D}^0}) = \frac{N_{\text{ch jet,det}}^{\mathrm{c}\to\mathrm{D}^0}(p_{\mathrm{T,det}}^{\mathrm{D}^0})}{N_{\text{ch jet,truth}}^{\mathrm{c}\to\mathrm{D}^0}(p_{\mathrm{T,det}}^{\mathrm{D}^0})}. \tag{B.1}$$

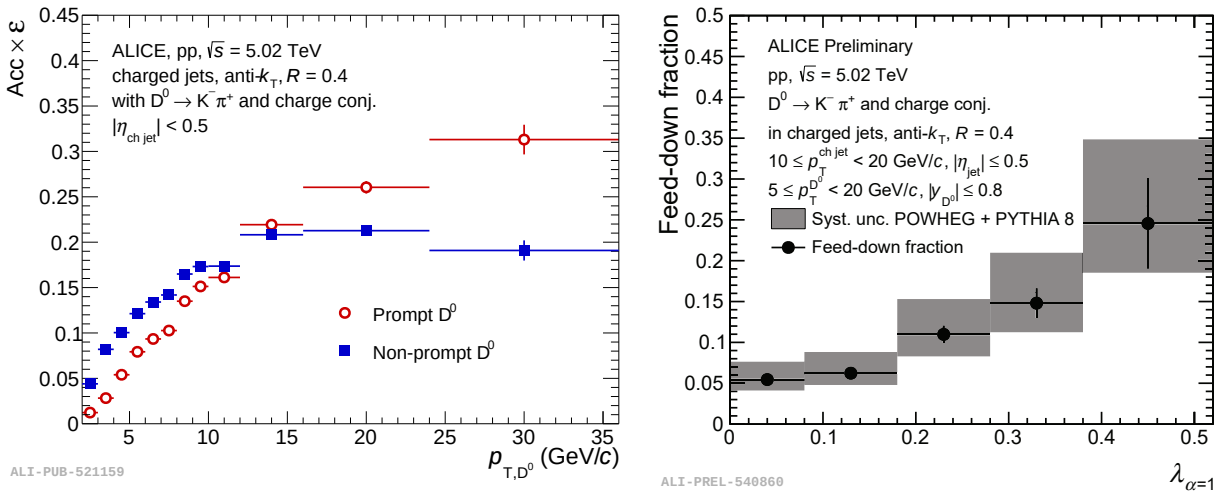


Figure B.3: *Left:* prompt (red) and non-prompt (blue) D^0-meson jet acceptance × efficiency as a function of detector-level $p_{\mathrm{T}}^{\mathrm{D}^0}$, as evaluated with PYTHIA 8 + GEANT3 [317]. *Right:* fraction of jet angularities λ_1 coming from the feed-down of a B meson, as calculated with POWHEG + PYTHIA 8 and folded to detector level, versus the number of inclusive, background-subtracted D^0-tagged jets, as measured in data [303]. Systematic uncertainties (shaded bands) are discussed in Sect. B.4.

Using this value, the background-subtracted yield (e.g., the red-colored λ_1 curve in the right panel of Fig. B.2) can be corrected for the prompt D^0 finding efficiency,

$$N^{\mathrm{data}}(p_{\mathrm{T,det}}^{\text{ch jet}}, O_{\mathrm{det}}) = \sum_{p_{\mathrm{T,det}}^{\mathrm{D}^0}} \frac{N^{\mathrm{data,raw}}(p_{\mathrm{T,det}}^{\text{ch jet}}, O_{\mathrm{det}})}{\varepsilon^{\mathrm{c}\to\mathrm{D}^0}(p_{\mathrm{T,det}}^{\mathrm{D}^0})}, \tag{B.2}$$

where O is an arbitrary jet substructure observable distribution (e.g., λ_1).

Some number of D^0 mesons come from the decay of a B meson; these are referred to as "non-prompt" decays (see Fig. B.1). One can similarly calculate the tagging efficiency for non-prompt decays,

$$\varepsilon^{\mathrm{b}\to\mathrm{D}^0}(p_{\mathrm{T,det}}^{\mathrm{D}^0}) = \frac{N_{\text{ch jet,det}}^{\mathrm{b}\to\mathrm{D}^0}(p_{\mathrm{T,det}}^{\mathrm{D}^0})}{N_{\text{ch jet,truth}}^{\mathrm{b}\to\mathrm{D}^0}(p_{\mathrm{T,det}}^{\mathrm{D}^0})}, \tag{B.3}$$

by only considering the D^0 mesons coming from a B meson. The calculated prompt and non-prompt tagging efficiencies from this procedure, combined with detector acceptance effects, are shown in the left panel of Fig. B.3 as a function of $p_{\mathrm{T,det}}^{\mathrm{D}^0}$.

To fully correct for the presence of feed-down decays in the measured data, another subtraction procedure must be used. Since a data-driven measurement of the non-prompt D^0-tagged jet is not available, a simulation using POWHEG matched to PYTHIA 8 parton shower and hadronization models [44] (see Sect. 1.2.4) is used to estimate the feed-down

contribution. The NLO matrix elements in POWHEG produce the D^0 cross section with higher accuracy than the LO ones from PYTHIA alone. In order to use this separate MC production without running the full GEANT3 simulation, a "forward-folding" procedure is used to account for detector effects (see Sect. 3.3). The anchored PYTHIA 8 + GEANT3 simulation is utilized to create a four-dimensional *response matrix* (RM),

$$\mathrm{R}^{\mathrm{b}\to\mathrm{D}^0}\left(p_{\mathrm{T,det}}^{\text{ch jet}}, p_{\mathrm{T,truth}}^{\text{ch jet}}, O_{\mathrm{det}}, O_{\mathrm{truth}}\right). \tag{B.4}$$

This RM is then multiplied by the POWHEG + PYTHIA 8 prediction (i.e., the MC prediction is "folded") in order to obtain a detector-level distribution,

$$\begin{aligned} N^{\mathrm{b}\to\mathrm{D}^0}(p_{\mathrm{T,det}}^{\text{ch jet}}, O_{\mathrm{det}}) = \mathrm{R}^{\mathrm{b}\to\mathrm{D}^0}\left(p_{\mathrm{T,det}}^{\text{ch jet}}, p_{\mathrm{T,truth}}^{\text{ch jet}}, O_{\mathrm{det}}, O_{\mathrm{truth}}\right) \times \\ \sum_{p_{\mathrm{T,truth}}^{\mathrm{D}^0}} \left[\frac{\varepsilon^{\mathrm{b}\to\mathrm{D}^0}(p_{\mathrm{T,truth}}^{\mathrm{D}^0})}{\varepsilon^{\mathrm{c}\to\mathrm{D}^0}(p_{\mathrm{T,truth}}^{\mathrm{D}^0})} N_{\mathrm{POWHEG}}^{\mathrm{b}\to\mathrm{D}^0}(p_{\mathrm{T,truth}}^{\text{ch jet}}, O_{\mathrm{truth}}, p_{\mathrm{T,truth}}^{\mathrm{D}^0})\right] \end{aligned} \tag{B.5}$$

where $\varepsilon^{\mathrm{b}\to\mathrm{D}^0}(p_{\mathrm{T,truth}}^{\mathrm{D}^0})$ and $\varepsilon^{\mathrm{c}\to\mathrm{D}^0}(p_{\mathrm{T,truth}}^{\mathrm{D}^0})$ are calculated in the same way as given in Eq. B.1 and B.3, but in intervals of $p_{\mathrm{T,truth}}^{\mathrm{D}^0}$ instead of $p_{\mathrm{T,det}}^{\mathrm{D}^0}$ since the distributions are at truth level. The extra factor of $\varepsilon^{\mathrm{c}\to\mathrm{D}^0}(p_{\mathrm{T,truth}}^{\mathrm{D}^0})$ is required to account for the equivalent scaling that occurred on the "hidden" feed-down component in Eq. B.2; these distributions will be subtracted below. Equation B.5 also includes corrections due to *kinematic efficiencies* of the response matrix, which are a minor effect and for simplicity are not explicitly shown; for details on the kinematic efficiency, see Sect. 3.3.3.

The number of D^0-tagged jets coming from non-prompt B meson decays relative to the number of inclusive (prompt plus non-prompt) decays is known as the *feed-down fraction*. An example of the feed-down fraction for λ_1 is shown in the right panel of Fig. B.3, which is obtained by taking the ratio of $N^{\mathrm{b}\to\mathrm{D}^0}(p_{\mathrm{T,det}}^{\text{ch jet}}, O_{\mathrm{det}})$ as calculated above and dividing by $N^{\mathrm{data}}(p_{\mathrm{T,det}}^{\text{ch jet}}, O_{\mathrm{det}})$ (before correcting for the feed-down effect).

The background-subtracted data distributions can finally be corrected at detector level by using a combination of the prompt efficiency-corrected data (Eq. B.2) and the estimated value of $N^{\mathrm{b}\to\mathrm{D}^0}(p_{\mathrm{T,det}}^{\text{ch jet}}, O_{\mathrm{det}})$ from POWHEG + PYTHIA 8,

$$N^{\mathrm{c}\to\mathrm{D}^0}(p_{\mathrm{T,det}}^{\text{ch jet}}, O_{\mathrm{det}}) = N^{\mathrm{data}}(p_{\mathrm{T,det}}^{\text{ch jet}}, O_{\mathrm{det}}) - N^{\mathrm{b}\to\mathrm{D}^0}(p_{\mathrm{T,det}}^{\text{ch jet}}, O_{\mathrm{det}}), \tag{B.6}$$

which is the final prompt distribution that will be unfolded for detector effects, using a RM produced using only jets with prompt $\mathrm{c}\to\mathrm{D}^0$ mesons,

$$\mathrm{R}^{\mathrm{c}\to\mathrm{D}^0}\left(p_{\mathrm{T,det}}^{\text{ch jet}}, p_{\mathrm{T,truth}}^{\text{ch jet}}, O_{\mathrm{det}}, O_{\mathrm{truth}}\right). \tag{B.7}$$

Parameter	Central Value	Systematic Variations
m_{b}	4.75 GeV/c^2	4.5, 5.0 GeV/c^2
PDF	CT10nlo (11000)	CT10nlo (21200)
(μ_{F}, μ_{R})	(1, 1)	(0.5, 0.5), (0.5, 1), (1, 0.5), (2, 2), (2, 1), (1, 2)

Table B.2: Parameters of the POWHEG + PYTHIA 8 simulations which are varied to estimate uncertainty on the D^0-tagged jet contribution from feed-down decays of B mesons [303].

B.4 Systematic uncertainties

In the D^0-tagged jet analysis, in addition to the systematic uncertainties discussed for inclusive jets in Sect. 3.4 additional uncertainties are assigned for the topological selection criteria, the invariant mass signal extraction, and the B meson feed-down correction.

The procedure used to select the D^0 meson adds additional systematic uncertainty to the measurement. Uncertainty due to the topological selection criteria is attributed by varying the selection criteria in the vicinity of the values given in Table B.1. Four variations were performed, with three sets of looser and one set of stricter cuts, which are compared to the nominal settings. The root mean square of these four ratios is taken as a systematic uncertainty.

There is also an uncertainty on the feed-down contribution coming from B meson decays, as the distributions obtained from POWHEG + PYTHIA 8 MC (discussed in Sect. B.3) may have some deviation from real B meson decays. In order to assess this systematic uncertainty, the MC simulation is performed with different choices of the quark mass m_{b}, the factorization scale μ_{F}, the renormalization scale μ_{R}, and the PDF [303]. Table B.2 lists these parameters, their central values, and the systematic variations which are considered. The resulting varied distributions are compared to the nominal result, with the maximum spread in their ratios taken as the systematic uncertainty.

Finally, an uncertainty is introduced when extracting the D^0 yield from the invariant mass distributions. Several variations are performed to estimate this:

- The bandwidth of the signal and background regions are varied by $\pm 2\sigma$, with six combinations probing the extremes of these variations.
- The background template is changed from an exponential decay to a linear fit, to account for shape bias.
- The Gaussian fit on the D^0 mass peak is varied from having a free mean peak value (default) to being fixed at the accepted value $m_{\mathrm{D}^0} = 1.864$ GeV/c^2 [3] as well as the value predicted by MC.
- The upper and lower limits of the Gaussian fit are also varied by 0.09 and 0.03 GeV/c^2, respectively.

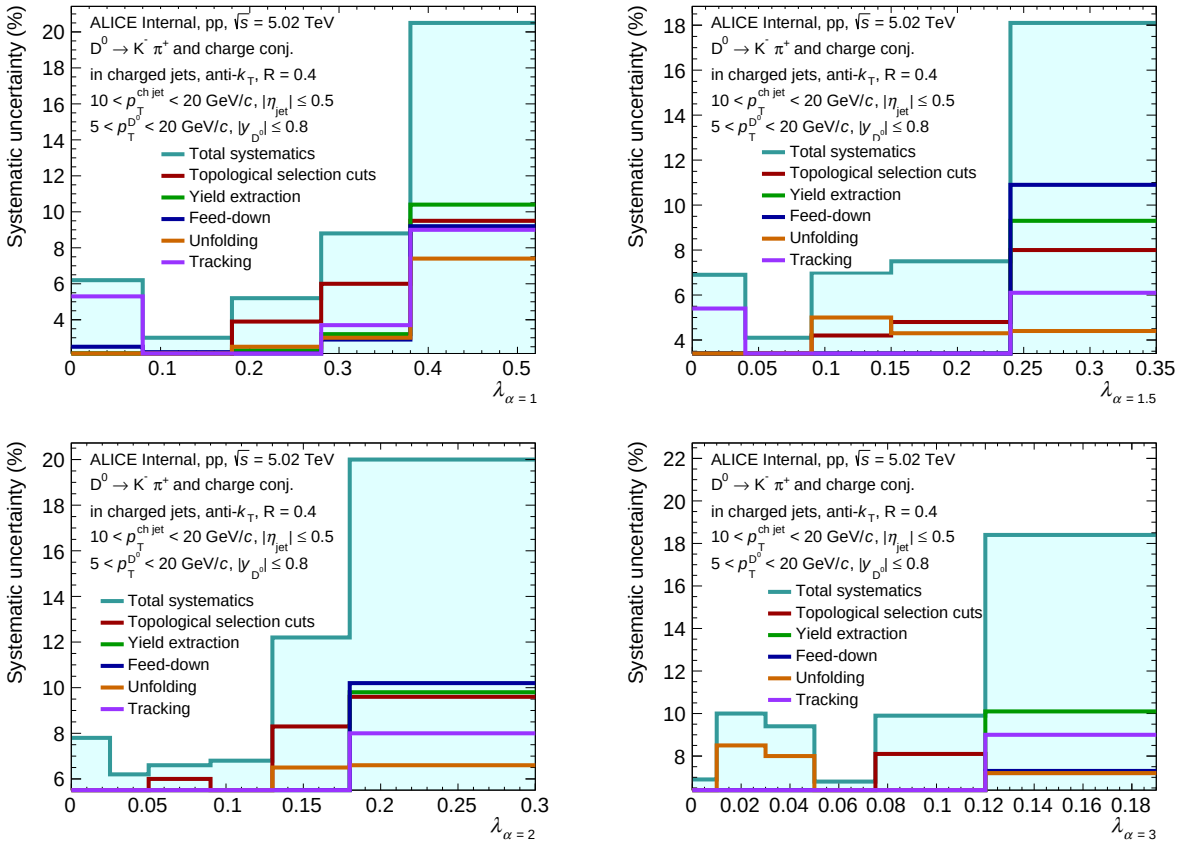


Figure B.4: Total systematic uncertainty of D^0-tagged jet angularities λ_α for $R = 0.4$ charged-particle jets in pp collisions, along with individual sources. The total uncertainty is taken as the quadrature sum of all contributions. [303]

- The invariant mass distributions are rebinned both by a factor of 2 and 5 to estimate bias introduced by binning.

The root mean square of all of these variations, as compared to the nominal result, is taken as a systematic uncertainty.

Systematic uncertainties for the D^0-tagged jet analyses are shown in Fig. B.4. The uncertainty in the last bin has a large statistical covariance, though the systematic uncertainty is kept large to maintain a conservative estimate.