

Czech Technical University in Prague  
Faculty of Nuclear Sciences and Physical Engineering  
Department of Physics



# **Inclusive production of $b$ jets in collisions of $p$ - $Pb$ and $pp$ in ALICE**

Doctoral thesis

*Ing. Artem Isakov*

Ph.D. programme: Applications of Natural Sciences  
Branch of study: Nuclear Engineering  
Supervisor: RNDr. Filip Křížek, Ph.D.

Prague, February 2023

**Thesis Supervisor:**

RNDr. Filip Křížek, Ph.D.  
Faculty of Nuclear Sciences and Physical Engineering  
Czech Technical University in Prague  
Břehová 7, Praha 1, 115 19  
Czech Republic

Copyright © February 2023 Ing. Artem Isakov

# Declaration

I hereby declare I have written this doctoral thesis independently and quoted all the sources of information used in accordance with methodological instructions on ethical principles for writing an academic thesis. Moreover, I state that this thesis has neither been submitted nor accepted for any other degree.

In Prague, February 2023

.....  
Ing. Artem Isakov

# Bibliografický záznam

*Autor:* Ing. Artem Isakov

*Název práce:* **Inkluzivní produkce b jetů ve srážkách p-Pb a pp v experimentu ALICE**

*Studijní program:* Aplikace přírodních věd

*Studijní obor:* Jaderné inženýrství

*Školitel:* RNDr. Filip Křížek, Ph.D., Katedra fyziky, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze

*Akademický rok:* 2022/2023

*Počet stran:* 140

*Klíčová slova:* Srážky těžkých iontů, b jety, b kvark, zhášení jetu, efekty studené jaderné hmoty, experiment ALICE, ITS

# Bibliographic entry

*Author:* Ing. Artem Isakov

*Title:* **Inclusive production of b jets in collisions of p–Pb and pp in ALICE**

*Degree programme:* Applications of Natural Sciences

*Field of study:* Nuclear Engineering

*Sort of project:* Doctoral thesis

*Supervisor:* RNDr. Filip Křížek, Ph.D., Department of Physics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague

*Academic year:* 2022/2023

*Number of pages:* 140

*Key words:* Heavy-ion collisions, b jets, b quark, jet quenching, cold nuclear matter effects, ALICE experiment, ITS

# Abstrakt

Relativně velká hmotnost b kvarků z nich dělá vhodný nástroj pro studium kvarkového-gluonového plazmatu vznikajícího při ultrarelativistických srážkách těžkých iontů. Díky této vlastnosti dochází k produkci b-kvarků převážně v počátečních tvrdých procesech a její výtěžek lze vypočítat pomocí poruchové kvantové chromodynamiky. Krásné kvarky lze tedy považovat za ideální sondy, které pronikají vytvořeným médiem, a lze je využít ke zkoumání vlastností energetických ztrát partonů v médiu a efektů studené jaderné hmoty (CNM) v závislosti na hmotnosti. Vzhledem k tomu, že produkce b-jetů ve srážkách p-Pb nese informaci o změnách partonové distribuční funkce v jádře vůči protonu, představují tato měření důležitou referenci pro budoucí měření v těžších srážkových systémech.

Experiment ALICE na Velkém hadronovém urychlovači (LHC) v CERN využil svých vynikajících rekonstrukčních vlastností částic k přesné rekonstrukci jetů a identifikaci vrcholů rozpadu b-hadronů posunutých o stovky mikrometrů od vrcholu primární interakce. Během druhé LHC kampaně byl nabrán rozsáhlý soubor událostí ze srážek p-Pb a pp při  $\sqrt{s_{NN}} = 5.02$  TeV. Změřená data umožnila sledovat produkci b-jetů až do relativně nízkých hodnot příčné hybnosti, což přináší cenná data o velikosti CNM efektů v oblasti, která je důležitá pro testování různých modelů energetických ztrát ve srážkách Pb-Pb.

Předkládaná práce také pojednává o nedávné modernizaci vnitřního dráhového detektoru experimentu ALICE (ITS), která výrazně zlepšila možnosti identifikace b-jetů. V dizertaci je zejména rozebrán autorův příspěvek k vývoji software Quality Control, který monitoruje funkčnost detektoru ITS.

# Abstract

A relatively large mass of b quarks makes them a convenient tool to study the quark-gluon plasma produced in ultra-relativistic heavy-ion collisions. Thanks to this feature, b-quark production occurs mostly in initial hard scatterings and is calculable using perturbative quantum chromodynamics. Thus, beauty quarks can be considered as ideal penetrating probes of the created medium and can be utilized to investigate the mass-dependent properties of in-medium parton energy loss and cold nuclear matter effects (CNM). Since b-jet production in p–Pb collisions carries information about the changes in Parton Distribution Function in nuclei with respect to protons, these measurements represent an important baseline for future measurements in heavier collision systems.

The ALICE experiment at the Large Hadron Collider (LHC) at CERN exploits its excellent particle-tracking capabilities that allow for precise jet reconstruction and identification of beauty-hadron decay vertices displaced hundreds of micrometers from the primary interaction vertex. A large amount of data of p–Pb and pp collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV recorded during the LHC Run 2 makes it possible to follow the beauty-jet production down to relatively low transverse momenta, constraining CNM effects in a range relevant for testing different energy-loss models in Pb–Pb collisions.

The presented thesis also discusses the recent upgrade of the ALICE Inner Tracking System (ITS) that will significantly improve b-jet tagging capabilities. The emphasis is put on the corresponding Quality Control software, which monitors the performance of ITS sensors.

# Acknowledgements

I would like to thank my supervisor Filip Křížek for his patience, the guidance provided throughout this work, and support that had a significant impact on the course of my Ph.D. degree. Also, I want to express my gratitude to Robert Vertessi, Hadi Hassan, and Ashik Iqbal Sheikh for their invaluable help. Additionally, I would like to thank the ALICE ITS team, especially Ivan Ravasenga, for his help with the service task. Last but not least, I want to acknowledge my parents for inspiring me to pursue physics.



# List of Tables

1.1	Basic quark properties [22]. . . . .	4
3.1	The number of minimum bias events in the measured data sample before ( $N_{\text{events}}^{\text{raw}}$ ) and after ( $N_{\text{events}}^{\text{phys.sel.}}$ ) vertex and pileup cuts for pp and p-Pb collisions systems. Note that the quoted number of events after the selection is lower than the number of MB events reported in the text, because for the latter we also accounted for the MB events without reconstructed vertex. . .	32
3.2	List of $p_{\text{T,hard}}$ bins used in the MC simulation of the pp and p-Pb data. . .	33
3.3	List of POWHEG simulation settings with the default values highlighted. . .	46
3.4	Relative size of statistical and systematic uncertainties of b-jet cross section spectra measured in pp and p-Pb collisions and of the corresponding nuclear modification factor. The uncertainties are quoted for three $p_{\text{T,chjet}}$ bins. An additional uncertainty from the normalization by the integrated luminosity [128], [129] is quoted in the last row. . . . .	62
4.1	Layout of the upgraded ITS. The numbers are taken from [103]. . . . .	83
4.2	Sensor requirements for the IB and the OB. Taken from [103]. . . . .	84

# List of Figures

1.1	Energy scale dependence of strong coupling constant from measurements and theory. Taken from [26]. . . . .	6
1.2	Example of jets created in proton-proton collisions as reconstructed by the CMS experiment at the LHC. Taken from [31]. . . . .	7
1.3	Visualization of jet areas taken by $k_T$ (left) and anti- $k_T$ (right) jets. Here $y$ stands for rapidity and $\varphi$ for azimuthal angle. Taken from [30]. . . . .	8
1.4	The CT18 parton distribution functions obtained by the CTEQ global analysis at $Q = 2$ GeV and $Q = 100$ GeV for for quarks and gluons. Taken from [40]. . . . .	10
1.5	Ratios of nuclear and vacuum parton distribution functions at $Q^2 = 10$ GeV <sup>2</sup> . The nuclear PDFs correspond to EPPS16, EPS09, and DSSZ parametrizations. Here $R_V^{\text{Pb}}$ and $R_S^{\text{Pb}}$ correspond to valence quarks and sea quarks in lead. Taken from [41]. . . . .	10
1.6	The phase diagram of the strongly interacting matter. Different regions of the diagram can be reached by means of collisions of heavy ions. The figure presents regions of interest probed by different high-energy physics facilities. Taken from [50]. . . . .	12
1.7	Left: Average energy loss of a bottom quark in 0–7.5% central collisions at the LHC. Right: Average energy loss of a light quark in 0–5% central collisions at RHIC. Both are taken from [64]. . . . .	13
1.8	The nuclear modification factor $R_{AA}$ for different high- $p_T$ particles as a function of centrality. Different markers correspond to: charged hadrons (h), charmed mesons (D), charmonium (J/ $\Psi$ ), and beauty hadrons (B) [70].	15
1.9	Left: Transverse momentum dependence of the nuclear modification factor $R_{\text{pPb}}$ of charged particles measured in minimum-bias p–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV in comparison to data on the nuclear modification factor $R_{\text{PbPb}}$ from central Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76$ TeV. Taken from [80]. Right: $R_{\text{dAu}}(p_T)$ of charged hadrons for minimum bias and central d–Au collisions, and central Au–Au collisions measured by the STAR collaboration at 200 GeV/ $c$ . Taken from [81]. . . . .	17
1.10	Left: The PHENIX $R_{\text{dAu}}$ for $R = 0.3$ anti- $k_T$ jets measured in minimum bias d–Au collisions at $\sqrt{s_{\text{NN}}} = 200$ GeV. The data are compared with calculations which account for EPS09 nuclear PDFs but no energy loss and a model which assumes parton energy loss. Right: The PHENIX measurements of the nuclear modification factor in different centrality intervals. Taken from [85]. . . . .	18

1.11	$R_{pPb}$ as a function of $p_T$ for anti- $k_T$ $R = 0.4$ jets in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV in central (stars), mid-central (diamonds) and peripheral (crosses) events. Each panel corresponds to a different rapidity range. Positive rapidity corresponds to the proton-going direction and negative rapidity to the lead-going direction. Taken from [76]. . . . .	19
1.12	Top: ALICE $R_{pPb}$ of electrons from heavy-flavour hadron decays at 5 TeV and 8.16 TeV [89]. Bottom: The b jet $R_{pA}^{PYTHIA}$ as a function of jet $p_T$ together with a pQCD prediction [11]. . . . .	20
1.13	Left: Comparison of jet-particle $v_2$ and inclusive charged-particle $v_2$ in p–Pb and Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Figure taken from [7]. Right: $v_2$ of charged-particles as a function of particle $p_T$ as measured by ATLAS in p–Pb collisions at $\sqrt{s_{NN}} = 8.16$ TeV. Taken from [59]. . . . .	20
1.14	The ratios of the splitting-angle probability for $D^0$ -meson jets to inclusive jets in pp collisions at $\sqrt{s} = 13$ TeV. The data was compared with PYTHIA 8 and SHERPA simulations. The panels correspond to various $E_{Radiator}$ : the left panel is for $5 < E_{Radiator} < 10$ GeV (left panel), the middle panel is for $10 < E_{Radiator} < 20$ GeV and the right panel shows the ratio for $20 < E_{Radiator} < 35$ GeV. Taken from [32]. . . . .	21
1.15	$n_{SD}$ (right) distributions of prompt $D^0$ -tagged jets compared to inclusive jets for $15 < p_T^{jet\ ch} < 30$ GeV/c in pp collisions at $\sqrt{s} = 13$ TeV. Model/data ratios are shown in the bottom panels for PYTHIA 8 and POWHEG + PYTHIA 6 simulations. Taken from [96]. . . . .	22
1.16	Fraction of jets associated with b-quark fragmentation among inclusive jets as measured by ATLAS in pp collisions at $\sqrt{s} = 5.02$ TeV and CMS in pp collisions at $\sqrt{s} = 7$ TeV. The data measured by ATLAS are compared with a calculation by PYTHIA 8. Taken from [97]. . . . .	23
1.17	Basic diagrams which describe production of b quarks. Left: flavor creation process, middle: flavor excitation process, right: gluon splitting process. . . . .	24
1.18	Share of b jets produced by GSP (in cyan), FEX (in yellow) and FCR (in green) processes as predicted by HERWIG for pp collisions at $\sqrt{s} = 14$ TeV. The transverse momentum of produced b jets is denoted $p_t$ . Taken from [98]. . . . .	24
2.1	Scheme of the ALICE detector during the LHC Run 2. Taken from [100]. . . . .	26
2.2	Impact parameter resolution vs particle $p_T$ . Left: in the longitudinal direction. Right: in the transverse plane. Taken from [104], [105]. . . . .	27
2.3	Specific energy loss ( $dE/dx$ ) in the TPC vs. particle momentum in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV. The lines show the parametrizations of the expected mean energy loss. Taken from [13]. . . . .	29
2.4	Event reconstruction flow. Taken from [103] . . . . .	30
3.1	Distributions of primary event vertices along the beam axis after applying event selection and pile-up rejection criteria for pp and p–Pb data. . . . .	33
3.2	Simulated detector-level inclusive $p_T$ spectrum of charged-particle anti- $k_T$ jets with $R = 0.4$ obtained with and without outlier rejection in pp collisions at $\sqrt{s_{NN}} = 5$ TeV. Jet $p_T$ is corrected for the mean underlying event (UE) density. . . . .	34

3.3	Left: Track reconstruction efficiency for pp and p–Pb data as a function of track $p_T$ . Right: The mean relative resolution of $p_T^{-1}$ for hybrid tracks as a function of track $p_T$ in pp collisions. The red line represents a fit of the trend. . . . .	35
3.4	The sketch of the secondary vertex. Taken from [118] and changed. . . . .	36
3.5	2D distributions of $SL_{xy}$ and $\sigma_{SV}$ obtained with PYTHIA+EPOS simulation of p–Pb events. $p_{T, \text{chjet}}$ was set to be larger than 10 GeV/ $c$ . . . . .	37
3.6	Distributions of the tagging discriminators used in the SV method for the pp data set, $SL_{xy}$ (left), and $\sigma_{SV}$ (right) for b jets, c jets, and light-flavor jets as obtained from the PYTHIA 8 detector-level MC simulation. . . . .	37
3.7	Distributions of the tagging discriminators used in the SV method, $SL_{xy}$ (left) and $\sigma_{SV}$ (right), for the measured pp and p–Pb data sets. . . . .	38
3.8	Mis-tagging efficiency of light flavor and c jets versus b-jet selection efficiency when varying $SL_{xy}$ from 3 to 10 while keeping $\sigma_{SV}$ fixed at 0.04 cm for p–Pb data in $30 < p_{T, \text{chjet}} < 40$ GeV/ $c$ . The results obtained with the default selection of b-jet tagging cuts are analogous. . . . .	38
3.9	Distribution of $\sigma_{SV}$ in real p–Pb data and the corresponding PYTHIA+EPOS detector-level simulations for three different jet $p_T$ bins. Distributions are normalized by their integral. In this comparison, $SL_{xy}$ was set to be greater than 7. . . . .	39
3.10	Distribution of $SL_{xy}$ in real p–Pb data and the corresponding PYTHIA+EPOS detector-level simulations for three different jet $p_T$ bins. Distributions are normalized by their integral. In this comparison, $\sigma_{SV}$ was set to be less than 0.03 cm. . . . .	39
3.11	Example of raw inclusive $p_T$ spectra of b jets for $SL_{xy} > 7$ when varying the selection on $\sigma_{SV}$ . Left: pp data, right: p–Pb data. Size of bin is 1 GeV/ $c$ . . . . .	40
3.12	Tagging and mistagging efficiencies of beauty, charm, and light-flavor jets for the SV method in pp (solid markers) and p–Pb (open markers) collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The data are plotted as a function of jet transverse momentum. . . . .	41
3.13	Invariant mass distribution of the most displaced secondary vertex in jets with $20 < p_{T, \text{chjet}}^{\text{reco}} < 25$ GeV/ $c$ tagged with the default selection $SL_{xy} > 7$ and $\sigma_{SV} < 0.03$ cm, for pp and p–Pb collisions. The measured data (black points) are fitted with the detector-level MC templates corresponding to beauty, charm, and light-flavor jets to assess purity of the b-jet candidate sample. . . . .	44
3.14	Results of the closure test comparing b-jet sample purity for pp (left) and p–Pb (right) collisions estimated by the template fit method applied to the detector-level MC simulation and the true purity. The input detector-level MC data are based on PYTHIA (pp) and PYTHIA+EPOS (p–Pb) simulation. . . . .	45
3.15	The $\chi^2/N$ values as obtained for each POWHEG setting variation. The $x$ -axis labels provide the POWHEG parameters which were varied. The other parameters were considered to be fixed at their default values. The settings consistent with the data-driven purities are marked red. The settings failing the test are blue. . . . .	46

3.16	Purity of the b-jet candidates selected with the SV method when using the default tagging cuts for pp (left) and p–Pb (right) collisions. The purity was estimated with the data-driven template fit method (red points) and the POWHEG-simulation approach. The POWHEG scale variations accepted by the statistical analysis are colored green, the rejected ones are gray. . . . .	47
3.17	Rebinned raw spectrum of b jets as obtained for the pp (left) and p–Pb (right) data. . . . .	48
3.18	Top: Particle-level distribution of charged-particle anti- $k_T$ $R = 0.4$ b jets, which provides normalization for the instrumental response matrix. Bottom: Instrumental response matrix. . . . .	50
3.19	Top: Projections of the response matrix on the $p_{T,\text{chjet}}^{\text{det}}$ axis for the pp and p–Pb data in different $p_{T,\text{chjet}}^{\text{part}}$ bins. Bottom: Projections of the response matrix on the $p_{T,\text{chjet}}^{\text{det}}$ axis for the b-jet enhanced and the default simulated pp data in different $p_{T,\text{chjet}}^{\text{part}}$ bins. . . . .	51
3.20	Top: $\delta p_T$ distribution as obtained by the embedding and by the random cone method. Middle: Comparison of the $\delta p_T$ distributions obtained for different ranges of embedded particle momentum. Bottom: $p_T$ matrix obtained from embedded templates. Bin size is $1 \text{ GeV}/c \times 1 \text{ GeV}/c$ . . . . .	53
3.21	Top: Combined response matrix for b jets as obtained from Eq. (3.16). The matrix has a bin size of $1 \text{ GeV}/c \times 1 \text{ GeV}/c$ . Bottom: The same matrix after rebinning with (3.14) and (3.15). . . . .	54
3.22	Results of the closure test when unfolding PYTHIA detector-level b-jet spectrum in pp collisions. Upper left: comparison of the raw, unfolded, refolded, and prior b-jet spectrum. Upper right: $d_i$ vector components. The chosen optimal regularization parameter was 5. Bottom left: Ratio of the refolded unfolding solutions to the input raw spectrum corrected for SV tagging efficiency and purity. Bottom right: Ratio of the unfolded spectrum to the true spectrum. In the bottom panels, different markers correspond to different choices of the regularization parameter $i$ . . . . .	56
3.23	Results of the p–Pb SVD unfolding closure test using PYTHIA data. The optimal choice of the regularization parameter was 5. Meaning of the individual panels is analogous to Figure 3.22. . . . .	57
3.24	QA plots for the SVD unfolding of the measured pp b-jet spectrum. Meaning of the individual panels is analogous to Figure 3.22. The right bottom panel shows a ratio of the unfolded solutions to the prior spectrum. The primary analysis spectrum was obtained with $i = 4$ . . . . .	59
3.25	QA plots for the SVD unfolding of the measured p–Pb b-jet spectrum. Meaning of the individual panels is analogous to Figure 3.22. The right bottom panel shows a ratio of the unfolded solutions to the prior spectrum. The primary analysis spectrum was obtained with $i = 5$ . . . . .	60
3.26	Ratios of the fully-corrected b-jet spectra obtained with the varied and the PA SV tagging setting for pp (left) and p–Pb (right) data. The top panels show the results obtained for $\sigma_{\text{SV}}$ variations. The bottom panels show analogous plots for $SL_{xy}$ variations. Statistical errors on the data points are not shown since the numerator and denominator are highly correlated. . . . .	63
3.27	The ratio of fully corrected b-jet spectra obtained with the SV tagging corrections based on various POWHEG settings to the PA spectrum for pp (left) and p–Pb (right) collisions. . . . .	64

3.28	Ratio of the fully corrected spectrum obtained with 4% lower tracking efficiency over the primary analysis spectrum for the pp (left) and p-Pb (right) data. . . . .	65
3.29	Ratio of the fully corrected b-jet spectra obtained with PA smearing and with the modified $p_T$ smearing which accounts differences in track $p_T$ spectra as a function of azimuth. The left panel corresponds to pp data, the right is the p-Pb analysis. . . . .	66
3.30	Ratio of the fully corrected b-jet spectrum obtained using the varied unfolding setting and the PA setting. Left: pp data. Right: p-Pb data. . . .	68
3.31	Ratio of the fully corrected b-jet spectrum obtained with the varied raw spectrum binning over the primary analysis binning for pp (left) and p-Pb (right) data. . . . .	68
3.32	Ratio of the fully corrected b-jet spectrum obtained with varied POWHEG prior spectrum and the PA spectrum for pp (left) and p-Pb (right) data. .	68
3.33	Ratios of fully corrected b-jet spectra obtained with varied $\delta p_T$ estimator and combined response matrix truncation to the PA spectrum. “Full ReMx” label denotes a response matrix which has the same range as the PA matrix. . . . .	69
3.34	DCA distributions of charged hybrid tracks in pp (left) and p-Pb data (right). The real data are fitted with a sum of two simulated templates corresponding to the DCA distribution of physical primaries and the DCA distribution of secondary tracks. . . . .	70
3.35	Impact of variation in secondary particle admixture on the fully corrected b-jet spectra. Left: pp. Right: p-Pb. See text for more details. . . . .	71
3.36	Ratio of the invariant mass distributions of the tagged strange particle SVs and of all light-flavor tagged SVs. . . . .	71
3.37	Top panel: the fully corrected b-jet cross section spectrum from pp and p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Gray boxes represent the total systematic uncertainties. The bottom panel shows the individual sources of systematic uncertainties. . . . .	72
3.38	The nuclear modification factor $R_{pPb}$ of the inclusive charged-particle anti- $k_T$ $R = 0.4$ b jets as a function of $p_T$ together with individual sources of systematic uncertainties. Gray boxes represent the total systematic uncertainties. . . . .	73
3.39	Comparison of the $p_T$ differential production cross section of charged-particle b jets measured in pp and p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV using the IP and SV methods. Systematic and statistical uncertainties are shown as boxes and error bars, respectively. . . . .	74
3.40	The nuclear modification factor $R_{pPb}$ of the inclusive charged-particle anti- $k_T$ $R = 0.4$ b jets as a function of $p_T$ from the IP and SV method. . . . .	74
3.41	Top panels: The combined differential production cross-section of charged-particle anti- $k_T$ $R = 0.4$ b jets in pp and p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The data are compared with a NLO pQCD prediction by the POWHEG dijet tune with PYTHIA 8 fragmentation [77, 78]. Additional normalization uncertainty due to luminosity is quoted separately. Bottom panels: ratio of the theoretical calculations to the data. . . . .	76

3.42	The fraction of charged-particle anti- $k_T$ $R = 0.4$ b jets measured in pp and p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, compared with POWHEG NLO pQCD calculations with PYTHIA 8 fragmentation. See text for more details.	77
3.43	The combined nuclear modification factor $R_{pPb}^{b\text{-jet}}$ compared with the calculation by the POWHEG dijet tune with the PYTHIA 8 fragmentation. Systematic and statistical uncertainties are shown as boxes and error bars, respectively.	77
3.44	Left: The nuclear modification factor $R_{pPb}$ for charged-particle b jets measured by the ALICE experiment compared with the b-jet measurement from the CMS experiment [11]. Right: The nuclear modification factor of b jets compared to that of untagged jets from Ref. [6].	78
4.1	Left: Impact parameter resolutions for the old ITS (Pb-Pb 2010 data) and the upgraded ITS. Taken from [143]. Right: Stand-alone tracking efficiency for the old ITS and the upgraded ITS. Taken from [143].	81
4.2	Scheme of the Upgraded ITS system. Taken from [145].	82
4.3	Structure of an IB (left) and OB (right) stave of the upgraded ITS. Taken from [146].	82
4.4	Cross section of a MAPS made by the TowerJazz 0.18 $\mu\text{m}$ technology, taken from [143].	83
4.5	The Computing architecture of ALICE in Run 3. Taken from [151].	85
5.1	Scheme of the ITS QC system architecture in Run 3. Blocks with turquoise color correspond to software developed especially for ITS. Blocks with Calibration and Noisy Pixel are depicted in distinct colors since they are not parts of the physics data-taking workflow.	88
5.2	Example of an activation function of a single ALPIDE pixel. $Q_{\text{THR}}$ corresponds to a charge which activates the pixel with a 50% probability. Taken from [156].	92
5.3	Examples of MOs provided by QC Threshold Task for the non-tuned detector. Left: The average threshold value in each ALPIDE sensor in Layer 0. The $z$ -axis corresponds to the mean threshold activation charge in DAC units calculated from all pixels of the given ALPIDE sensor. The lower $Q_{\text{THR}}$ value of the Stave 8 can be explained by different VCASN values for this stave compared to others. Right: The number of dead pixels in each ALPIDE sensor of Layer 0.	93
5.4	Left: Distribution of the threshold charge in pixels of the Stave 10 of Layer 0. Right: Distribution of the average threshold value in the chips of the Stave 10 in Layer 0.	94

- 5.5 Examples of MOs produced by the ITS Cluster Task as they are seen by an ALICE QC shifter. Top left: maximum value of cluster occupancy per each ITS stave. Each triangle corresponds to a separate stave. Data were collected during the pp pilot beam with  $\sqrt{s} = 900$  GeV. Top right: average cluster occupancy for each chip in Layer 0. The central chips demonstrate higher occupancy since they are closer to the primary vertex located around  $z = 0$  cm. Bottom left: average cluster size for each chip in Layer 0. Chips at the borders of the stave have larger cluster sizes because of a larger number of inclined tracks in this region. Bottom right: cluster size distribution for all clusters in Layer 0. The last three plots were obtained for the run 526532 (pp collisions at  $\sqrt{s} = 13.6$  TeV, interaction rate 500 kHz, framing rate 202 kHz). . . . . 95
- 5.6 Example of MOs produced by the ITS Track Task as they were seen by the central ALICE QC shifter during the run 526532 (pp collisions at  $\sqrt{s} = 13.6$  TeV, interaction rate 500 kHz, framing rate 202 kHz). Top left: Distribution of tracks as a function of azimuthal angle and pseudorapidity. The histogram is normalized per the number of vertices. Top right: Distribution of the number of clusters per track reconstructed in the ITS. The majority ITS tracks have clusters in all 7 layers. The second peak at 4 clusters corresponds to the tracks which have hits only in the OB, while the other contributors come from tracks having hits in OB + IB. Bottom left: Distribution of ITS-tracks primary vertex coordinates in the plane perpendicular to the beam direction. Bottom right: Vertex distribution in the direction of the beam. . . . . 96
- 5.7 Example of MOs produced by the ITS Calibration task for monitoring of ITHR (left), VCASN (middle), and Threshold (right) for the ITS Inner Barrel. Each bin corresponds to one ALPIDE sensor. Values on the  $z$ -axis of the ITHR and VCASN plots are expressed in DAC units, while the  $z$ -axis of the threshold plot is in the electron charges. The solid lines provide the separation between the ITS layers. . . . . 97
- 5.8 Top panels: ITS standalone tracking efficiency as a function of generated particle  $p_T$  (left) and  $\eta$  (right). Bottom panels: fake-track rate as a function of generated particle  $p_T$  and  $\eta$ . The confidence intervals were handled by the TEfficiency class of the CERN ROOT framework [158] using the Clopper–Pearson method [159]. The distributions were obtained using the ALICE MC production cycle anchored to a real data run of pp collisions at  $\sqrt{s} = 13.6$  TeV. . . . . 99
- 5.9 Time trend of average cluster occupancy per layer for physical runs with pp collisions at  $\sqrt{s} = 13.6$  TeV. Each point corresponds to a single GOOD quality run. All runs are in chronological order. The  $x$ -axis gives dates corresponding to the beginning of data collection campaigns. The interaction rate was estimated with the FT0 Cherenkov array detector. . . . . 100
- 5.10 Average cluster occupancy as a function of IR for different ITS Layers in real data and MC simulation. The MC predictions are based on Pythia 8 + GEANT 3 simulation performed by I. Ravasenga. . . . . 101
- 5.11 Time trend of average cluster size per layer for physical pp collisions at  $\sqrt{s} = 13.6$  TeV. The figure is analogous to Fig. 5.9. . . . . 102



5.12	Average cluster size as a function of IR for different ITS Layers in real data and MC simulation. The MC predictions are based on Pythia 8 + GEANT 3 simulation performed by I. Ravasenga. . . . .	102
5.13	Time trend of the position of the primary seed vertex which is used in further steps of global tracking for pp collisions at $\sqrt{s} = 13.6$ TeV. The figure is analogous to Fig. 5.9. . . . .	103
5.14	Time trend of an average number of clusters per track for physical pp collisions at $\sqrt{s} = 13.6$ TeV. The figure is analogous to Fig. 5.9. . . . .	103

# List of Acronyms

**CNM** cold nuclear matter. vi, vii, 1, 16, 78, 105

**ITS** Inner Tracking System. ix, xv–xvii, xx, 2, 3, 32, 79–86, 88–91, 93–103, 105

**LHC** Large Hadron Collider. x, xi, 2, 7, 13, 14, 16, 23, 25, 26, 29, 85, 99, 101

**NLO** next to leading order. xiv, xv, 5, 23, 75–78, 104, 105

**PDF** parton distribution function. x, 1, 9, 10, 16, 18, 42, 76

**pQCD** perturbative QCD. xi, xiv, xv, 11, 20, 21, 75–78, 105

**QCD** Quantum Chromodynamics. xix, 2, 4–7, 9, 11, 12, 21

**QGP** quark-gluon plasma. 1, 11–13, 16, 18, 25, 79

**RHIC** Relativistic Heavy Ion Collider. x, 12, 13, 16

**SV** Secondary Vertex. xii–xiv, xx, 31, 36–41, 43, 47, 52, 55, 56, 61, 63, 64, 70, 74, 75

**TPC** Time Projection Chamber. xi, 25–30, 32, 35, 80, 85, 98

# Contents

<b>Abstract</b>	<b>vii</b>
<b>Acknowledgements</b>	<b>viii</b>
<b>List of Tables</b>	<b>ix</b>
<b>List of Figures</b>	<b>x</b>
<b>List of Acronyms</b>	<b>xviii</b>
<b>Introduction</b>	<b>1</b>
<b>1 State of the art</b>	<b>4</b>
1.1 Standard Model and QCD Lagrangian . . . . .	4
1.2 Jet algorithms . . . . .	6
1.3 Factorization theorem . . . . .	9
1.4 QCD matter phase diagram . . . . .	11
1.5 Parton energy loss in proton–nucleus collisions . . . . .	13
1.6 Dead-cone effect . . . . .	14
1.7 Experimental overview of small collision systems . . . . .	15
<b>2 ALICE experiment</b>	<b>25</b>
2.1 Detector system overview . . . . .	26
2.1.1 ITS . . . . .	26
2.1.2 TPC . . . . .	27
2.1.3 VZERO . . . . .	28
2.2 Trigger System . . . . .	29
2.3 Tracking Procedure . . . . .	29
<b>3 Analysis of b jets</b>	<b>31</b>
3.1 Data sets and Monte Carlo productions . . . . .	31
3.2 Jet reconstruction . . . . .	34
3.3 b-jet tagging based on secondary vertex reconstruction . . . . .	36
3.4 Correction of the b-jet spectra on tagging performance . . . . .	40
3.4.1 b-jet tagging efficiency . . . . .	40
3.4.2 b-jet tagging purity . . . . .	41
3.5 Unfolding . . . . .	47
3.6 Unfolding of b-jet spectra . . . . .	48
3.7 MC closure tests of the unfolding procedure . . . . .	55
3.8 Unfolding of the measured b-jet spectrum . . . . .	58

3.9	$p_T$ differential b-jet cross section spectrum from the SV method . . . . .	61
3.10	Systematic Uncertainties . . . . .	61
3.10.1	Tagging of b jets . . . . .	63
3.10.2	Purity of b-jet candidates . . . . .	64
3.10.3	Tracking efficiency . . . . .	65
3.10.4	Track $p_T$ resolution . . . . .	66
3.10.5	Unfolding systematics . . . . .	67
3.10.6	$\delta p_T$ estimator and combined response matrix truncation . . . . .	69
3.10.7	Fraction of secondary particles in MC . . . . .	69
3.10.8	Summary of systematic uncertainty calculations . . . . .	72
3.11	Results . . . . .	74
3.11.1	Combining the results of the IP and SV method . . . . .	75
3.11.2	b-jet cross-section . . . . .	75
3.11.3	Fraction of the b-jets . . . . .	76
3.11.4	The b-jet nuclear modification factor $R_{pPb}^{b\text{-jet}}$ . . . . .	77
3.12	Summary . . . . .	78
<b>4</b>	<b>ALICE ITS Upgrade</b>	<b>79</b>
4.1	ITS Upgrade design objectives . . . . .	80
4.2	Layout of the Upgraded ITS . . . . .	81
4.3	MAPS for the ALICE ITS Upgrade . . . . .	82
4.4	ALICE computing model . . . . .	84
4.4.1	Calibration . . . . .	86
4.4.2	Quality Control software . . . . .	86
4.4.3	Condition and Calibration Data Base . . . . .	87
<b>5</b>	<b>ITS QC system in Run 3</b>	<b>88</b>
5.1	ITS QC system architecture . . . . .	88
5.1.1	Overview of the ITS QC tasks . . . . .	90
5.2	Development of the ITS QC tasks . . . . .	91
5.2.1	ITS QC Threshold Task for the commissioning stage . . . . .	91
5.2.2	ITS QC Cluster Task . . . . .	94
5.2.3	ITS QC Track Task . . . . .	95
5.2.4	ITS QC Calibration Task . . . . .	97
5.2.5	ITS Monte Carlo Tracking QA task . . . . .	97
5.3	ITS2 performance at the beginning of Run 3 . . . . .	99
5.3.1	Performance of the Clustering . . . . .	100
5.3.2	Performance of the ITS Tracking . . . . .	102
<b>6</b>	<b>Summary</b>	<b>104</b>
<b>A</b>	<b>List of publications</b>	<b>106</b>
<b>B</b>	<b>List of conferences</b>	<b>108</b>
	<b>Bibliography</b>	<b>120</b>

# Introduction

Beauty and charm quarks can be produced in ultra-relativistic collisions with a large square of transferred four-momentum,  $Q^2$ . Quarks having high initial virtuality reduce their energy by gluon emissions and form a spray of collimated hadrons called a jet. A theoretical description of such processes in proton-proton collisions can be successfully performed using a factorization theorem. According to that theorem, a collision process can be expressed as a convolution of parton distribution functions (PDF), a short-distance parton-level cross-section, and a fragmentation function.

The factorization can also be used for more complex collision systems such as proton-nucleus. In this case, the influence of various cold nuclear matter (CNM) effects is incorporated in the so-called nuclear PDF while the cross-section and hadronization terms remain unchanged [1]. The CNM effects encompass the differences between parton momentum distributions in a bound and a free nucleon. The perturbative approach cannot describe these effects, therefore their magnitude is usually assessed with experimental methods and quantified using the nuclear modification factor, which is defined as the ratio of the particle or jet yield measured in proton-nucleus collisions and the expected yield that would be obtained from a superposition of the corresponding number of independent pp collisions [2].

It is still unclear whether jet production in proton-nucleus collisions at LHC energies is affected by hot nuclear matter effects emerging from an interaction of the partonic shower with quark-gluon plasma (QGP) [2]. Such an interaction causes the jet quenching effect, which leads to energy losses of produced partons, the redistribution of energy inside a jet, and suppression in the measured nuclear modification factor of high-transverse momentum ( $p_T$ ) hadrons and jets [3]. While experiments are still searching for signatures of jet quenching in proton-nucleus collisions [4]–[6], such small collision systems exhibit other signatures of the QGP formation, for instance, collective phenomena [7] and strangeness enhancement [8].

Jets associated with open heavy-flavor hadrons are likely to stem from heavy-flavor quark fragmentation and can be thus used to study mass and flavor dependence of cold and hot nuclear matter effects [9]. Vacuum fragmentation of heavy-flavor quarks is known

to be affected by the so-called dead-cone effect, which suppresses the emission of gluons to a forward cone [10]. The cone angle is proportional to the quark mass divided by its energy, therefore the effect should be more pronounced for slow heavy quarks. However, it is unknown whether the presence of a QCD medium modifies the dead cone radiation pattern. These and related problems make the b jets intriguing objects of interest.

Recently, the CMS collaboration reported the nuclear modification factor of b jets in p–Pb collisions [11]. The data spans over the transverse momentum range 50 – 400 GeV/ $c$ , where one can expect only marginal susceptibility to the dead cone effect [9]. This invites the ALICE detector [12], [13] with its excellent tracking capabilities to push the lower  $p_T$  limit of this measurement down.

The measurements of low- $p_T$  open-heavy-flavor hadrons are also one of the key targets of the ALICE physics program in Run 3 and Run 4 [14], [15]. This program primarily motivated the recent major upgrade of the ALICE Inner Tracking System (ITS) detector. The upgraded ITS is an entirely new, low material budget, seven-layer silicon tracker based on monolithic active pixel sensors ALPIDE [16]. The project required extensive work on the hardware and software. Thus, in addition to performing the b-jet analysis of LHC Run 2 data, I participated in many tasks related to the ITS upgrade, such as the characterization of ALPIDE sensors, simulation benchmarks, and, in particular, the development of the Quality Control software for the ITS in the ALICE O<sup>2</sup> framework [17].

The list of goals that will be addressed in this thesis is presented below:

1. The analysis of inclusive transverse momentum spectra of charged-particle b jets in p–Pb and pp collisions at  $\sqrt{s_{NN}} = 5.02$  TeV based on the data collected by ALICE in 2016 and 2017, respectively. The evaluation of the nuclear modification factor and the fraction of inclusive b jets among untagged jets.
2. The development of the Quality Control software for the upgraded ITS.

The thesis structure is as follows: the first chapter includes basic facts about QCD, jet reconstruction algorithms, and collinear factorization. Furthermore, it reviews recent untagged and heavy-flavor jet measurements in small collision systems to put the b-jet analysis in a broader context. The second chapter describes the ALICE detector during the LHC Run 2 and discusses its capabilities for heavy-flavor measurements. The third chapter is devoted to the analysis of charged-particle b jets in p–Pb and pp collisions at  $\sqrt{s_{NN}} = 5.02$  TeV and provides a discussion of the results. The fourth chapter deals with the upgrade project of the Inner Tracking System. The subsequent chapter discusses my contributions to this project, particularly the Quality Control software development. The thesis is concluded with a summary of the achieved results.

## Statement of author's contribution

The author's contributions to the material presented in this thesis can be summarized as follows:

1. I have performed the analysis of b jets produced in p–Pb and pp collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV, which utilized the SV method to tag b jets. I am the principal author of the corresponding paper by the ALICE collaboration [18].
2. I have developed several ITS QC tasks (Cluster, Track, Threshold, and Calibration Tasks) and served as the coordinator of ITS Data Quality activities during Run 3 in the second half of 2022.

Besides that, I was also involved in other projects related to the ITS2 upgrade, such as the benchmarking of ITS simulation [19] and testing radiation hardness of sensors [20] for ITS2. Currently, I am also involved in the subsequent upgrade project of the inner barrel of the ITS detector, called ITS3. In this project, I contributed to the development of the software used for sensor characterization, and I participated in the test beams at DESY and CERN [21] and in the studies of radiation hardness of sensor prototypes that were carried out at the Nuclear Physics Institute of the Czech Academy of Sciences. Finally, I was also co-supervising the bachelor thesis “Characterization of monolithic active pixel sensors in ALICE” of Nikita Bikmetov.

The list of publications and the list of conferences are provided in Appendix A and B, respectively.

# Chapter 1

## State of the art

### 1.1 Standard Model and QCD Lagrangian

Quantum Chromodynamics (QCD) is the fundamental quantum-field theory of strong interaction. QCD describes interactions of particles that possess a charge of the strong interaction – color. Such particles are quarks and gluons with spins 1/2 and 1, respectively. Each quark flavor thus exists in three color states denoted as red, blue, and green. Other basic properties of quarks are shown in Table 1.1. There are eight different color states of massless gluons, which combine color with an anti-color. The strong interaction between colored particles is mediated by gluon exchange. The Lagrangian of the QCD arises from a Lagrangian of a free-fermion field by requiring its local gauge invariance with respect to the  $SU(3)_c$  symmetry, where the subscript  $c$  stands for color,

$$\mathcal{L}_{QCD} = \sum_f \bar{\psi}_{fi} (i\gamma^\mu D_{\mu ij} - m_f \delta_{ij}) \psi_{fj} - \frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu}. \quad (1.1)$$

Here  $\psi_{fi}$  is a local quantum field operator associated to a quark field with a flavor  $f$ , color  $i$ , and a bare mass  $m_f$ ,  $\gamma^\mu$  are the four Dirac matrices,  $D_{\mu ij}$  is the covariant derivative,  $F^a$  is the gluon-field tensor, and  $\delta_{ij}$  is a Kronecker delta. Greek letters denote Lorentz indices. Einstein's summation convention is assumed. The covariant derivative can be written as:

Name	up	down	charm	strange	top	bottom
Symbol	u	d	c	s	t	b
Mass (MeV/ $c^2$ )	$2.16^{+0.49}_{-0.26}$	$4.67^{+0.48}_{-0.17}$	$1270^{+2}_{-2}$	$93^{+13}_{-5}$	$172760^{+300}_{-300}$	$4180^{+3}_{-2}$
Electrical Charge ( $e$ )	$+\frac{2}{3}$	$-\frac{1}{3}$	$+\frac{2}{3}$	$-\frac{1}{3}$	$+\frac{2}{3}$	$-\frac{1}{3}$

Table 1.1: Basic quark properties [22].



$$D_{\mu ij} = \partial_\mu \delta_{ij} - ig_s t_{ij}^a A_\mu^a, \quad (1.2)$$

where  $g_s$  is the QCD coupling constant,  $t^a$  are the eight  $3 \times 3$  generators of the  $SU(3)_c$  color group and  $A^a$  are eight gluon fields. The gluon field tensor is defined as:

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g_s f^{abc} A_\mu^b A_\nu^c. \quad (1.3)$$

Here  $f^{abc}$  are the structure constants of the  $SU(3)$  group [22], defined through the commutation relations as  $[t^a; t^b] = i f^{abc} t^c$ .

The last term in (1.3) allows for a direct coupling of a gluon with other gluons. This property is not present in Quantum Electrodynamics (QED) [22] because photons do not possess any charge and hence cannot directly interact with each other. This self-interaction in QCD reflects the non-abelian nature of QCD, due to the nonzero commutation relations between the generators of  $SU(3)_c$ .

The fine structure constant of the strong interaction,  $\alpha_s = g_s^2/4\pi$ , becomes in next to leading order (NLO) a function of the four-momentum transferred  $Q^2$ :

$$\alpha_s(Q^2) = \frac{\alpha_s(\Lambda^2)}{1 + \alpha_s(\Lambda^2) \frac{33-2N_f}{12\pi} \ln\left(\frac{Q^2}{\Lambda^2}\right)}, \quad (1.4)$$

where  $\Lambda$  stands for the scale at which the fine-structure constant should be determined experimentally, and  $N_f$  is equal to the number of quark flavors. Figure 1.1 shows  $\alpha_s$  as a function of the scale as measured in experiments and as determined from the QCD. The running of the  $\alpha_s$  results in two limit scenarios known as the *color confinement* [23] and *asymptotic freedom*. Asymptotic freedom is a property of the QCD theory when the strong coupling constant tends to zero for high values of  $Q^2$ . In the limit of vanishing  $\alpha_s$ , quarks and gluons behave like non-interacting particles. On the other hand, when  $Q^2$  tends to zero,  $\alpha_s$  rapidly grows, which prevents isolating individual quarks and gluons. Therefore, they will be bound into objects with a neutral color charge – hadrons. Thereby the color confinement is established. In this limit, QCD processes cannot be calculated with perturbative expansions. Static properties of hadrons, such as their mass spectrum, can be nevertheless computed using non-perturbative lattice QCD [24], [25].

Another distinction between QCD and QED can be found in the configuration of field lines, which appear to be different due to gluon self-interaction and color confinement. The electrostatic field reaches an infinite distance, whereas the color field is constrained to a flux tube interconnecting two color charges. The potential energy between a heavy-flavor quark – anti-quark pair is known to be described by the Cornell potential [27]:

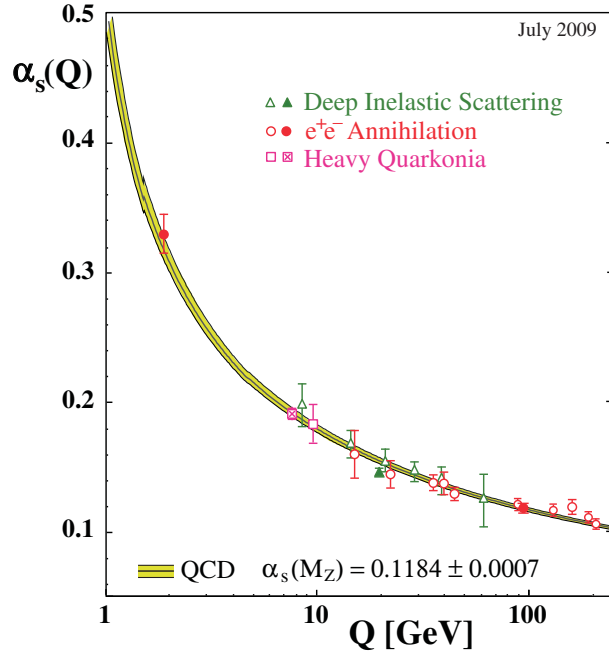


Figure 1.1: Energy scale dependence of strong coupling constant from measurements and theory. Taken from [26].

$$V(r) = -\frac{a}{r} + br, \quad (1.5)$$

where  $a$  and  $b$  are model parameters and  $r$  is the distance between both color charges. The linear term in this potential reflects color confinement. If one tries to separate two color-charged objects, it will increase the potential energy since the coupling strength between the two interacting quarks increases with their distance. At some point, it will become more energetically preferable to create a new  $q\bar{q}$  pair from the QCD vacuum. Hence, the color strings that interconnect quarks and gluons that underwent a hard scattering process will stretch until they reach a point of break up when they produce a shower of hadrons known as a jet. This picture of the fragmentation process is behind the Lund fragmentation model [28] implemented in the PYTHIA 8 event generator [29].

## 1.2 Jet algorithms

High energy partons produced in the hard scattering processes rapidly fragment, hadronize and create a collimated spray of particles, which detector systems can observe in the form of hadrons or directed energy flow. Fig. 1.2 illustrates an event display with jets. Consistent implementation of the jet concept in experiments and theory requires that jets are defined through algorithms [30]. A jet algorithm provides a set of rules to group particles such that the kinematic properties of the resulting jet would relate to the corresponding

properties of the initial parton. Modern jet sequential recombination algorithms measure the distance between two particles/protojets, decide whether they should be combined, and provide a rule on how their four-momenta should be combined, the so-called recombination scheme. These jet algorithms exhibit the property of infrared and collinear (IRC) safety, which guarantees that the jet definition is consistent between experiment and QCD theory calculations. The IRC safety means that the configuration of final state jets is insensitive to low energy particle emissions and collinear splittings [30].

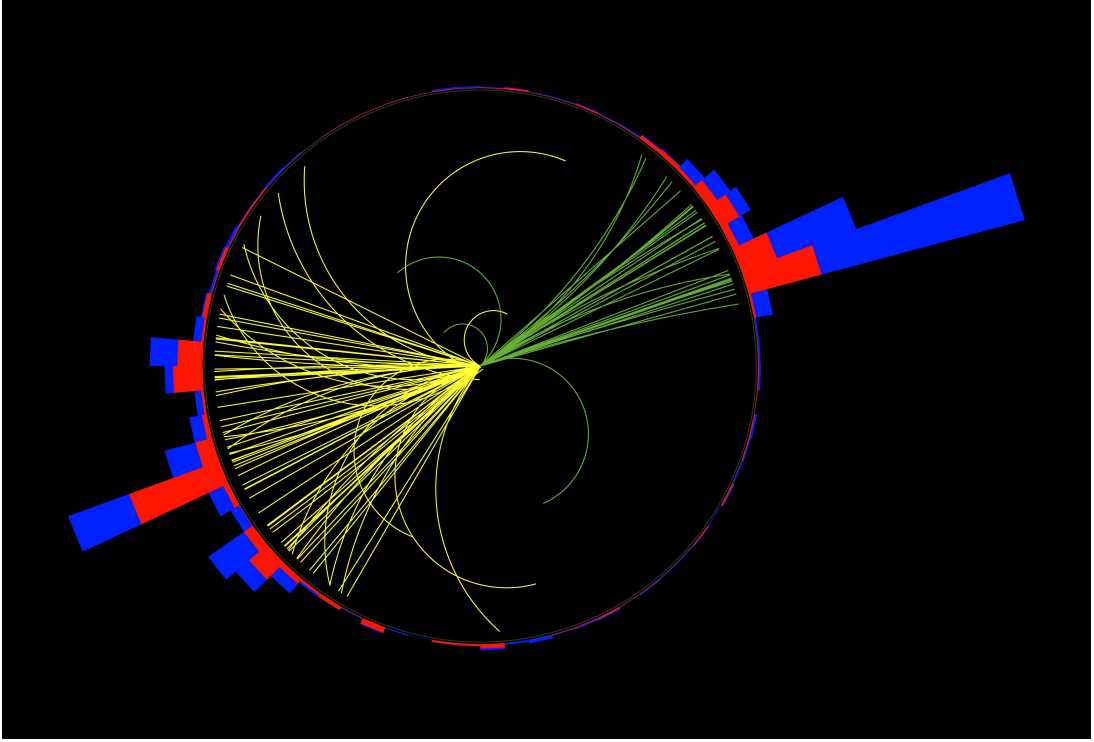


Figure 1.2: Example of jets created in proton-proton collisions as reconstructed by the CMS experiment at the LHC. Taken from [31].

A sequential recombination algorithm proceeds as follows [30]:

1. For each particle  $i$  calculate the distance to the beam ( $d_{iB}$ ) and distance to each particle  $j$  ( $d_{ij}$ ):

$$d_{ij} = \min(p_{T,i}^{2p}, p_{T,j}^{2p}) \frac{\Delta_{ij}^2}{R^2}, \quad \text{where } \Delta_{ij}^2 = (y_i - y_j)^2 + (\varphi_i - \varphi_j)^2, \quad (1.6)$$

$$d_{iB} = p_{T,i}^{2p}, \quad (1.7)$$

where  $p_T$ ,  $y_i$ , and  $\varphi_i$  are the transverse momentum, rapidity, and azimuthal angle of particle  $i$ , further,  $R$  is a resolution parameter, which sets the jet size, and  $p$  is a parameter defining a type of the algorithm.

2. Find minimum of  $d_{iB}$  and  $d_{ij}$ .

3. If  $d_{ij}$  is the minimum, create a new particle as a combination of  $i$  and  $j$ , remove the particles  $i$  and  $j$  from the list of particles and proceed to step 1.
4. Otherwise,  $d_{iB}$  is the minimum, and  $i$  is a final state jet, remove it from the list of particles and proceed to step 1.
5. Continue with the steps until all particles are assigned to final state jets.

Sequential algorithms differ by a value of the parameter  $p$  in Eq. (1.6). Accordingly, the three most popular sequential algorithms are:

- $k_T$  algorithm ( $p = 1$ ) begins the clusterization procedure from soft particles. This algorithm property is often used to assess underlying event density in heavy-ion collisions.
- Cambridge/Aachen algorithm ( $p = 0$ ) accounts for the angular distance between two particles and ignores the magnitude of their transverse momentum. This algorithm is primarily used in studies involving the angular structure of jets, such as measurements of a dead-cone effect [32] or jet substructure [33], since it respects the angular ordering of the partonic shower.
- Anti- $k_T$  algorithm ( $p = -1$ ) preferably clusters high transverse momentum particles. This algorithm is, therefore, less susceptible to the underlying event. Reconstructed high- $p_T$  anti- $k_T$  jets usually have regular conical shapes in the  $y - \varphi$  plane, see Figure 1.3.

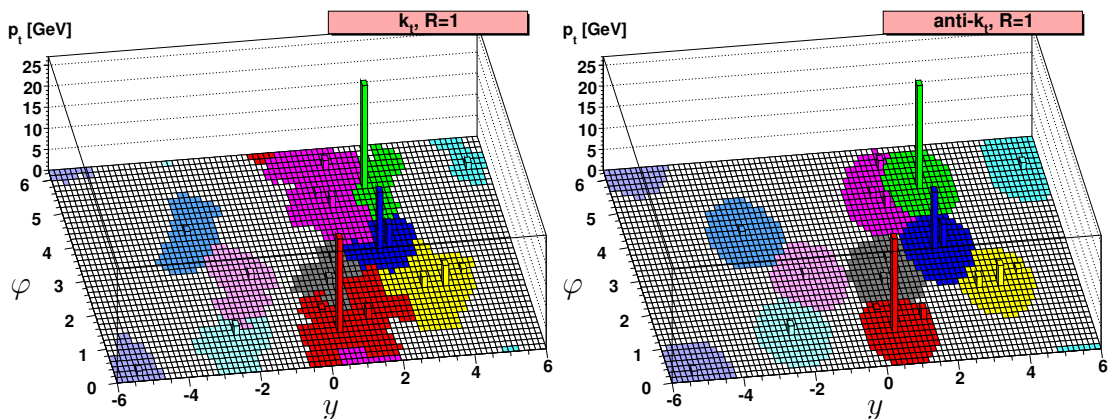


Figure 1.3: Visualization of jet areas taken by  $k_T$  (left) and anti- $k_T$  (right) jets. Here  $y$  stands for rapidity and  $\varphi$  for azimuthal angle. Taken from [30].

The above-mentioned jet algorithms are implemented in the FastJet [34] package. FastJet also allows measuring jet area in the  $y - \varphi$  plane by injection of a large number

of ultra-soft ghost particles. Due to the IRC safety of the sequential recombination algorithms, these particles do not change the configuration of final state jets. The area is proportional to the number of ghost particles associated with the jet. The concept of jet area is helpful when correcting the jet energy scale for the contribution of the underlying event, i.e., particles originating from processes unrelated to the given hard scattering producing the jet of interest. In the first-order jet energy scale can be corrected by means of the formula:

$$p_{\text{T, jet}}^{\text{reco}} = p_{\text{T, jet}}^{\text{raw}} - \rho \times A_{\text{jet}}, \quad (1.8)$$

where  $p_{\text{T, jet}}^{\text{raw}}$  is the original uncorrected jet transverse momentum,  $A_{\text{jet}}$  is a jet area, and  $\rho$  is the mean density of the underlying events estimated on the event by event basis. One of the classical methods to calculate  $\rho$  is based on jets reconstructed using the  $k_{\text{T}}$  algorithm. In each event, the  $k_{\text{T}}$  jets are sorted by the value of a ratio  $p_{\text{T, jet}}^{\text{raw}}/A_{\text{jet}}$  and the median value is taken as the estimate of  $\rho$  [35].

### 1.3 Factorization theorem

The QCD factorization theorem [36] was introduced to interconnect processes that occur at the perturbative-QCD level with the production of hadrons. According to the factorization theorem, the cross-section for a proton-proton collision producing a hadron  $h$  in a final state can be written with the following equation:

$$\begin{aligned} d\sigma^{pp \rightarrow h+X} = \sum_{abcd} \int dx_a \int dx_b \int dz f_{a/p}(x_a, Q^2, \mu_f) f_{b/p}(x_b, Q^2, \mu_f) \times \\ \times d\sigma^{a+b \rightarrow c+d}(Q^2, \mu_f, \mu_R) D_{c/h}(z, Q^2, \text{FS}), \end{aligned} \quad (1.9)$$

where  $\sum_{abcd}$  includes all incoming and outgoing parton species, and  $\mu_R/\mu_f$  denotes the renormalization/factorization scale introduced in the renormalization of ultraviolet divergences. The factorization divides the whole collision process into three stages, which are assumed to be well separated in time:

1. Initial Evolution - is described by Parton Distribution Functions (PDF)  $f_{a/p}(x_a, Q^2, \mu_f)$  which gives the probability to find a parton of a type  $a$  within the proton, carrying a fraction  $x_a$  of the proton's momentum [37]. Parton distribution functions are determined by interactions of partons at large distances and are thus incalculable in perturbative QCD. The form of the PDF depends on the factorization scale  $\mu_f$  meaning maximal virtuality of the quark, which is introduced during solving

collinear singularities in the evolution equation. Since the factorization theorem claims the universality of PDFs, they can be extracted from experimental data like deep-inelastic scattering experiments and used to predict new processes. Some examples showing parton distribution functions can be found in Fig. 1.4. In the case of more complex collision processes, such as a proton interaction with a nucleus, parton distribution functions get modified to encompass various cold nuclear matter effects [38]. The ratio of these nuclear PDFs w.r.t. proton PDFs is shown in Fig. 1.5. Let us note that nPDFs are suffering from large uncertainties in the region of small Bjorken  $x$ . However, it can be seen that for  $x < 10^{-2}$  nuclear PDFs are suppressed (nuclear shadowing), in the region  $x \approx 10^{-1}$  there is an enhancement (anti-shadowing), which is followed by another depletion (EMC effect) and rise towards  $x = 1$  due to nuclear Fermi motion [39].

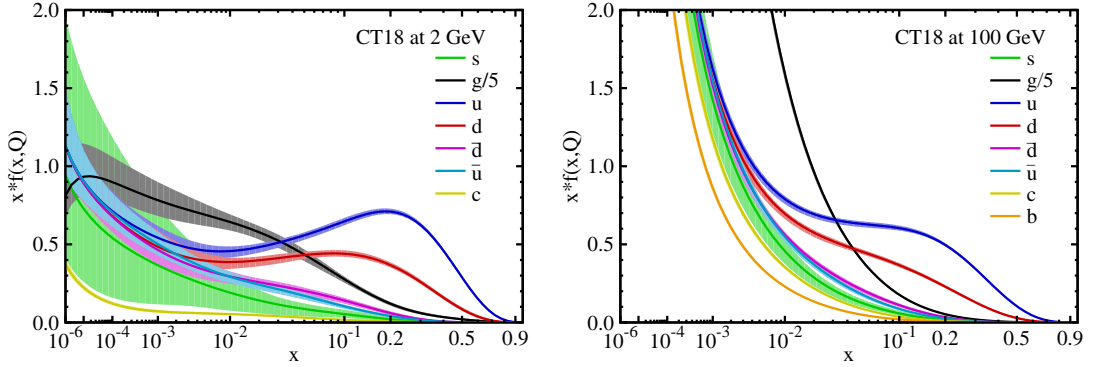


Figure 1.4: The CT18 parton distribution functions obtained by the CTEQ global analysis at  $Q = 2$  GeV and  $Q = 100$  GeV for for quarks and gluons. Taken from [40].

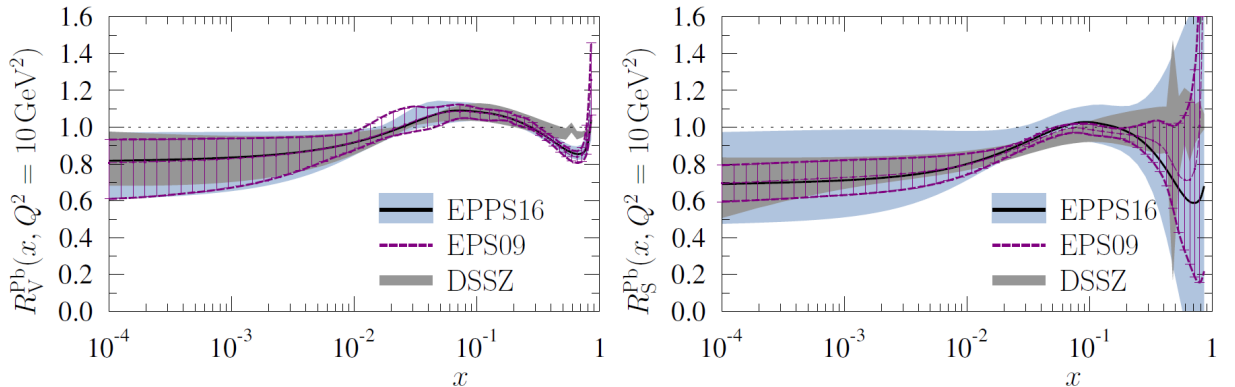


Figure 1.5: Ratios of nuclear and vacuum parton distribution functions at  $Q^2 = 10$  GeV<sup>2</sup>. The nuclear PDFs correspond to EPPS16, EPS09, and DSSZ parametrizations. Here  $R_V^{Pb}$  and  $R_S^{Pb}$  correspond to valence quarks and sea quarks in lead. Taken from [41].

2. Hard scattering refers to processes dominated by short-distance interactions. This

means that for this stage, the influence of other partons in the nucleon is neglected, and the cross-section of the partonic process  $\sigma^{a+b \rightarrow c+d}$  is calculated at a fixed order of perturbation theory as if the partons were real particles.

3. The final-state partons produced in hard-scattering processes will decrease their high initial virtuality by gluon radiation and will split into  $q\bar{q}$  pairs. This process continues until the virtuality reaches a value of the order of the scale  $\Lambda_{\text{QCD}} \approx 200\text{MeV}$  [42]. At this point, the produced partons hadronize and turn into a collimated spray of hadrons observed as a jet. Processes that involve hadronization of the parton shower cannot be described with pQCD, and therefore various models are utilized. The independent fragmentation model [43] successfully describes hadronization processes with a fragmentation function  $D_{c/h}(z, Q^2, \text{FS})$ , which provides the probability that a parton of type  $c$  will fragment through the factorization scheme FS into a hadron of type  $h$  carrying a fraction  $z$  of its momentum. There is also a possibility of defining an analog of fragmentation functions for jets. The so-called *jet functions* can be constructed using the soft collinear effective theory [44], [45]. Parton fragmentation can also be affected by jet quenching, which occurs in hot and dense strongly interacting matter [46]. Therefore, let us briefly mention a couple of facts about hot QCD matter, the experimental signatures attributed to QGP production, and the energy losses of jets in the QGP.

## 1.4 QCD matter phase diagram

Current knowledge about the phase diagram of the strongly interacting matter is summarized in Fig. 1.6. The phase diagram is shown as a function of the baryochemical potential  $\mu_B$  and the thermodynamic temperature  $T$ . The ordinary nuclear matter is located at low temperatures and  $\mu_B \approx 940$  MeV. When this matter is heated up, a hadron gas phase is formed. An increase in temperature is expected to transform the hadron gas into the phase called QGP, in which quarks and gluons escape their confinement in hadrons [47]. In the vicinity of  $\mu_B = 0$ , the lattice QCD predicts that such a transition occurs at  $T \sim 155$  MeV and has a cross-over character [48]. For finite values of  $\mu_B$ , one needs to rely on phenomenological models based on chiral perturbation theory, which predicts that the phase transition is of the first order and ends at a critical point [49]. The exact location of this critical point is not known.

It is believed that the early Universe evolved from its initial hot QGP phase following a trajectory close to  $\mu_B = 0$  [51]. A finite volume of this matter can also be produced in a laboratory by colliding ultra-relativistic heavy nuclei [52]. Figure 1.5 shows several trajectories that follow the evolution of the medium produced in such collisions. Tra-

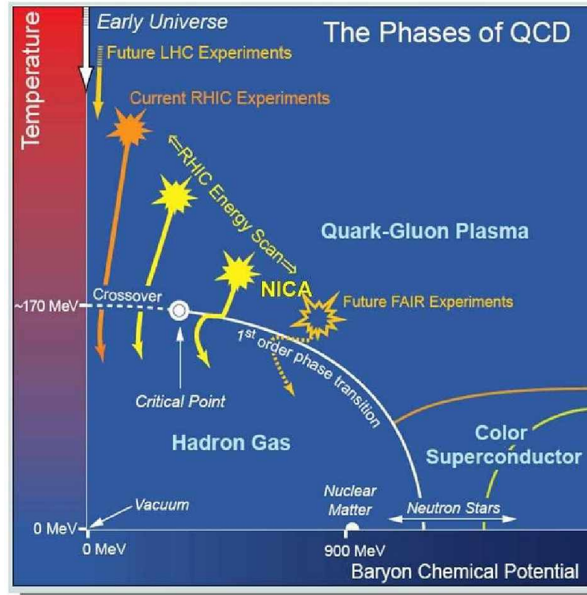


Figure 1.6: The phase diagram of the strongly interacting matter. Different regions of the diagram can be reached by means of collisions of heavy ions. The figure presents regions of interest probed by different high-energy physics facilities. Taken from [50].

jectories start at the point where the QCD matter in the collision zone reaches a local thermal equilibrium in the QGP. Hydro models predict that this stage occurs relatively early, typically on a time scale of  $1 \text{ fm}/c$  [2]. The volume expands, cools, and undergoes hadronization, chemical freezeout, and kinematic freezeout.

Since QGP lives only for a few  $\text{fm}/c$ , properties of the QGP medium have to be inferred by means of indirect observables. The observables can be divided into those which study:

- bulk particles that carry information on the hydrodynamical properties of the produced medium and its chemical and kinematic freezeout, and
- hard probes, such as jets, which are produced in the early collision stage by high- $Q^2$  processes and modified by in-medium interactions. Since jet production is known with high precision in pp collisions, observed modifications can be used to learn about parton-medium interactions [30].

Based on the comparison of the models with the experimental data, it is possible to extract the macroscopic characteristics of the produced medium and its evolution [53]–[55]. The most remarkable success of this approach was the discovery of the formation of hot and dense strongly interacting matter in ultra-relativistic collisions at RHIC [56]. Experiments have shown that the produced QGP behaves like a perfect and strongly coupled liquid [57] with large vorticity [58]. The produced QGP largely dissipates the energy of penetrating partonic showers by means of radiative and collisional processes [3]



resulting in *jet quenching*. This QGP contrasts with the original expectation that QGP has a form of non-interacting gas of quarks and gluons [2].

The emergence of QGP in collisions of heavy ions is also related to a question whether there is some minimal system size for QGP formation. It is still unknown whether smaller systems, such as proton–nucleus collisions, produce a sufficient volume of QGP, if any. Although the final states of these systems exhibit collectivity [7], [59], a possible jet-quenching signal is below the sensitivity of current experiments [4], [60]–[62].

## 1.5 Parton energy loss in proton–nucleus collisions

Parton fragmentation is affected by energy losses induced by interactions with the medium, which can occur through inelastic (radiative energy loss) and elastic (collisional energy loss) processes. The structure of energy loss varies depending on parton mass and energy: the radiative energy loss dominates at high energies, while the elastic energy loss prevails at lower energies [63]. In the case of heavy quarks, one also needs to consider the dead-cone effect, which suppresses gluon emissions at forward angles [10]. The comparison between radiative and collisional energy loss for light and bottom quarks is shown in Fig. 1.7.

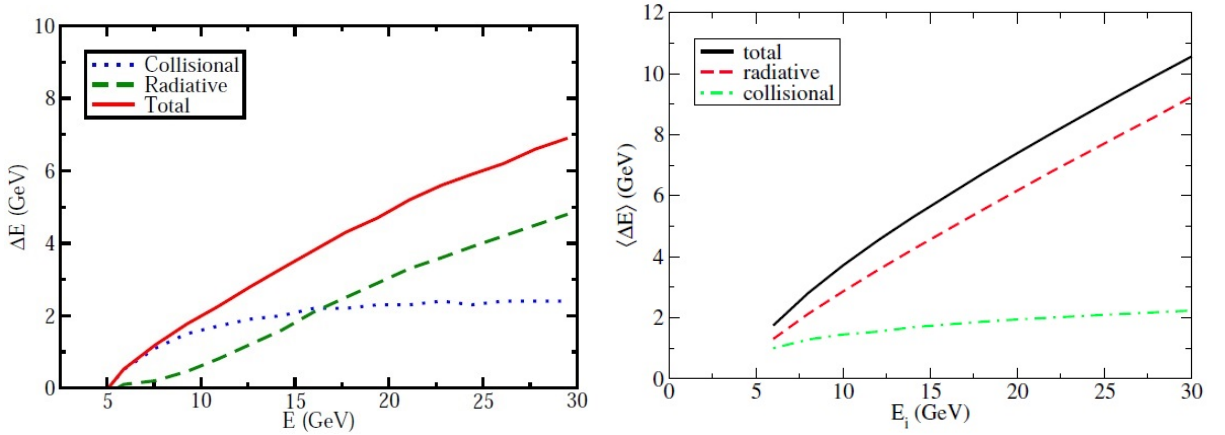


Figure 1.7: Left: Average energy loss of a bottom quark in 0–7.5% central collisions at the LHC. Right: Average energy loss of a light quark in 0–5% central collisions at RHIC. Both are taken from [64].

In the BDMPs approximation [65], the elastic energy loss of a parton in QGP,  $\Delta E$ , is quadratically proportional to the length of the traversed path  $L$  in the medium,  $\Delta E = \hat{q}L^2$ . Here  $\hat{q}$  denotes a transport coefficient which quantifies the mean transverse momentum kick squared  $\langle k_T^2 \rangle$  acquired by a parton per mean free path length  $\lambda$  due to interactions with the medium constituents  $\hat{q} = \frac{\langle k_T^2 \rangle}{\lambda}$ . Based on energy density arguments, paper [66]

estimates that in the LHC energy regime, the  $\hat{q}$  for p–Pb collisions is seven times smaller than the  $\hat{q}$  for Pb–Pb. The size of  $\hat{q}$  for Pb–Pb at LHC energies was evaluated by the JET Collaboration [54]. They found a value of  $(1.9 \pm 0.7)$  GeV<sup>2</sup>/fm. This can be compared to  $\hat{q}$ , which was extracted for 0.02 GeV<sup>2</sup>/fm cold nuclear matter [67], and which is two orders of magnitude smaller. Thus, it could be expected that the size of jet quenching should be substantially smaller in p–Pb systems than in Pb–Pb. Indeed, the ALICE collaboration has used the hadron-jet correlation technique to measure the medium-induced energy transport outside of charged particle jets with a cone  $R = 0.5$  in Pb–Pb [68], which was found to be  $\Delta E = (8 \pm 2)$  GeV/ $c$ . On the other hand, an analogous measurement for central p–Pb collisions could set just an upper limit on this energy transport which was found to be less than 0.4 GeV/ $c$  at 90% confidence level for energy transport out of the  $R = 0.4$  cone [4].

## 1.6 Dead-cone effect

The dead-cone effect denotes the specific radiation pattern of heavy quarks, where the phase space of soft-gluon emission is limited to the region of the relatively large radiation angle  $\theta > \frac{m_q}{E_q}$ . Initially, it was predicted for vacuum gluonstrahlung [10], later this effect was also applied to medium-induced radiation taking into account multiple rescattering of the radiated gluon in the medium [69]. This phenomenon leads to a smaller energy loss for heavy quarks compared to lighter ones; experimentally, it is reflected in the ordering of nuclear-modification factors  $R_{AA}^{\text{light quark}} < R_{AA}^{\text{charm}} < R_{AA}^{\text{beauty}}$  [70]. This is illustrated in Fig. 1.8, which shows the significant variation in  $R_{AA}$  of particles with different masses in Pb–Pb collisions. The corresponding pattern for a small collision system is expected to be less pronounced due to the lower density of color charges in the medium [69]. The magnitude of the dead-cone effect is limited by the quark energy greater than the characteristic value  $E_{\text{HQ}} = M\sqrt{\hat{q}L^3}$ , which depends on the medium size ( $L$ ), quark mass ( $M$ ), and the gluon transport coefficient ( $\hat{q}$ ) [69]. This results in the mass independence of energy loss at higher parton energies (80 – 250 GeV) in Pb–Pb collisions [71].

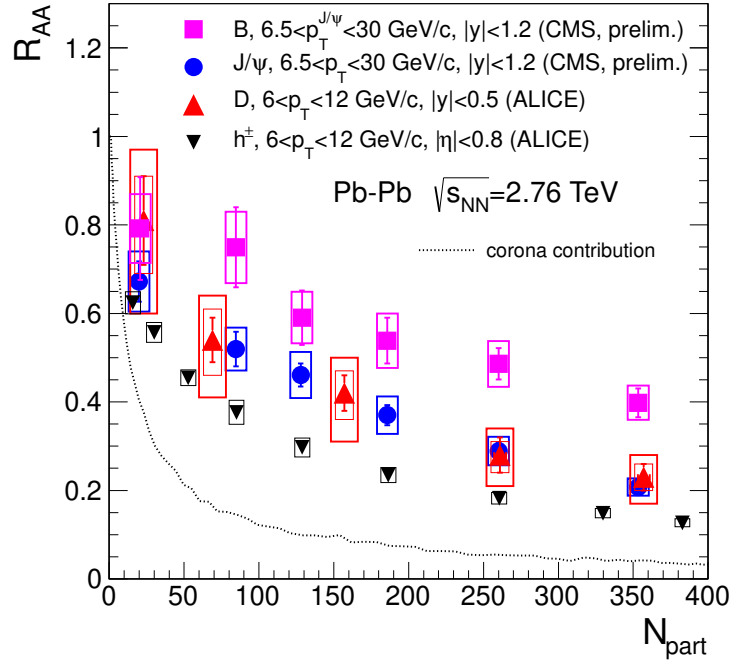


Figure 1.8: The nuclear modification factor  $R_{AA}$  for different high- $p_T$  particles as a function of centrality. Different markers correspond to: charged hadrons (h), charmed mesons (D), charmonium ( $J/\Psi$ ), and beauty hadrons (B) [70].

## 1.7 Experimental overview of small collision systems

### Measurements of p–Pb collisions

Modifications of the spectra of high  $p_T$  hadrons and jets due to hot and/or cold nuclear matter effects are traditionally quantified by means of the nuclear modification factor, which is defined for proton-nucleus collisions as:

$$R_{pA}(p_T) = \frac{1}{\langle T_{pA} \rangle} \frac{1/N_{\text{evt}} dN_{pA}/dp_T}{d\sigma_{pp}/dp_T}. \quad (1.10)$$

Here  $1/N_{\text{evt}} dN_{pA}/dp_T$  is the differential distribution of the hadron or jet yield per event measured in proton-nucleus collisions as a function of transverse momentum,  $d\sigma_{pp}/dp_T$  is the corresponding differential cross section measured in pp collisions, and  $T_{pA}$  is the nuclear thickness function that accounts for the geometry of the collision.  $T_{pA}$  is estimated using the Glauber model [72]–[74], which relates the multiplicity measured in some detector to the geometry of the collision (number of participants, impact parameter).

For small collision systems, it is challenging to determine the nuclear overlap function for centrality-biased collisions with sufficient accuracy. Geometry information, encoded in spectator and/or participant multiplicity, is smeared by large fluctuations coming from

produced jets and their fragmentation [75], [76]. In addition, there is a non-negligible correlation between different parts of phase space due to conservation laws [77]. Therefore, the imposition of centrality bias does not necessarily bring stricter limits on the jet quenching signal [78].

For minimum bias events, the Glauber model nevertheless provides an exact formula for the nuclear modification factor [72], which can be expressed as:

$$R_{pA} = \frac{1}{A} \frac{d\sigma_{pA}/dp_T}{d\sigma_{pp}/dp_T}, \quad (1.11)$$

where  $A$  is the number of nucleons in the nucleus and  $d\sigma_{pA}/dp_T$  is the particle or jet production cross section in the proton-nucleus collision.

The p–A and d–A collisions are an important reference for high- $p_T$  hadrons and jets production in A–A collisions. Production in A–A can be influenced by jet quenching as well as initial cold nuclear matter effects, making stand-alone measurements of large systems insufficient to disentangle the net effect of QGP and CNM. The p–A / d–A systems have a significantly lower number of binary collisions, and it can be expected that CNM effects will prevail. In this context, it is also important to study the energy dependence of nuclear modification factors. Measurements of  $R_{pA}$  for hadrons and jets at the LHC and RHIC are sensitive to different regions of the PDFs. Whereas in the LHC energy regime, PDFs are dominated by gluons, and in the RHIC energy regime, they are dominated by quarks.

Figure 1.9 shows the nuclear modification factors of inclusive charged hadrons in different collision systems measured by ALICE at  $\sqrt{s_{NN}} = 5$  TeV and measured by the STAR experiment at  $\sqrt{s_{NN}} = 200$  GeV. The observed magnitude of suppression in central Pb–Pb collisions is about the same as that in central Au–Au collisions. However, the corresponding energy loss at the LHC is larger, since because the  $p_T$  spectra are harder. The left figure also shows nuclear modification factors for direct photons and heavy weak bosons W and Z, which are compatible with unity, suggesting that our understanding of the collision geometry in Pb–Pb is under control, as none of these probes is expected to be affected by QGP. In addition to that, Fig. 1.9 demonstrates that the nuclear modification factor of inclusive hadrons measured in minimum-bias p–Pb and d–Au collisions is compatible with unity for hadron  $p_T > 8$  GeV/c, suggesting that the magnitude of possible jet quenching is below the precision of current measurements. Nuclear modification factors for  $p_T$  less than 5–7 GeV/c are affected by the Cronin effect and flow [79].

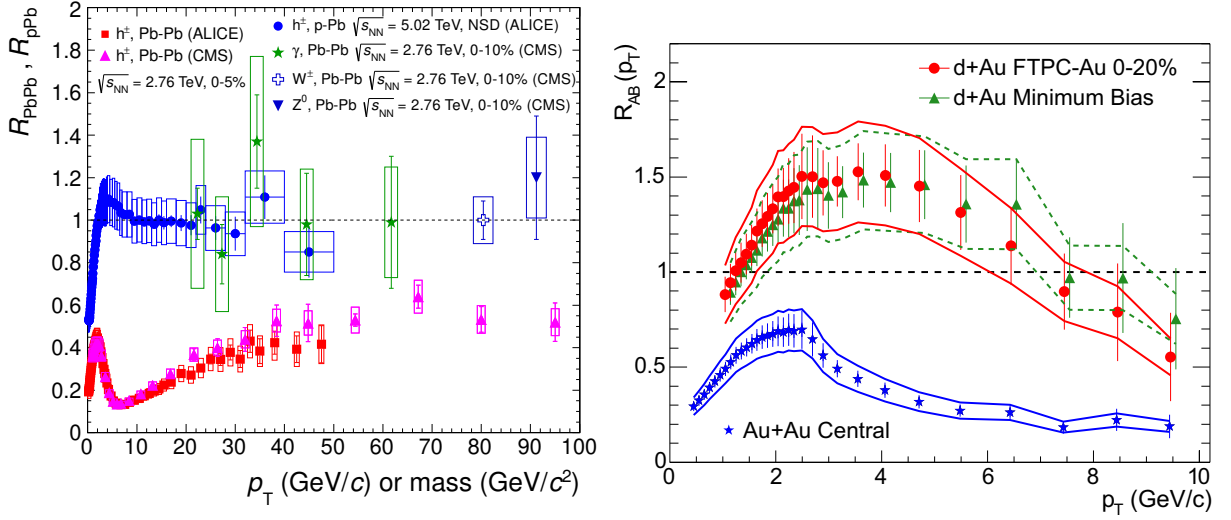


Figure 1.9: Left: Transverse momentum dependence of the nuclear modification factor  $R_{pPb}$  of charged particles measured in minimum-bias p–Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV in comparison to data on the nuclear modification factor  $R_{PbPb}$  from central Pb–Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV. Taken from [80]. Right:  $R_{dAu}(p_T)$  of charged hadrons for minimum bias and central d–Au collisions, and central Au–Au collisions measured by the STAR collaboration at 200 GeV/c. Taken from [81].

The nuclear modification factor was measured similarly for jets in minimum-bias collisions. As can be seen in Figure 1.10, the data were also found to be compatible with no suppression [6], see also the results of other experiments [4], [5], [82], [83]. However, the PHENIX experiment also published more differential analysis, where events were sorted into several centrality bins, as can be seen in the right panel of Fig. 1.10. They found suppression of jets in central collisions and surprising enhancement for peripheral collisions. However, such data interpretation is controversial, as peripheral collisions should provide a spectrum compatible with pp collisions. ATLAS also reported a similar analysis for p–Pb, where the nuclear modification was also analyzed as a function of rapidity and where one sees a pronounced difference between the direction of the proton and the direction of the lead, see Figure 1.11. In order to understand this behavior, ATLAS performed a reference pp measurement [84]. In this analysis, the authors studied the correlation between transverse energy flow in the forward ATLAS calorimeter and jet production using a specific collision system, where one of the protons represented a nucleon in the Pb (target) ion, and the other proton served as the projectile. The obtained results show that if the hard scattering process involves a large  $x$  parton in a target Pb nucleon, the corresponding nucleon remnant has lower longitudinal energy, leading to a reduction in transverse energy flow at large pseudorapidity. Therefore, the centrality estimator based

on the calorimeter measurement is biased.

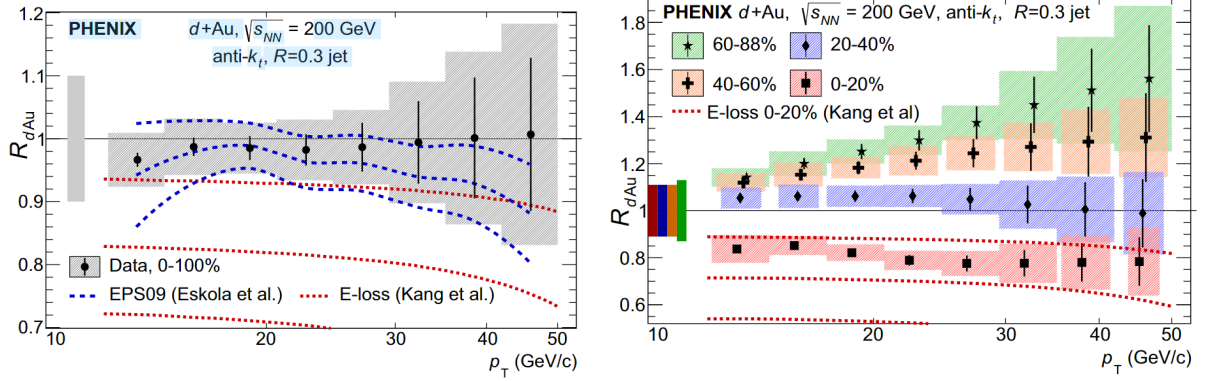


Figure 1.10: Left: The PHENIX  $R_{dAu}$  for  $R = 0.3$  anti- $k_T$  jets measured in minimum bias d–Au collisions at  $\sqrt{s_{NN}} = 200$  GeV. The data are compared with calculations which account for EPS09 nuclear PDFs but no energy loss and a model which assumes parton energy loss. Right: The PHENIX measurements of the nuclear modification factor in different centrality intervals. Taken from [85].

Heavy-flavor probes can serve as a possible tool to extend jet quenching measurements. Production of heavy charm and bottom quarks can occur only in the initial collisions due to their high mass (1.3 and 4.2 GeV/ $c^2$ , respectively) in contrast to the light quarks, which can also be produced thermally. Due to the low rate of annihilation and a lifetime significantly longer than QGP evolution, the heavy-flavor quarks experience the complete evolution of the system.

Measurements of beauty hadrons are typically exploiting decay channels of b quarks which can lead to production of hadrons (e.g. non-prompt J/ $\Psi$ )[86], [87] or undergo a semi-leptonic decay [88]. Heavy-flavor yields reconstructed by both approaches demonstrate expected behavior showing significant suppression for Pb–Pb for  $p_T > 5$  GeV/ $c$ . The measurements of  $R_{pA}$  for different HF probes do not show any sign of modification by nuclear-matter effects; see Figure 1.12.

Despite the absence of a sign of jet quenching effect, proton-nucleus collisions exhibit in small collision systems other signatures of the QGP formation, such as strangeness enhancement [8] and collective phenomena observed by ALICE [7], and CMS [90] at  $\sqrt{s_{NN}} = 5.02$  TeV, and by ATLAS at  $\sqrt{s_{NN}} = 8.16$  TeV [59]. These phenomena can arise from the presence of a QGP droplet or complex effects occurring in the vacuum of quantum chromodynamics [66]–[69]. ALICE has recently reported a two-particle correlation analysis that reveals positive  $v_2$  for jet fragments; see the left panel of Figure 1.13. Similar results can also be observed in the  $v_2$  measurements of high  $p_T$  charged hadrons

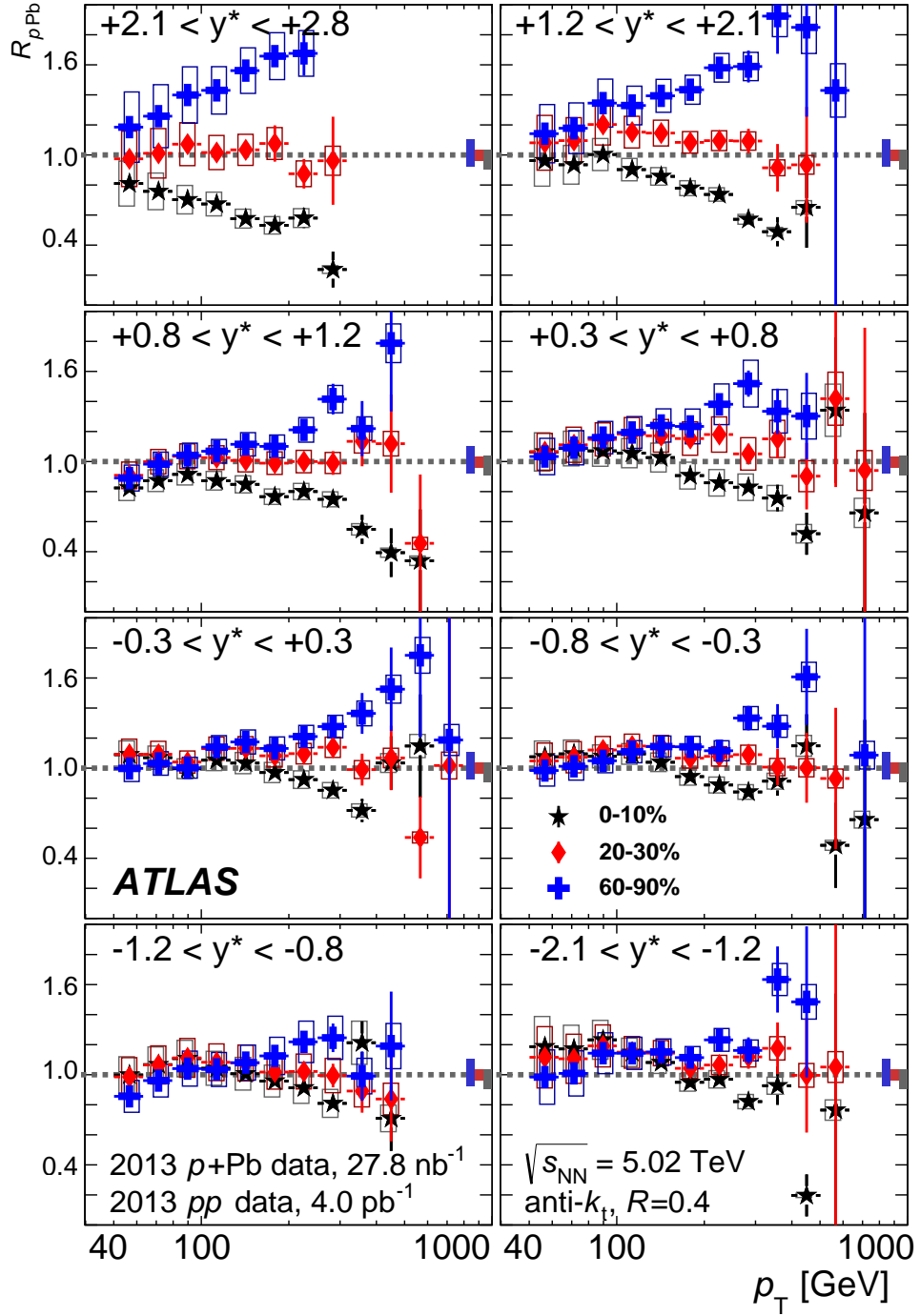


Figure 1.11:  $R_{pPb}$  as a function of  $p_T$  for anti- $k_T$   $R = 0.4$  jets in p-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV in central (stars), mid-central (diamonds) and peripheral (crosses) events. Each panel corresponds to a different rapidity range. Positive rapidity corresponds to the proton-going direction and negative rapidity to the lead-going direction. Taken from [76].

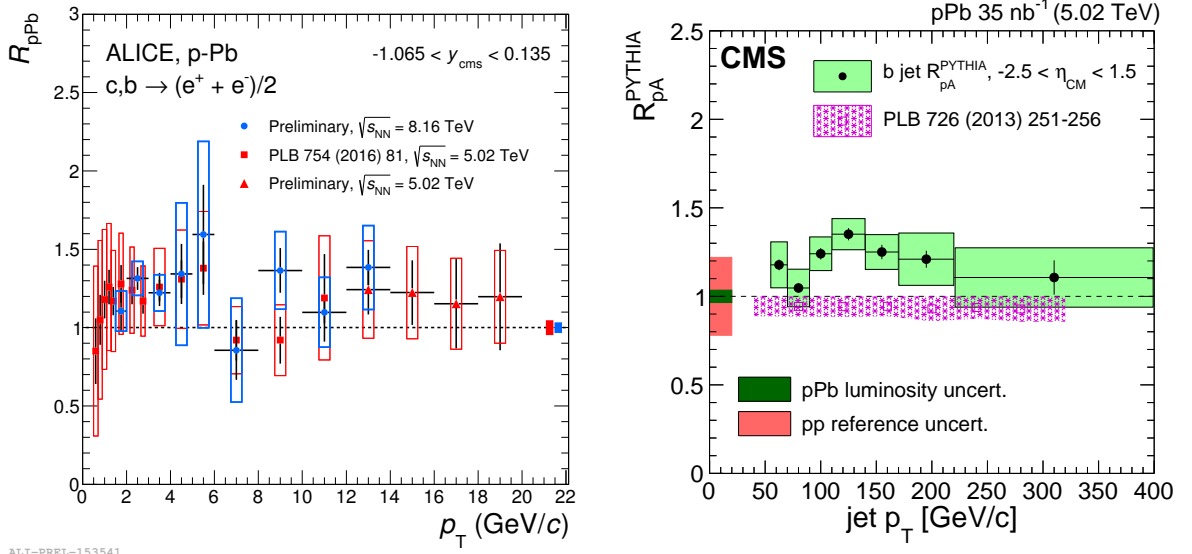


Figure 1.12: Top: ALICE  $R_{pPb}$  of electrons from heavy-flavour hadron decays at 5 TeV and 8.16 TeV [89]. Bottom: The b jet  $R_{pA}^{PYTHIA}$  as a function of jet  $p_T$  together with a pQCD prediction [11].

by ATLAS; see the right panel of Figure 1.13. These  $v_2$  measurements resemble an effect of path-length-dependent energy loss, which is, however, not reflected by suppression in  $R_{pPb}$  measurements.

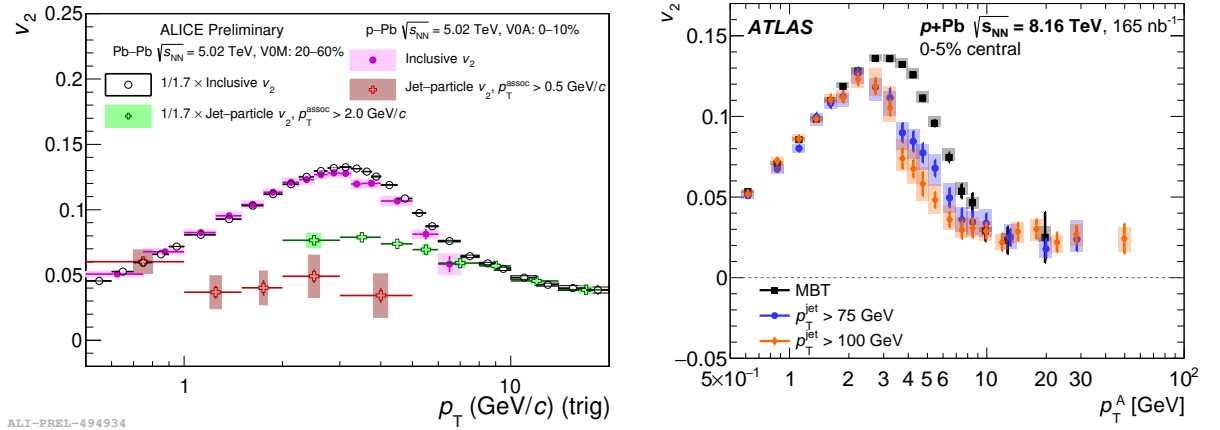


Figure 1.13: Left: Comparison of jet-particle  $v_2$  and inclusive charged-particle  $v_2$  in p-Pb and Pb-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV. Figure taken from [7]. Right:  $v_2$  of charged-particles as a function of particle  $p_T$  as measured by ATLAS in p-Pb collisions at  $\sqrt{s_{NN}} = 8.16$  TeV. Taken from [59].



## Measurements of pp collisions

Besides being the baseline for final-state effects studies in Pb–Pb and p–Pb systems, pp collisions constitute an important test bench for various QCD phenomena, such as the dependence of charm fragmentation functions on a collision system [91], multi-parton interactions [92], and possible final-state effects in high-multiplicity pp collisions [93]. State-of-the-art pQCD calculations successfully describe particle and jet production in pp collisions [94], [95]. This invites us to use pp collisions to investigate more complex observables, such as jet substructure measurements. The application of iterative declustering on jets associated with a charm hadron has recently led to the first direct observation of the dead cone effect in QCD [32], see Fig. 1.14. The figure shows a ratio of opening-angle distributions between a subleading prong and a branch associated with a charm quark or an inclusive parton, radiator. One can see suppression of forward emissions from the charm quark, which vanishes when the radiator energy increases.

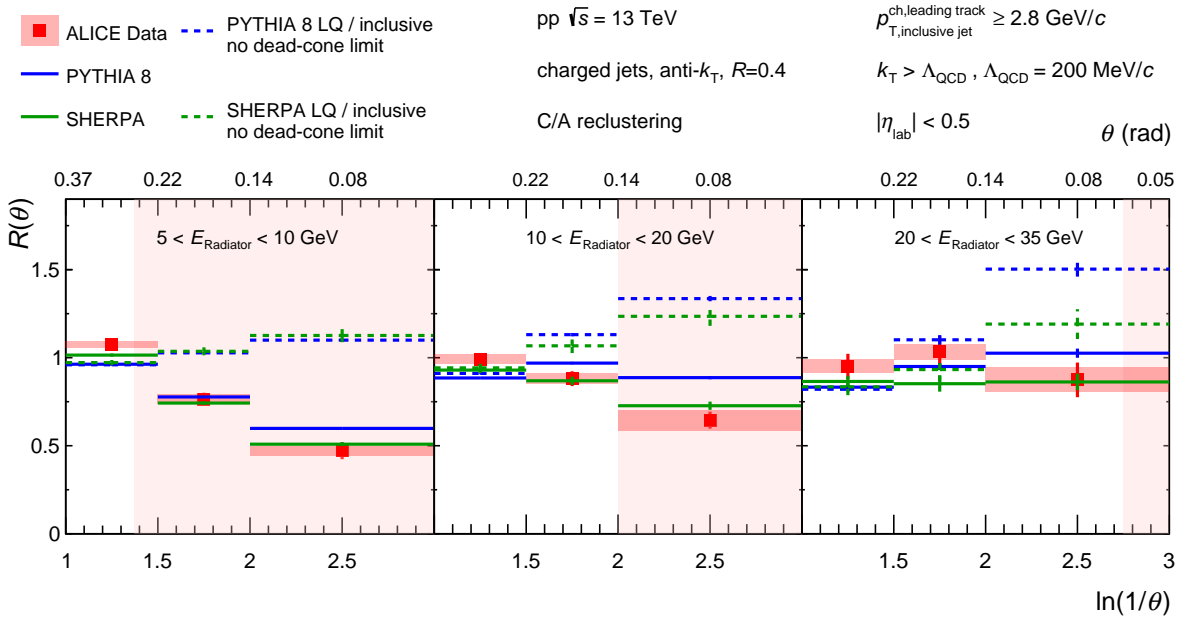


Figure 1.14: The ratios of the splitting-angle probability for  $D^0$ -meson jets to inclusive jets in pp collisions at  $\sqrt{s} = 13$  TeV. The data was compared with PYTHIA 8 and SHERPA simulations. The panels correspond to various  $E_{\text{Radiator}}$ : the left panel is for  $5 < E_{\text{Radiator}} < 10$  GeV (left panel), the middle panel is for  $10 < E_{\text{Radiator}} < 20$  GeV and the right panel shows the ratio for  $20 < E_{\text{Radiator}} < 35$  GeV. Taken from [32].

The impact of the dead-cone effect on heavy-flavor jet fragmentation was also observed in the measurements of groomed-jet substructure for charm jets [96]. In this work, authors studied only perturbative splittings selected by means of the soft-drop grooming condition:

$$z = \frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} < z_{\text{cut}}, \quad (1.12)$$

where  $p_{T,1}$  and  $p_{T,2}$  are the transverse momenta of the leading and subleading prongs of the splitting and  $z_{\text{cut}}$  was set equal to 0.1. Figure 1.15 shows the average number of perturbative splittings which passed the soft drop condition when unwinding the history of the clusterization process. Jets associated with  $D^0$  hadron exhibit lower  $n_{\text{SD}}$  on average than inclusive jets. This is in line with the expectation that heavy-flavor quarks radiate less due to the dead cone.

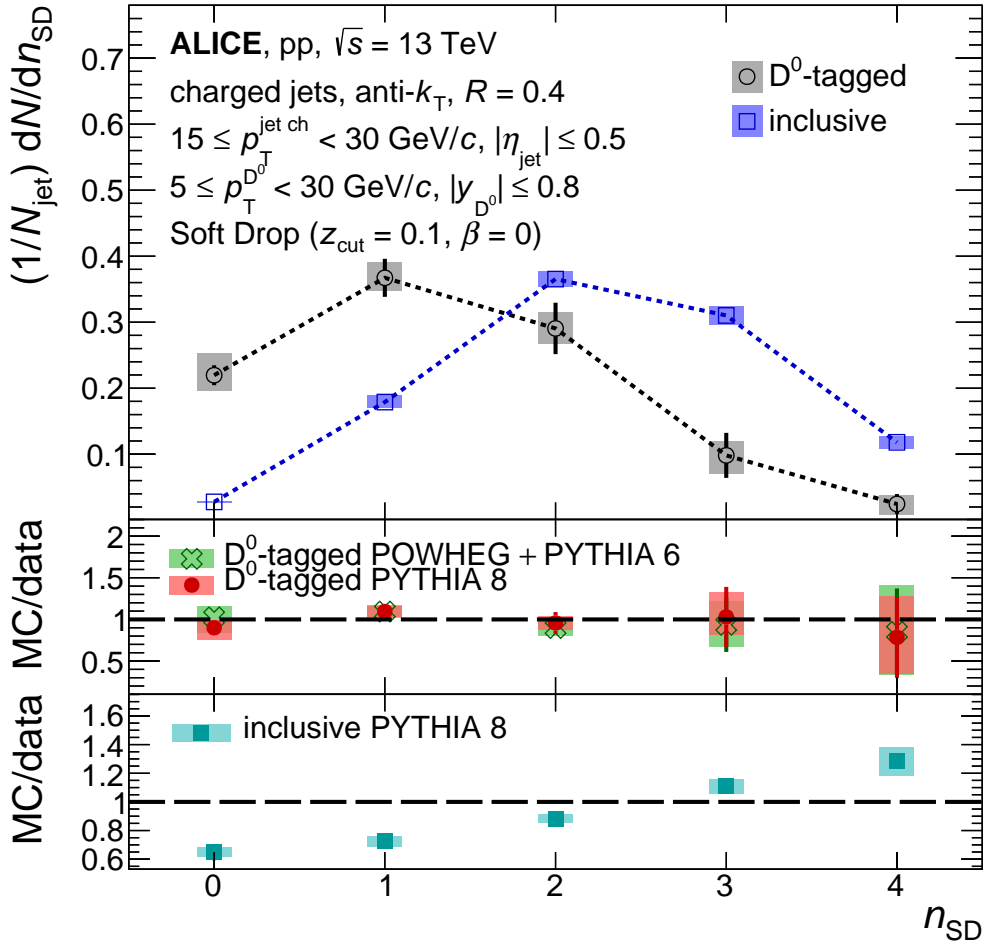


Figure 1.15:  $n_{\text{SD}}$  (right) distributions of prompt  $D^0$ -tagged jets compared to inclusive jets for  $15 < p_T^{\text{jet ch}} < 30$  GeV/c in pp collisions at  $\sqrt{s} = 13$  TeV. Model/data ratios are shown in the bottom panels for PYTHIA 8 and POWHEG + PYTHIA 6 simulations. Taken from [96].

The abovementioned measurements call for further studies of mass-dependent effects with heavy-flavor probes, particularly the jets associated with b-quark fragmentation. However, these studies are challenging due to a small fraction of HF-tagged jets in the inclusive jet sample; see recent ATLAS studies of b jets [97] providing the fraction of b jets in p-p collisions at  $\sqrt{s} = 5$  TeV at the level of 4%, see Fig 1.16. This implies the importance of high luminosity and high  $\sqrt{s}$  of collisions. These conditions are met at the LHC, where one can measure beauty production with significant statistics.

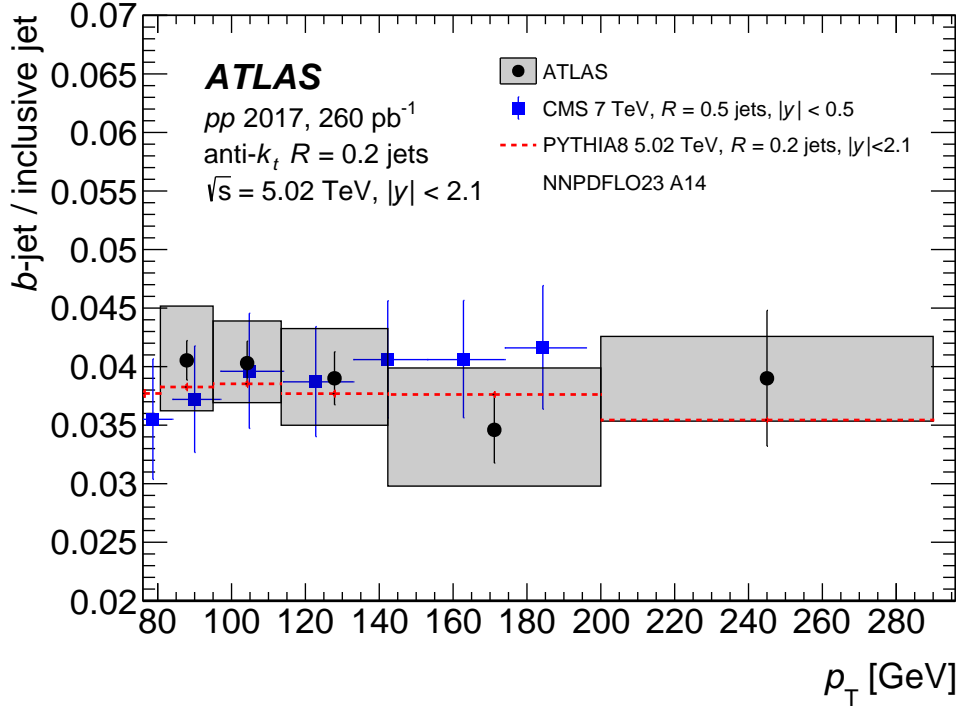


Figure 1.16: Fraction of jets associated with b-quark fragmentation among inclusive jets as measured by ATLAS in pp collisions at  $\sqrt{s} = 5.02$  TeV and CMS in pp collisions at  $\sqrt{s} = 7$  TeV. The data measured by ATLAS are compared with a calculation by PYTHIA 8. Taken from [97].

The most common processes that contribute to the production of b-quarks are Flavor Creation (FCR), Flavor Excitation (FEX), and Gluon Splitting (GSP), which are described by the Feynman diagrams shown in Fig. 1.17. As can be seen in Fig. 1.18, in the LHC energy regime, the LO channel (FCR) produces a smaller b-jet yield compared to the NLO channels (FEX and GSP), signaling that the NLO channels receive a substantial enhancement from collinear logarithms [98].

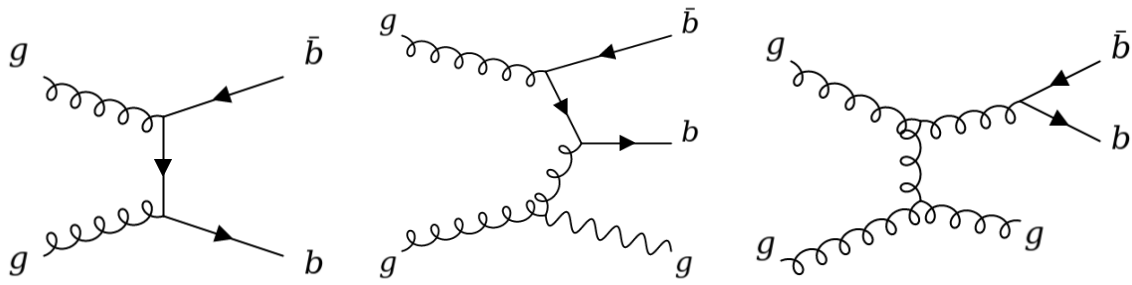


Figure 1.17: Basic diagrams which describe production of b quarks. Left: flavor creation process, middle: flavor excitation process, right: gluon splitting process.

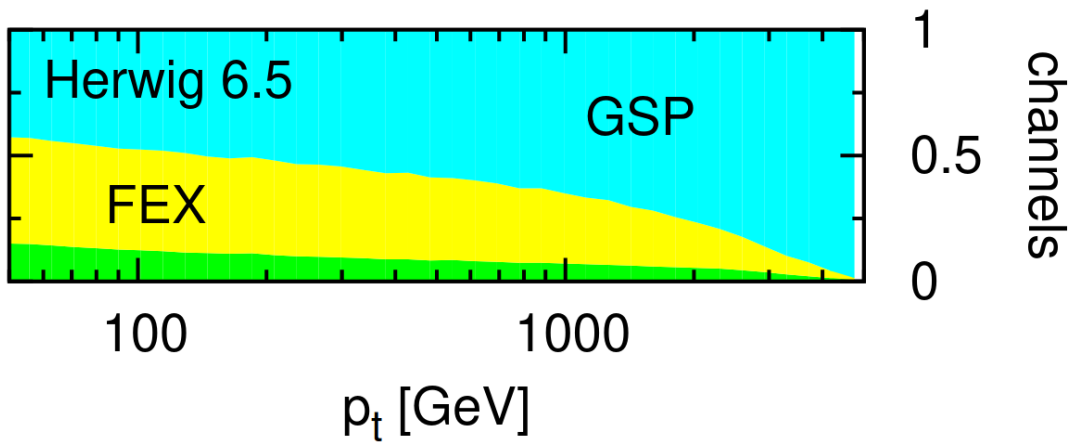


Figure 1.18: Share of b jets produced by GSP (in cyan), FEX (in yellow) and FCR (in green) processes as predicted by HERWIG for pp collisions at  $\sqrt{s} = 14$  TeV. The transverse momentum of produced b jets is denoted  $p_t$ . Taken from [98].

# Chapter 2

## ALICE experiment

A Large Ion Collider Experiment (ALICE) [12] is one of the four large-scale experiments at the Large Hadron Collider (LHC). ALICE focuses on the physics of ultra-relativistic heavy-ion (HI) collisions and measurements of QGP properties. Consequently, the ALICE detector needs to have high granularity and be capable to cope with a high particle density in central Pb–Pb collisions where the charged particle multiplicity is of the order of  $dN_{\text{ch}}/d\eta \approx 1600$  [99].

Figure 2.1 shows a schematic view of the ALICE detector. The ALICE detector consists of 18 different subsystems that form a central barrel, a muon arm, and forward detectors. The central barrel comprises the main systems required for track reconstruction, reconstruction of the primary collision vertex, and particle identification. The central barrel spans over a pseudorapidity range of  $-0.9 < \eta < 0.9$  and a full azimuth and consists of the following detection subsystems: Inner Tracking System (ITS), Time Projection Chamber (TPC), Transition Radiation Detector (TRD), Time of Flight (TOF) detector, HMPID, and the electromagnetic calorimeters EMCal and PHOS. The muon arm covers a pseudorapidity range of  $-2.5 < \eta < -4.0$ , and it is used to measure muons coming from decays of open heavy-flavor hadrons and quarkonia. Forward and backward pseudorapidity is covered by detectors providing trigger signals and event characterization complementary to the central barrel measurements. ALICE has excellent particle identification capabilities that utilize measurements of ionization energy loss  $-dE/dx$ , time-of-flight, Cherenkov radiation, electromagnetic calorimetry, muon filters, and topological decay reconstruction [12].

Compared to the ATLAS [101] and CMS [102] experiments, the ALICE central region is distinguished by its low material budget. In particular, the value of the radiation length necessary for a particle to enter the calorimeter is 10–12%  $X_0$  for ALICE and  $\approx 40\%$   $X_0$  for ATLAS and CMS [103]. ALICE achieves such a low material budget by utilizing a gas mixture for a material of the active volume of the TPC. Another important ALICE

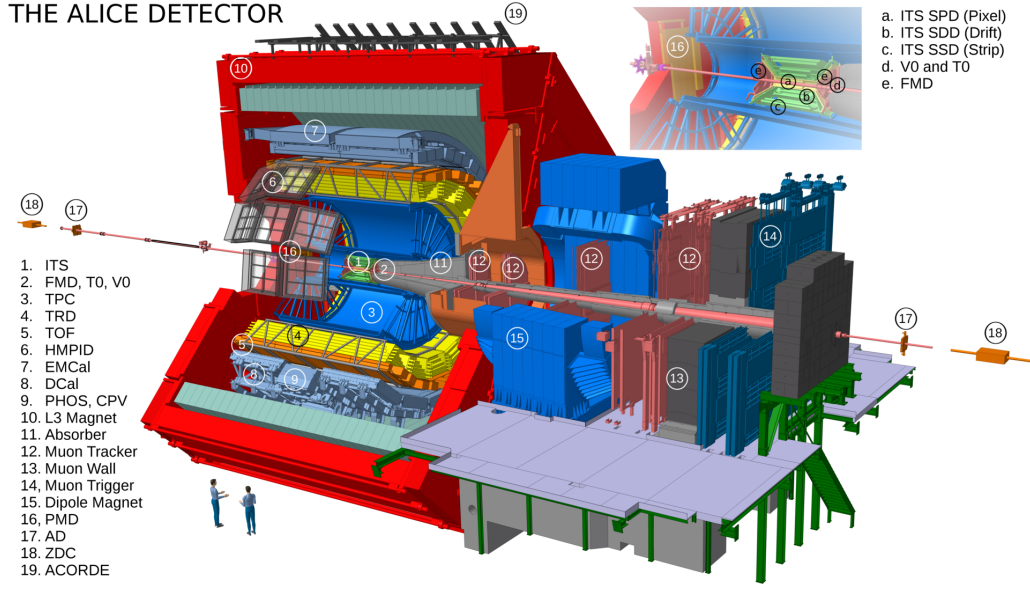


Figure 2.1: Scheme of the ALICE detector during the LHC Run 2. Taken from [100].

feature is a lower magnetic field intensity of 0.5 T when compared to ATLAS or CMS magnets, which have 2 T and 3.8 T, respectively. The magnetic field in ALICE is created by a solenoidal magnet L3. All these design features allow for track reconstruction from 150 MeV/ $c$  with good particle identification capabilities up to 20 GeV/ $c$  [12].

## 2.1 Detector system overview

The b-jet analysis uses tracks reconstructed from space points measured by the ITS and TPC detectors and event selection based on the VZERO scintillator array. These systems will be described in more detail below.

### 2.1.1 ITS

The ITS [12] is the detector closest to the collision point. The ITS covers the radial range from 4 to 43 cm and consists of six cylindrical layers designed with three different silicon sensor technologies. The two innermost ITS layers operate in a region with the highest radiation level where the track density can reach up to 50 tracks/cm<sup>2</sup> [12].

The first two layers of the ITS are equipped with the Silicon Pixel Detector (SPD), which provides precise position measurement for hits necessary for secondary vertex reconstruction. The SPD sensor comprises a two-dimensional pixel matrix of reverse-biased silicon detector diodes. The sensor matrix includes  $256 \times 160$  pixels having a pitch of  $50 \times 425 \mu\text{m}^2$  and a thickness of 200  $\mu\text{m}$ . The SPD sensor has a resolution of 12  $\mu\text{m}$  in the plane traverse to the beam axis and about 100  $\mu\text{m}$  along the beam axis. The material

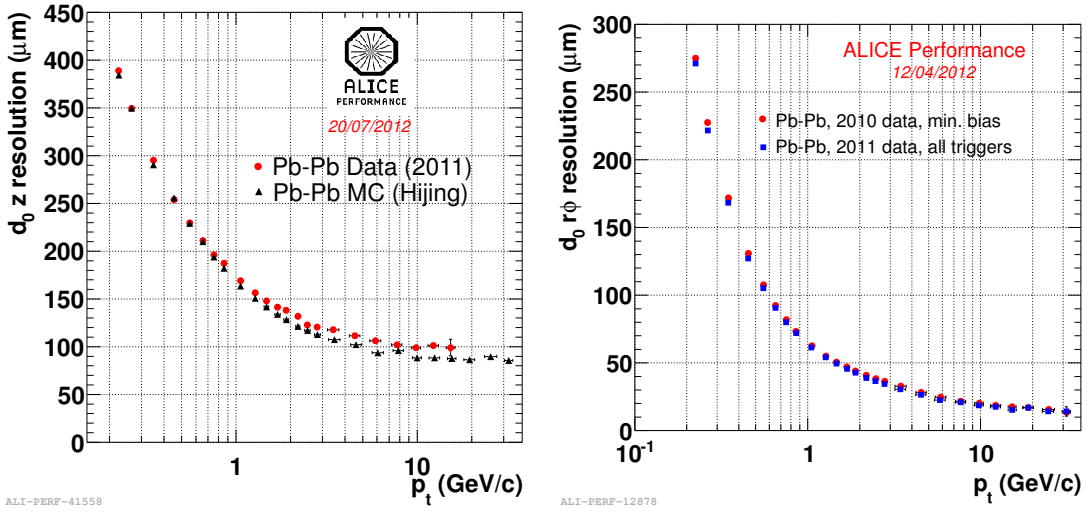


Figure 2.2: Impact parameter resolution vs particle  $p_T$ . Left: in the longitudinal direction. Right: in the transverse plane. Taken from [104], [105].

budget of the SPD is approximately 1.14%  $X_0$  per layer.

The following two ITS layers constitute silicon drift detector technology (SDD) with a material budget of 1.13%  $X_0$  and 1.26%  $X_0$ . SDD sensors demonstrate good multi-track capability and provide two out of four  $-dE/dx$  samples needed for ITS particle identification. The average resolution of the SDD track is 35  $\mu\text{m}$  and 25  $\mu\text{m}$  along the beam axis and in the transverse plane, respectively [12].

The outermost two layers are formed by the Silicon Strip Detector (SSD) with a material budget of 0.86%  $X_0$  per layer only. SSD layers are crucial for merging ITS and TPC tracks and provide two  $dE/dx$  samples for PID of particles with transverse momenta below 200  $\text{MeV}/c$ . The spatial resolution of the SSD system in the direction along the beam is 820  $\mu\text{m}$  and 20  $\mu\text{m}$  in the transverse plane [12].

The ITS performs tracking of low- $p_T$  particles and improves the  $p_T$  resolution of the central barrel tracking in combination with TPC. Another function of the ITS is to improve the pointing resolution towards the interaction point and to locate the primary vertex with a resolution better than 100  $\mu\text{m}$  [12]. These parameters are essential for the secondary decay vertex reconstruction, especially for the heavy-flavor hadron studies based on decay topology.

## 2.1.2 TPC

The Time Projection Chamber [106] is the main tracking device in ALICE. The active volume of the TPC contains 88  $\text{m}^3$  of Ne- $\text{CO}_2$ - $\text{N}_2$  gas in the mixture and covers a radial span of 0.85 m to 2.5 m and  $-2.5 \text{ m}$  to  $2.5 \text{ m}$  along the beam axis. The material budget of the TPC is 3.5%  $X_0$  only. The TPC measures the primary ionization of charged particles

passing through an active volume of the detector. Electrons and ions produced drift in an electric field established by a central electrode with a drift voltage of 100 keV in the direction of the front/back parts of the detector (electrons) or the central electrode (ions). It results in an amplification of the electron signal by a factor of 7000–8000 and a maximum drift time of electrons of about 100  $\mu$ s. The low drift speed leads to a pileup of several events in the active area during pp collisions at a rate of 100 kHz. Therefore, the TPC is optimized to suppress possible space-charge effects. The gating grid limits the backflow of ions from the amplification region to the drift region. The maximum opening frequency of the gating grid is 3.5 kHz and is the main limiting factor for the readout rate in pp collisions. Another limitation is the bandwidth of the readout electronics in Pb–Pb collisions, since the size of a TPC event significantly varies for different collision systems. For example, it equals 1 MB for pp and exceeds 70 MB for a central Pb–Pb collision. At the same time, the ITS Pb–Pb event size will be a few hundred kB only.

In addition to the significant role in track reconstruction, TPC provides a data sample for PID. The left side of Figure 2.3 shows the TPC measurements of the specific energy loss  $dE/dx$  as a function of particle momentum. The black lines show the mean energy loss predicted by the Bethe-Bloch formula for different particle species. This method of particle identification is very limited in the range of applications in the case of particle-by-particle identification. For particles with a momentum greater than 1 GeV/ $c$ , one can observe the intersection of the lines for the pion and kaon, making a PID decision in that region ambiguous. Therefore, yields of different particle species are extracted on a statistical basis with a template fit method. This approach can be applied to particles with transverse momenta up to 20 GeV/ $c$ . The TOF can enhance the PID capability of the TPC.

### 2.1.3 VZERO

The VZERO detector consists of two arrays of scintillator counters, covering the pseudo-rapidity ranges  $2.8 < \eta < 5.1$  (VZERO-A) and  $-3.7 < \eta < -1.7$  (VZERO-C) [107]. Both VZERO-A and VZERO-C are segmented into 32 individual counters. The VZERO-A detector is located 340 cm from the vertex on the side opposite the muon spectrometer. The VZERO-C is fixed 90 cm from the vertex to the front side of the hadronic absorber. The VZERO provides minimum-bias (MB) triggers in pp, and Pb–Pb collisions via time coincidence of a VZERO-A and VZERO-C signal. The signal from VZERO is further used to trigger on event multiplicity. VZERO timing information is also used to reject beam-gas events in the TPC [12].



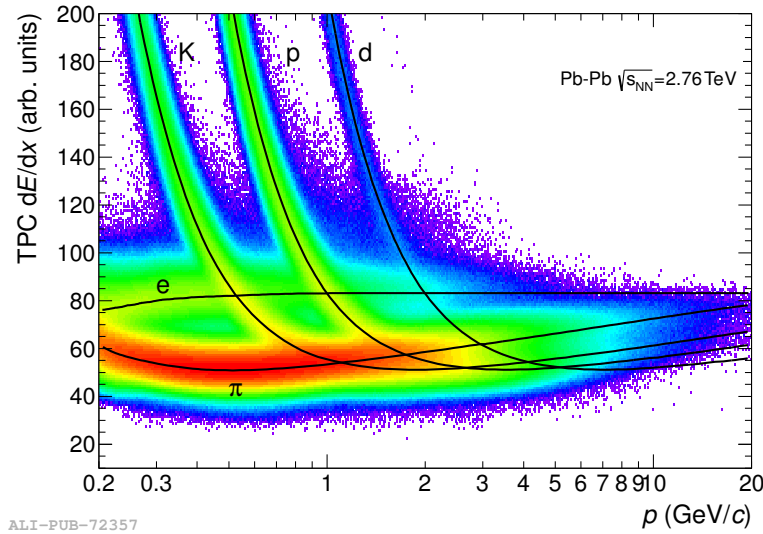


Figure 2.3: Specific energy loss ( $dE/dx$ ) in the TPC vs. particle momentum in Pb–Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV. The lines show the parametrizations of the expected mean energy loss. Taken from [13].

## 2.2 Trigger System

The Central Trigger Processor (CTP) [108] is responsible for handling the ALICE trigger system. It assumes the synchronization of information from all the triggering detectors and transfers the resulting trigger signal back to the detectors to initiate their readout. The CTP has a three-level scheme. First, the CTP processes the L0 signals from the VZERO and T0 detectors, coming after  $0.9 \mu\text{s}$  after the collision. The L0 signals strobe fast detectors and initiate the BUSY signal for all detectors in a used detector partition. Partitioning allows for the simultaneous operation of different groups of ALICE detectors. Signals in L1 have higher latency up to  $6.5 \mu\text{s}$  due to the longer computing time or propagation time. The L1 trigger setups transfer the event data to a multi-event buffer. The final CTP level corresponds to the TPC drift time and processes that occurred after  $100 \mu\text{s}$  from the beginning of the interaction, since the TPC is the slowest ALICE detector. After reaching the signal from the L2 trigger, the CTP sends the data to the Data Acquisition system (DAQ) and the High-Level Trigger system (HLT). In addition to that, the CTP can suppress the background using data from the LHC bunch-filling scheme.

## 2.3 Tracking Procedure

The reconstruction of the tracks in the central barrel is based on the points of the trajectory space measured by ITS and TPC [109]. Figure 2.4 describes the tracking procedure as

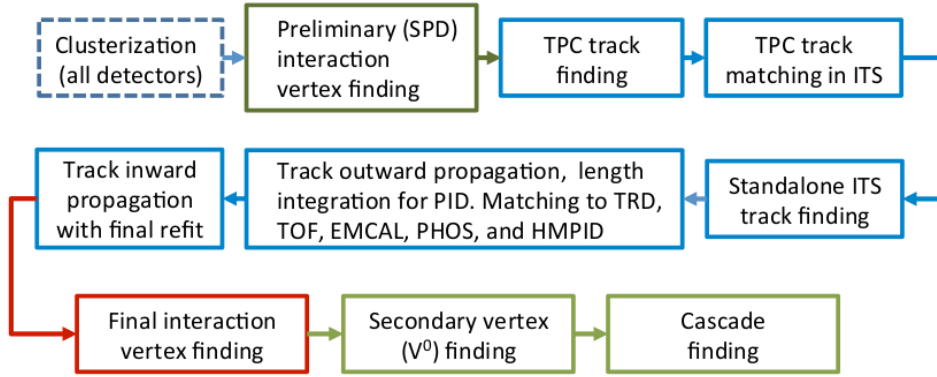


Figure 2.4: Event reconstruction flow. Taken from [103]

a sequence of steps. In the first step, each detector performs a clustering of measured hits. Then the cluster information from the SPD is used in a preliminary vertex reconstruction. The tracking procedure utilizes a Kalman filter with the inward-outward-inward scheme:

- The first tracking step starts with searching for short-track candidates at the larger TPC radii and proceeds radially inwards. After reaching the TPC edge, the algorithm matches TPC tracks to SPD clusters, which results in a refit of tracks.
- In the next step, the remaining ITS clusters are used to form standalone ITS tracks. The track reconstruction in this step proceeds radially outwards. This procedure builds tracks through the dead regions of the TPC or below the TPC momentum cutoff, resulting in an extension of the  $p_T$  track spectrum towards low momenta.
- During this step, a track is assigned with the most probable mass (determined by the TPC) and prolonged by additional points from TRD, TOF, EMCAL, PHOS, and HMPID. This information is used for PID purposes, not for track fitting.
- The final step includes the recalculation of primary vertices and the search for secondary decay vertices or cascade decays.

The performance of the ALICE tracking will be discussed in more detail in Section 3.2.

# Chapter 3

## Analysis of b jets

In this analysis, I have measured the inclusive  $p_T$  spectrum of jets associated with b-hadron fragmentation (b jets). The b-jet production was studied in collisions of p-Pb and pp at  $\sqrt{s_{NN}} = 5.02$  TeV. In my analysis, the tagging of b-jet candidates was based on the Secondary Vertex (SV) method [18]. This method uses the track constituents of a jet to construct possible candidates for displaced secondary decay vertices of b hadrons. In parallel, an independent analysis of b jets was performed on the same data sets based on the Impact Parameter (IP) method [110]. The author of this latter analysis was Hadi Hassan. Both analyzes were carried out independently and resulted in fully corrected inclusive  $p_T$  differential cross-section spectra of b jets and the corresponding nuclear modification factor. Although the IP method generally provided better b-jet tagging efficiency, the SV method was shown to be more stable at low  $p_T$ . Subsequently, the results of both approaches were combined using the BLUE method [111] for the final results to improve accuracy. The SV analysis was conducted in cooperation with Robert Vértési, Ashik Sheikh, and Filip Křížek. Therefore, wherever applicable, I will specify the contribution of the other members of our analysis team.

### 3.1 Data sets and Monte Carlo productions

#### Measured data

The analysis was performed on data recorded by ALICE during the runs with collisions of p-Pb and pp at  $\sqrt{s_{NN}} = 5.02$  TeV that took place in 2016 and 2017, respectively. The beam energies in the proton and lead beams during p-Pb collisions were asymmetric. The proton had 4 TeV, and the nucleons in the lead ions had 1.59 TeV. This resulted in a rapidity shift of the nucleon-nucleon center-of-mass system by  $\Delta y = 0.465$  in the direction of the proton beam in the laboratory frame.

Events were triggered using the ALICE minimum bias (MB) trigger, which requires

a coincidence of time signals from the VZERO-A and VZERO-C scintillator arrays. The integrated luminosities of the analyzed MB events were as follows  $\mathcal{L}_{\text{pPb}} = 298 \pm 11 \mu\text{b}^{-1}$  and  $\mathcal{L}_{\text{pp}} = 18.9 \pm 0.4 \text{ nb}^{-1}$  corresponding to  $6.24 \times 10^8$  and  $9.68 \times 10^8$  MB events, respectively. A larger fraction of the data (65.4% and 56% for pp and p–Pb, respectively) was taken with the so-called FAST L1 trigger cluster, and the rest was recorded with the “slow trigger”. The FAST trigger cluster performs data-taking with excluded ITS SDD due to its large maximum drift time  $\approx 4.3 - 6.3 \mu\text{s}$  [12]. In the “slow trigger” regime, the SDD was included in the data-taking. However, to have a data set compatible in data quality with the data triggered by the FAST trigger cluster, track reconstruction was performed without considering SDD data.

The position of the primary vertex along the beam direction was constrained to  $|z_{\text{vtx}}| < 10 \text{ cm}$ , providing uniform detector coverage in terms of acceptance and tracking efficiency in TPC and ITS. The primary vertex distributions in the pp and p–Pb collisions after event selection can be seen in Figure 3.1. The event selection further removed pile-up events, which were of the order of 1% and 0.5% for pp and p–Pb collisions, respectively [18]. Pile-up events were identified with multiple primary vertices, which were reconstructed using an algorithm that utilizes track segments formed by SPD clusters [13]. The total number of events in each collision system before and after event selection is quoted in Table 3.1.

Data set	$N_{\text{events}}^{\text{raw}}$	$N_{\text{events}}^{\text{phys.sel.}}$
pp	$1.2 \cdot 10^9$	$9.2 \cdot 10^8$
p–Pb	$8.4 \cdot 10^8$	$6.1 \cdot 10^8$

Table 3.1: The number of minimum bias events in the measured data sample before ( $N_{\text{events}}^{\text{raw}}$ ) and after ( $N_{\text{events}}^{\text{phys.sel.}}$ ) vertex and pileup cuts for pp and p–Pb collisions systems. Note that the quoted number of events after the selection is lower than the number of MB events reported in the text, because for the latter we also accounted for the MB events without reconstructed vertex.

### Monte Carlo productions

Measured data are affected by various instrumental effects that influence the purity and efficiency of the b-jet tagging, the energy scale of reconstructed jets, etc. These effects were investigated by means of Monte Carlo (MC) simulations in which particles from the initial generator (particle-level simulation) are transported through a detailed GEANT3 model [112] of the ALICE detector, digitized to raw hits, which were subsequently reconstructed as real data (detector-level simulation). These detector-level simulations are tuned to reflect the status of the ALICE detector at the time of data taking.

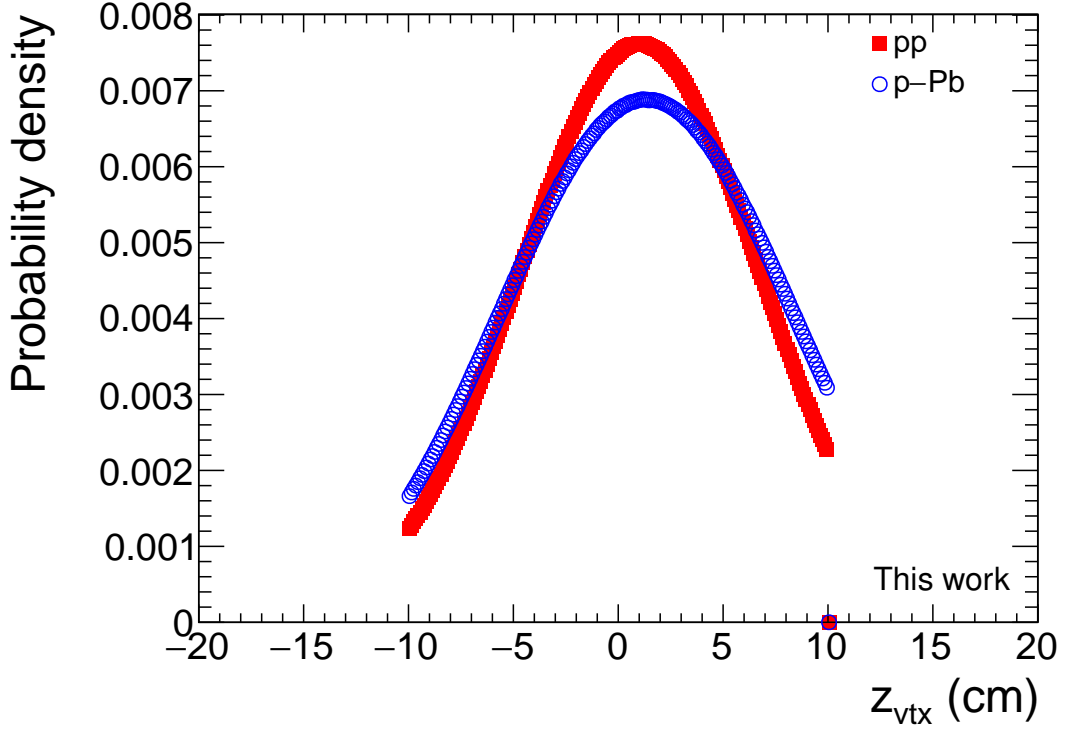


Figure 3.1: Distributions of primary event vertices along the beam axis after applying event selection and pile-up rejection criteria for pp and p–Pb data.

The MC simulations for the p–Pb system were based on PYTHIA 6 [113] events embedded into the underlying events simulated by the EPOS event generator [114]. Similarly, the simulations of the pp system were based on the PYTHIA 8 Monash tune [115]. The data were simulated in three variants: default settings, enhanced b-quark production, and enhanced c-quark production. To generate events in phase-space regions with low statistics, the data sets were generated in six  $p_{T,\text{hard}}$  bins, which restrict the range of transverse momenta in the rest frame of the scattering process; see Table 3.2. The range is set by the PYTHIA parameters `PhaseSpace:pTHatMin` and `PhaseSpace:pTHatMax`. The contributions of these bins were then summed up as weighted by the cross-section of the bin divided by the corresponding number of trials [29].

Consecutive bin number	1	2	3	4	5	6
$p_{T,\text{hard}}$ (GeV/ $c$ )	5–10	10–18	18–30	30–50	50–70	$\geq 70$

Table 3.2: List of  $p_{T,\text{hard}}$  bins used in the MC simulation of the pp and p–Pb data.

However, this weighting sometimes gives a large weight to rare events that emerge in low  $p_T$  hard bins. Therefore, it was necessary to introduce a function that rejects events in which jets have an unusually large transverse momentum compared to  $p_{T,\text{hard}}$  of the

given event. The function rejected events for which  $p_{T,\text{ch,jet}} > 4 \times p_{T,\text{hard}}$ . The resulting effect on the jet  $p_T$  spectrum can be seen in Figure 3.2.

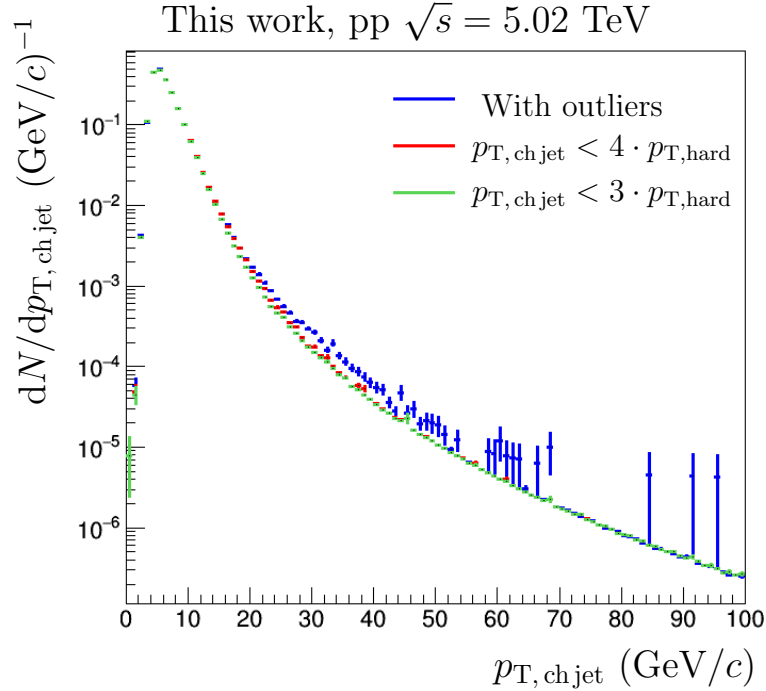


Figure 3.2: Simulated detector-level inclusive  $p_T$  spectrum of charged-particle anti- $k_T$  jets with  $R = 0.4$  obtained with and without outlier rejection in pp collisions at  $\sqrt{s_{\text{NN}}} = 5$  TeV. Jet  $p_T$  is corrected for the mean underlying event (UE) density.

Additional simulations were used to estimate instrumental effects on jets reconstructed in p–Pb. These simulations used only the PYTHIA part of the original PYTHIA+EPOS event. Such a configuration allows one to separate instrumental effects from momentum smearing due to underlying events. For this purpose, a dedicated analysis task was written.

## 3.2 Jet reconstruction

The analysis uses so-called hybrid tracks [13] to achieve azimuthal uniformity in the tracking at midrapidity. Hybrid tracks consist of good quality global tracks with at least one hit in the SPD and complementary tracks without SPD signals (3.5% of the sample). The momentum resolution of complementary tracks is improved by constraining the track origin to the primary vertex. The described data set has a tracking efficiency for primary tracks with  $p_T > 1$  GeV/ $c$  between 70–86% depending on the transverse momentum, and the momentum resolution is about 0.7% at  $p_T = 1$  GeV/ $c$ , 1.6% at  $p_T = 10$  GeV/ $c$ , and 4% at  $p_T = 50$  GeV/ $c$ , for more details, see Fig. 3.3. Track momenta smearing due to the finite momentum resolution of the detector was estimated

from the corresponding covariance matrix element of the Kalman fit and was expressed as the relative  $p_T$  resolution,  $\sigma(p_{T,\text{track}}^{-1})/p_{T,\text{track}}^{-1}$ . The right panel of Fig. 3.3 demonstrates the mean value of the relative  $p_T$  resolution as the function of  $p_{T,\text{track}}$ . The data was fitted with a function  $a + b \cdot \sqrt{p_T + c}$ , where  $a$ ,  $b$ , and  $c$  are parameters and  $p_{T,\text{track}}$  is in GeV/ $c$ .

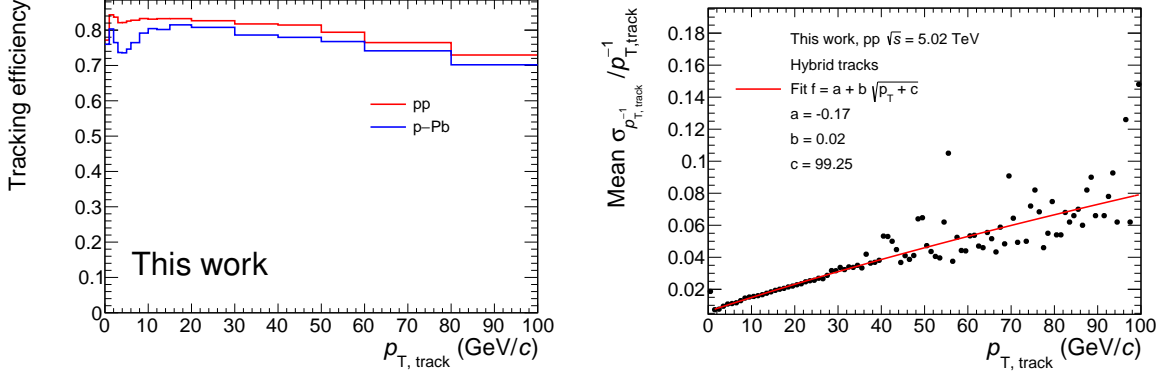


Figure 3.3: Left: Track reconstruction efficiency for pp and p-Pb data as a function of track  $p_T$ . Right: The mean relative resolution of  $p_T^{-1}$  for hybrid tracks as a function of track  $p_T$  in pp collisions. The red line represents a fit of the trend.

Jets were reconstructed using the infrared and collinear safe anti- $k_T$  algorithm implemented in the FastJet package. As input to the jet clustering algorithm, we have used a sample of charged-particle tracks with transverse momentum greater than 0.15 GeV/ $c$  and pseudorapidity constrained to  $|\eta_{\text{track}}| < 0.9$ . Momentum determination for such tracks is considered not accurate enough [4]. Jets having a constituent with  $p_{T,\text{track}} > 100$  GeV/ $c$  were rejected from further analysis. Jet four-momentum is calculated using the boost-invariant  $p_T$  recombination scheme assuming zero mass of all particles [34]. The jet resolution parameter was set to  $R = 0.4$ , providing the jet cone large enough to include most of the initial parton momentum [116]. Pseudorapidity of jets was constrained to  $|\eta_{\text{jet}}| < 0.9 - R = 0.5$  to select only the jets fully comprised in the TPC acceptance. Finally, the reconstructed transverse momentum of the jets was corrected for the underlying event using equation (1.8), where  $\rho$  was estimated on an event-by-event basis using the estimator introduced by CMS [117].

The tagging of b-jets was based on good quality tracks having at least one hit in the SPD. For such tracks, the spatial resolution of the impact parameter with respect to the primary vertex is better than 75  $\mu\text{m}$  for charged-particle tracks with transverse momentum  $p_T > 1$  GeV/ $c$  and better than 20  $\mu\text{m}$  for tracks with  $p_T > 20$  GeV/ $c$  [13].

### 3.3 b-jet tagging based on secondary vertex reconstruction

Most common beauty hadrons such as  $B^0$ ,  $B^\pm$ , and  $\Lambda_b$  have relatively long lifetimes due to the weak character of the corresponding decay process. Considering the high initial momentum of particles produced in a hard-scattering process, this feature of the beauty hadrons results in a significant displacement between the secondary and primary vertices, and it can be used for b-jet tagging.

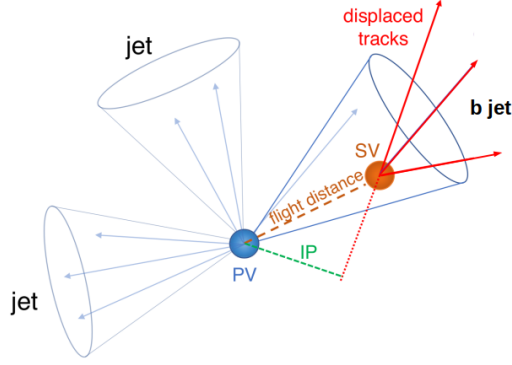


Figure 3.4: The sketch of the secondary vertex. Taken from [118] and changed.

Figure 3.4 shows a sketch of a jet that contains constituents from a secondary decay vertex. The Secondary Vertex (SV) algorithm identifies the candidates for the b-jet by imposing selection criteria on the most displaced SV selected from all possible 3-prong combinations made of constituents in the given jet. The selection criteria are as follows:

- The cut on the maximal value of the dispersion of the reconstructed secondary vertex,  $\sigma_{SV} = \sqrt{d_x^2 + d_y^2 + d_z^2}$ . Here  $d_i$  denotes the distance of the closest approach of each SV constituent to the SV itself.
- The cut on the minimal value of the significance of the SV displacement from the primary vertex,  $SL_{xy} = L_{xy}/\sigma_{L_{xy}}$ . Here,  $L_{xy}$  is the distance between the primary vertex and the SV, and  $\sigma_{L_{xy}}$  is the uncertainty of the  $L_{xy}$  measurement.  $L_{xy}$  is a Lorentz invariant quantity since it is measured in the plane transverse to the beam direction.

The default parameters of the b-jet tagging were chosen as follows:  $SL_{xy} > 7$  and  $\sigma_{SV} < 0.03$  cm. Variations of  $SL_{xy}$  and  $\sigma_{SV}$  cuts were considered for the corresponding systematic uncertainty estimates and were selected as follows:  $0.02 \leq \sigma_{SV} \leq 0.05$  cm,  $3 \leq SL_{xy} \leq 9$ . Limits on the variation of the tagging cuts were selected using the  $p_{T,jet}$ -integrated 2D distributions of  $SL_{xy}$  and  $\sigma_{SV}$  shown in Figure 3.5. The light-flavor



component is significantly suppressed when  $SL_{xy} > 3$ , while the probability of charm suppression decreases slowly. On the one hand, by setting the  $SL_{xy}$  limit to higher values, we can achieve a higher purity of the selected b-jet sample, but at the same time, it will result in a lower tagging efficiency. Similar behavior is also visible for  $\sigma_{SV}$ , suggesting setting the upper limit to  $\sigma_{SV} \leq 0.05$  cm. Figure 3.6 compares the  $SL_{xy}$  and  $\sigma_{SV}$  distributions for jets having different flavors as obtained from the PYTHIA 8 Monash simulations at the detector level [115] after constraining  $SL_{xy}$  and  $\sigma_{SV}$ . It is also interesting to inspect how the mistagging efficiency of c-jets and light-flavor jets correlates with the b-jet tagging efficiency; see Fig. 3.7. Note that the corresponding distributions for both systems are essentially the same. Figure 3.8 presents the mistagging efficiency of c jets and light-flavor jets versus b-jet tagging efficiency. The default SV tagging cuts were set to suppress the yield of the light-flavor jets by a factor of 100.

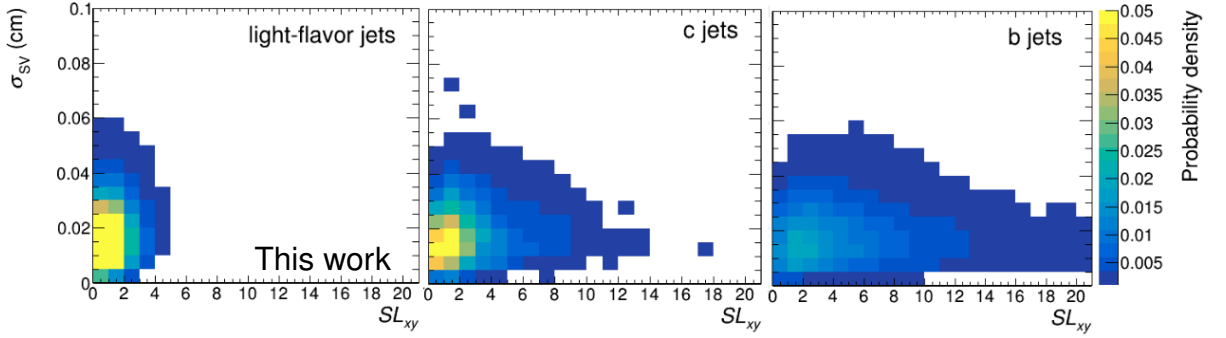


Figure 3.5: 2D distributions of  $SL_{xy}$  and  $\sigma_{SV}$  obtained with PYTHIA+EPOS simulation of p–Pb events.  $p_{T, \text{ch, jet}}$  was set to be larger than 10 GeV/c.

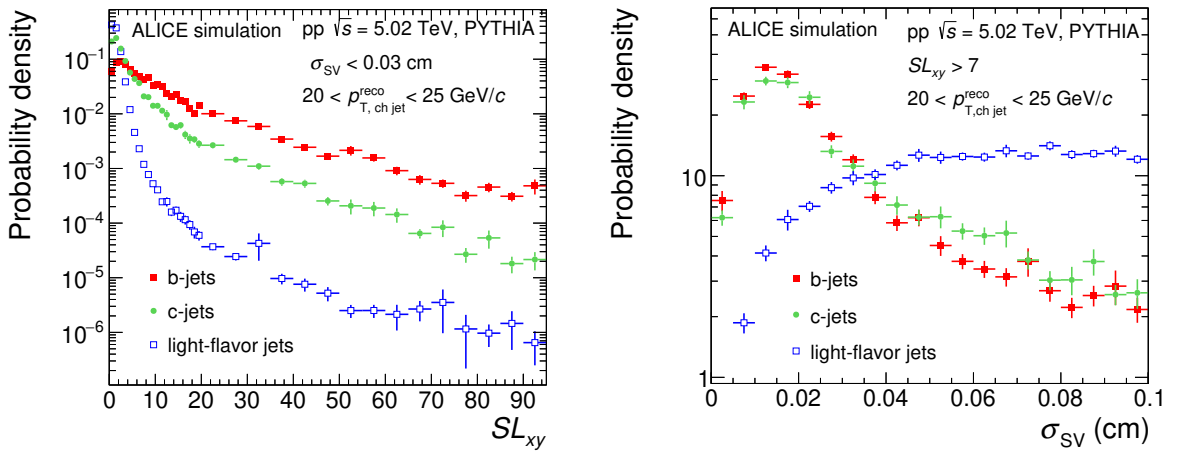


Figure 3.6: Distributions of the tagging discriminators used in the SV method for the pp data set,  $SL_{xy}$  (left), and  $\sigma_{SV}$  (right) for b jets, c jets, and light-flavor jets as obtained from the PYTHIA 8 detector-level MC simulation.

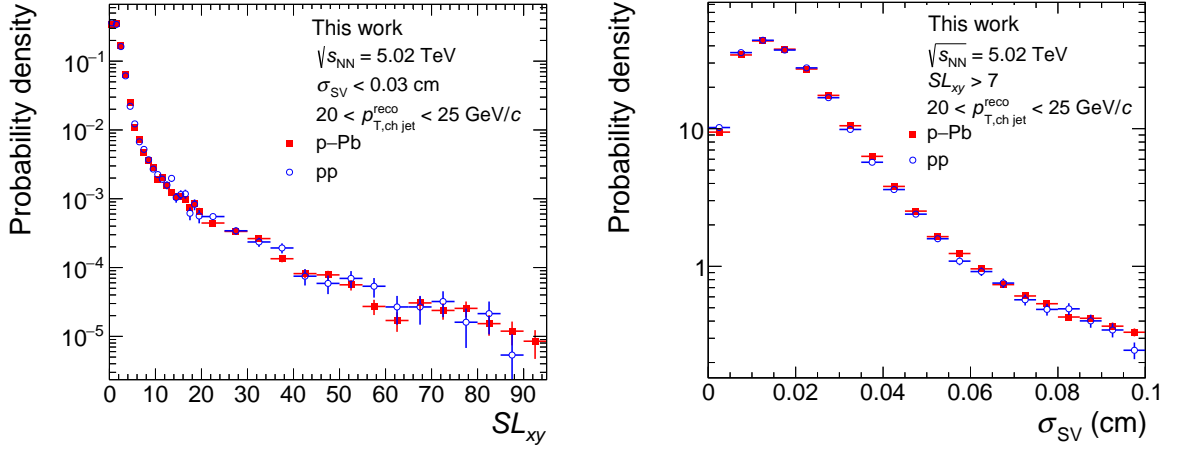


Figure 3.7: Distributions of the tagging discriminators used in the SV method,  $SL_{xy}$  (left) and  $\sigma_{SV}$  (right), for the measured pp and p-Pb data sets.

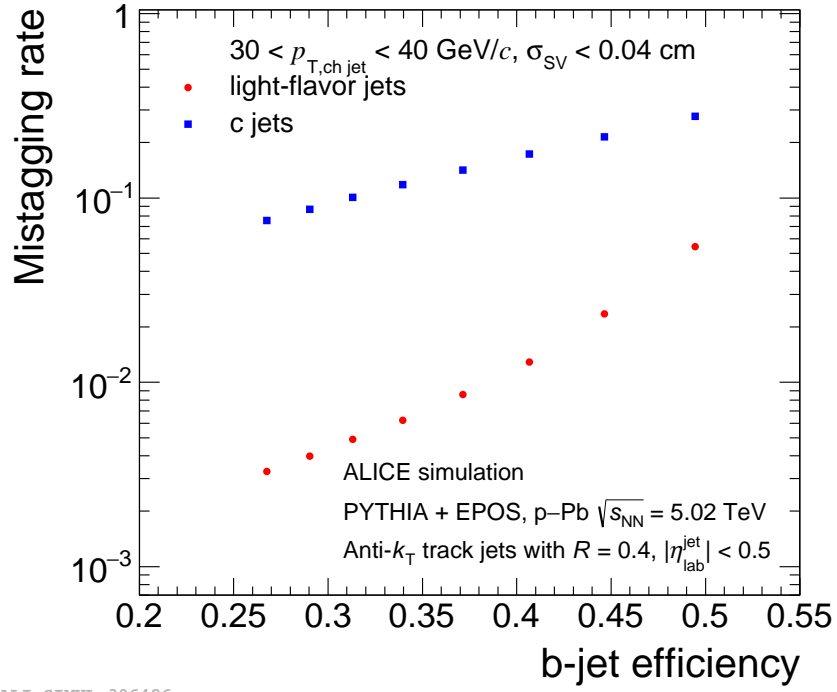


Figure 3.8: Misstaging efficiency of light flavor and c jets versus b-jet selection efficiency when varying  $SL_{xy}$  from 3 to 10 while keeping  $\sigma_{SV}$  fixed at 0.04 cm for p-Pb data in  $30 < p_{T, \text{ch jet}} < 40$  GeV/c. The results obtained with the default selection of b-jet tagging cuts are analogous.

Figures 3.10 and 3.9 illustrate how the ALICE detector-level PYTHIA+EPOS simulation reproduces the measured distributions of the SV tagging observables in the p–Pb system. In general, the MC reproduces well the trend seen in the real data. An analogous comparison for the pp system looks similar.

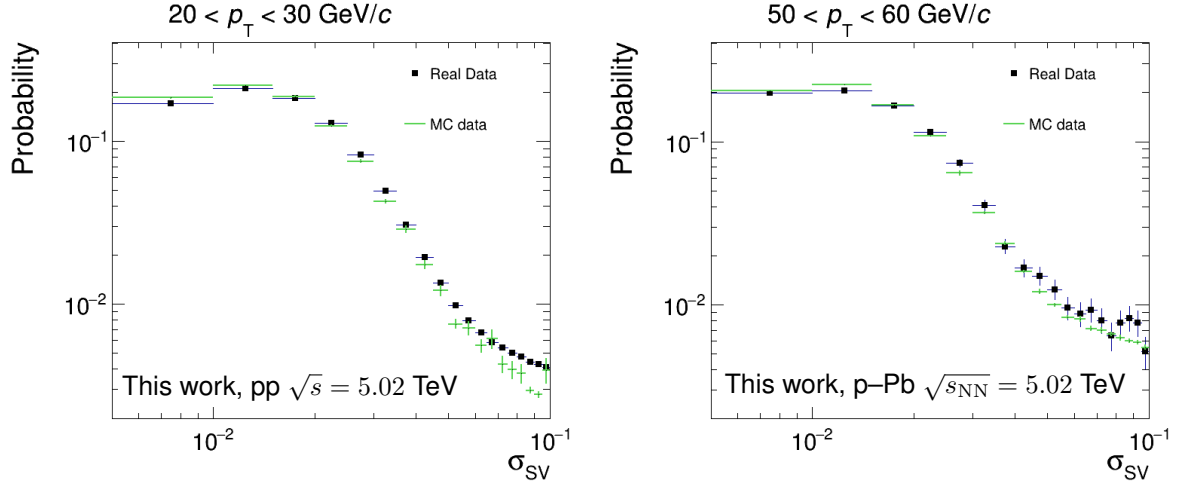


Figure 3.9: Distribution of  $\sigma_{SV}$  in real p–Pb data and the corresponding PYTHIA+EPOS detector-level simulations for three different jet  $p_T$  bins. Distributions are normalized by their integral. In this comparison,  $SL_{xy}$  was set to be greater than 7.

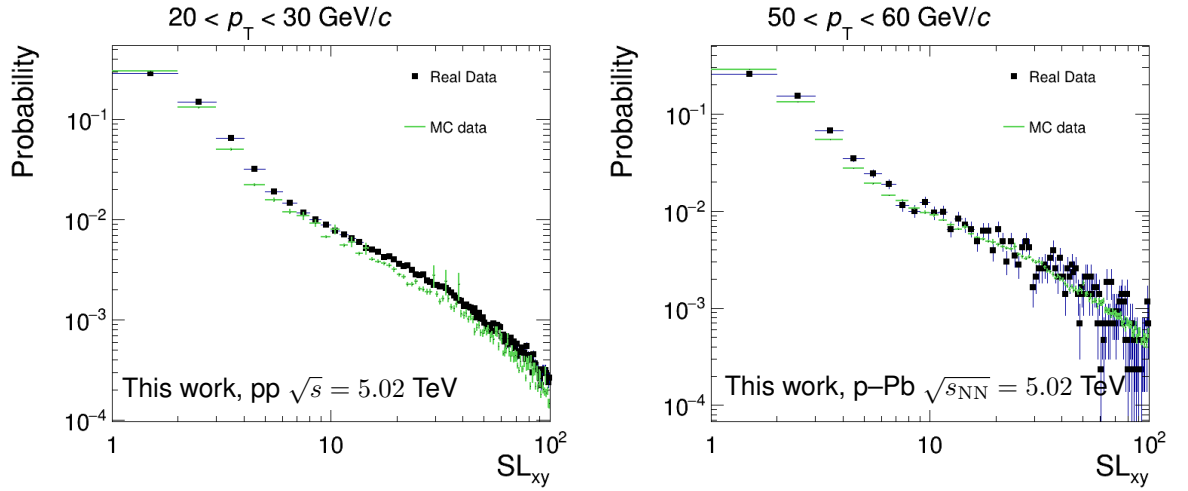


Figure 3.10: Distribution of  $SL_{xy}$  in real p–Pb data and the corresponding PYTHIA+EPOS detector-level simulations for three different jet  $p_T$  bins. Distributions are normalized by their integral. In this comparison,  $\sigma_{SV}$  was set to be less than 0.03 cm.

### 3.4 Correction of the b-jet spectra on tagging performance

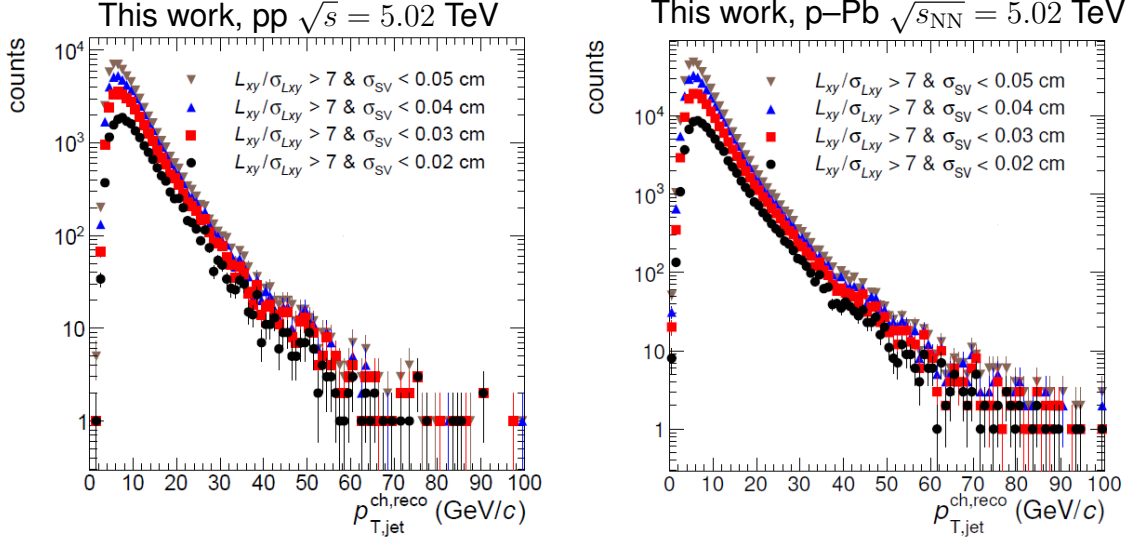


Figure 3.11: Example of raw inclusive  $p_T$  spectra of b jets for  $SL_{xy} > 7$  when varying the selection on  $\sigma_{SV}$ . Left: pp data, right: p–Pb data. Size of bin is 1 GeV/c.

The raw inclusive  $p_T$  spectra of b-jet candidates in pp and p–Pb for  $SL_{xy} > 7$  and  $\sigma_{SV} < 0.02 - 0.05$  cm are shown in Figure 3.11. In the next step, the spectra were rebinned according to the scheme:

$$(0, 5, 10, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 100) \text{ GeV}/c \quad (3.1)$$

in order to provide statistically significant content in each bin. Raw spectra of b-jet candidates,  $dN^{\text{tagged}}/dp_{T,\text{chjet}}^{\text{reco}}$ , were corrected for the SV tagging efficiency,  $\varepsilon_b$ , and b-jet purity,  $P_b$ , using the formula:

$$\frac{dN_{\text{raw}}^{\text{bjet}}}{dp_{T,\text{chjet}}^{\text{reco}}} = \frac{dN^{\text{tagged}}}{dp_{T,\text{chjet}}^{\text{reco}}} \cdot \frac{P_b}{\varepsilon_b}. \quad (3.2)$$

The resulting raw spectra of b jets were later corrected for jet energy scale smearing due to background fluctuations and instrumental effects using the unfolding technique.

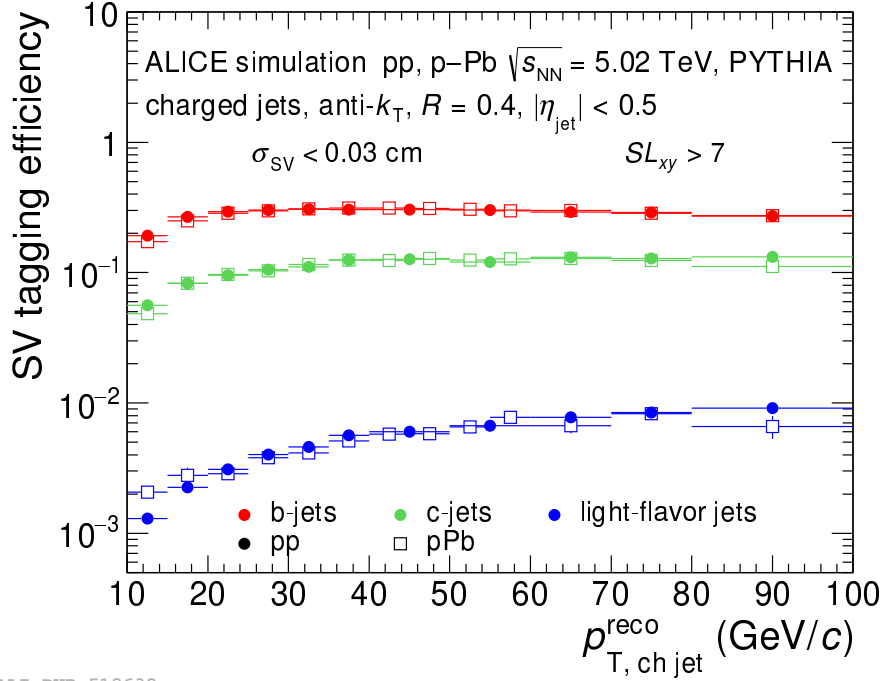
#### 3.4.1 b-jet tagging efficiency

In this analysis, a jet was defined as a b jet if a particle-level b-hadron momentum points to the jet cone. An analogous definition was also used for c jets; the remaining jets were declared light-flavor jets. The tagging efficiency is the probability that a given tagging algorithm correctly identifies a jet originating from a b quark as a b jet. Similarly, one

can also define mistagging efficiency as the probability that a jet originating from a charm quark or a light-flavor parton will be falsely tagged as a b jet. The SV tagging efficiency was calculated using the equation:

$$\varepsilon_i(p_{T,\text{ch, jet}}^{\text{reco}}) = \frac{N_i^{\text{tagged}}}{N_i^{\text{total}}}. \quad (3.3)$$

Here  $i$  is a jet flavor (c, b, or light-flavor),  $N_i^{\text{tagged}}$  is the number of jets tagged with the flavor  $i$ , and  $N_i^{\text{total}}$  is the total number of jets with the flavor  $i$  before the b-jet tagging. Jet flavor information was obtained from the PYTHIA and the PYTHIA+EPOS simulations for pp and p-Pb data. The efficiency of tagging and mistagging as a function of the jet  $p_T$  is shown in Figure 3.12.



ALI-PUB-518632

Figure 3.12: Tagging and mistagging efficiencies of beauty, charm, and light-flavor jets for the SV method in pp (solid markers) and p-Pb (open markers) collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV. The data are plotted as a function of jet transverse momentum.

### 3.4.2 b-jet tagging purity

The SV tagging algorithm can falsely identify light or charm flavor jets as b jets. Given the higher production cross-section of these jets, they can significantly contaminate the b-jet candidate's spectra. Hence, the b-jet candidate spectrum needs to be corrected for purity by Eq. (3.1). The purity of the b-tagged sample is defined as a fraction of true b-jet candidates among the total number of tagged jets:

$$P_b(p_{T,\text{chjet}}^{\text{reco}}) = \frac{N_{\text{bjet}}^{\text{tagged}}}{N^{\text{tagged}}}. \quad (3.4)$$

Here  $N_{\text{bjet}}^{\text{tagged}}$  is the number of tagged true b jets and  $N^{\text{tagged}}$  is the total number of tagged jets.

In this work, the purity of the b-jet tagging was calculated using two methods. The first one is based on particle-level jet spectra calculated by POWHEG, and the second one is based on a data-driven template fit method. The advantage of the latter method is that it does not depend on the corresponding production cross-sections. Unfortunately, the data-driven template fit method demands large statistics, and, in our case, it fails to provide purity for jets with transverse momentum greater than 40 GeV/ $c$ . In turn, the method based on POWHEG simulation can give reasonable purity estimations over a wide momentum range. However, its drawback is that the results of POWHEG calculations significantly vary depending on the model parameters, such as the quark masses and the renormalization and factorization scales, which cannot be accessed directly. Therefore, the purity estimation by POWHEG is based on the choice of these model settings. Therefore, the set of plausible POWHEG variations was determined by comparing the purities obtained with both methods. Ashik Sheikh generated the POWHEG data that we have been using.

### **b-jet tagging purity from the POWHEG-based method**

In this method, b-jet tagging purity in the pp data was assessed with the help of particle-level POWHEG HVQ simulations [119] with the CTEQ6M PDF set [120]. In the case of the p-Pb system, the EPS09 nPDF set [121] was applied, and the rapidity shift was considered. The POSitive Weight-Hardest Emission Generator (POWHEG) provides a Next-to-Leading order calculation of  $p_T$  spectra of b and c jets at the particle level. Subsequently, the spectra were folded with a response matrix that accounted for the jet  $p_T$  smearing due to local background fluctuations and instrumental detector effects to obtain the corresponding detector-level spectra. These response matrices will be described in detail in Chapter 3.6. The remaining spectrum of light-flavor jets was assessed by subtracting the smeared b-jet and c-jet spectra from the measured raw inclusive untagged jet  $p_T$  spectrum. The resulting purity of the selected b-jet sample is then equal to:

$$P_b = \frac{\varepsilon_b N_b}{\varepsilon_b N_b + \varepsilon_c N_c + \varepsilon_{\text{lf}}(N_{\text{incl}} - N_b - N_c)}, \quad (3.5)$$

where  $\varepsilon_b, \varepsilon_c, \varepsilon_{\text{lf}}$  are the tagging and mistagging efficiencies of the beauty, charm, and light-flavor jets, respectively. The estimated raw spectra of the b and c jets are denoted  $N_b$  and  $N_c$ , respectively. Finally, the raw inclusive spectrum of the measured untagged jets

is labeled  $N_{\text{incl}}$ .

### b-jet tagging purity from the data-driven template fit method

An invariant mass distribution of the tagged 3-prong SV depends on the jet flavor. In our case, we have reconstructed the invariant mass distribution, assuming that all 3 track constituents are pions. The measured SV invariant mass distribution was parameterized by a linear combination of simulated invariant mass distributions of the most displaced SV found in b, c and light-flavor jets:

$$n_{\text{SV}}(m_{\text{inv}}) = P_{\text{b}} \cdot T_{\text{b}}(m_{\text{inv}}) + P_{\text{c}} \cdot T_{\text{c}}(m_{\text{inv}}) + P_{\text{lf}} \cdot T_{\text{lf}}(m_{\text{inv}}), \quad (3.6)$$

$$\text{where } 1 = P_{\text{b}} + P_{\text{c}} + P_{\text{lf}}. \quad (3.7)$$

Here  $n_{\text{SV}}(m_{\text{inv}})$  is the fraction of SVs measured in the given invariant mass bin and  $T_{\text{b}}, T_{\text{c}}, T_{\text{lf}}$  are the b-jet, c-jet, and light-flavor-jet SV invariant mass templates. These MC templates were obtained from the MC data using the same topological cuts imposed on the measured data. The templates were normalized to have the integral equal to unity. The parameters  $P_{\text{b}}, P_{\text{c}}$ , and  $P_{\text{lf}}$  give a fraction of the given quark flavor in the measured distribution; thus  $P_{\text{b}}$  is the purity of the selected b jet sample. The value of these parameters can be acquired by fitting the measured SV invariant mass distribution with Eq. (3.6). To solve this task, we have designed a dedicated fit function using the TMinuit package [122], which allows us to fit the SV invariant mass distributions in the considered jet  $p_{\text{T}}$  bins and to account for statistical errors of the MC templates. This puts less weight on particular invariant mass bins in the MC, where statistical errors were still high even after the outlier removal procedure. The following  $\chi^2$  function was minimized by TMinuit:

$$\chi^2 = \sum_{\text{i}} \frac{(n_{\text{SV}} - P_{\text{b}} \cdot T_{\text{b}} - P_{\text{c}} \cdot T_{\text{c}} - P_{\text{LF}} \cdot T_{\text{LF}})^2}{\sigma_{n_{\text{SV},\text{i}}}^2 + (\sigma_{T_{\text{b},\text{i}}} \cdot P_{\text{b}})^2 + (\sigma_{T_{\text{c},\text{i}}} \cdot P_{\text{c}})^2 + (\sigma_{T_{\text{lf},\text{i}}} \cdot P_{\text{lf}})^2}, \quad (3.8)$$

where  $\sigma_{n_{\text{SV},\text{i}}}$  is the statistical uncertainty in the i-th bin of the measured spectrum of the secondary vertex invariant mass. The statistical uncertainties  $\sigma_{T_{\text{b},\text{i}}}, \sigma_{T_{\text{c},\text{i}}}, \sigma_{T_{\text{lf},\text{i}}}$  correspond to each jet flavor and are obtained from the MC templates. The sum runs over invariant mass bins.

Two examples of data-driven TMinuit template fits in the measured pp and p-Pb data sets can be seen in Fig. 3.13. The data are well described by the sum of the MC templates in the  $p_{\text{T, ch jet}}^{\text{reco}}$  range from 10 to about 40 GeV/c, but for larger momenta, the fit procedure fails due to lack of statistics.

We have carried out a closure test to benchmark the accuracy of the purity estimates

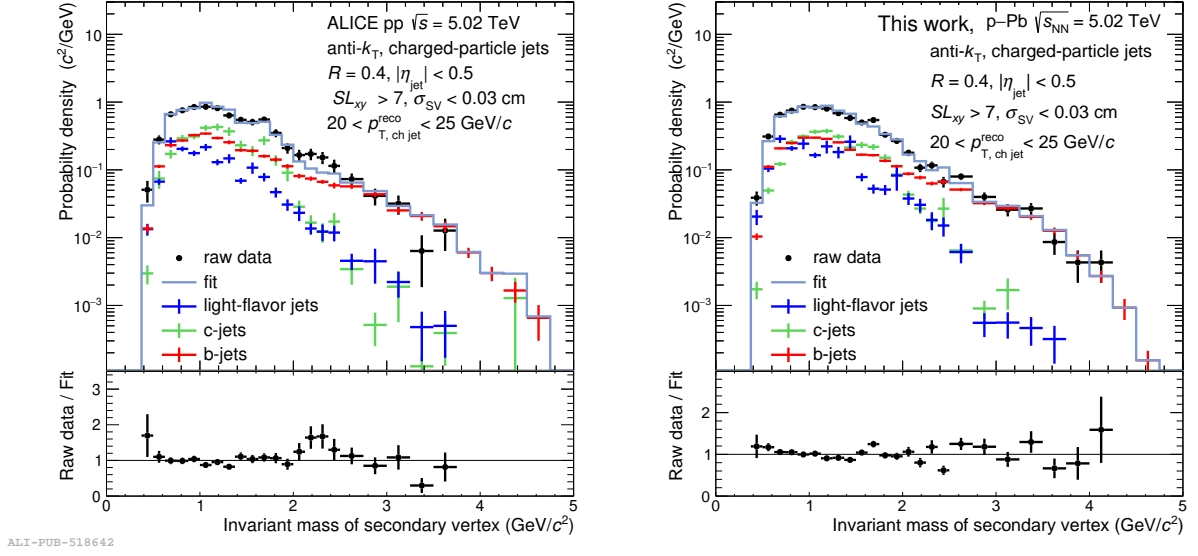


Figure 3.13: Invariant mass distribution of the most displaced secondary vertex in jets with  $20 < p_{T, \text{ch jet}}^{\text{reco}} < 25 \text{ GeV}/c$  tagged with the default selection  $SL_{xy} > 7$  and  $\sigma_{\text{SV}} < 0.03 \text{ cm}$ , for pp and p-Pb collisions. The measured data (black points) are fitted with the detector-level MC templates corresponding to beauty, charm, and light-flavor jets to assess purity of the b-jet candidate sample.

provided by this method. In this test, we used the detector-level simulation to obtain the b-jet purity using the data-driven template fit method by Eq. (3.8) and compared the result with the true b-jet purity obtained from Eq. (3.6). Templates and raw detector-level invariant mass spectra were taken from statistically independent samples. The obtained distribution of the b-jet purity as a function of  $p_{T, \text{ch jet}}^{\text{reco}}$  is shown in Fig. 3.14. The data-driven method provides purity compatible with the true value for both collision systems, indicating that this method can be applied to real data.



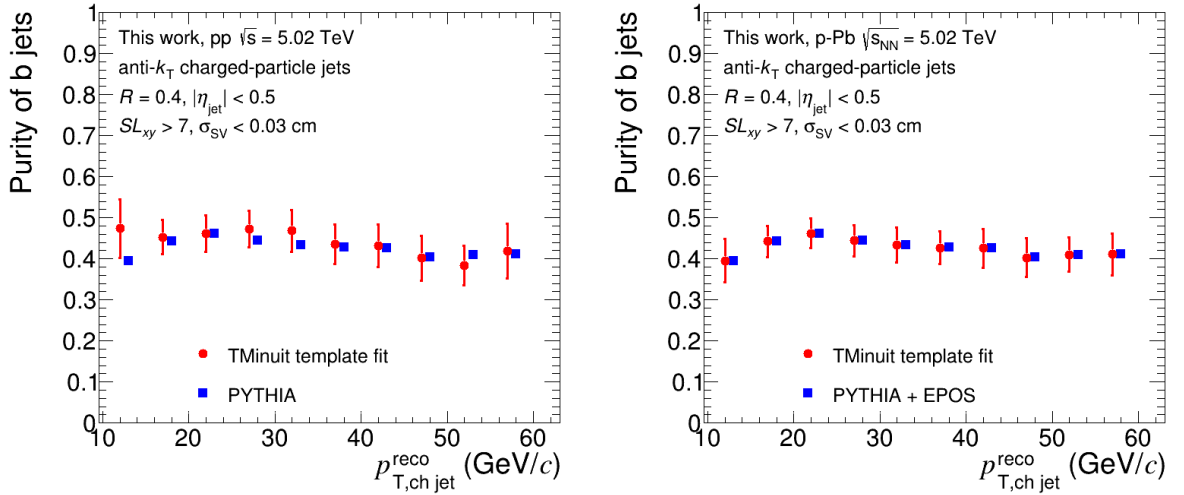


Figure 3.14: Results of the closure test comparing b-jet sample purity for pp (left) and p–Pb (right) collisions estimated by the template fit method applied to the detector-level MC simulation and the true purity. The input detector-level MC data are based on PYTHIA (pp) and PYTHIA+EPOS (p–Pb) simulation.

### Hybrid method

The selection of POWHEG tunes, which provide statistically compatible purity estimates with the data-driven method, was made by Robert Vértési. In his approach, he quantified the difference between the corresponding b-jet purities in terms of  $\chi^2$ . For the given choice of  $SL_{xy}$  and  $\sigma_{SV}$ , the  $\chi^2$  was calculated as follows:

$$\chi_i^2 = \sum_{\text{bins}} \frac{(P_{b,i}^{\text{POWbc}} - P_{b,i}^{\text{data driven}})^2}{\sigma_{\text{POWbc},i}^2 + \sigma_{\text{data driven},i}^2}. \quad (3.9)$$

Here, the sum runs over all  $p_T$  bins where the data-driven method converged; the index  $i$  corresponds to a specific combination of the b-tagger setting:  $\sigma_{SV} < V$  and  $SL_{xy} > U$ , where  $V \in \{0.02, 0.03, 0.04, 0.05\}$  cm and  $U \in \{5, 6, 7, 8, 9\}$ . These  $\chi_i^2$  values were then combined to form a discriminator, which can be expressed as follows:

$$\frac{\chi^2}{N} = \frac{\sum_i \chi_i^2}{\sum_i N_i}. \quad (3.10)$$

The number of degrees of freedom  $N$  in the  $\chi^2$  distribution is equal to the number of data points used to calculate  $\chi^2$ .

Table 3.3 gives the default value of the POWHEG model parameter and the range considered in the calculations  $\chi^2$ . The values obtained from  $\chi^2/N$  analysis for the variations considered in the POWHEG settings are shown in Fig. 3.15. The smallest  $\chi^2/N$  value was obtained using the POWHEG setting where the factorization scale is 0.5, the

renormalization scale is 1, and the masses of b and c quarks have their default value. Figure 3.16 compares the purity of the b-jet sample obtained for the default tagging cuts with the data-driven template fit method and the POWHEG simulation-based approach. The POWHEG settings having the value  $\chi^2/N < 10$  are considered plausible (marked green). The other POWHEG settings that did not pass  $\chi^2/N$  selection are marked gray. The selected set of plausible POWHEG-based purities was further used to estimate the corresponding systematic uncertainty.

parameter	explanation	values
$m^c$	charm quark mass (GeV/c)	1.3 <b>1.5</b> 1.7
$m^b$	beauty quark mass (GeV/c)	4.5 <b>4.75</b> 5.0
$\mu_f^c$	charm factorization scale	0.5 <b>1</b> 2
$\mu_r^c$	charm renormalization scale	0.5 <b>1</b> 2
$\mu_f^b$	beauty factorization scale	0.5 <b>1</b> 2
$\mu_r^b$	beauty renormalization scale	0.5 <b>1</b> 2

Table 3.3: List of POWHEG simulation settings with the default values highlighted.

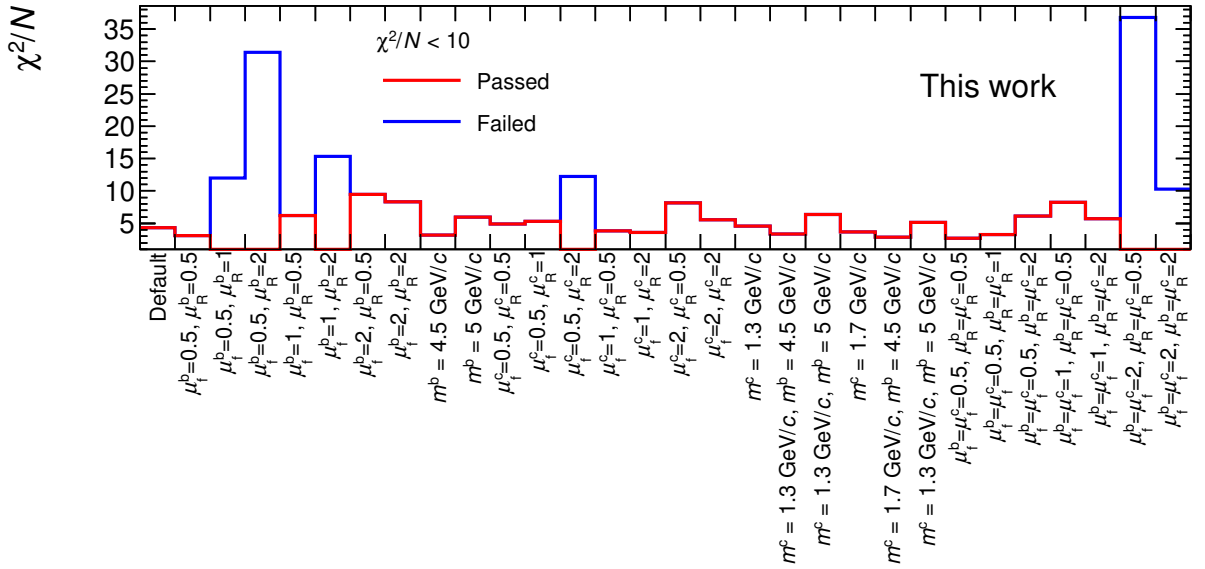


Figure 3.15: The  $\chi^2/N$  values as obtained for each POWHEG setting variation. The  $x$ -axis labels provide the POWHEG parameters which were varied. The other parameters were considered to be fixed at their default values. The settings consistent with the data-driven purities are marked red. The settings failing the test are blue.

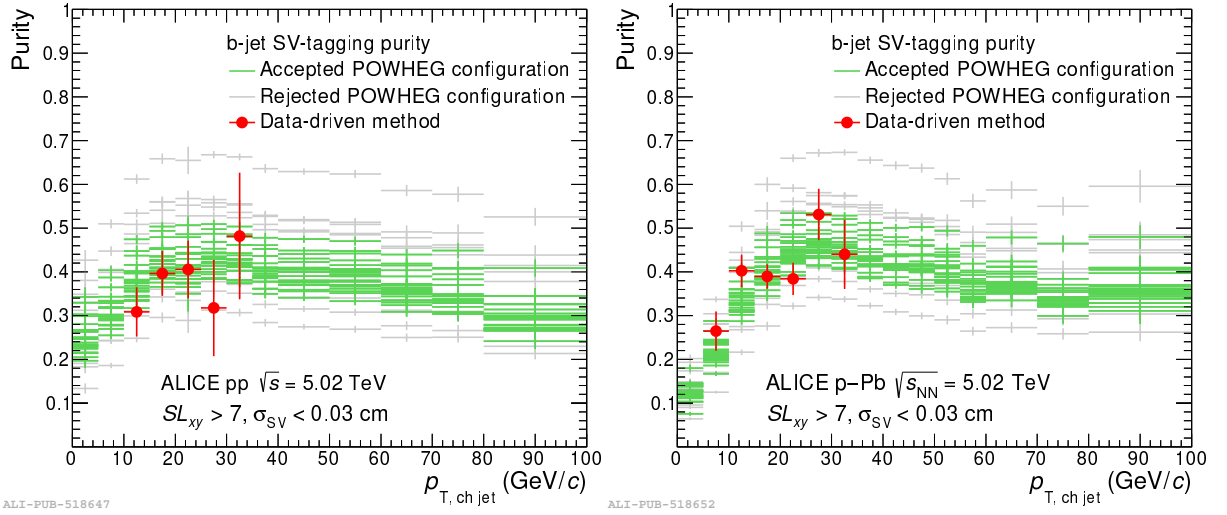


Figure 3.16: Purity of the b-jet candidates selected with the SV method when using the default tagging cuts for pp (left) and p–Pb (right) collisions. The purity was estimated with the data-driven template fit method (red points) and the POWHEG-simulation approach. The POWHEG scale variations accepted by the statistical analysis are colored green, the rejected ones are gray.

### 3.5 Unfolding

The b-jet spectrum corrected for SV tagging efficiency and purity is still affected by jet energy scale shift and smearing stemming from various instrumental effects due to finite detector efficiency and resolution and local background fluctuations with respect to the mean underlying event density. For a general continuous spectrum, one usually assumes that the smearing can be described as:

$$m(y) = \int A(x, y)t(x)dx. \quad (3.11)$$

Here  $t(x)$  is a true spectrum,  $m(y)$  is a measured smeared spectrum, and  $A(x, y)$  is some response function that incorporates the detector and background effects. In this analysis, the measured and true spectra correspond to the detector and particle-level  $p_{T, \text{ch jet}}$  spectra of the b jets, respectively. The finite bin size  $p_T$  of the measured data requires discretization of the above equation, which can be rewritten as:

$$m_j = \sum_i A_{ji}t_i. \quad (3.12)$$

Here,  $m_j$  and  $t_i$  are the content of  $j$ -th and  $i$ -th bin in the histograms representing the measured and true spectrum, and  $A$  denotes the response matrix. Assuming that the

inverse matrix  $A^{-1}$  exists, one can obtain the true spectrum as:

$$t_i = \sum_j A_{ij}^{-1} m_j. \quad (3.13)$$

Explicit calculation of the true spectrum with (3.13) requires finding the inverted response matrix  $A$ . However, the response matrix suffers from statistical fluctuations and is often singular; therefore, standard numerical linear algebra methods cannot find the inverted matrix [123]. Instead, more sophisticated approximate methods have to be used to acquire the fully corrected spectrum. Such methods include unfolding algorithms that apply regularization to suppress non-physical solutions. Basic unfolding methods are implemented in the RooUnfold [124] package. This work considers unfolding methods based on the Bayesian theorem [125] and the SVD decomposition [123].

## 3.6 Unfolding of b-jet spectra

### Raw spectrum

Figure 3.17 shows the raw spectrum corrected for efficiency and purity of b jets that were rebinned according to the following scheme to improve the stability of the unfolding procedure:

$$\{10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 100\} \text{ GeV}/c. \quad (3.14)$$

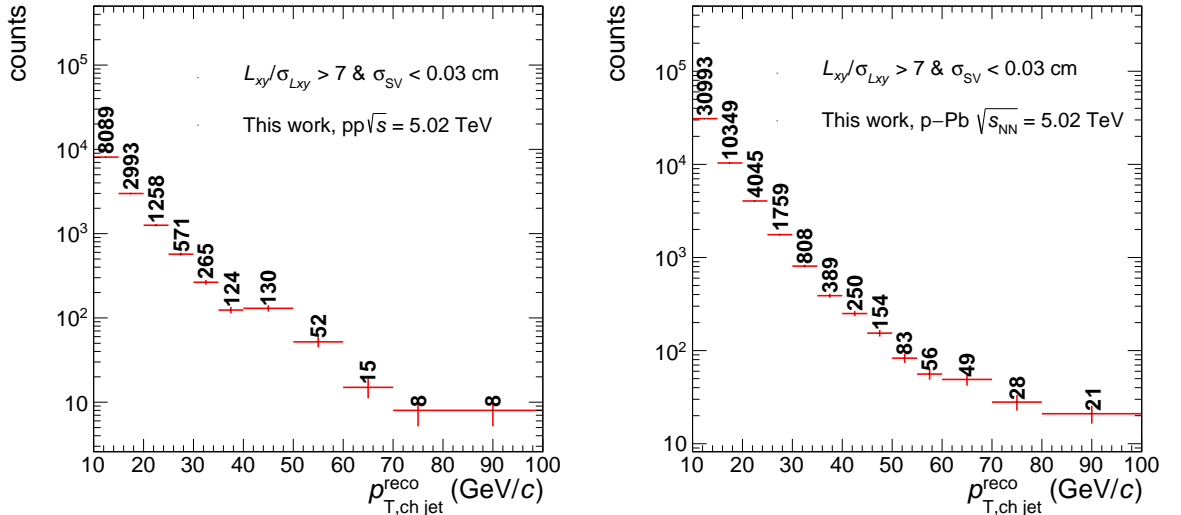


Figure 3.17: Rebinned raw spectrum of b jets as obtained for the pp (left) and p-Pb (right) data.

### Prior spectrum

Unfolding procedures need a guess of the expected true spectrum, the so-called prior. In this work, we have used as a primary analysis prior spectrum the b-jet spectrum generated by the particle-level PYTHIA+POWHEG simulation with the default POWHEG setting discussed in Chapter 3.4.2. The prior spectrum and the resulting unfolded solution have the same binning scheme, which was as follows:

$$\{0, 5, 10, 20, 30, 40, 50, 60, 70, 80, 100, 200\} \text{ GeV}/c. \quad (3.15)$$

The binning includes overflow and underflow bins to account for the migration of jet yield from the  $p_T$  regions that are not measured to the  $p_T$  region of the raw data.

### Response matrix

We assume that the response matrix can be decomposed into a product of two matrices that describe the smearing of the momentum due to the fluctuations of the underlying event ( $\delta p_T$ ) and instrumental effects ( $M$ ) [68]:

$$A(p_{T,\text{ch jet}}^{\text{part}}, p_{T,\text{ch jet}}^{\text{reco}}) = \delta p_T(p_{T,\text{ch jet}}^{\text{det}}, p_{T,\text{ch jet}}^{\text{reco}}) \times M(p_{T,\text{ch jet}}^{\text{part}}, p_{T,\text{ch jet}}^{\text{det}}). \quad (3.16)$$

Here  $p_{T,\text{ch jet}}^{\text{part}}$  denotes the transverse momentum of jets at the particle level,  $p_{T,\text{ch jet}}^{\text{det}}$  is the jet  $p_T$  smeared by instrumental effects, and  $p_{T,\text{ch jet}}^{\text{reco}}$  is the jet  $p_T$  smeared by local UE fluctuations and detector effects.

In the case of pp collisions, the instrumental response matrix was constructed from the detector-level simulation of PYTHIA 8 events with an enhanced b-jet sample. In the case of p–Pb, we used the analogous PYTHIA 6 data extracted from PYTHIA+EPOS events, see Section. 3.1. The response matrix was filled for jets where we found an exact geometrical match between a jet at the particle and detector levels. The angular distance of the corresponding jet axes was required to be less than 0.25 [126]. The instrumental matrix for the pp and p–Pb data can be found in Figure 3.18. The figure also shows the particle-level b-jet spectra used to normalize the response matrices. The response matrices obtained from both collision systems are compatible, as can be seen in the top panel of Figure 3.19, which illustrates projections of the response matrices on the  $p_{T,\text{ch jet}}^{\text{det}}$  axis in different  $p_{T,\text{ch jet}}^{\text{part}}$  bins. The bottom panels of Figure 3.19 show that the instrumental response matrices for b jets and light-flavor jets are similar.

The  $\delta p_T$  matrix, which describes jet  $p_T$  smearing due to local background fluctuations, was evaluated using the embedding method. In this approach, we selected events with an SV-tagged jet that met the standard topological cut in  $SL_{xy}$  and  $\sigma_{\text{SV}}$  and embedded a

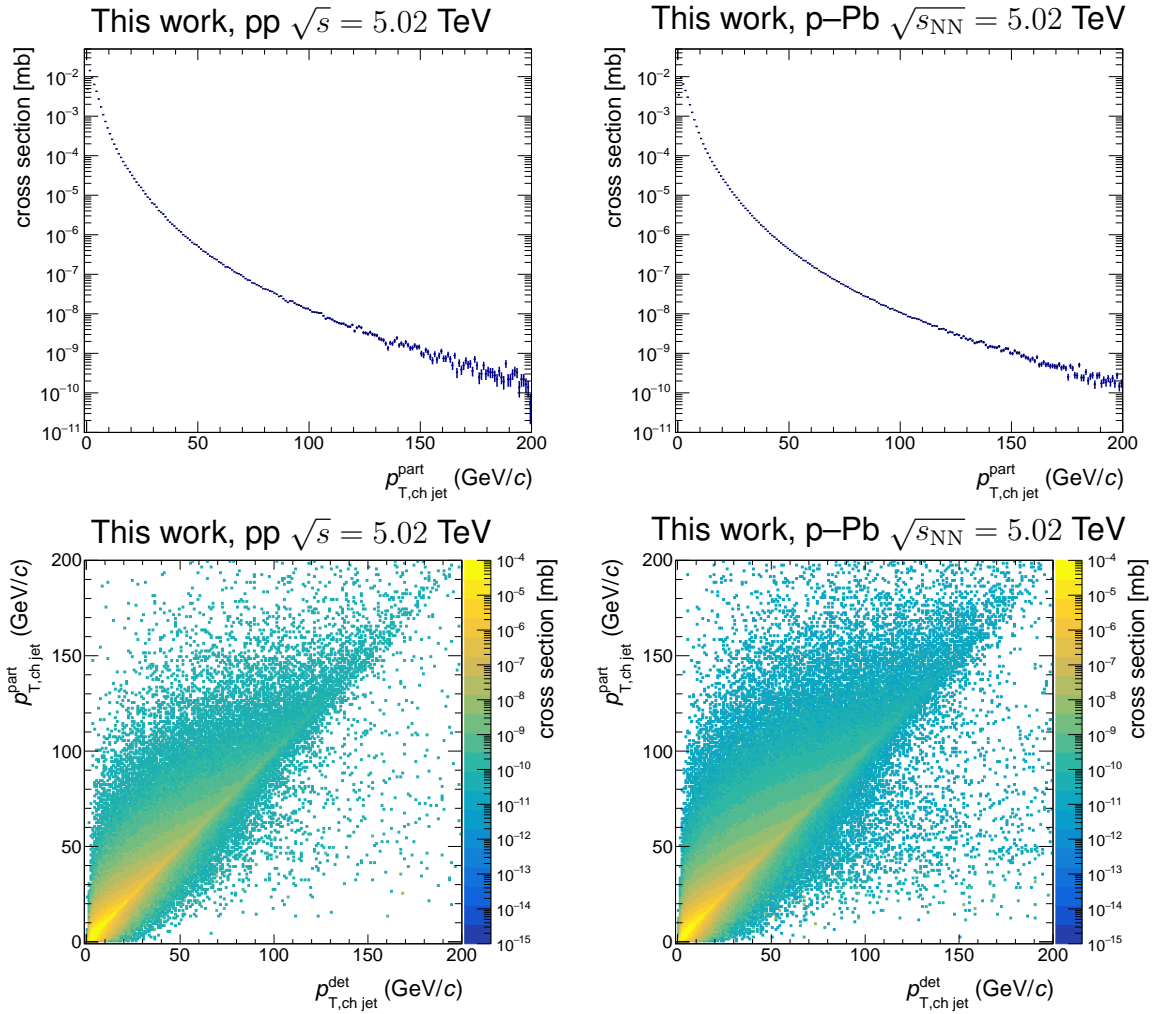


Figure 3.18: Top: Particle-level distribution of charged-particle anti- $k_T$   $R = 0.4$  b jets, which provides normalization for the instrumental response matrix. Bottom: Instrumental response matrix.

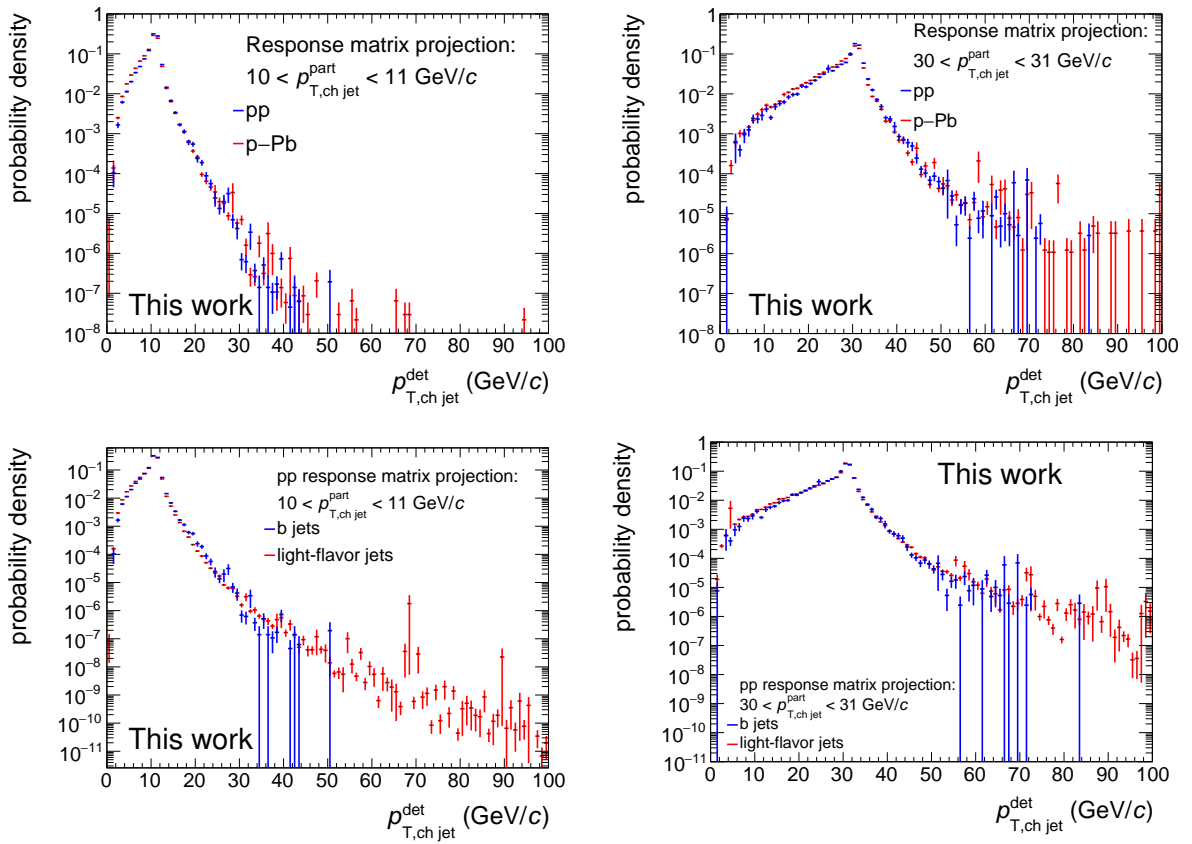


Figure 3.19: Top: Projections of the response matrix on the  $p_{T,chjet}^{det}$  axis for the pp and p-Pb data in different  $p_{T,chjet}^{part}$  bins. Bottom: Projections of the response matrix on the  $p_{T,chjet}^{det}$  axis for the b-jet enhanced and the default simulated pp data in different  $p_{T,chjet}^{part}$  bins.

single particle in such an event. The  $p_T$  of the embedded track was randomly sampled from a uniform distribution in the range of 0 to 100 GeV/ $c$ . The embedded track was placed perpendicularly to the SV tagged jet in the azimuth and had the same pseudorapidity. This geometrical configuration implies that the embedded track should be located in the region dominated by the underlying event. By rerunning the anti- $k_T$  jet clusterization procedure in a given event, a jet will be aggregated around this embedded particle. The magnitude of underlying event fluctuations with respect to the mean background can be estimated as:

$$\delta p_T = p_{T,\text{chjet}}^{\text{raw}} - \rho A_{\text{jet}} - p_{T,\text{emb}}. \quad (3.17)$$

Here  $p_{T,\text{chjet}}^{\text{raw}}$  is a transverse momentum of the charged-particle jet containing the embedded track,  $\rho$  is a background density estimator introduced by CMS [35], and  $p_{T,\text{emb}}$  is the transverse momentum of the embedded track.

The middle panels of Fig. 3.20 compare the  $\delta p_T$  distributions obtained for various choices of  $p_{T,\text{emb}}$ . The shape of the distribution slightly differs for  $p_{T,\text{emb}} < 10$  GeV/ $c$  and for  $p_{T,\text{emb}} > 10$  GeV/ $c$ . In order to account for these differences, we have considered different  $\delta p_T$  templates for different  $p_{T,\text{emb}}$  ranges. The  $\delta p_T$  templates were parameterized by a smooth fit function for  $\delta p_T > 5$  GeV/ $c$  to suppress statistical fluctuations. Below 5 GeV/ $c$  the template was based on the original histogram. Finally, the  $\delta p_T$  matrix is constructed from the  $\delta p_T$  histograms normalized by their integral by shifting them along the diagonal; see the bottom panel of Figure 3.20.

For systematic studies, we also estimated the size of local background fluctuations using an alternative random cone method. In this approach, we generated a cone with  $R_{\text{cone}} = 0.4$  in a random position in the  $\eta - \varphi$  plane in each event containing a tagged b-jet candidate. The cone was placed to avoid overlap with the b-tagged jet, with the leading and the subleading jets in the event. The momentum smearing is calculated using the formula:

$$\delta p_T^{\text{RC}} = \sum_{\text{tracks}} p_{T,\text{track}}^{\text{RC}} - \rho \pi R_{\text{cone}}^2, \quad (3.18)$$

where the sum runs over all tracks found inside the random cone. Only tracks with an angular distance from the cone center,  $\Delta R = \sqrt{(\eta_{\text{track}} - \eta_{\text{cone}})^2 + (\varphi_{\text{track}} - \varphi_{\text{cone}})^2}$ , less than  $R_{\text{cone}}$  are selected. The comparison of the original  $\delta p_T$  distributions obtained with both methods is shown in the upper panel of Figure 3.20. Both methods provide compatible  $\delta p_T$  distributions. Figure 3.21 shows the combined response matrix for the b jets calculated from Eq. (3.16). This matrix is used for our primary analysis.



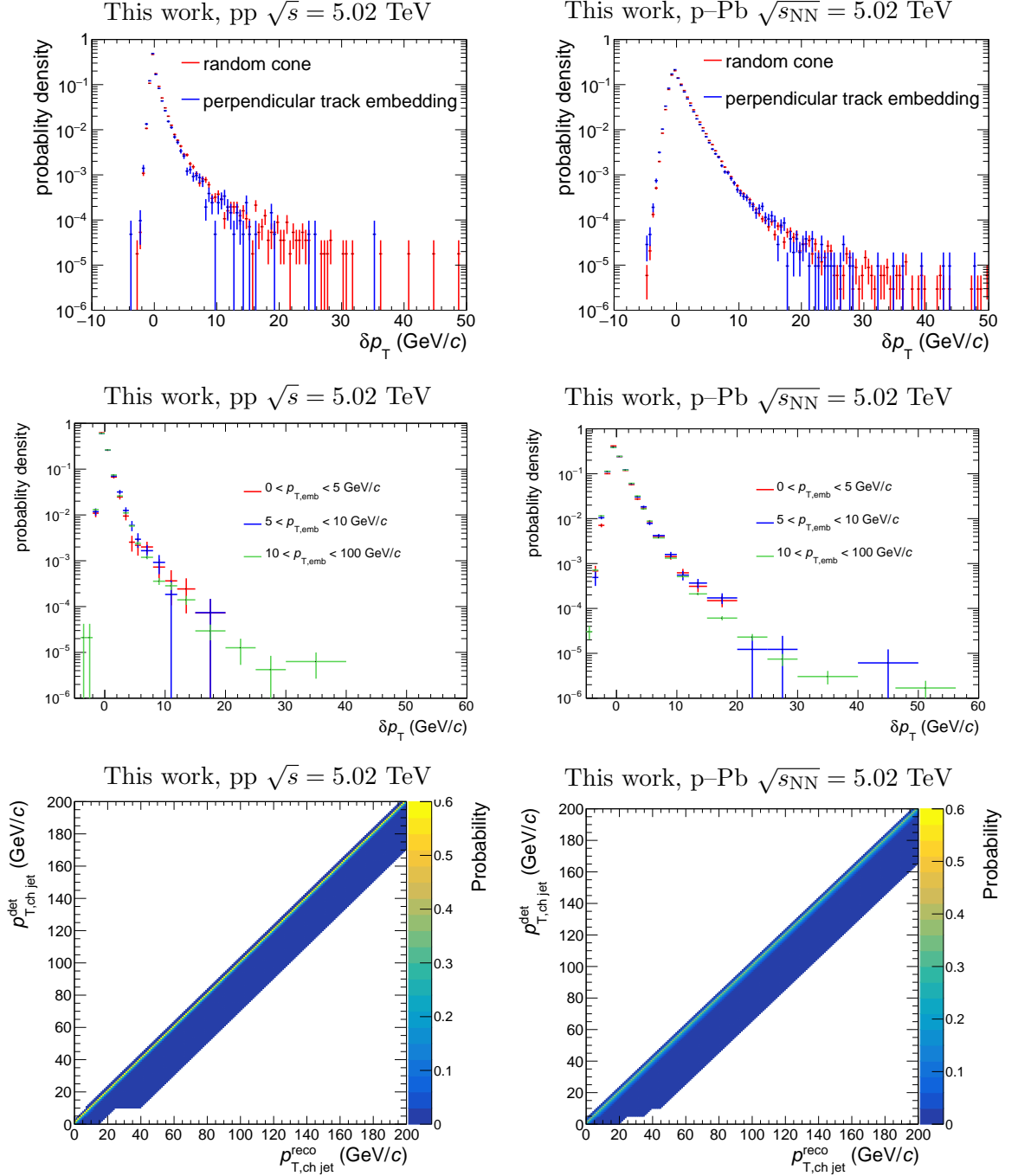


Figure 3.20: Top:  $\delta p_T$  distribution as obtained by the embedding and by the random cone method. Middle: Comparison of the  $\delta p_T$  distributions obtained for different ranges of embedded particle momentum. Bottom:  $p_T$  matrix obtained from embedded templates. Bin size is  $1 \text{ GeV}/c \times 1 \text{ GeV}/c$ .

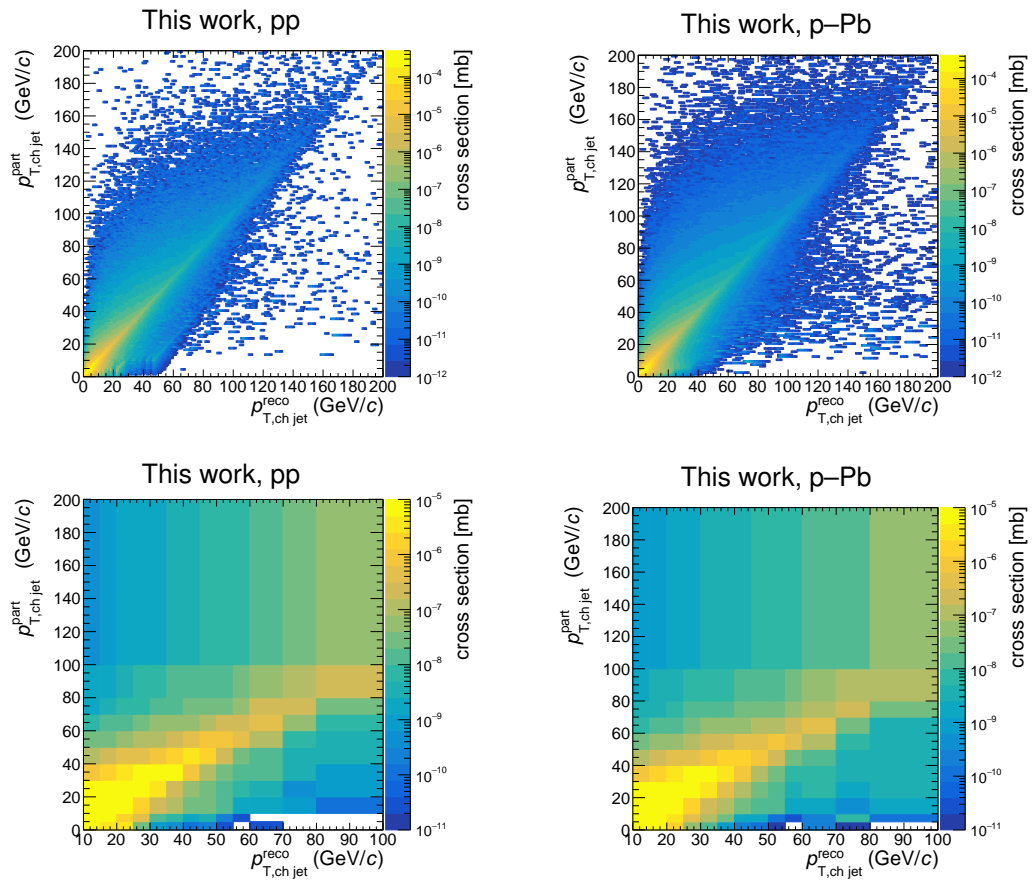


Figure 3.21: Top: Combined response matrix for b jets as obtained from Eq. (3.16). The matrix has a bin size of  $1 \text{ GeV}/c \times 1 \text{ GeV}/c$ . Bottom: The same matrix after rebinning with (3.14) and (3.15).

### 3.7 MC closure tests of the unfolding procedure

The closure test of the unfolding procedure verifies the self-consistency of the correction procedure using MC data. Namely, it tests whether the fully corrected detector-level spectrum of b-jet candidates is compatible with the true particle-level b-jet spectrum. The list of inputs for the closure test was the following:

- **Raw spectrum of b-jets** which was obtained from the detector-level PYTHIA simulation by applying the primary analysis SV tagging selections and correction on the SV tagging purity and efficiency. The raw spectrum was binned according to the binning scheme (3.14).
- **True spectrum of b-jets** which was the corresponding particle-level b-jet spectrum of the same PYTHIA sample rebinned according to the scheme (3.15).
- **Prior spectrum**, which was a statistically independent true spectrum of particle-level b jets.
- $\delta p_T$  distribution which was obtained from the detector-level PYTHIA simulation.
- **Instrumental response matrix** which was obtained from a statistically independent detector-level simulation with an enhanced b-quark fraction, i.e., the same matrix used for the real data unfolding.

The results of the closure test are shown in Figure 3.22. The raw spectrum was unfolded using the SVD unfolding method with the regularization parameter  $i = 5$  for both collision systems. The optimal choice of the regularization parameter was selected based on the  $|d_i|$  vector distribution [123], see the top right panel of Figure 3.22. The fully corrected spectrum is compatible with the true spectrum within  $\approx 5\%$  for the pp and p–Pb spectra. A subsequent test also shows that multiplication of the unfolded spectrum with the response matrix produces a spectrum compatible with the raw spectrum; see the bottom left panel of Fig. 3.22. A closure test performed for the Bayesian unfolding gives a similar level of agreement.

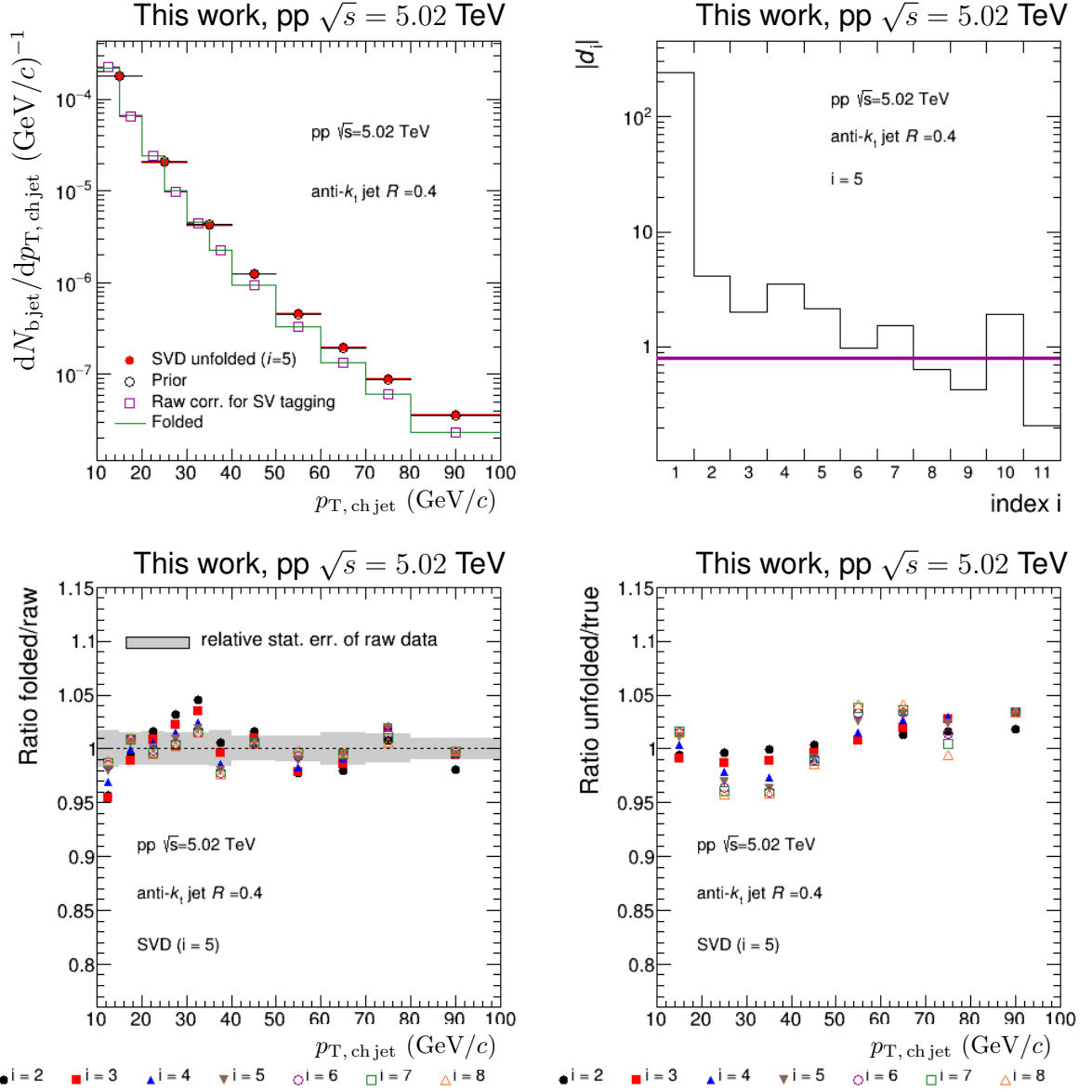


Figure 3.22: Results of the closure test when unfolding PYTHIA detector-level b-jet spectrum in pp collisions. Upper left: comparison of the raw, unfolded, refolded, and prior b-jet spectrum. Upper right:  $d_i$  vector components. The chosen optimal regularization parameter was 5. Bottom left: Ratio of the refolded unfolding solutions to the input raw spectrum corrected for SV tagging efficiency and purity. Bottom right: Ratio of the unfolded spectrum to the true spectrum. In the bottom panels, different markers correspond to different choices of the regularization parameter  $i$ .

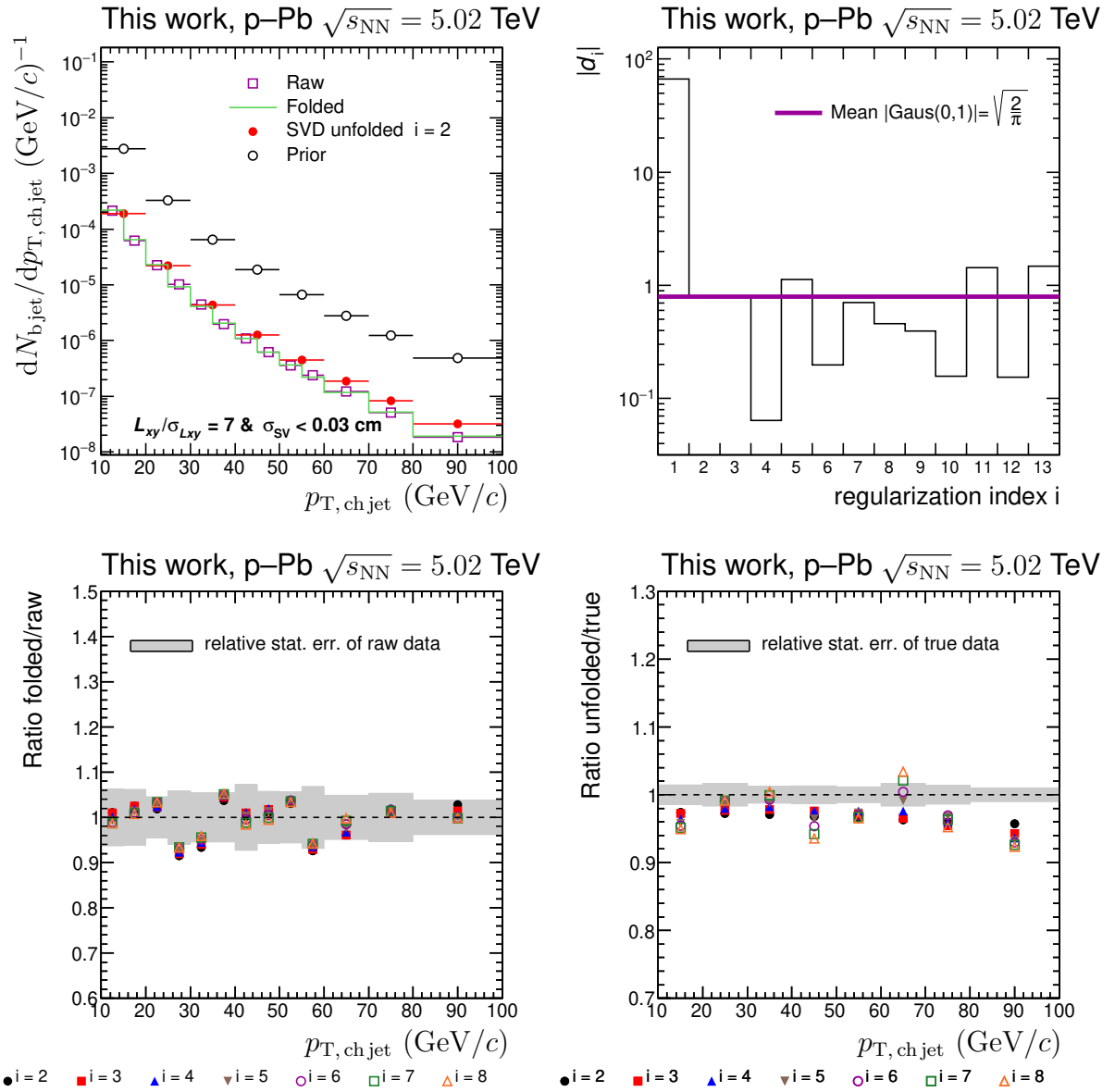


Figure 3.23: Results of the p-Pb SVD unfolding closure test using PYTHIA data. The optimal choice of the regularization parameter was 5. Meaning of the individual panels is analogous to Figure 3.22.

### 3.8 Unfolding of the measured b-jet spectrum

For the unfolding of the measured b-jet spectrum in pp and p-Pb, we have likewise used the SVD unfolding with the PYTHIA particle-level b-jet spectrum prior. The QA plots for the SVD unfolding with the primary analysis (PA) setting are shown in Figs. 3.24 and 3.25 for pp and p-Pb collisions, respectively. For both collision systems, the unfolding results converge for regularization parameters  $i$  greater than or equal to 4. Based on the  $|d_i|$  vector distribution, the optimal values of the regularization parameter were chosen to be 4 and 5 for pp and p-Pb, respectively. This choice of regularization parameters also provides a refolded spectrum compatible with statistical uncertainties with the raw input spectrum. The ingredients of the unfolding procedure (algorithm, prior, binning, regularization parameter) were varied to assess systematic uncertainties; more details about the calculation of systematic uncertainties can be found in Chapter 3.10.

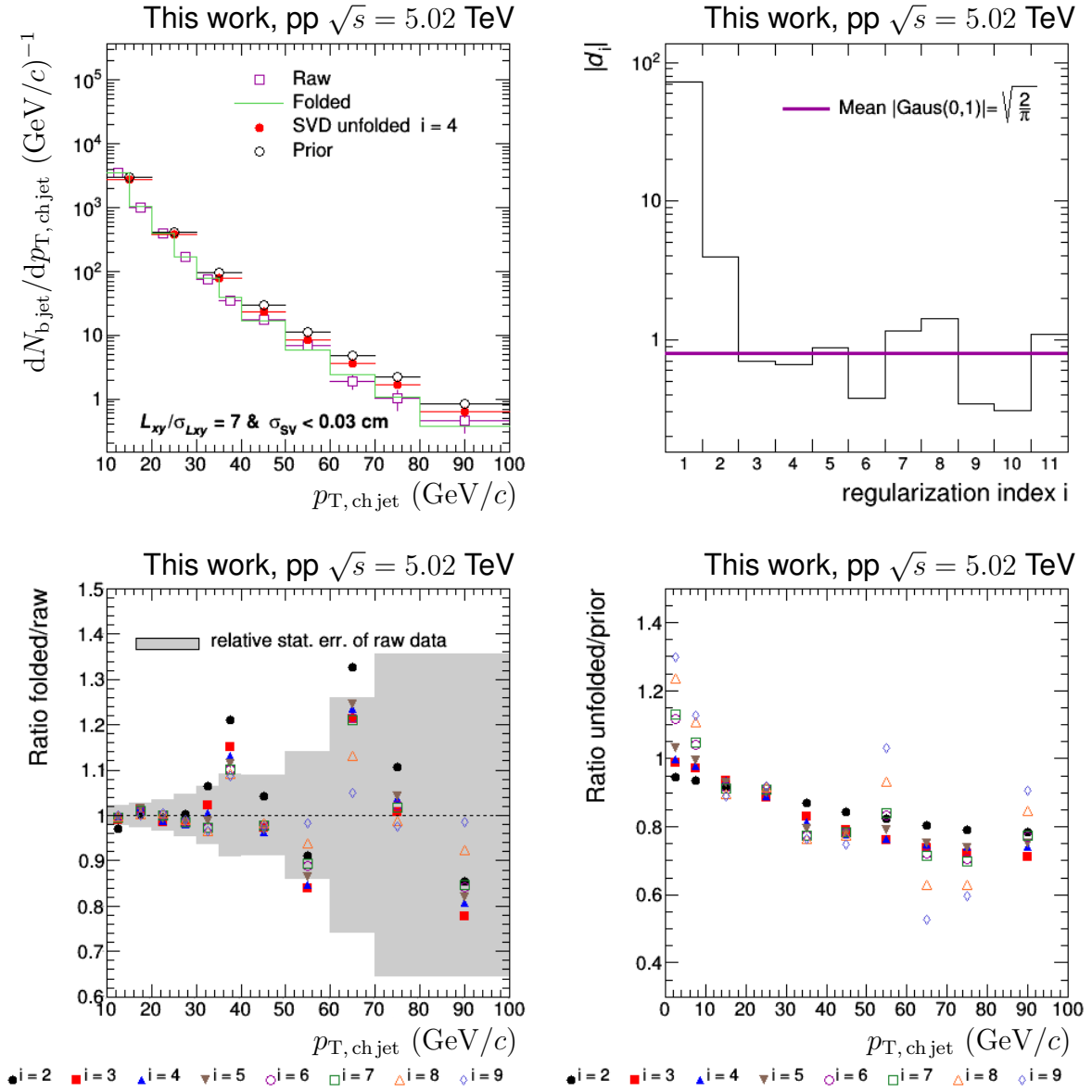


Figure 3.24: QA plots for the SVD unfolding of the measured pp b-jet spectrum. Meaning of the individual panels is analogous to Figure 3.22. The right bottom panel shows a ratio of the unfolded solutions to the prior spectrum. The primary analysis spectrum was obtained with  $i = 4$ .

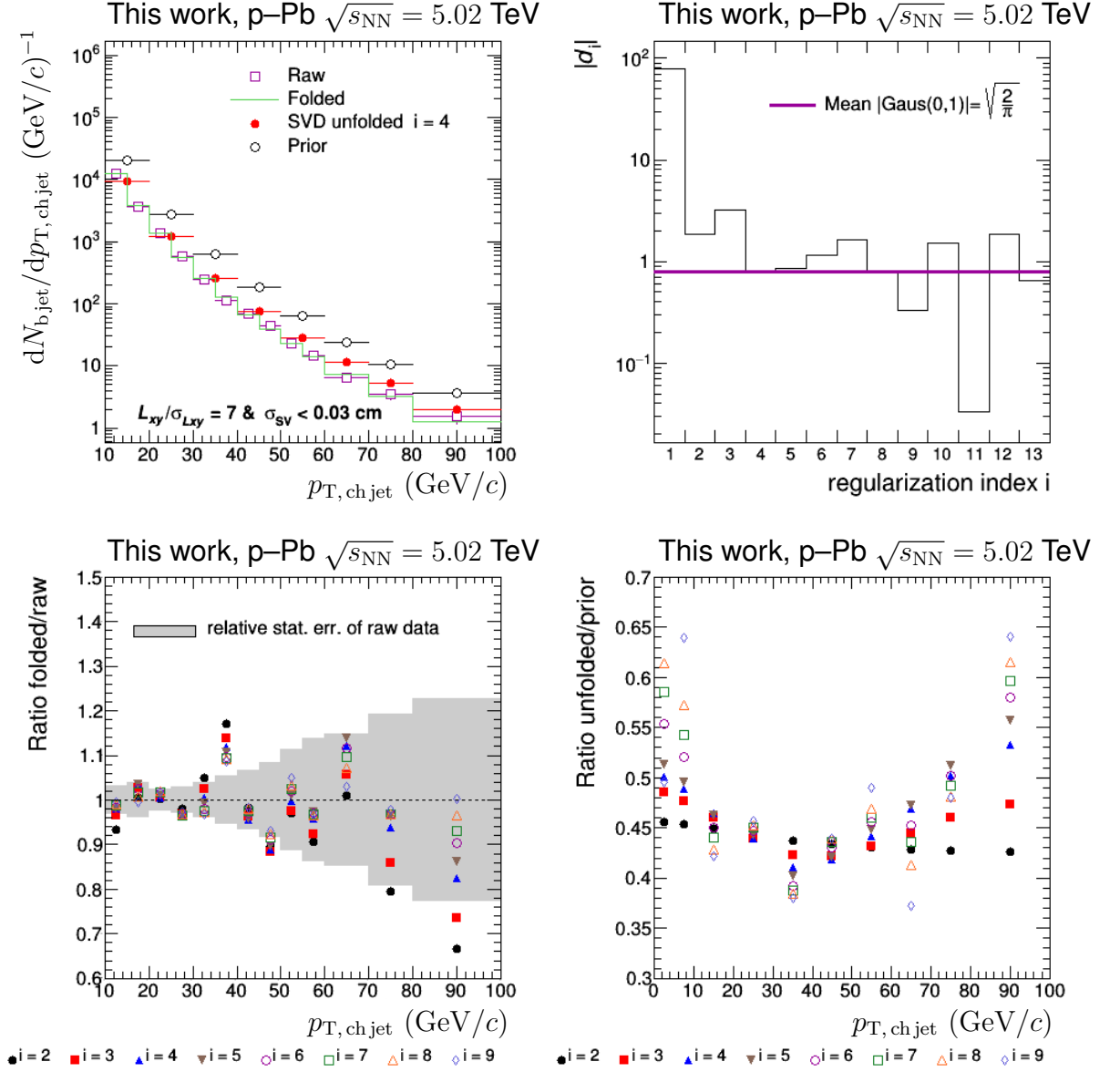


Figure 3.25: QA plots for the SVD unfolding of the measured p-Pb b-jet spectrum. Meaning of the individual panels is analogous to Figure 3.22. The right bottom panel shows a ratio of the unfolded solutions to the prior spectrum. The primary analysis spectrum was obtained with  $i = 5$ .



### 3.9 $p_T$ differential b-jet cross section spectrum from the SV method

After the unfolding, the fully corrected was converted to the cross-section:

$$\frac{d^2\sigma^{\text{bjet}}}{dp_{T,\text{ch jet}}d\eta_{\text{ch jet}}} = \frac{1}{\mathcal{L}} \frac{d^2N_{\text{unfolded}}^{\text{bjet}}}{dp_{T,\text{ch jet}}d\eta_{\text{ch jet}}}, \quad (3.19)$$

where  $\mathcal{L}$  represents the integrated luminosity of the collected data. This luminosity can be calculated as  $\mathcal{L} = N \cdot \varepsilon_{\text{MB}}/\sigma_{\text{MB}}$ . Here  $\sigma_{\text{MB}}$  is the cross section of minimum bias events registered by the ALICE V0 detector [127], [128]. The uncertainties related to  $\sigma_{\text{MB}}$  are discussed in Ref. [129] for pp collisions and in Ref. [128] for p-Pb collisions.  $N$  is the number of minimum bias events with a reconstructed vertex within  $|z_{\text{vtx}}| < 10$  cm and  $\varepsilon_{\text{MB}}$  is a correction for the minimum bias events when the true vertex was between  $\pm 10$  cm, but the vertex was not reconstructed. The size of  $\varepsilon_{\text{MB}}$  was estimated to be 1.047 and 1.015 for the pp and p-Pb data, respectively.

### 3.10 Systematic Uncertainties

The extent to which the fully corrected b-jet spectra and the corresponding nuclear modification factor depend on the particular choice of the analysis setting is assessed by evaluating systematic uncertainties. The main sources of systematic uncertainties are listed in Table 3.4. In the following sections, we discuss the applied systematic variations in the analysis procedure in more detail.

$p_{T, \text{chjet}}$ interval		10–20 GeV/ $c$	40–50 GeV/ $c$	80–100 GeV/ $c$
Statistical uncertainty	pp	1.9	5.2	18.4
	p–Pb	1.9	3.6	12.9
	$R_{\text{pPb}}^{\text{bjet}}$	2.6	6.4	22.5
Tagging of b jets	pp	+0.9/ – 2.8	+3.4/ – 6.5	+6.8/ – 13.4
	p–Pb	+3.4/ – 1.6	+4.6/ – 8.6	+6.0/ – 15.3
	$R_{\text{pPb}}^{\text{bjet}}$	+2.5/ – 2.2	+5.0/ – 5.3	+10.7/ – 13.8
Purity of b-jet candidates	pp	+13.0/ – 21.8	+16.4/ – 16.8	+21.8/ – 17.3
	p–Pb	+13.1/ – 21.0	+11.9/ – 16.3	+21.1/ – 15.7
	$R_{\text{pPb}}^{\text{bjet}}$	+5.2/ – 9.4	+5.2/ – 6.6	+8.2/ – 9.7
Tracking efficiency	pp	7.9	8.4	9.5
	p–Pb	6.7	9.2	8.6
	$R_{\text{pPb}}^{\text{bjet}}$	1.4	1.0	1.3
Track $p_T$ resolution	pp	+1.2/ – 1.2	+3.9/ – 3.9	+6.0/ – 6.0
	p–Pb	+3.3/ – 3.3	+4.5/ – 4.5	+5.3/ – 5.3
	$R_{\text{pPb}}^{\text{bjet}}$	+2.1/ – 2.2	+0.6/ – 0.6	+0.8/ – 0.7
Unfolding	pp	+7.2/ – 0.9	+1.0/ – 1.9	+27.1/ – 6.1
	p–Pb	+9.5/ – 5.6	+0.5/ – 4.5	+11.4/ – 14.3
	$R_{\text{pPb}}^{\text{bjet}}$	+2.5/ – 5.4	+3.1/ – 4.7	+4.4/ – 15.9
$\delta p_T$ estimators and matrix truncation	pp	+0.0/ – 5.4	+0.0/ – 10.3	+3.2/ – 0.0
	p–Pb	+0.0/ – 3.1	+0.0/ – 5.4	+3.1/ – 0.0
	$R_{\text{pPb}}^{\text{bjet}}$	+2.6/ – 0.0	+6.7/ – 0.0	+1.4/ – 0.0
Fraction of secondary particles in MC	pp	+1.6/ – 0.0	+2.4/ – 0.0	+2.9/ – 0.0
	p–Pb	+4.1/ – 0.0	+5.4/ – 0.0	+7.8/ – 0.0
	$R_{\text{pPb}}^{\text{bjet}}$	+0.0/ – 2.5	+0.0/ – 3.0	+0.0/ – 5.1
Total systematic uncertainty	pp	+17.0/ – 24.0	+19.3/ – 22.8	+37.4/ – 25.4
	p–Pb	+18.6/ – 23.2	+17.2/ – 22.2	+28.0/ – 28.1
	$R_{\text{pPb}}^{\text{bjet}}$	+7.3/ – 11.6	+10.4/ – 10.2	+14.3/ – 23.8
Normalization uncertainty	pp	2.34		
	p–Pb	3.7		
	$R_{\text{pPb}}^{\text{bjet}}$	4.37		

Table 3.4: Relative size of statistical and systematic uncertainties of b-jet cross section spectra measured in pp and p–Pb collisions and of the corresponding nuclear modification factor. The uncertainties are quoted for three  $p_{T, \text{chjet}}$  bins. An additional uncertainty from the normalization by the integrated luminosity [128], [129] is quoted in the last row.

### 3.10.1 Tagging of b jets

The choice of SV tagging cuts influences the reconstructed spectrum of the b-jet candidates and the SV tagging purity and efficiency values. In this study, we have estimated the effect of the SV tagging cuts on the resulting fully corrected b-jet spectrum. We have repeated the reconstruction and correction procedure, keeping one tagging parameter fixed at its default value while varying the other parameter. The PA setting of the SV tagging cuts was  $SL_{xy} > 7$  and  $\sigma_{SV} < 0.03$  cm. During this test,  $SL_{xy}$  was changed from 6 to 9, and  $\sigma_{SV}$  was ranged from 0.02 to 0.05 cm. The ratios of the fully corrected b-jet spectra obtained with the varied SV tagging cuts and the PA cuts are shown in Fig. 3.26. Since the variations of the tagging cuts are not independent, the assigned systematic uncertainties were set to the maximal upper and lower deviations in the given  $p_{T,\text{ch jet}}$  bin.

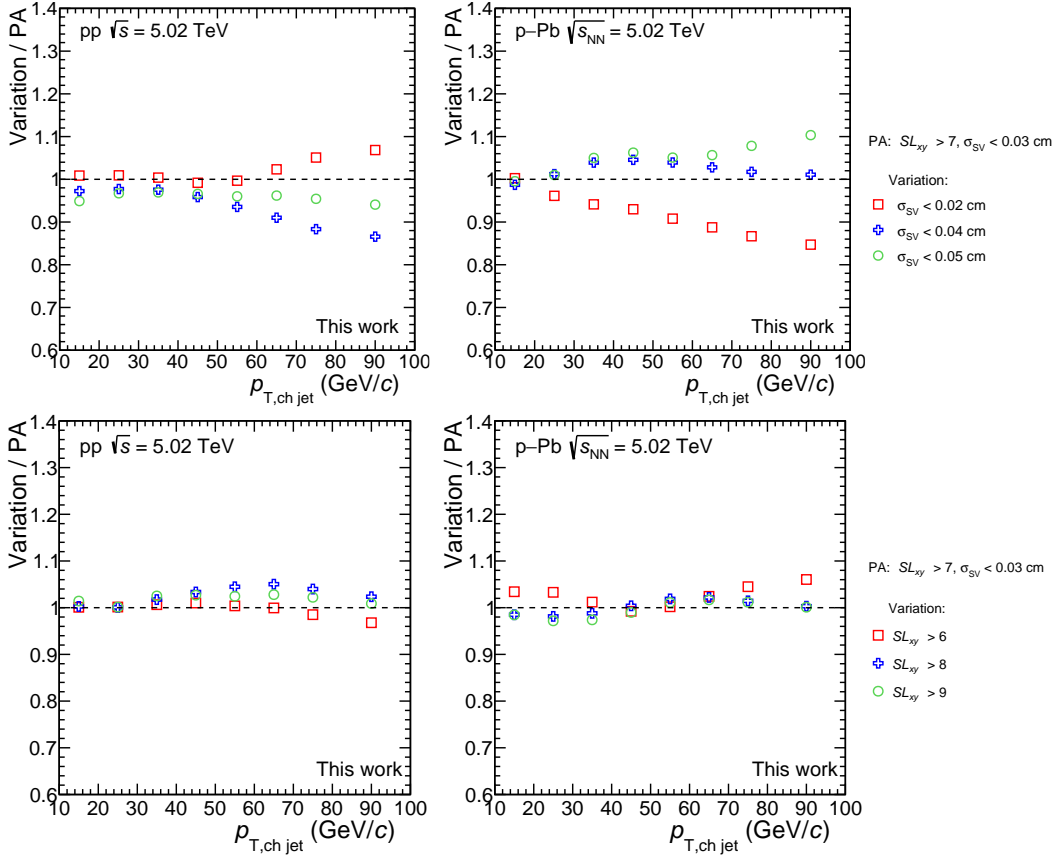


Figure 3.26: Ratios of the fully-corrected b-jet spectra obtained with the varied and the PA SV tagging setting for pp (left) and p-Pb (right) data. The top panels show the results obtained for  $\sigma_{SV}$  variations. The bottom panels show analogous plots for  $SL_{xy}$  variations. Statistical errors on the data points are not shown since the numerator and denominator are highly correlated.

### 3.10.2 Purity of b-jet candidates

The purity of tagged b-jet candidates estimated with the POWHEG-simulation-based method strongly depends on the POWHEG parameters listed in the Tab. 3.3. The table gives the default POWHEG value of the given parameter and the range considered for the systematic variation study. The effect of POWHEG scale variation was assessed by repeating the SV b-jet analysis with cross-variations of the quark masses, the renormalization, and the factorization scales from Table 3.3. The ratio of the corresponding fully corrected b jet spectra to the primary analysis spectrum can be seen in Fig. 3.27. The envelope of the observed spectrum variations defined the final uncertainty.

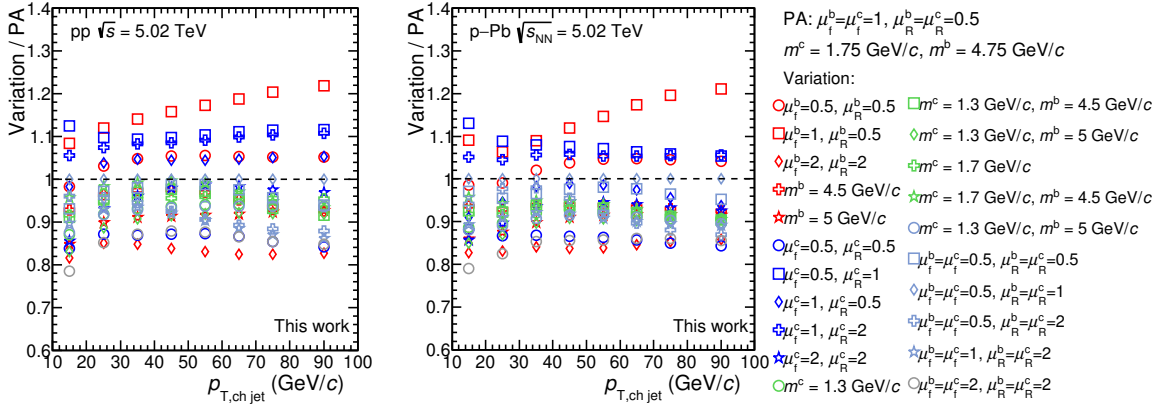


Figure 3.27: The ratio of fully corrected b-jet spectra obtained with the SV tagging corrections based on various POWHEG settings to the PA spectrum for pp (left) and p-Pb (right) collisions.

### 3.10.3 Tracking efficiency

The inefficiency of the ALICE track reconstruction of 4% [130] affects the jet energy scale of the reconstructed charged-particle jets and represents one of the important sources of systematic uncertainties in this analysis. To estimate the impact of this variation on the fully corrected spectrum of the b jets, we have constructed an alternative response matrix that accounted for a 4% lower number of tracks on the detector level. The ratio of fully corrected b-jet spectra obtained with the alternative response matrix and the PA response matrix is shown in Fig. 3.28. It is assumed that a 4% variation towards higher tracking efficiency would affect the results symmetrically.

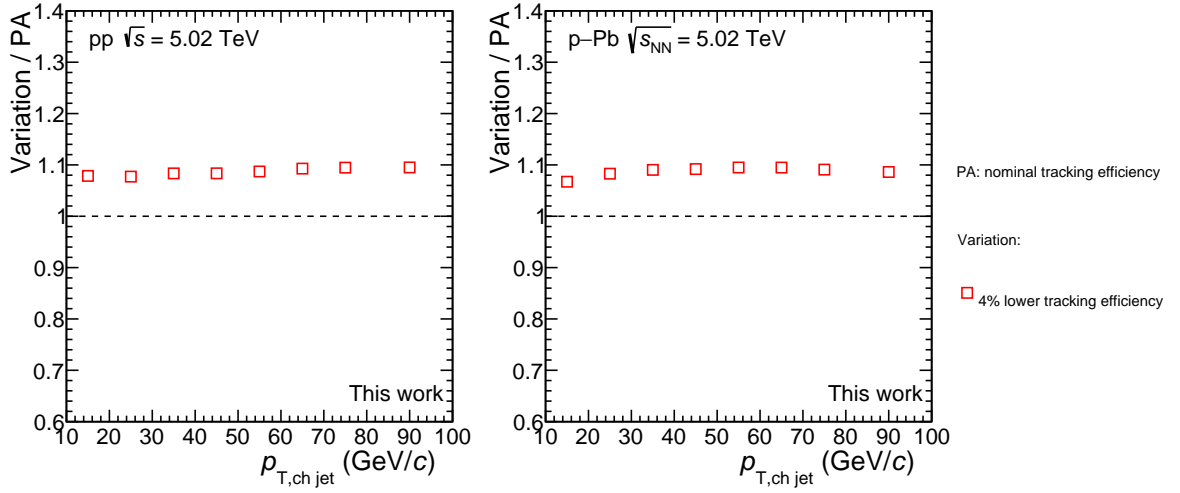


Figure 3.28: Ratio of the fully corrected spectrum obtained with 4% lower tracking efficiency over the primary analysis spectrum for the pp (left) and p-Pb (right) data.

### 3.10.4 Track $p_T$ resolution

The instrumental response matrix accounts for the smearing of track momenta due to the finite momentum resolution of the detector. The size of the track  $p_T$  smearing can be estimated from the corresponding element of covariance matrix of the Kalman fit. This element carries information about relative  $p_T$  resolution,  $\sigma(p_T^{-1})/p_T^{-1}$ . The mean value of this resolution can be seen in the right panel of Fig. 3.3. The impact of track  $p_T$  resolution uncertainty on the fully corrected spectrum of b jets was investigated by means of a PYTHIA 8 Tune 5 simulation which was used to generate pp collisions at  $\sqrt{s} = 5.02$  TeV. The generated tracks were filtered using the track reconstruction efficiency given by the left panel of Fig. 3.3. Subsequently, their transverse momentum was smeared based on (i) the PA setting, corresponding to the  $\sigma(p_T^{-1})/p_T^{-1}$  distribution estimated from the corresponding covariance matrix element of the Kalman fit [131], or (ii) combined smearing due to  $\sigma(p_T^{-1})/p_T^{-1}$  and dispersion that reflects azimuthal variation in the  $1/p_T$  spectrum of tracks [132]. Both sets of PYTHIA particles were used to construct instrumental response matrices, which were further folded with the  $\delta p_T$  matrix and used to unfold the measured b-jet spectra. The relative difference between the resulting b-jet spectra is then taken as systematic uncertainty; see Fig. 3.29. The uncertainty in track  $p_T$  smearing affects the b-jet spectra at the level of 5%.

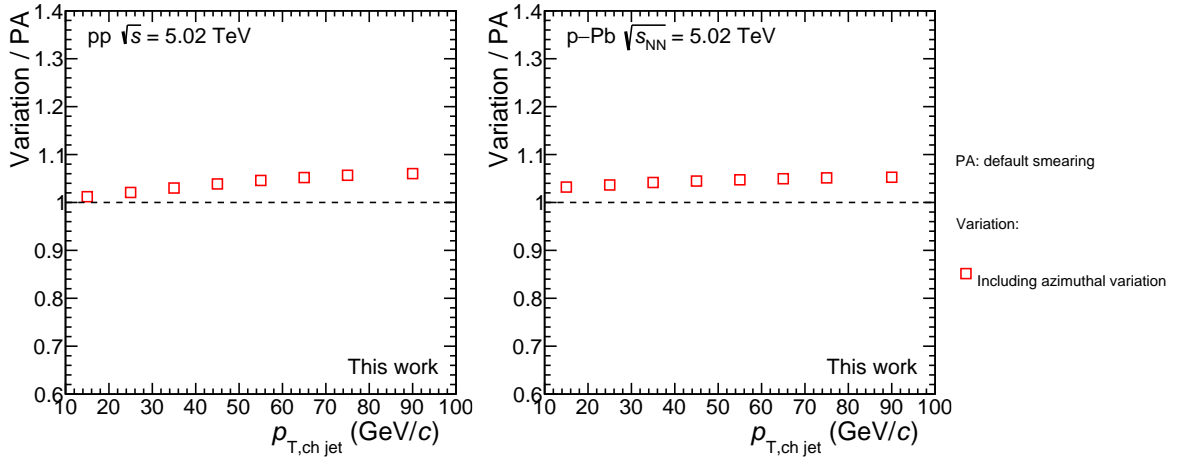


Figure 3.29: Ratio of the fully corrected b-jet spectra obtained with PA smearing and with the modified  $p_T$  smearing which accounts differences in track  $p_T$  spectra as a function of azimuth. The left panel corresponds to pp data, the right is the p-Pb analysis.

### 3.10.5 Unfolding systematics

The primary analysis is done using the SVD method with the regularization parameter  $i = 4$  for the pp and  $i = 5$  for the p-Pb data, and prior obtained from the POWHEG+PYTHIA simulation with the default setting. In this systematic study, we have considered the following variations of unfolding inputs:

- Choice of raw spectrum binning. Two binning schemes were considered:
  - (10, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 100) – PA binning
  - (10, 14, 19, 23, 27, 31, 35, 39, 49, 61, 71, 100) – varied binning

The value of the regularization parameter  $i$  was optimized for each binning scheme separately. The result of this test is shown in Fig. 3.31.

- Choice of the regularization parameter  $i$ : the SVD unfolding regularization parameter was varied  $\pm 1$  with respect to the PA value. The result of this test is shown in Fig. 3.30.
- Choice of the unfolding method: the unfolding was repeated with the same prior spectrum, but instead of the SVD method, we used the Bayesian unfolding with the optimal number of iterations equal to 4. The result of this test is shown in Fig. 3.30.
- Choice of the prior spectrum. The PA was used as a prior the particle-level b-jet spectrum generated by POWHEG with PYTHIA fragmentation. POWHEG was run using its default setting. For variation, the prior was altered with eight other POWHEG spectra, which were calculated using different regularization and factorization scale values; see Fig. 3.32.

In order to suppress the impact of statistical fluctuations on the systematic uncertainties related to the unfolding, Robert Vértési carried out pseudo-random experiments. In this approach, we make several statistically equivalent copies of the input PA raw spectrum of b jets using Poissonian resampling of each bin content in the original spectrum. In the next step, the copies are corrected with the PA setting and varied setting, and both spectra' ratios are calculated. The systematic uncertainty for the given variation is obtained as the mean of these ratios over all randomized copies. Since the studied unfolding uncertainties are not independent, the resulting systematics were obtained by taking their maximum deviation in each bin. The resulting systematic uncertainties are considered to be asymmetric.

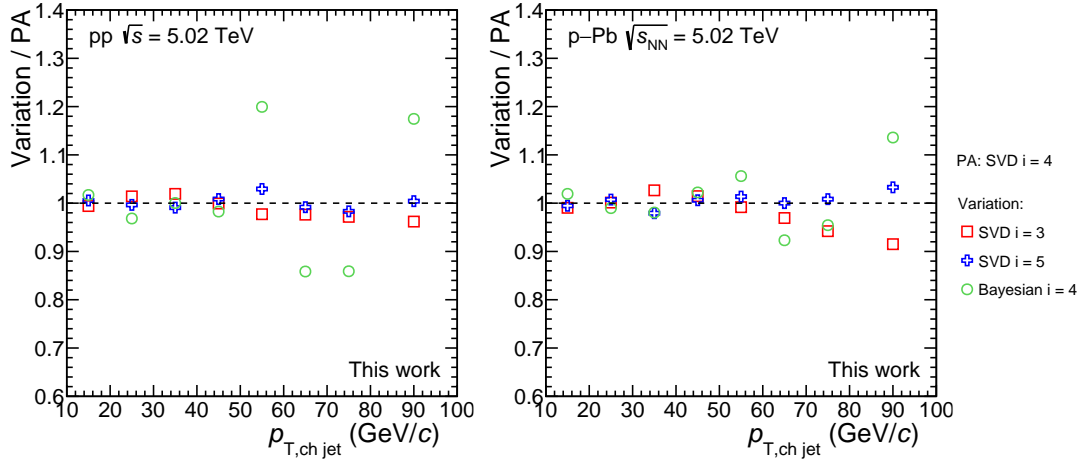


Figure 3.30: Ratio of the fully corrected b-jet spectrum obtained using the varied unfolding setting and the PA setting. Left: pp data. Right: p-Pb data.

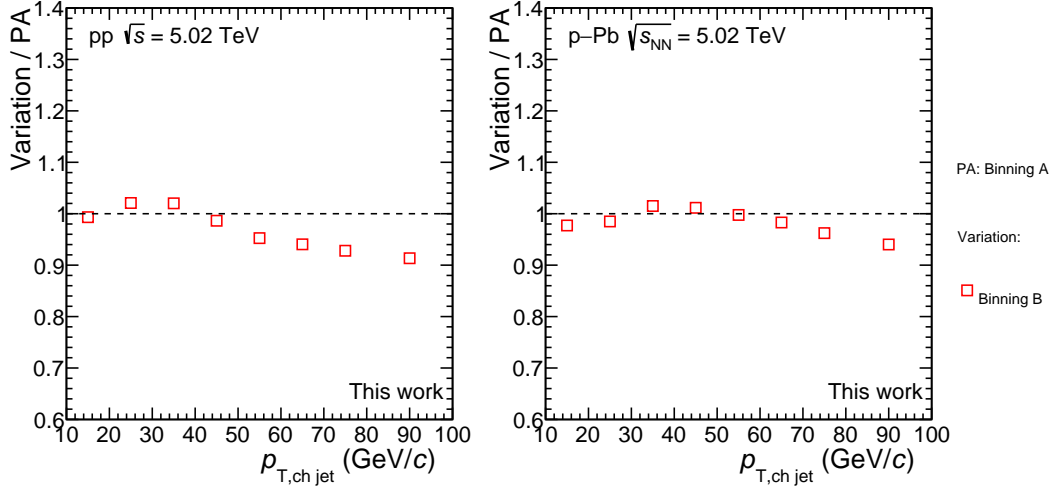


Figure 3.31: Ratio of the fully corrected b-jet spectrum obtained with the varied raw spectrum binning over the primary analysis binning for pp (left) and p-Pb (right) data.

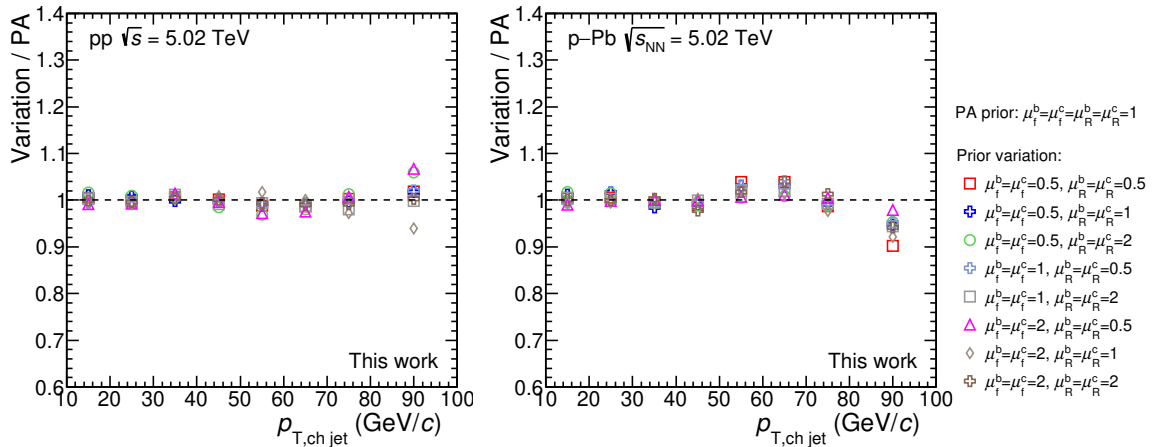


Figure 3.32: Ratio of the fully corrected b-jet spectrum obtained with varied POWHEG prior spectrum and the PA spectrum for pp (left) and p-Pb (right) data.



### 3.10.6 $\delta p_T$ estimator and combined response matrix truncation

The combined response matrix describes momentum smearing due to local background fluctuations and instrumental effects. We have investigated how the fully corrected b-jet spectrum changes when one changes the default PA  $\delta p_T^{\text{emb}}$  matrix (obtained by track embedding) to the  $\delta p_T^{\text{RC}}$  matrix obtained from the random cone method. In addition, we have also studied the sensitivity of the fully corrected spectrum to a possible inaccurate description of the low- $p_T$  region in the instrumental response matrix by the PYTHIA simulation. In this case, the particle level  $p_T$  axis of the instrumental matrix was truncated at 5 GeV/ $c$  from below (Trunc ReMx). Figure 3.33 shows the relative deviations of the corresponding fully corrected b-jet spectrum w.r.t. PA. The resulting uncertainty is obtained as an envelope of the variations.

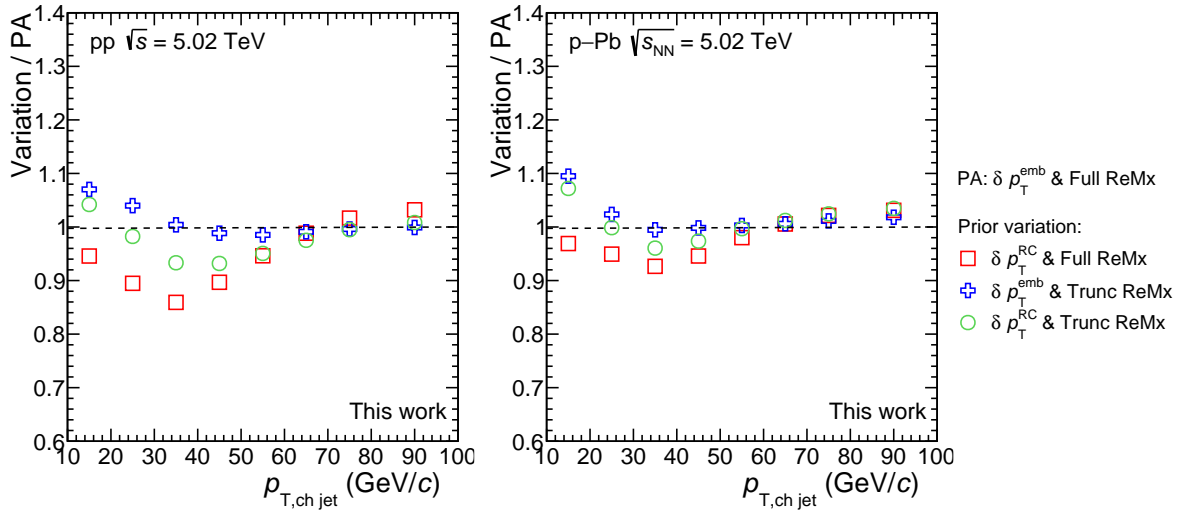


Figure 3.33: Ratios of fully corrected b-jet spectra obtained with varied  $\delta p_T$  estimator and combined response matrix truncation to the PA spectrum. “Full ReMx” label denotes a response matrix which has the same range as the PA matrix.

### 3.10.7 Fraction of secondary particles in MC

The instrumental response matrix also accounts for the jet energy scale shift as a result of the admixture of secondary particles. The size of this admixture depends on the event generator used to generate the instrumental matrix. According to the detector-level simulation, secondary particles carry about 4.2% and 3.6% of jet  $p_T$  for pp and p-Pb data, respectively. In order to check to what extent our simulation describes the fractions of physical primaries [133] and secondary tracks in the real data, we have parameterized the measured distance of closest approach (DCA) distributions of reconstructed hybrid tracks with a linear combination of simulated detector-level DCA templates for physical

primaries and secondary tracks; see Fig. 3.34. The free parameters of the fit were the scaling factors of the templates.

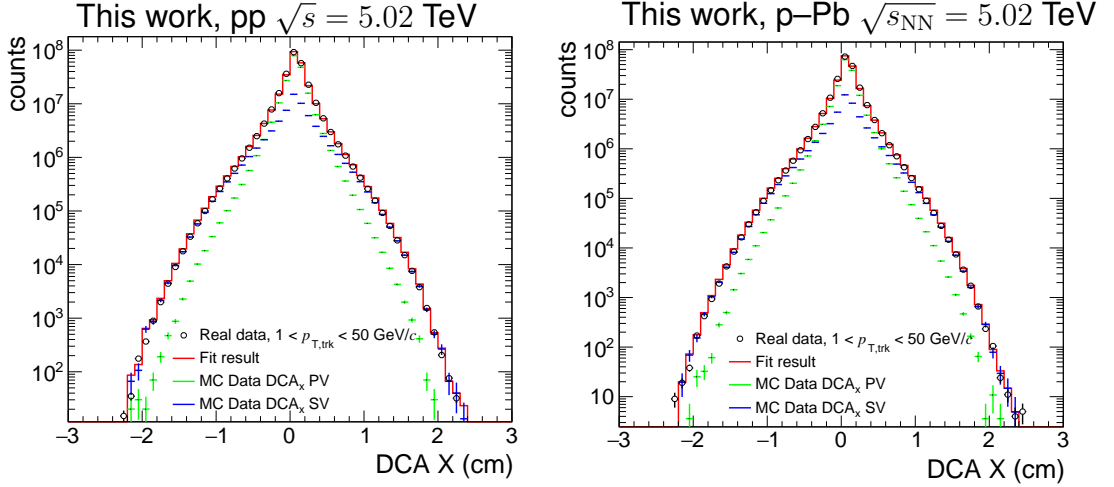


Figure 3.34: DCA distributions of charged hybrid tracks in pp (left) and p-Pb data (right). The real data are fitted with a sum of two simulated templates corresponding to the DCA distribution of physical primaries and the DCA distribution of secondary tracks.

This study revealed that the secondary track fraction is 11.2% (26%) lower in the measured pp (p-Pb) data compared to the MC simulation. The real fraction of the jet  $p_{T,\text{ch,jet}}$  carried by the secondary tracks in the detector-level MC should be smaller than  $4.2\% \cdot 0.888 = 3.72\%$  and  $3.6\% \cdot 0.74 = 2.66\%$  for the pp and p-Pb data, respectively.

In the next step, we have fitted the fully corrected b-jet spectra with a sum of two exponential functions:

$$f(p_{T,\text{ch,jet}}) = \sum_{i=1}^2 a_i \cdot \exp[\xi \cdot p_{T,\text{ch,jet}}/b_i], \quad (3.20)$$

where  $a_i$  and  $b_i$  are fit parameters. The parameter  $\xi$  accounts for the shift due to different admixture of secondary tracks. In the case of pp collisions, we first assumed that the parameter  $\xi$  was equal to  $1.0372/1.042 = 0.995$  and obtained the fit parameters  $a_i$  and  $b_i$ . These parameters were then used to evaluate (3.20) also for  $\xi = 1$ , which would correspond to the case when the fraction of secondary tracks would be correctly described in the MC. The uncertainty attributed to systematics is then given by the ratio of both functions and is shown in Fig. 3.35. We evaluated the systematics for the p-Pb spectrum analogously. The resulting systematic uncertainty is assumed to be asymmetric.

It is well known that the PYTHIA MC event generator often has problems reproducing strangeness production in pp collisions [134]–[136]. Therefore, we have investigated to what extent this might impact the obtained MC light-flavor SV invariant mass templates used in the data-driven method to extract the purity of the tagged b-jet candidates. In

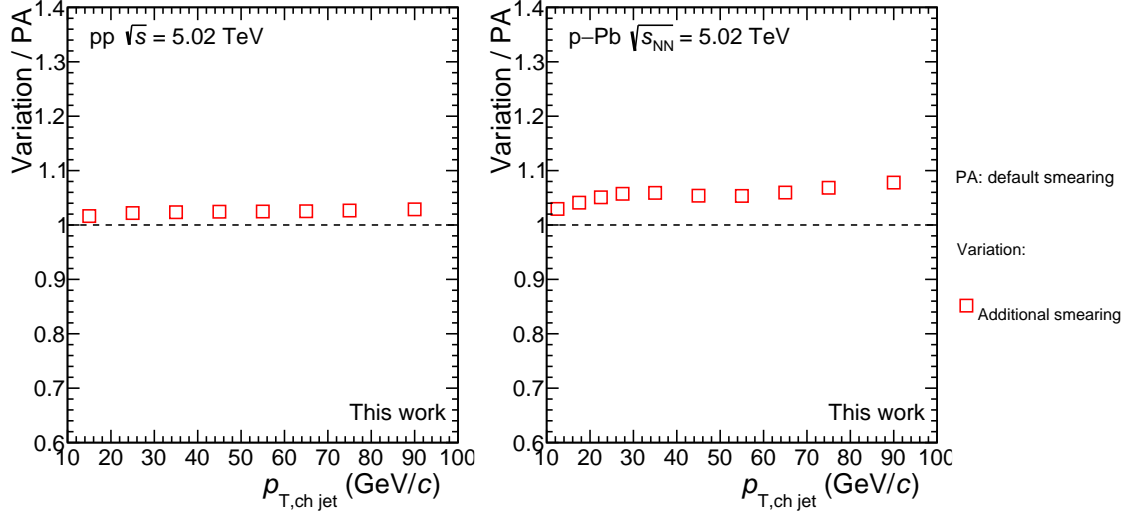


Figure 3.35: Impact of variation in secondary particle admixture on the fully corrected b-jet spectra. Left: pp. Right: p-Pb. See text for more details.

the MC simulation, we traced down which tagged light-flavor SVs stem from decays of  $\Lambda$ ,  $K_S^0$  or multi-strange baryons, and we have found that portion of such cases among all tagged light-flavor SVs is less than 1%, see Fig. 3.36. Thus, the impact of the possible variation in the yield of strange particles on the extracted b-jet purity was considered negligible.

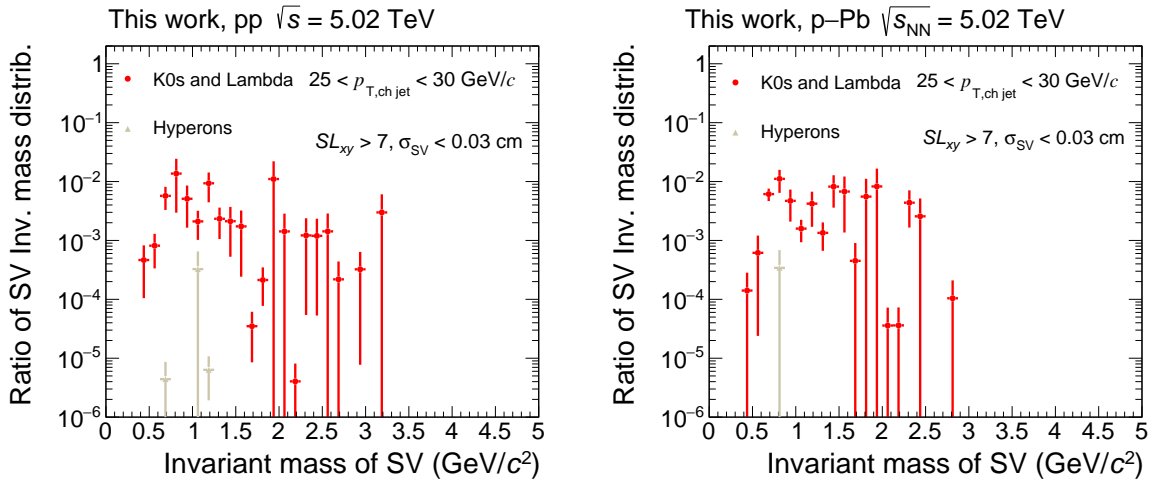


Figure 3.36: Ratio of the invariant mass distributions of the tagged strange particle SVs and of all light-flavor tagged SVs.

### 3.10.8 Summary of systematic uncertainty calculations

The main sources of systematic uncertainties, as listed in Table 3.4 were considered independent and were added in quadrature to obtain the total systematic uncertainty; see Fig. 3.37.

When evaluating the systematic uncertainty of the nuclear modification factor  $R_{p\text{Pb}}$ , systematic uncertainties of the numerator and denominator spectrum partly cancel depending on the correlation between pp and p–Pb uncertainties. In the case of correlated systematic uncertainties, variation was performed in the same direction, and the systematic was obtained as the difference between the PA and varied  $R_{p\text{Pb}}$ . Such an approach was used for b-jets tagging, the purity of b-jet candidates, tracking efficiency, unfolding and  $\delta p_T$  estimators, and matrix truncation uncertainties. The remaining uncertainties of track  $p_T$  resolution and the fraction of secondary particles in MC were calculated as an envelope of 4 possible combinations of variations with respect to PA spectra. Summary of systematic uncertainties for  $R_{p\text{Pb}}$  can be found in Fig. 3.38.

The statistical and systematic uncertainties of the untagged-jet cross sections used in calculating the charged-particle b-jets fraction were assumed to be uncorrelated with the b-jets sample. For more details, follow Sec. 3.11.3.

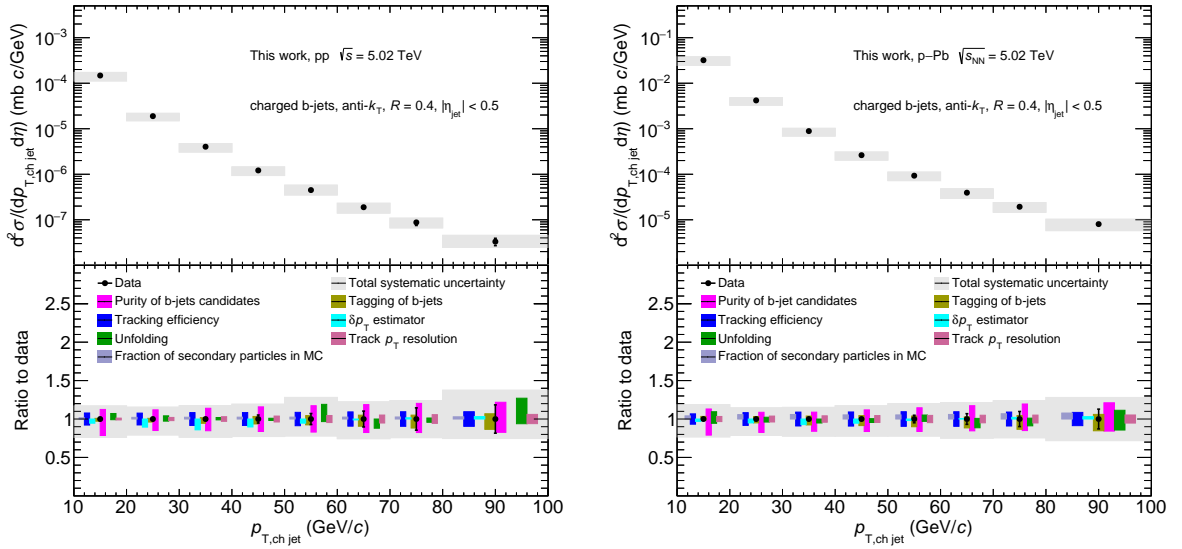


Figure 3.37: Top panel: the fully corrected b-jet cross section spectrum from pp and p–Pb collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV. Gray boxes represent the total systematic uncertainties. The bottom panel shows the individual sources of systematic uncertainties.

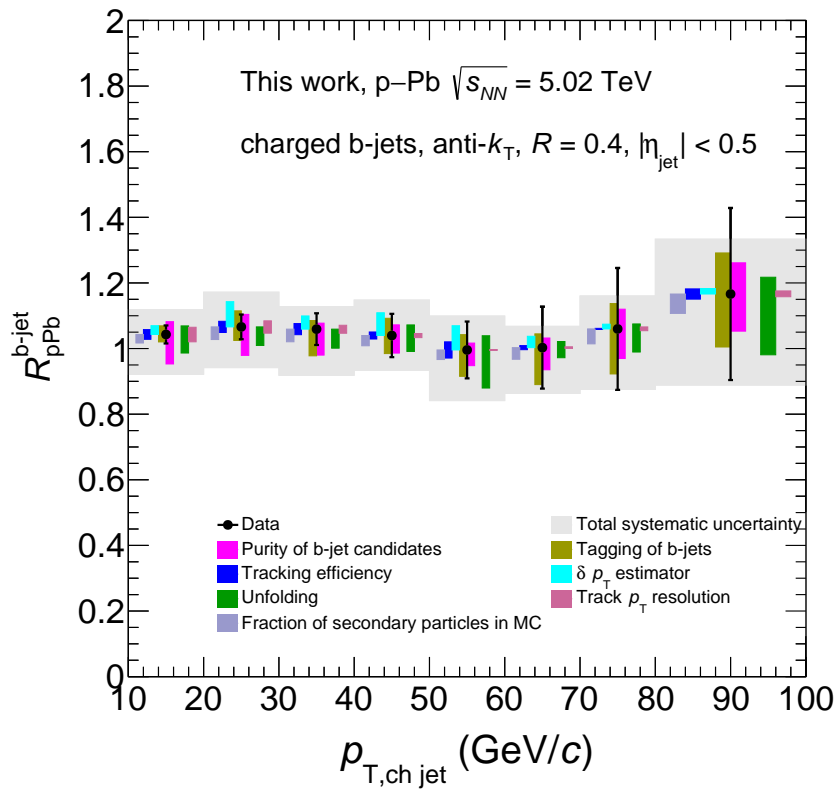


Figure 3.38: The nuclear modification factor  $R_{pPb}$  of the inclusive charged-particle anti- $k_T$   $R = 0.4$  b jets as a function of  $p_T$  together with individual sources of systematic uncertainties. Gray boxes represent the total systematic uncertainties.

### 3.11 Results

Figure 3.39 presents the  $p_T$ -differential production cross-section of b jets obtained from the IP and SV analyses in pp and p-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV. Both methods provide compatible results. Let us note that in the first  $p_T$  bin, the IP method could not extract the purity and efficiency of tagged b-jet candidates; thus, the corresponding data point from the IP method is missing.

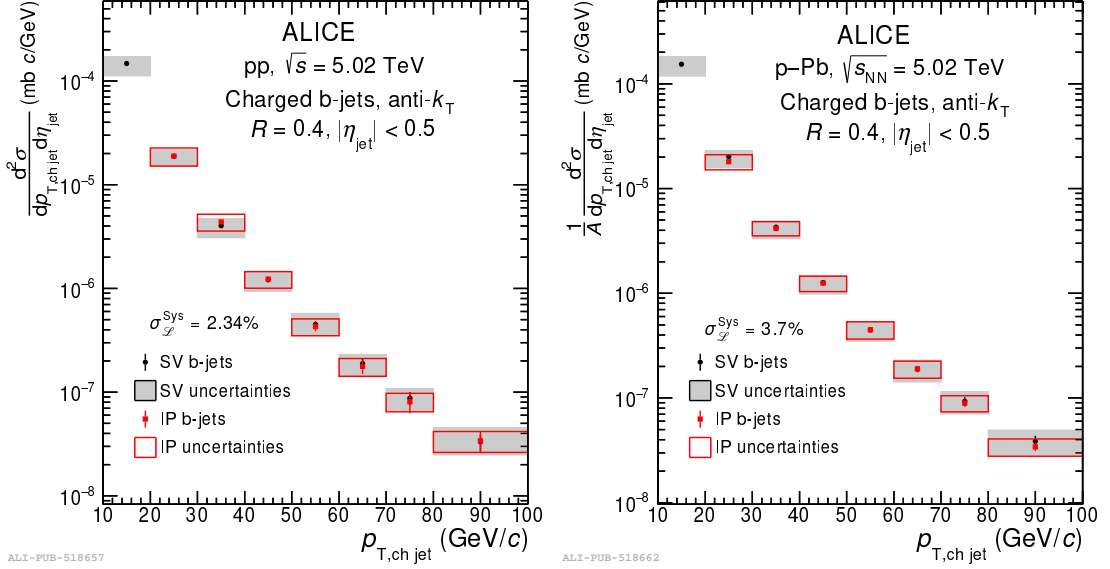


Figure 3.39: Comparison of the  $p_T$  differential production cross section of charged-particle b jets measured in pp and p-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV using the IP and SV methods. Systematic and statistical uncertainties are shown as boxes and error bars, respectively.

The nuclear modification factor of charged-particle b-jets was calculated using Eq. (1.11). Figure 3.40 shows the  $R_{pPb}^{b\text{-jet}}$  obtained with the SV and IP methods. The nuclear modification factor from both methods is consistent with uncertainties.

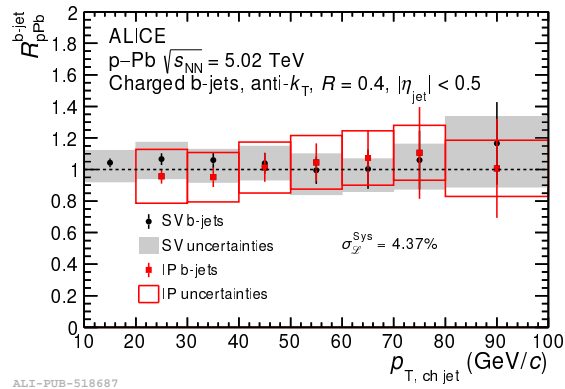


Figure 3.40: The nuclear modification factor  $R_{pPb}$  of the inclusive charged-particle anti- $k_T$   $R = 0.4$  b jets as a function of  $p_T$  from the IP and SV method.

### 3.11.1 Combining the results of the IP and SV method

The  $p_T$ -differential b-jet production cross sections obtained from the IP and SV methods were combined by Hadi Hassan using the Best Linear Unbiased Estimator (BLUE) method [111], [137]. This method combines different measurements of the same physical quantity, where the uncertainties of individual measurements are correlated. The resulting spectrum and its uncertainty were calculated as follows:

$$\hat{\theta} = \hat{\theta}_{\text{SV}} \cdot \frac{\sigma_{\text{IP}}^2 - \rho\sigma_{\text{SV}}\sigma_{\text{IP}}}{\sigma_{\text{SV}}^2 - 2\rho\sigma_{\text{SV}}\sigma_{\text{IP}} + \sigma_{\text{IP}}^2} + \hat{\theta}_{\text{IP}} \cdot \frac{\sigma_{\text{SV}}^2 - \rho\sigma_{\text{SV}}\sigma_{\text{IP}}}{\sigma_{\text{SV}}^2 - 2\rho\sigma_{\text{SV}}\sigma_{\text{IP}} + \sigma_{\text{IP}}^2}, \quad (3.21)$$

$$\sigma = \sqrt{\frac{\sigma_{\text{SV}}^2\sigma_{\text{IP}}^2(1 - \rho^2)}{\sigma_{\text{SV}}^2 - 2\rho\sigma_{\text{SV}}\sigma_{\text{IP}} + \sigma_{\text{IP}}^2}}, \quad (3.22)$$

where  $\hat{\theta}$ ,  $\hat{\theta}_{\text{SV}}$ ,  $\hat{\theta}_{\text{IP}}$  and  $\sigma$ ,  $\sigma_{\text{SV}}$ ,  $\sigma_{\text{IP}}$  are fully corrected spectrum and systematic uncertainties for the combined, SV, and IP b-jet spectra. The parameter  $\rho$  denotes the correlation coefficient between analyzes. We have assumed that systematic uncertainties from tagging and purity extraction are uncorrelated ( $\rho = 0$ ), and the tracking efficiency,  $p_T$  resolution uncertainty and contamination by secondary tracks were treated as fully correlated ( $\rho = 1$ ). The statistical uncertainty is partially correlated because the same data set was used in both analyzes, and the corresponding correlation coefficient was obtained as:

$$\rho_{\text{stat}} = \frac{\text{Cov}(\text{IP}, \text{SV})}{\sigma_{\text{SV}}^{\text{stat}}\sigma_{\text{IP}}^{\text{stat}}} \quad \text{with} \quad \text{Cov}(\text{IP}, \text{SV}) = \frac{\sigma_{\text{IP}}^{\text{stat}2}\sigma_{\text{SV}}^{\text{stat}2}}{\sigma_{\text{SV}\cap\text{IP}}^{\text{stat}2}}, \quad (3.23)$$

where  $\sigma_{\text{SV}}^{\text{stat}}$  and  $\sigma_{\text{IP}}^{\text{stat}}$  are the statistical uncertainties of the jet samples from the SV and IP methods, and  $\sigma_{\text{SV}\cap\text{IP}}^{\text{stat}}$  is the statistical uncertainty of the merged sample. The correlation coefficient for statistical uncertainty was found to be 0.35 (0.27) for pp (p–Pb). Background fluctuations and unfolding uncertainties were assumed to be partially correlated between both methods, therefore, an arbitrarily chosen correlation coefficient value of  $\rho = 0.5$  with consistency checks equal to 0 and 1 were used. The resulting systematic uncertainties of other parameters were obtained with the same approach and were negligible.

Since the SV has asymmetric errors while the IP has symmetric errors, the IP and upper SV and IP and lower SV were merged separately. The final result from the two combined spectra is obtained from the bin with larger uncertainties.

### 3.11.2 b-jet cross-section

The obtained combined b-jet cross-section was compared by Robert Vertesi and Hadi Hassan with the NLO pQCD calculations by the POWHEG dijet tune with PYTHIA

8 fragmentation, see Fig. 3.41. The uncertainties of the POWHEG calculation were obtained by variation of  $\alpha_s$  in the range 0.111–0.123, variations of the renormalization and factorization scales by 0.5–2, and variation of the PDFs type (CT14NLO parton distribution function for pp and EPPS16 nPDF for p–Pb). The uncertainties of CT14NLO and EPPS16 were assessed using the Hessian prescript from Eq. (53) of Ref. [138]. One can see that the measured b-jet cross section is consistent with the model predictions within the experimental and theoretical uncertainties.

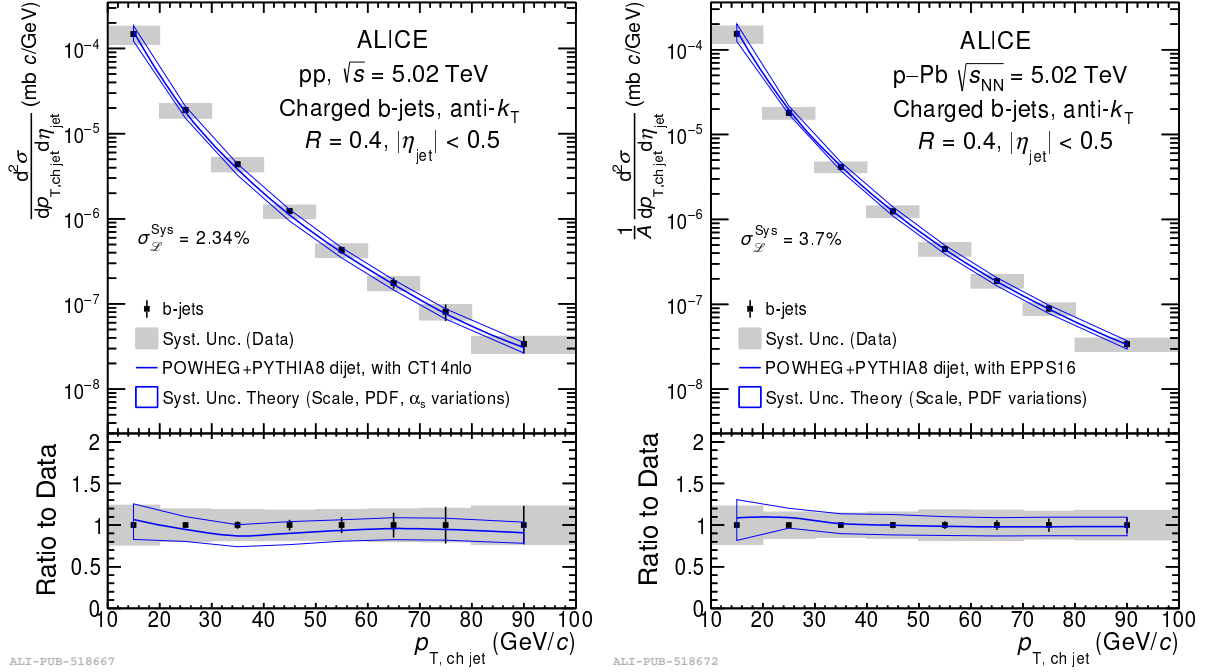


Figure 3.41: Top panels: The combined differential production cross-section of charged-particle anti- $k_T$   $R = 0.4$  b jets in pp and p–Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV. The data are compared with a NLO pQCD prediction by the POWHEG dijet tune with PYTHIA 8 fragmentation [77, 78]. Additional normalization uncertainty due to luminosity is quoted separately. Bottom panels: ratio of the theoretical calculations to the data.

### 3.11.3 Fraction of the b-jets

The fraction of the b-jets can be defined as a ratio of the cross-section of b-jets and inclusive untagged charged-particle jets. In our case, we took the inclusive untagged charged-particle spectra for pp and p–Pb from the ALICE papers Ref. [128] and Ref. [129], respectively. These spectra were measured for statistically independent data sets. Figure 3.42 shows the comparison of the b-jet fraction of this analysis and the calculation using the POWHEG dijet tune with PYTHIA 8 fragmentation [139], [140]. The POWHEG calculation for p–Pb collisions considered the EPPS16 nPDFs. The results of the b-jet fraction provided by this analysis are compatible with the model predictions within uncertainties.



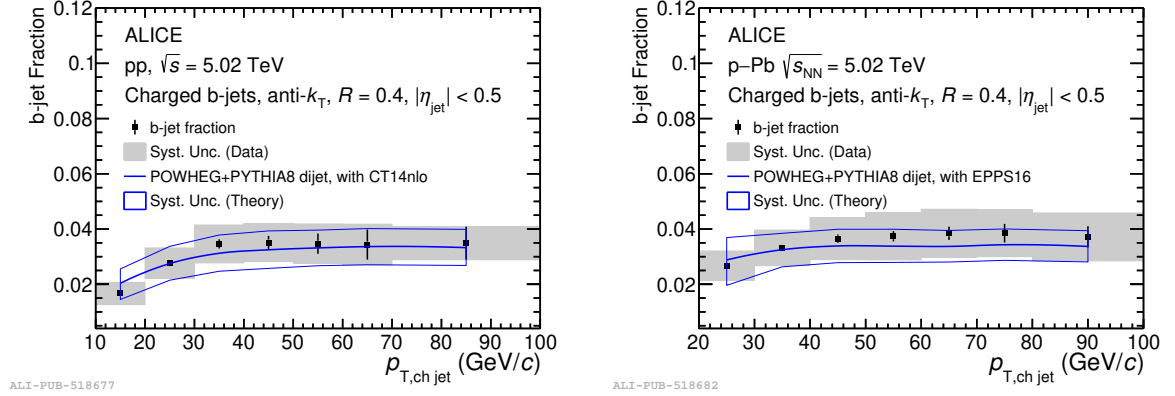


Figure 3.42: The fraction of charged-particle anti- $k_T$   $R = 0.4$  b jets measured in pp and p–Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV, compared with POWHEG NLO pQCD calculations with PYTHIA 8 fragmentation. See text for more details.

### 3.11.4 The b-jet nuclear modification factor $R_{pPb}^{b\text{-jet}}$

Figure 3.43 shows the combined b-jet  $R_{pPb}^{b\text{-jet}}$  compared to the POWHEG dijet tune with PYTHIA 8 b-jet fragmentation calculation. The measurement of  $R_{pPb}^{b\text{-jet}}$  is compatible with unity within uncertainties, suggesting that there are no strong nuclear matter effects present in b-jet production at midrapidity in p–Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV. The calculation describes the data within the uncertainties.

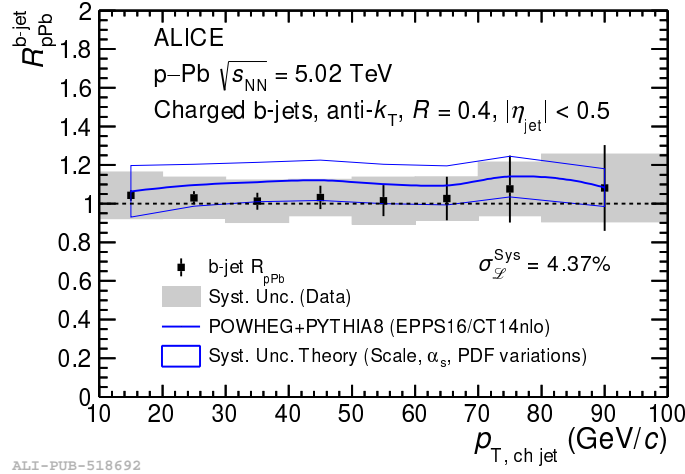


Figure 3.43: The combined nuclear modification factor  $R_{pPb}^{b\text{-jet}}$  compared with the calculation by the POWHEG dijet tune with the PYTHIA 8 fragmentation. Systematic and statistical uncertainties are shown as boxes and error bars, respectively.

The left panel of Figure 3.44 compares the  $R_{pPb}^{b\text{-jet}}$  measurements of ALICE and CMS. The CMS measured b jets calorimetrically, i.e., including the neutral component [11]. There is, therefore, a shift in the energy scale between both measurements. The ALICE  $p_{T,\text{ch jet}}$  of 10 GeV/c approximately corresponds to  $p_{T,\text{full jet}}$  of 15 GeV/c. The ALICE and

CMS results are compatible in the region of overlap. Let us also note that the CMS data have an additional scaling uncertainty of 22%, since their pp reference was taken from PYTHIA.

The right panel of Figure 3.44 shows the nuclear modification factor of the charged-particle b jets compared to that of untagged charged-particle jets from ALICE Ref. [6]. Both nuclear modification factors are consistent with unity. This suggests that jets in the given  $p_{T,\text{ch,jet}}$  range may only be subject to mild cold nuclear matter effects, which are below the sensitivity of our measurement. The flavor of the initial parton does not seem to play a significant role, though we need to keep in mind that a large fraction of the b jets stems from gluon splitting.

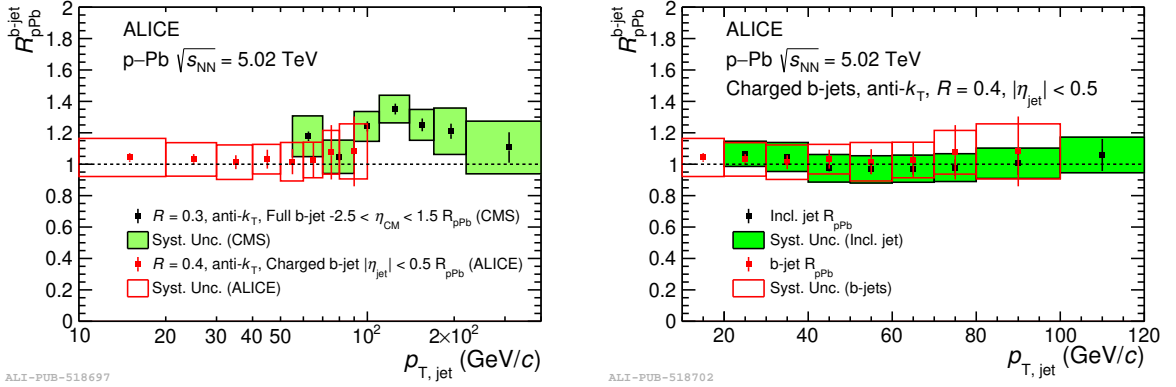


Figure 3.44: Left: The nuclear modification factor  $R_{pPb}$  for charged-particle b jets measured by the ALICE experiment compared with the b-jet measurement from the CMS experiment [11]. Right: The nuclear modification factor of b jets compared to that of untagged jets from Ref. [6].

### 3.12 Summary

This study reports measurements of the charged-particle b-jet spectra in pp and p–Pb collisions at  $\sqrt{s_{NN}} = 5$  TeV in central rapidity in the  $p_T$  range from 10 to 100 GeV/c. The analysis extends the data range previously reported by CMS down to lower transverse momenta. The performed analysis showed that the b-jets constitute only about 2–4% of inclusive untagged jets in pp and p–Pb collision systems. This fraction is consistent with the model predictions provided by the POWHEG+PYTHIA simulation. The nuclear modification factor for charged-particle b jets is found to be consistent with unity within the current precision of the experiment, suggesting that the b-jet production in central p–Pb collisions at  $\sqrt{s_{NN}} = 5$  TeV does not seem to be affected by CNM effects or jet-quenching. The b-jet  $R_{pPb}$  measurements are described by NLO pQCD POWHEG calculations with PYTHIA 8 fragmentation within uncertainties.

# Chapter 4

## ALICE ITS Upgrade

Since 2018, the ALICE detector has been undergoing a major upgrade [141] to be able to cope with ten times increased luminosities in Run 3 and Run 4. The upgraded detector has to be able to take Pb–Pb data at the readout rate of 100 kHz (in Run 2, it was 500 Hz). Throughout Run 3 and Run 4, ALICE is expected to collect 100 times more Pb–Pb events ( $13 \text{ nb}^{-1}$ ) than Run 1 and Run 2. Another motivation for the ALICE upgrade is physics goals that the ALICE collaboration would like to address: in Run 3 and Run 4, the ALICE collaboration will focus mainly on quantifying QGP properties through detailed measurements of heavy-flavor hadrons and low-mass di-leptons, with an opportunity for other studies like precise measurements of light nuclei, anti-nuclei, and hyper-nuclei production. These observables cannot be efficiently triggered. Therefore, ALICE will take all data in a continuous readout mode and apply event selection later in an asynchronous mode [142]. The ALICE upgrade strategy follows two different paths: improving the efficiency of collision vertex reconstruction and tracking at low  $p_T$ , and increasing the readout rate.

The upgrade affected the following subsystems:

- A new beryllium beam pipe having a radius of 19.2 mm and 0.8 mm thick walls was installed. The new pipe allows for moving the first layer of ITS closer to the interaction point and improving track reconstruction of low- $p_T$  particles.
- A new Inner Tracking System (ITS2) formed by seven layers of silicon monolithic active pixel sensors (MAPS) replaced the old ITS. The new ITS covers the pseudorapidity range  $-1.2 < \eta < 1.2$ . The upgraded ITS is expected to have improved pointing resolution, tracking efficiency, and  $p_T$  resolution for low  $p_T$  particles [143].
- A new Muon Forward Tracker covering  $-3.26 < \eta < -2.45$  was instrumented in front of the absorber of the Muon Spectrometer to provide additional tracking

planes. These planes will improve the pointing resolution of muon tracks to the primary vertex, and the invariant mass resolution of dimuon pairs [14].

- Upgrade of the Time Projection Chamber (TPC) readout, consisting of the replacement of the wire chambers of the electron detector with Gas Electron Multiplier (GEM) detectors and the electronics update allowing for a continuous readout [144].
- Upgrade of the readout electronics of the Transition Radiation Detector (TRD), Time Of Flight detector (TOF), and Muon Spectrometer for high-rate operation and data taking in the continuous readout mode [14].
- Upgrade of the forward trigger detectors [14].
- Upgrade of the online systems and offline reconstruction and the analysis framework (ALICE O<sup>2</sup>).

## 4.1 ITS Upgrade design objectives

Measurements of low-momentum heavy-flavor hadrons or low-mass dileptons require a low-material budget inner silicon tracker, which allows for accurate reconstruction of displaced secondary decay vertices, and provides highly efficient detection of low transverse momentum particles. ITS2 was designed to improve vertex reconstruction capabilities and tracking at low  $p_T$ . The main focus was on the following:

- **Impact Parameter resolution** which was improved by a factor of 5 along the beam axis and by a factor of 3 in the transverse plane w.r.t. the old ITS, as illustrated in Figure 4.1 (left). This performance was achieved by reducing the radius of the beam pipe from 29.8 to 19.2 mm and moving the inner layer of the ITS closer to the interaction point. The resolution of the impact parameter at low  $p_T$  is mainly limited by multiple Coulomb scattering, so reducing the material budget is an efficient way to improve the efficiency and resolution of the tracking. The inner barrel of the new ITS has a material budget of 0.3%  $X_0$  per layer only, and the outer barrel has 1.0%  $X_0$  per layer. Here, advantageous features of the ALPIDE monolithic active-pixel sensor were exploited, which do not require extensive cooling due to a highly optimized electrical power distribution scheme and which could be thinned down to 50  $\mu\text{m}$  in the inner barrel layers (to be compared with the 350  $\mu\text{m}$  thick SPD sensors).
- **Tracking efficiency and  $p_T$  resolution at low  $p_T$ .** The updated ITS increases the number of silicon sensor layers from 6 to 7. They all employ the same silicon pixel

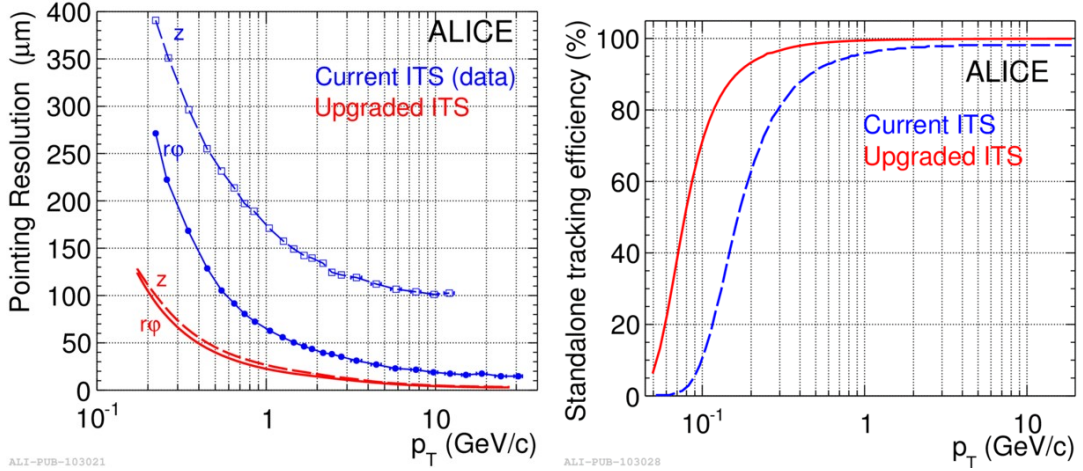


Figure 4.1: Left: Impact parameter resolutions for the old ITS (Pb–Pb 2010 data) and the upgraded ITS. Taken from [143]. Right: Stand-alone tracking efficiency for the old ITS and the upgraded ITS. Taken from [143].

technology based on monolithic active-pixel sensors. The achieved improvement in tracking efficiency is important for particles with  $p_T < 1$  GeV/ $c$  as illustrated in the right panel of Fig. 4.1.

- **Readout rate** could be increased from 500 Hz (Run 2) to 100 kHz for the Pb–Pb collision system and up to 200 kHz for pp [145].
- **Maintenance.** The updated ITS has a possibility of fast insertion or removal of components for maintenance during end-of-year stops. This is facilitated by placing all services, such as voltage supply, readout, and cooling system, on one side of the detector.
- **Radiation hardness.** Throughout Run 3 and Run 4, the ALPIDE sensors of the innermost layer will obtain a Total Ionization Dose (TID) of 270 krad and a Non-Ionizing Energy Losses (NIEL)  $1.7 \times 10^{12}$  1 MeV  $n_{eq}$   $cm^{-2}$ . The increased radiation tolerance of the ALPIDE sensors is required because of the increased luminosity and closer positioning of the first ITS layer to the beam. The project requires the sensors to sustain radiation loads that are ten times higher than expected while keeping detection efficiency greater than 99% and the value of the fake-hit rate less than  $10^{-6}$  hits/pixel/event [19].

## 4.2 Layout of the Upgraded ITS

Figure 4.2 shows a schematic view of the new ITS. The ITS consists of seven concentric cylindrical layers equipped with the same chip design based on the monolithic active pixel

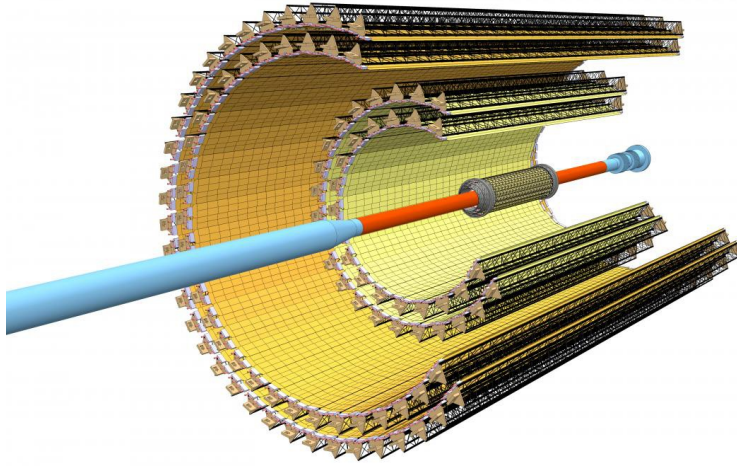


Figure 4.2: Scheme of the Upgraded ITS system. Taken from [145].

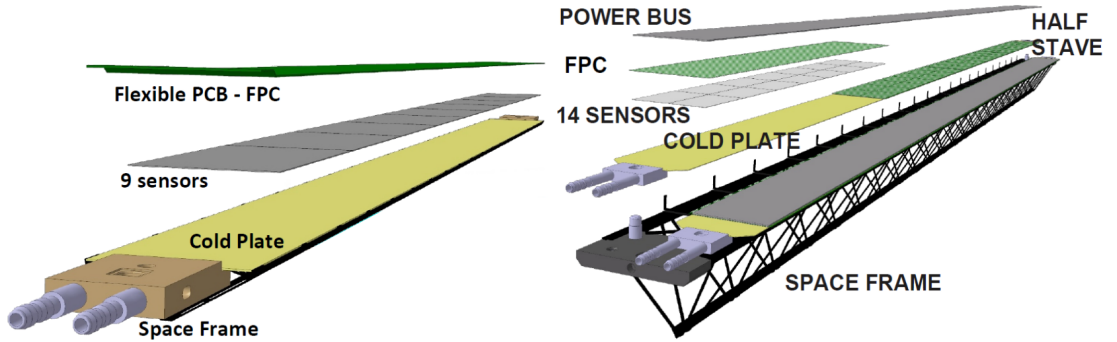


Figure 4.3: Structure of an IB (left) and OB (right) stave of the upgraded ITS. Taken from [146].

sensor technology. The first three layers of the new ITS form the Inner Barrel (IB), while the two middle and two outer layers make up the Outer Barrel (OB). The pseudorapidity range of the new ITS is  $|\eta| < 1.2$ , which corresponds to 90% of the most luminous region of the beam. More information on the geometry of the IB and OB layers can be found in Table 4.1. Each ITS layer is divided into mechanically independent segments called staves in the azimuthal direction. The stave comprises a support structure, cooling pipes, and a Hybrid Integrated Circuit (HIC). The IB HIC hosts one row of 9 ALPIDE sensors, whereas the OB HIC has two rows, each having 7 APLIDEs. The Staves for the Inner Barrel and the Outer Barrel are illustrated schematically in Fig. 4.3. The total detector surface amounts to  $10.3 \text{ m}^2$  containing about  $12.5 \times 10^9$  pixels with binary readout.

### 4.3 MAPS for the ALICE ITS Upgrade

To fulfill the ITS upgrade design objectives, the new ITS employs thin silicon chips with a small pixel pitch. It has been shown [143] that the requirements of the Technical Design

	Layer	Radius (mm)	Stave length (mm)	Number of Staves
Inner Barrel	0	23	271	12
	1	31		16
	2	39		20
Outer Barrel	3	194	843	24
	4	247		30
	5	353	1475	42
	6	405		48

Table 4.1: Layout of the upgraded ITS. The numbers are taken from [103].

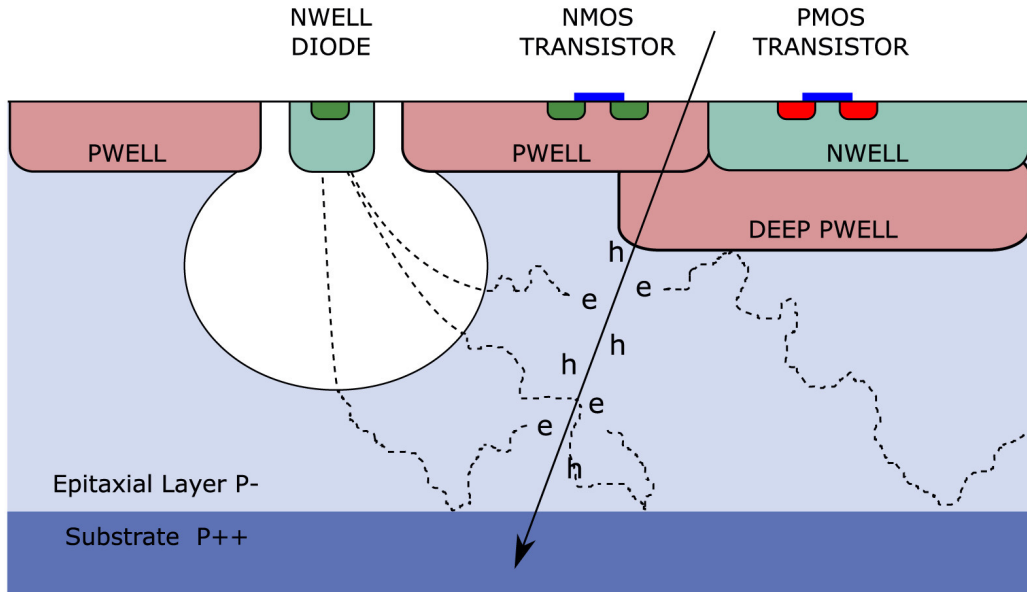


Figure 4.4: Cross section of a MAPS made by the TowerJazz 0.18  $\mu\text{m}$  technology, taken from [143].

Report (TDR) are met with monolithic active pixel sensors (MAPS), which integrate sensitive volume and readout electronics inside a single silicon matrix. The MAPS technology has been used in experimental physics for the last 20 years. The first large-scale detector based on this technology was the STAR PXL detector at RHIC [147], which was equipped with the ULTIMATE chips produced using the AMS 0.45  $\mu\text{m}$  technology. However, this detector did not meet the ALICE ITS upgrade requirements due to its low readout speed, high material budget values, and low resistance to radiation. As an alternative to this design, the CMOS TowerJazz technology [148] was chosen.

The CMOS technology by TowerJazz shows good radiation hardness capabilities due to the small transistor feature size of 0.18  $\mu\text{m}$  and the gate oxide thickness of about 3 nm [149], [150]. The next important feature of the chosen technology is the high number of metal layers available (up to six) that makes it possible to create an energy-efficient high-density electronic circuit and maximize the sensitive area. Figure 4.4 represents a

Parameter	Inner Barrel	Outer Barrel
Chip dimensions	15 mm × 30 mm	
Sensor thickness	50 μm	100 μm
Spatial resolution	5 μm	10 μm
Detection efficiency	> 99%	
Integration time	< 10 μs	
Fake-hit rate	< 10 <sup>-6</sup> event <sup>-1</sup> pixel <sup>-1</sup>	
Power density	< 300 mW/cm <sup>2</sup>	< 100 mW/cm <sup>2</sup>
Temperature	20 C° to 30 C°	
TID radiation hardness	2700 krad	100 krad
NIEL radiation hardness	1.7 × 10 <sup>13</sup> 1 MeV n <sub>eq</sub> cm <sup>-2</sup>	1 × 10 <sup>12</sup> 1 MeV n <sub>eq</sub> cm <sup>-2</sup>

Table 4.2: Sensor requirements for the IB and the OB. Taken from [103].

schematic cross-section of the ALPIDE sensor. A high-resistivity ( $\approx 1 \text{ k}\Omega / \text{cm}$ ) epitaxial p-type layer with a thickness of 18 – 40 μm is formed by the method of chemical deposition. The epitaxial layer, mostly only partially depleted, acts as a sensitive volume. The n-well diode and the active area of the p-type create the depleted region of the PN-junction. The reverse substrate bias voltage  $V_{\text{BB}}$  (0 to  $-6\text{V}$ ) can be used to operate the size of the depleted region and to increase the signal-to-noise ratio. Another characteristic feature of the ALPIDE sensor is the deep p-well, which prevents the n-well of PMOS transistors from collecting charges created by ionization. Table 4.2 quotes the requirements of the ITS Technical Design Report (TDR) [143] for silicon sensors in the IB and the OB.

The front-end circuit of the ALPIDE sensor features 8-bit digital-to-analog converters (DACs), allowing to configure and control of detector response. The most important DACs are  $V_{\text{CASN}}$  and  $I_{\text{THR}}$ , which can modify the shape and amplitude of the signal flow resulting in the threshold charge changes necessary to activate a pixel [103].

## 4.4 ALICE computing model

As discussed in Section 2, in Run 3 and Run 4, ALICE will need to cope with two orders of magnitude higher collision rates than in Run 2. The full analysis process of such a large data volume is a technological challenge for the detector systems and the online/offline computing system. Expected data throughput from Pb–Pb collisions in Run 3 is 3 TB/sec, roughly two orders of magnitude higher than in Run 1. A new design of the ALICE Computing Model for Run 3 was developed to minimize the computing system requirements for computing power and storage size. It reduces the detector data volume as early as possible during data flow. The new Computing Model introduces two stages of data reconstruction. In the first stage, ALICE processes the reconstruction data



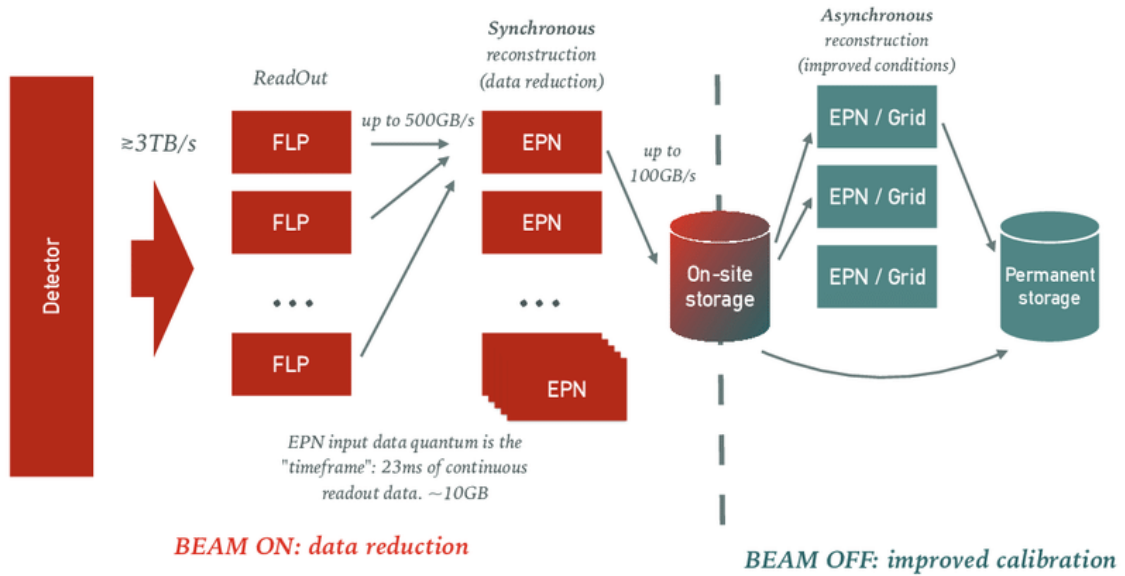


Figure 4.5: The Computing architecture of ALICE in Run 3. Taken from [151].

synchronously with the data acquisition. The second stage of the analysis is performed in asynchronous mode using the final calibration to achieve the required data quality. To cope with the difficulties of Run 3, the ALICE Collaboration, together with the FAIR Software Group at GSI, has developed a new software framework  $O^2$  [151]. The  $O^2$  system performs detector calibration and data reconstruction simultaneously with data collection [142]. The functional flow of the  $O^2$  hardware system follows the sequence of steps shown in Figure 4.5:

1. First Level Processors (FLP) are hosted by a farm consisting of 150 parallel computing units. It collects the detector data in the continuous readout mode or triggered by the MB trigger. Data are sent to FLP through the GBT front-end links and DDL optical links at a speed of 3 TB/sec [151].
2. FLP performs primary processing of the data fragments, which do not require the full detector output. That includes a first local calibration and detector patterning. These activities involve RAW data decoding, ITS and TPC cluster reconstruction, and further data compression by lossless algorithms. Local data processing on FLP provides average compression by a factor of 6. As a result, data from the detectors are formed into Sub-Time Frame (STF) data packets. STF contains data accumulated during a 20 ms long period around the given bunch crossing LHC clock tick.

3. Further STFs are distributed with a rate of 500 GB/s to the Event Processing Nodes (EPN). In this step, several STFs corresponding to the same time period are merged to form a Time-Frame (TF). EPN conducts a joint analysis of data from several detector systems and performs global calibration, track reconstruction, and cluster-track mapping. This procedure further compresses the data volume by a factor of 5 by rejecting all non-physical analysis clusters (noise, background from  $\delta$ -rays, and beam-gas collisions).
4. Fully compressed TFs are moved to on-site storage with speed up to 100 GB/s. Data are then archived in the O<sup>2</sup> Facility or Tier-1 data centers.
5. In the last step, one performs asynchronous global data reconstruction on EPN or the Grid, resulting in event summary data (ESD) files. The ESD contains lists of reconstructed tracks/particles and global event properties from all subdetectors. ESD files are permanently stored at the O<sup>2</sup> facility. ESD files are compressed to the Analysis Object Data (AOD) by filtering additional detector data for further analysis.

Apart from individual tasks dedicated to each computing step, the data processing flow contains shared activities integrated into the different processing stages on the FLPs and EPN. These processes include calibration and Quality Control tasks.

#### 4.4.1 Calibration

There are two phases of reconstruction and calibration: synchronous and asynchronous. The goal of synchronous calibration is to reduce the data volume as much as possible and maintain the required physics performance. The asynchronous calibration aims to provide analysis grade data for the entire detector, analyzable in terms of individual collisions.

Synchronous ITS calibration runs on the FLP unit during ITS standalone runs. ITS calibration will focus on identifying noisy/dead pixels and measurements of threshold activation charge and  $V_{\text{CASN}}$  plus  $I_{\text{THR}}$  values. A more detailed description of this procedure will follow in the following chapters. The information about the coordinates of noisy/dead pixels is transferred to the Condition and Calibration Data Base (CCDB) to exclude problematic pixels from further data processing.

#### 4.4.2 Quality Control software

Quality Control (QC) aims to provide a quick automatic quality assessment of data recording and handling processes, such as reconstruction and calibration during synchronous or asynchronous steps. This information helps shifters assess data quality during calibration

or data-reconstruction processes and decide whether data-taking should continue. Consequently, QC inherits the functionality of Quality Assurance (QA) and Data Quality Monitoring (DQM) from Run 1 and Run 2 [142].

Shifters and experts can access the QC results via a web application, which allows the display and manipulation of ROOT-based objects in a dynamic way. The system provides a RESTful API [152], which enables running most applications on the client side with JavaScript, so the user cannot affect other QC processes.

### 4.4.3 Condition and Calibration Data Base

The CCDB stores information related to detector calibrations, which is needed mainly for reconstruction and simulation purposes. The database design follows the architecture of the Offline Condition Data Base (OCDB) from Run 1 and Run 2 [153]. The content of the objects stored in the CCDB will be time-dependent since the data model in Run 3 and Run 4 will use time frames. Each object in the CCDB database should have associated metadata describing its main properties, such as the data-taking period to which they refer, the calibration type, the source, etc.

# Chapter 5

## ITS QC system in Run 3

### 5.1 ITS QC system architecture

The architecture of the ITS QC system is shown in Fig. 5.1. The ITS data quality assessment process includes two independent parts: *synchronous QC* that runs in parallel with data collection and *asynchronous QC* mainly used for the offline reconstruction of the full data sample. The final ITS run quality is assigned based on quality flags from both stages and provides important input to a discussion in the Data Preparation Group about run reconstruction, since the ITS is considered to be a crucial detector for the ALICE data quality in Run 3.

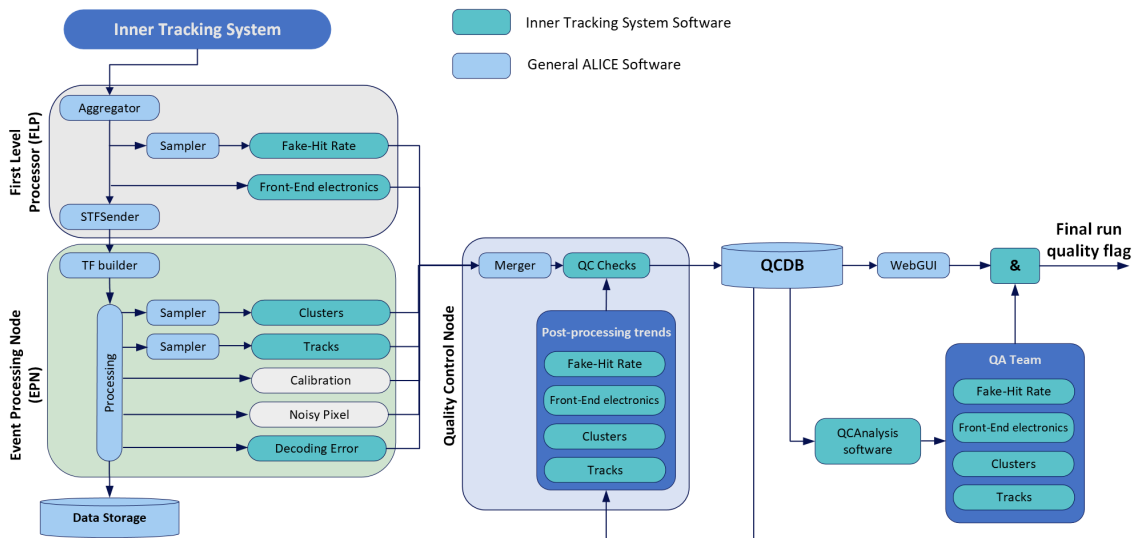


Figure 5.1: Scheme of the ITS QC system architecture in Run 3. Blocks with turquoise color correspond to software developed especially for ITS. Blocks with Calibration and Noisy Pixel are depicted in distinct colors since they are not parts of the physics data-taking workflow.

### The synchronous QC

This part of ITS QC supplements the standard ALICE data flow described in Chapter 4.4 with a set of user programs called *QC tasks*. Each QC task analyzes detector data in parallel with data-taking and generates monitoring objects (MO) such as ROOT histograms. A more detailed description of the ITS synchronous tasks will be provided in 5.1.1. The data flow in the synchronous QC can be described as follows:

1. Depending on the complexity of the QC analysis, the *Dispatcher* system distributes the complete or sampled detector data to the QC workflow. In the latter case, the pseudo-random selection of the data samples is carried out by the module called *Sampler*.
2. Since the synchronous QC analysis is executed on several computing machines in parallel, complete MOs are obtained by combining results from the several FLPs and EPNs with the *Merger* module running on a dedicated node.
3. In the next step, quality flags are attributed to combined MOs with *QC checks* algorithms by comparing the given MO with a reference value. For example, a QC task will notify a shifter if the given value is below or above a defined limit. In the future, it is planned that the QC framework will automatically modify the reference values according to the run and data-taking conditions.
4. Finally, MOs are sent to the CCDB for visualization with a web interface, where a shifter or detector expert can examine them to inspect the data quality. Furthermore, the data stored in the CCDB are used to create long-term trends of the detector parameters which are further handled by the *the Post-Processing Trends* module.

### The asynchronous QC

The asynchronous QC analysis is carried out by the ITS QA team using a special QC-Analysis software developed by Ivan Ravasenga [154]. This software gives an averaged summary of the detector parameters studied layer by layer for each run using QC data stored in CCDB. The QA group daily receives the run list for analysis and reports results through a ticketing system with a weekly report of observed problems. Such an approach introduces additional and more detailed control of detector data quality by inspecting the evolution of the data parameters over time.

### 5.1.1 Overview of the ITS QC tasks

QC tasks are user-defined algorithms that study various aspects of the detector data. The QC tracks are executed on the FLP and EPN machines depending on the specifics of inspected data. The available tasks in the ITS QC system are:

- QC Tasks run on the FLP machine:
  - **Fake-hit rate task** gives the level of fake hit rate for each layer during cosmic runs. This task can also be used during physics collisions to estimate the level of detector occupancy. The task estimates the number of noisy pixels and provides ALPIDE sensor hitmaps, which can be used to study occupancy anomalies in more detail.
  - **Front-End electronics task** provides information on detector readout performance in terms of the number and type of trigger flags, the payload size of the detector links, and the number of links that went into an error state.
- QC Tasks run on the EPN machine:
  - **Decoding Error task** gives an overview of the errors that occurred during data decoding.
  - **Cluster task** assesses the quality of the ITS clusters providing comprehensive information on the cluster size, topology, and occupancy. This task is described in more detail in Sect. 5.2.2.
  - **Track task** monitors the quality of the ITS standalone tracks. This will be discussed in more detail in Sect. 5.2.3.
  - **Calibration task** monitors the quality of the in-pixel discriminator thresholds and ALPIDE registers which are used to tune the thresholds for data taking. It is attached to calibration runs that are executed periodically at every beam dump to assess the calibration quality. Further information about this task will be given in Sect. 5.2.4.
  - **Noise Pixel Task** provides the number of noisy pixels using the full data sample. Values given by this task can be used to cross-check the FHR task.

The above mentioned synchronous QC tasks are accompanied by corresponding *QC Checks*, *Post-Processing Trend*, and *QCAnalysis* scripts.

## 5.2 Development of the ITS QC tasks

As a part of my service work, I have contributed to the development and maintenance of several ITS synchronous QC Tasks, including the Cluster, Track, and Calibration Tasks. Also, I have developed a QC Threshold task which measures the threshold charge of ALPIDE pixels and identifies dead pixels. Finally, I developed the Monte Carlo QA tool based on the QC workflow to benchmark ITS standalone tracking efficiency using simulated MC data. The developed software was used during the ITS commissioning and ALICE data-taking. In the second half of 2022, I was selected as the coordinator of the ITS QC group. In this period, ALICE collected the first data from Run 3.

### 5.2.1 ITS QC Threshold Task for the commissioning stage

The QC Threshold task collects the average in-pixel discriminator thresholds of every chip of the ITS and detects amount of dead pixels for each chip (pixels which did not register any hit for any threshold level). The calibration procedure employs the threshold task to mark problematic pixels that should be excluded in subsequent data-taking steps. This task was used find the optimal parameters for which ALPIDE sensors provided uniform response during ITS commissioning [155]. During this time, the QC threshold task was utilized in an offline mode using as input previously recorded raw data files. Later, it was replaced by a new online Calibration O<sup>2</sup>-based workflow used in Run 3 data-taking.

#### Measurement of threshold activation charge

To measure the in-pixel discriminator thresholds of the ITS pixels, a special calibration run called was performed to characterize each ITS pixel in terms of an activation function. In this run, a testing charge  $q_{\text{inj}}$  is injected into the signal processing chain of each pixel using a capacitor that is implemented in each pixel [103]. The probability of activating the pixel is obtained by counting the number of trials  $N_{\text{triggered}}$  when the given pixel is fired if a charge  $q_{\text{inj}}$  was injected; see Eq. (5.4).

$$r(q_{\text{inj}}) = \frac{N_{\text{triggered}}(q_{\text{inj}})}{N}. \quad (5.1)$$

By default, each charge is injected  $N = 50$  times. The pixel activation function can be obtained by varying  $q_{\text{inj}}$  from 1 to 50 digital-analog converter (DAC) units, where 1 DAC unit corresponds to a charge of 10 electrons. Assuming a Gaussian distribution of the noise, we can describe the pixel activation function by:

$$f(q_{\text{inj}}) = \frac{1}{2} \left[ 1 + \operatorname{erf} \left( \frac{q_{\text{inj}} - Q_{\text{THR}}}{\sqrt{2}\sigma} \right) \right], \quad (5.2)$$

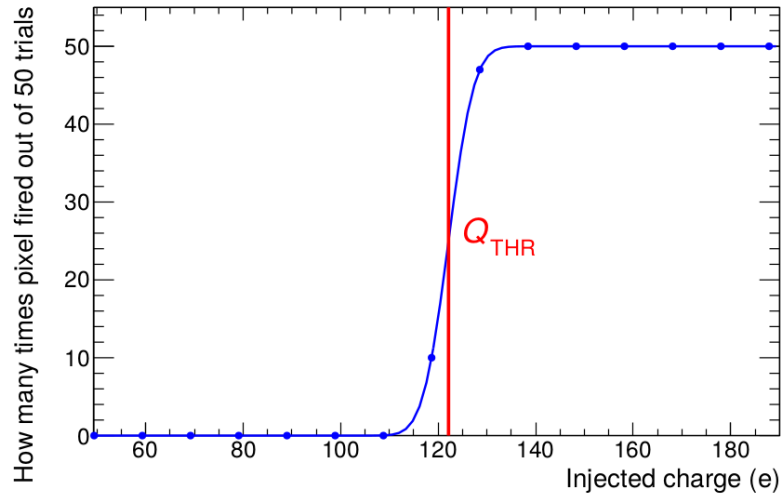


Figure 5.2: Example of an activation function of a single ALPIDE pixel.  $Q_{\text{THR}}$  corresponds to a charge which activates the pixel with a 50% probability. Taken from [156].

where  $Q_{\text{THR}}$  is the threshold charge value, which gives the amount of charge required to activate a pixel with a 50% probability. The parameter  $\sigma$  describes temporal noise. The error function [156] is defined as:

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (5.3)$$

An example of the fit of the activation function by Eq. (5.2) can be seen in Fig. 5.2.

A direct search for the function parameters by fitting each of the  $12 \cdot 10^9$  pixels would be impossible during synchronous data reconstruction because it would require large computational and memory resources. On the other hand, a geometric interpretation of the  $S$ -function allows us to estimate the threshold value without making the fit. The inspection of Figure 5.2 suggests that the value of the activation charge  $Q_{\text{THR}}$  can be assessed using a simple formula:

$$Q_{\text{THR}} = \left( 1 - \frac{N_{\text{triggered}}}{N_{\text{injected}}} \right) \cdot 50 \text{ DAC}. \quad (5.4)$$

The formula assumes that the amount of injected charge increased in steps 1, 2, ..., 50 DAC. Out of the total number of injection trials  $N_{\text{injected}}$  there were  $N_{\text{triggered}}$  cases registered when the given pixel was fired. The resulting threshold value obtained from Eq. (5.4) was filled to the corresponding monitoring objects of the QC Threshold Task.



### Workflow description

The stable version of the QC workflow that performs the data analysis in the synchronous data-taking mode was not released during the early stages of the ITS commissioning. As a temporary solution, recording detector RAW data to files and performing the asynchronous analysis later was proposed. For this purpose, Ruben Shahoyan developed a RAW data decoder followed by a specially designed data sampler that provided output variables in a form compatible with the QC. I have contributed to the development of a new sampling algorithm that significantly reduced the memory and computing time of the QC workflow. Later in the ITS upgrade project, this workflow was replaced by a more sophisticated  $O^2$  workflow. Both workflows mentioned above provide objects that carry information about activated pixels and a list of the ErrorIDs during digitization. The QC Threshold task processes these objects and fills Monitoring Objects (1D and 2D histograms) stored in the CCDB database for further analysis. For illustration, Figures 5.3 and 5.4 show a few histograms provided by this task during the ITS commissioning stage.

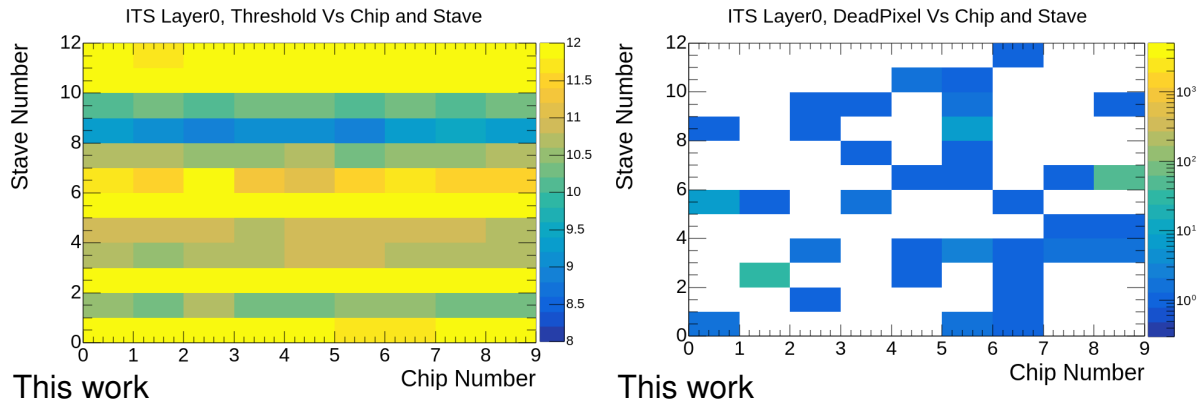


Figure 5.3: Examples of MOs provided by QC Threshold Task for the non-tuned detector. Left: The average threshold value in each ALPIDE sensor in Layer 0. The  $z$ -axis corresponds to the mean threshold activation charge in DAC units calculated from all pixels of the given ALPIDE sensor. The lower  $Q_{\text{THR}}$  value of the Stave 8 can be explained by different VCASN values for this stave compared to others. Right: The number of dead pixels in each ALPIDE sensor of Layer 0.

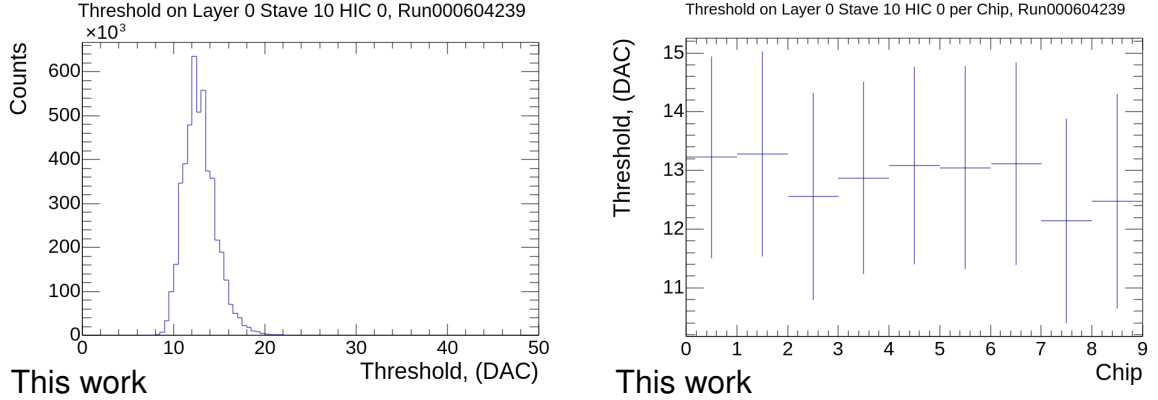


Figure 5.4: Left: Distribution of the threshold charge in pixels of the Stave 10 of Layer 0. Right: Distribution of the average threshold value in the chips of the Stave 10 in Layer 0.

### 5.2.2 ITS QC Cluster Task

The ITS QC cluster task evaluates performance of the clustering procedure which combines adjacent activated pixels to clusters. This task is executed on the EPN machines since the clusterizer and the ITS tracker are running in this environment due to the requirement of high computing performance. The large data flow was handled by multi-threading processing of data with the OpenMP library [157]. In addition, the sampling rate of data flow was decreased to 30%, which is a good compromise between the performance and data volume needed to evaluate the quality. Figure 5.5 shows examples of plots provided by the ITS Cluster Task during Run 3 pp collisions at  $\sqrt{s} = 13.6$  TeV. The shifter will be notified if an average cluster size or occupancy is different from expectations from MC simulations. In the future, it is planned to develop automatic tuning of the QC checks depending on the type of the run (heavy-ion or pp collisions), since the detector performance varies with the event multiplicity.

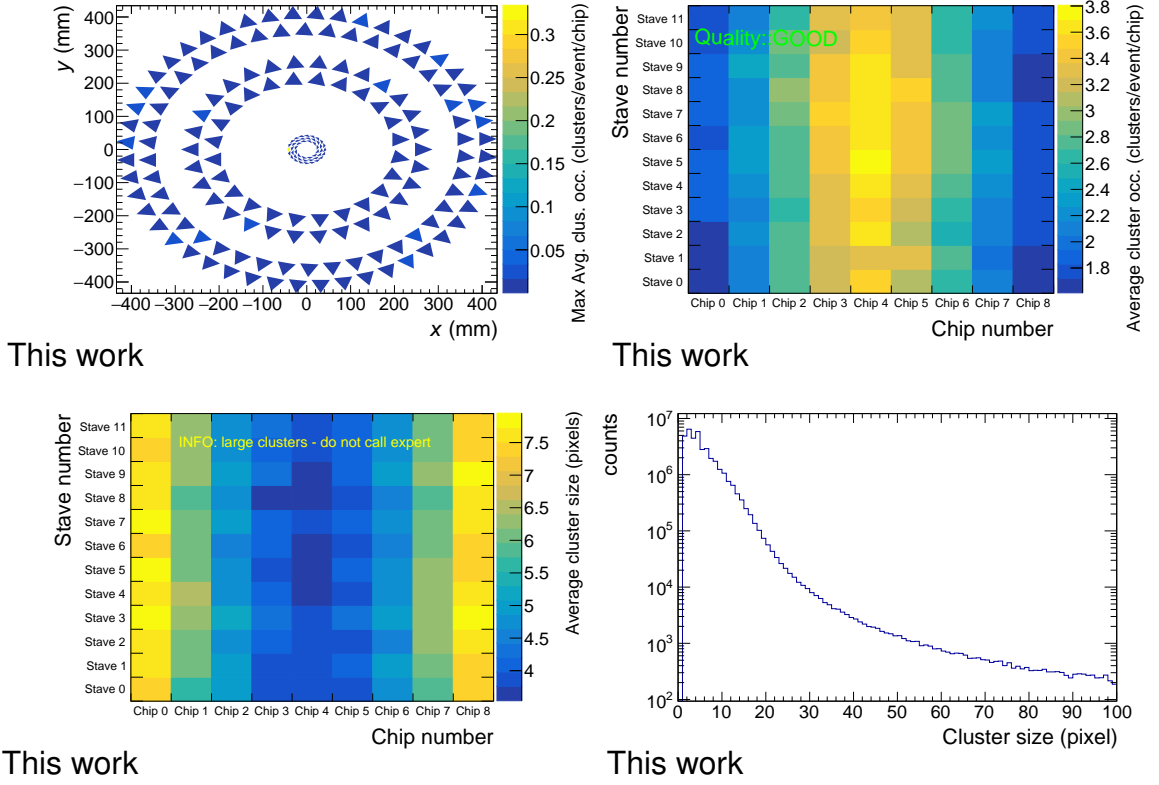


Figure 5.5: Examples of MOs produced by the ITS Cluster Task as they are seen by an ALICE QC shifter. Top left: maximum value of cluster occupancy per each ITS stave. Each triangle corresponds to a separate stave. Data were collected during the pp pilot beam with  $\sqrt{s} = 900$  GeV. Top right: average cluster occupancy for each chip in Layer 0. The central chips demonstrate higher occupancy since they are closer to the primary vertex located around  $z = 0$  cm. Bottom left: average cluster size for each chip in Layer 0. Chips at the borders of the stave have larger cluster sizes because of a larger number of inclined tracks in this region. Bottom right: cluster size distribution for all clusters in Layer 0. The last three plots were obtained for the run 526532 (pp collisions at  $\sqrt{s} = 13.6$  TeV, interaction rate 500 kHz, framing rate 202 kHz).

### 5.2.3 ITS QC Track Task

This task provides the quality assessment flags for the ITS standalone track reconstruction procedure. The ITS reconstruction workflow uses information about hits from several ITS layers and hence requires a preliminary merging of FLP outputs, therefore, the QC Track task runs on an EPN machine on the full data sample. This task provides histograms showing an angular distribution of tracks, the coordinates of primary reconstructed vertices, and other tracking characteristics such as the number of clusters used in the tracking, the number of ITS tracks per vertex, or the number of vertex contributors (i.e., tracklets

in the ITS IB) per readout frame. Figure 5.6 shows examples of plots provided by the ITS Track Task during Run 3 pp collisions at  $\sqrt{s} = 13.6$  TeV. The MOs mentioned above can detect shifts in the position of the primary vertex or anomalies in the tracking by analyzing the distribution of clusters per track. For example, in the pilot beam campaign, the QC ITS track task detected that the Interaction Point was displaced along the  $z$ -axis by 40 cm from the ideal position  $z = 0$  cm. This shift also was confirmed by other ALICE detector systems.

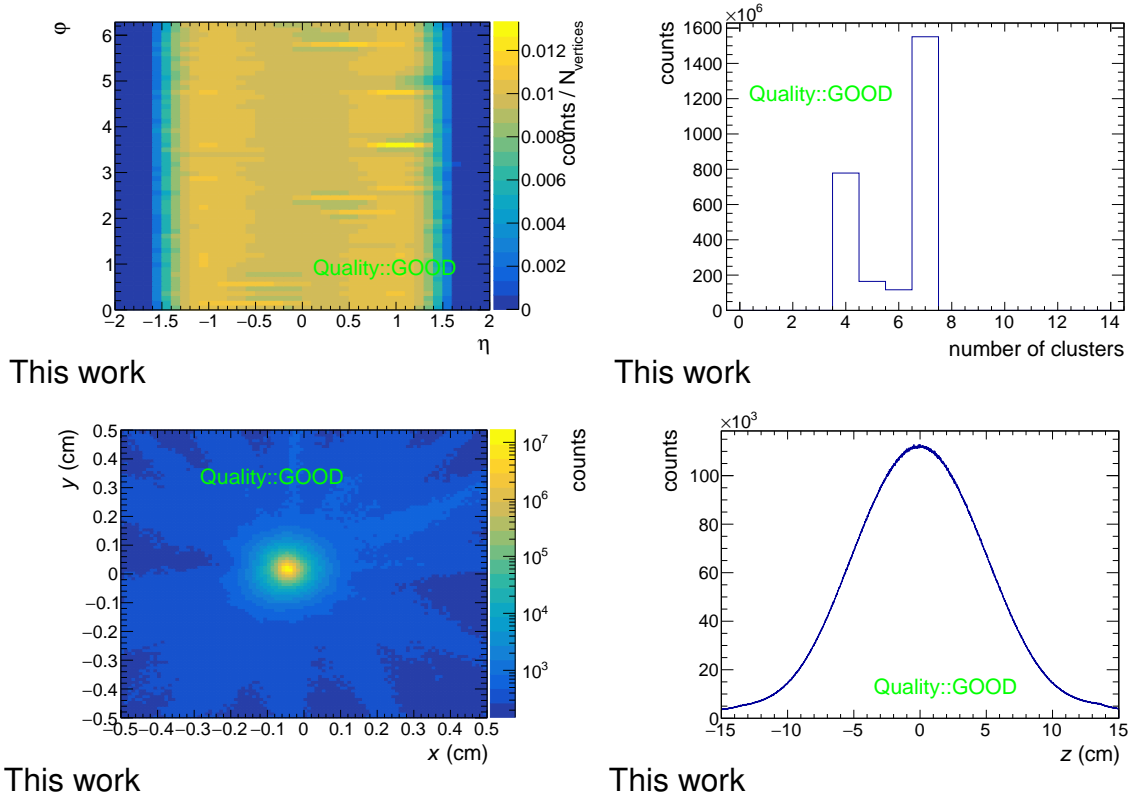


Figure 5.6: Example of MOs produced by the ITS Track Task as they were seen by the central ALICE QC shifter during the run 526532 (pp collisions at  $\sqrt{s} = 13.6$  TeV, interaction rate 500 kHz, framing rate 202 kHz). Top left: Distribution of tracks as a function of azimuthal angle and pseudorapidity. The histogram is normalized per the number of vertices. Top right: Distribution of the number of clusters per track reconstructed in the ITS. The majority ITS tracks have clusters in all 7 layers. The second peak at 4 clusters corresponds to the tracks which have hits only in the OB, while the other contributors come from tracks having hits in OB + IB. Bottom left: Distribution of ITS-tracks primary vertex coordinates in the plane perpendicular to the beam direction. Bottom right: Vertex distribution in the direction of the beam.

### 5.2.4 ITS QC Calibration Task

The QC calibration task is the final part of the ITS calibration workflow executed on EPN machines. This task is used to visualize the results of various ITS tests, including the number of dead and inefficient pixels and measurements of the ALPIDE DACs responsible for threshold adjustment ( $I_{\text{THR}}$ ,  $V_{\text{CASN}}$ ). An example of the MOs produced in the threshold calibration run are illustrated in Fig. 5.7. The provided example demonstrates the uniformity of the detector parameters when default tuning of the threshold to 100  $e^-$  is applied.

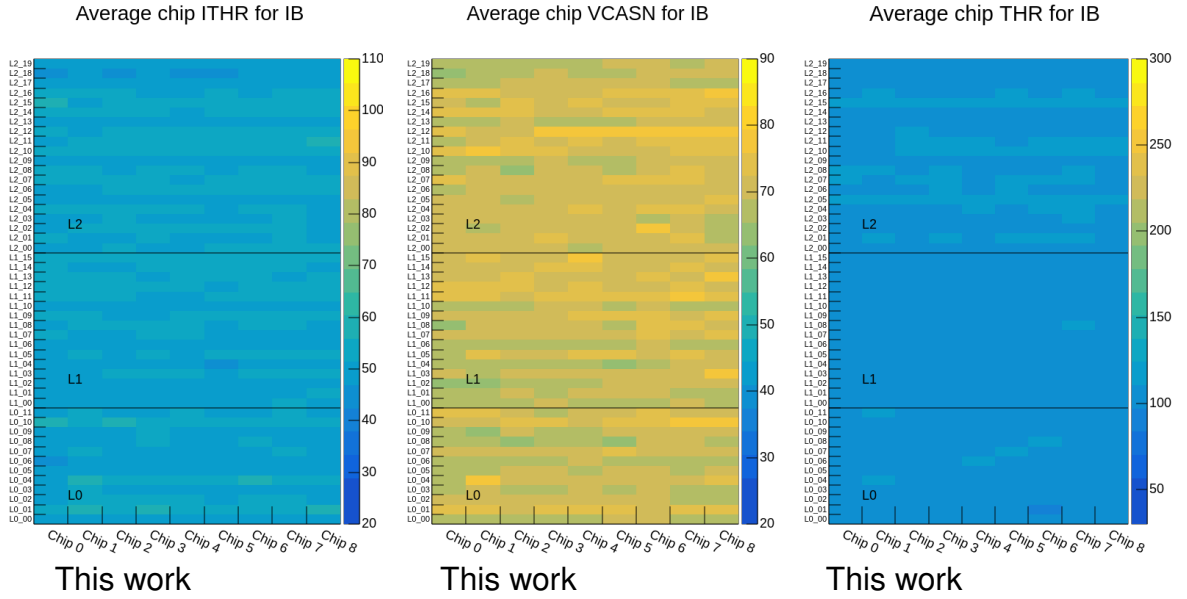


Figure 5.7: Example of MOs produced by the ITS Calibration task for monitoring of ITHR (left), VCASN (middle), and Threshold (right) for the ITS Inner Barrel. Each bin corresponds to one ALPIDE sensor. Values on the  $z$ -axis of the ITHR and VCASN plots are expressed in DAC units, while the  $z$ -axis of the threshold plot is in the electron charges. The solid lines provide the separation between the ITS layers.

### 5.2.5 ITS Monte Carlo Tracking QA task

The ITS MC Tracking QA task is used to estimate the efficiency of the ITS standalone track reconstruction as a function of the track  $p_T$ , and angular and geometric parameters using simulated events. The efficiency of the track reconstruction was defined as follows:

$$\text{eff}(x) = \frac{N_{\text{MC}}^{\text{reco}}(x)}{N_{\text{MC}}^{\text{gen}}(x)}, \quad (5.5)$$

where  $x$  represents  $p_T$ ,  $\varphi$ ,  $\eta$ ,  $z_{\text{vtx}}$  or  $r$ , the distance between the interaction point and the reconstructed vertex.  $N_{\text{MC}}^{\text{gen}}(x)$  corresponds to the number of generated MC particles,

$N_{\text{MC}}^{\text{reco}}(x)$  is the number of tracks reconstructed by the detector, i.e., those which pass certain selection criteria related to the geometry of the detector and the number of hits registered. The track reconstruction efficiency was calculated for MC particles having  $|\eta| < 1.2$  emitted from a vertex displaced less than 1 cm from the beam axis in the transverse plane. The particles were required to have a hit in each of the 7 ITS layers. The fake-track rate was calculated using a similar equation as Eq. 5.5 by taking instead of  $N_{\text{MC}}^{\text{reco}}$  the number of tracks with incorrectly assigned clusters. The ITS MC Tracking QA task was included in the general QA of the ALICE MC production cycles.

Performance of the ITS tracking as obtained from the ITS MC Tracking QA task is illustrated in Fig. 5.8. The current version of the ITS reconstruction workflow provides a tracking efficiency of approximately 99% for tracks with  $p_{\text{T, track}} > 1 \text{ GeV}/c$ , keeping the fake track rate at a level of less than 1%. The analogous high efficiencies are observed for  $\varphi$ ,  $z_{\text{vtx}}$ , and vertex displacement in the transverse plane. Also, this task provides plots with DCA and  $p_{\text{T, track}}$  spectra, which can be used to compare the ITS track reconstruction with the TPC tracking results.

The presented results do not demonstrate the final status of the tracking workflow, which will be further developed to be ready for heavy-ion runs. The developed ITS MC tracking QA task can be used for further optimization of the ITS tracking parameters.

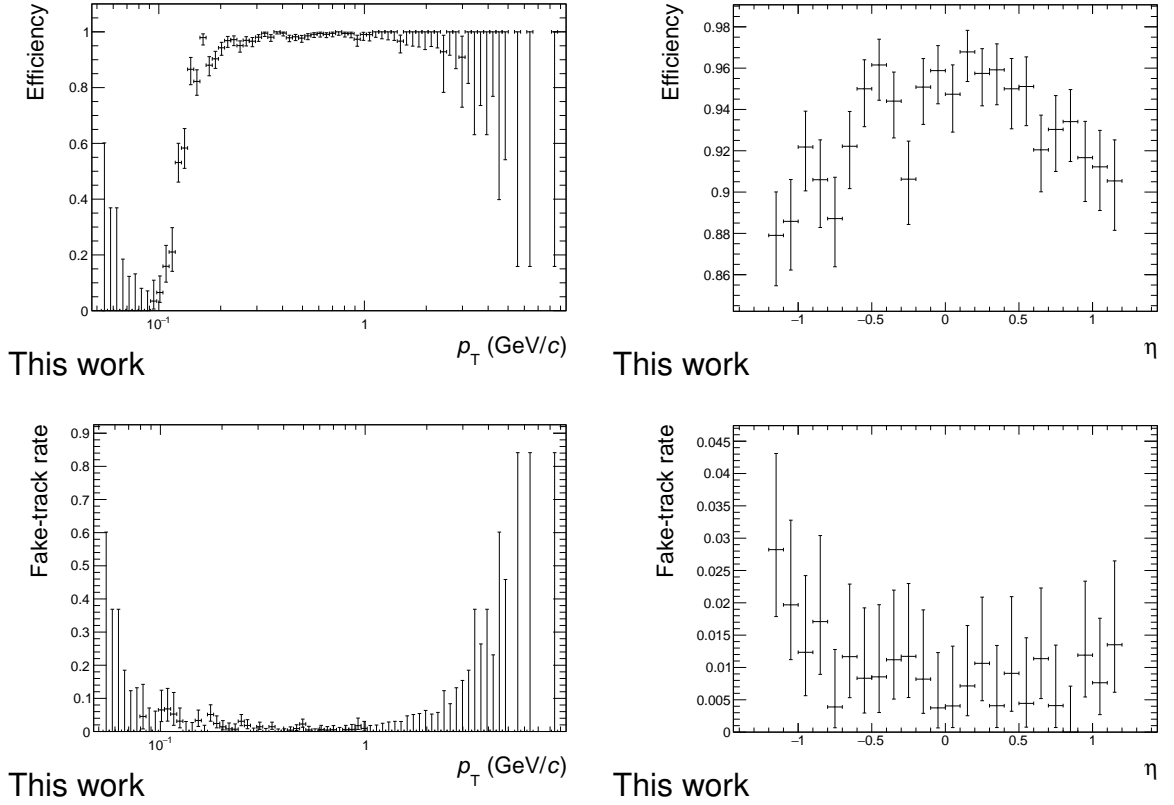
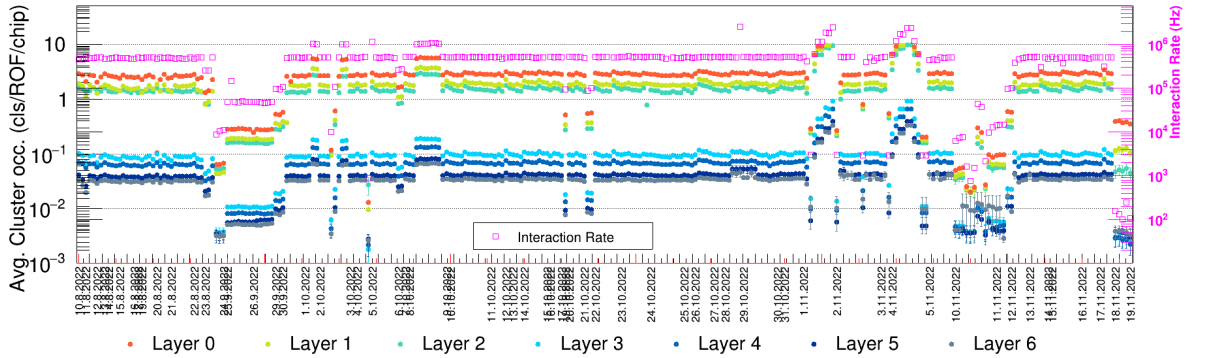


Figure 5.8: Top panels: ITS standalone tracking efficiency as a function of generated particle  $p_T$  (left) and  $\eta$  (right). Bottom panels: fake-track rate as a function of generated particle  $p_T$  and  $\eta$ . The confidence intervals were handled by the TEfficiency class of the CERN ROOT framework [158] using the Clopper–Pearson method [159]. The distributions were obtained using the ALICE MC production cycle anchored to a real data run of pp collisions at  $\sqrt{s} = 13.6$  TeV.

### 5.3 ITS2 performance at the beginning of Run 3

The initial stage of LHC Run 3 took place from 5 July 2022 until the technical stop at the end of November 2022. During this period, ALICE has recorded an integrated luminosity of  $18 \text{ pb}^{-1}$  of pp collisions at the top LHC energy of 13.6 TeV and a default interaction rate (IR) of 502 kHz. ITS has participated in 620 physics runs taking data at a framing rate of 202 kHz [160]. Besides that, ALICE has successfully passed through the stress test during the first-ever Pb–Pb collisions at  $\sqrt{s_{\text{NN}}} = 5.36$  TeV at a low interaction rate of  $\approx 10$  kHz at the end of 2022. This section will be dedicated to a review of the ITS performance during this period using data from the QC tasks developed by the author of this thesis. The analysis will be based on the MOs generated by synchronous ITS QC tasks from GOOD physical runs according to synchronous and asynchronous QA.

### 5.3.1 Performance of the Clustering

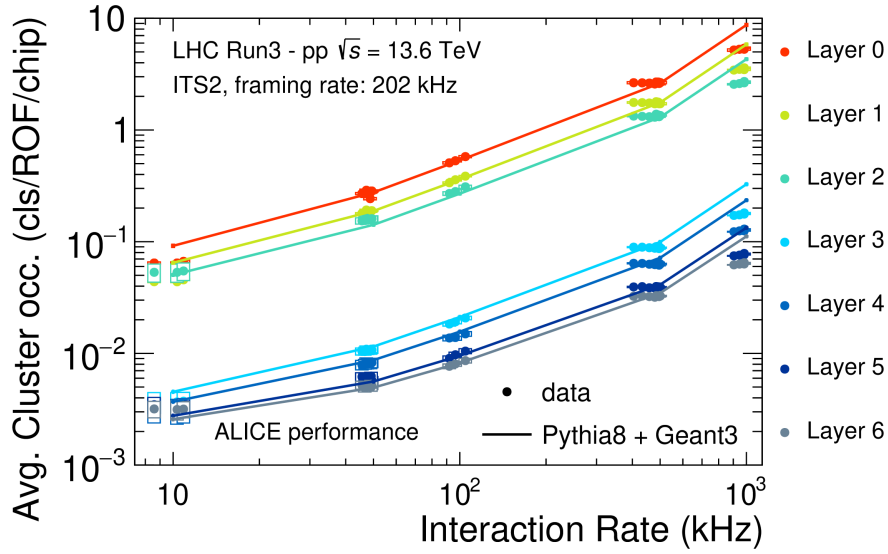


This work

Figure 5.9: Time trend of average cluster occupancy per layer for physical runs with pp collisions at  $\sqrt{s} = 13.6$  TeV. Each point corresponds to a single GOOD quality run. All runs are in chronological order. The  $x$ -axis gives dates corresponding to the beginning of data collection campaigns. The interaction rate was estimated with the FT0 Cherenkov array detector.

The long-term trend of cluster occupancy for each ITS layer in Fig. 5.9 demonstrates the stability of detector performance for the nominal interaction rate. The figure further shows that the ITS QC workflow is capable of data analysis up to interactions rates of the order of 3 MHz as tested during a stress test on 1.11.2022. The trending plot also shows a set of low IR Pb–Pb runs collected on 18.11.2022. The rate of hadronic interactions during this period was about 10 kHz. One can see that the detector occupancy strongly depends on the level of IR and decreases with the radial distance from the interaction point; the MC simulations well describe this trend, as can be seen in Fig. 5.10.

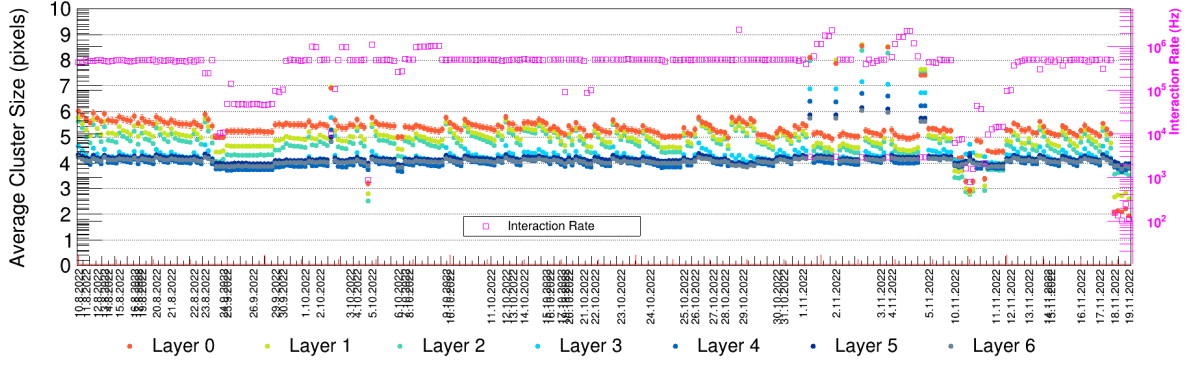




ALI-PERF-528510

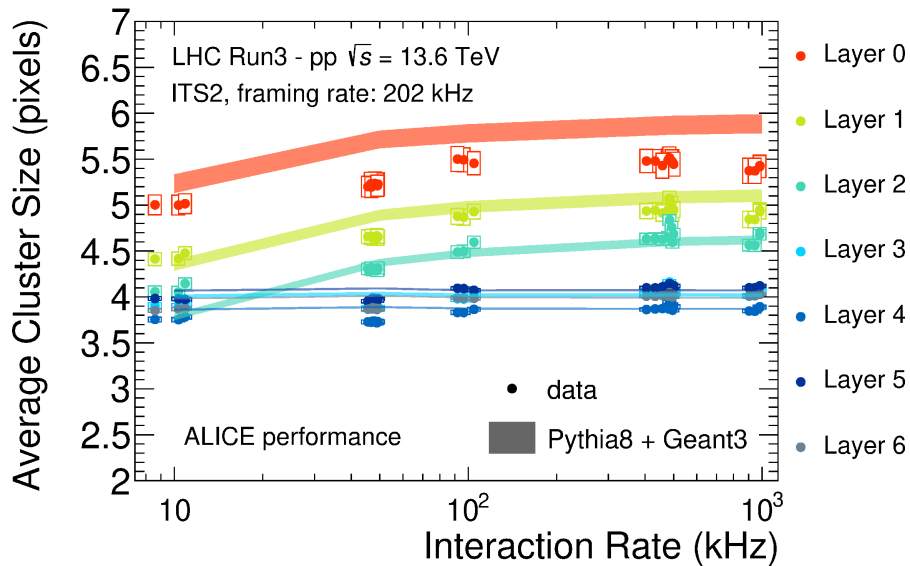
Figure 5.10: Average cluster occupancy as a function of IR for different ITS Layers in real data and MC simulation. The MC predictions are based on Pythia 8 + GEANT 3 simulation performed by I. Ravasenga.

In turn, a similar trend of the average cluster size is almost independent of the interaction rate; see Fig. 5.11. As shown in Fig. 5.12, MC simulation can also reproduce this behavior. The observed discrepancy between the MC and the measured data for the IB layers comes from the problematic noise estimation for the chips that are close to the interaction point. Figure 5.11 exhibits also has several effects that need to be highlighted. The first is the decrease in the average cluster size within one LHC fill, which can be explained by beam-gas collisions that create more inclined tracks, resulting in larger clusters [161]. The pressure of residual gas in the LHC beam lines decreases towards the end of each fill, making the beam-gas interactions less frequent. This also leads to the lower number of particles with large longitudinal momentum component and finally results in a smaller average cluster size. Notably, there are few LHC fills which do not follow the such pattern and have stable average cluster sizes over time (25.09.2022 and 8.10.2022). This can be explained by low IR collisions during these fills, so clusters recorded in these runs probably have a dominant noise nature with small cluster size. The second effect that can be spotted concerns the runs at the beginning of IR scan fills (see runs 1.11 - 5.11.2022). These runs exhibit a larger average cluster size. It is assumed that larger clusters are produced by particles from beam splashes during the TPC interaction rate scans. These particles again have large longitudinal momentum.



This work

Figure 5.11: Time trend of average cluster size per layer for physical pp collisions at  $\sqrt{s} = 13.6$  TeV. The figure is analogous to Fig. 5.9.



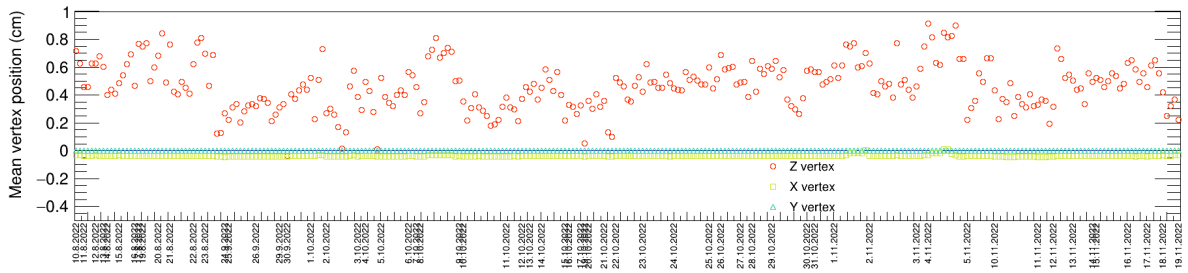
ALI-PERF-528522

Figure 5.12: Average cluster size as a function of IR for different ITS Layers in real data and MC simulation. The MC predictions are based on Pythia 8 + GEANT 3 simulation performed by I. Ravasenga.

### 5.3.2 Performance of the ITS Tracking

Figure 5.13 shows a reconstructed primary seed vertex position as a function of time. Here, one can see that the primary vertex has an average shift of 0.5 cm for the  $z$ -axis and almost no shift (several hundred  $\mu\text{m}$ ) for the  $x$  and  $y$  directions. This observable does not provide the actual position of the ALICE primary vertex, which will be obtained later during the global tracking, but it can be used as an immediate indication whether

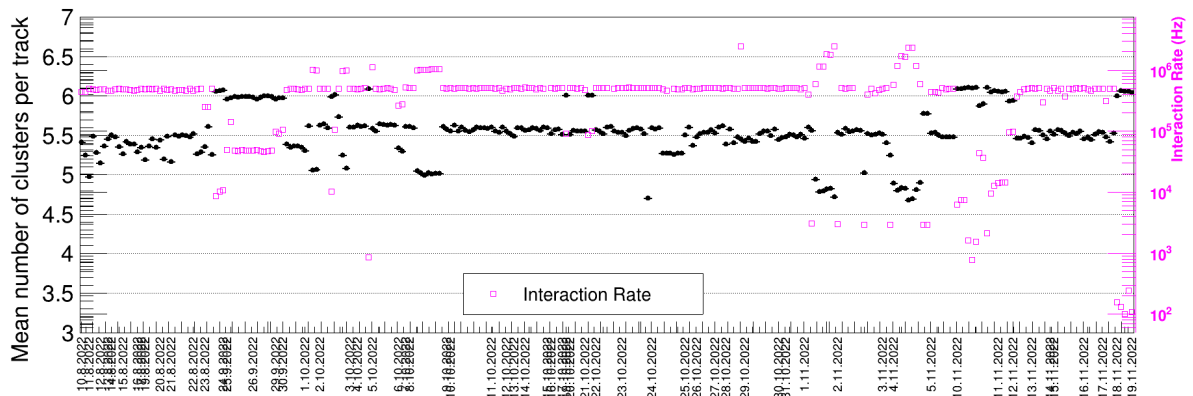
the stand-alone ITS vertexing is working stably.



This work

Figure 5.13: Time trend of the position of the primary seed vertex which is used in further steps of global tracking for pp collisions at  $\sqrt{s} = 13.6$  TeV. The figure is analogous to Fig. 5.9.

Another observable that characterizes ITS tracking is the average number of clusters used per track; see Fig. 5.14. The maximum number of clusters used in the tracking is defined by the number of layers equal to 7. In reality, the number of clusters per track can be smaller due to chips/modules/staves switched off during the run or tracks with vertices outside of the IB acceptance. Tracks with less than 7 clusters can also come from particles that decayed in the ITS and produced daughters, observed as tracks with a smaller number of clusters. Another possible situation when the number of clusters is smaller than 7 is the incorrect assignment of clusters to tracks, called a “fake” track. The frequency of such tracks increases with IR, so we see a smaller average number of clusters used in tracking for runs with higher IR. The QC trending plots for ITS tracking show stable tracking performance for the default beam and data-taking conditions.



This work

Figure 5.14: Time trend of an average number of clusters per track for physical pp collisions at  $\sqrt{s} = 13.6$  TeV. The figure is analogous to Fig. 5.9.

# Chapter 6

## Summary

A key question in the experimental study of the QGP is the limit of its formation in terms of the size of the initial colliding nuclei. Ultra-relativistic collisions of nuclei with a small number of nucleons exhibit some signatures, which are usually understood as a consequence of the QGP initial state. However, there is a lack of evidence that the production of high- $p_T$  hadrons and jets would be modified by jet quenching in these systems. Jet quenching is a necessary consequence of QGP formation, though its effects are expected to be small in small collision systems. A common signature of jet quenching in heavy-ion collisions is the suppression of inclusive hadron and jet yield measured in the nucleus–nucleus collisions compared to that expected by scaling the corresponding yield measured in pp collisions. This modification is usually quantified in terms of the nuclear modification factor. The value of this factor can nevertheless depend on the size of cold nuclear matter effects as well as on the mass of the initial parton. This motivates us to measure the nuclear modification factor for heavy-flavor probes such as b-jets, where the possible mass effects should be augmented.

The thesis is devoted to the measurements of charged-particle b-jet production in minimum bias pp and p–Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV published in [18]. The b jets were selected using the Secondary Vertex method. The spectrum of b-jet candidates was corrected to account for the purity and efficiency of the tagging algorithm; additionally, the smearing of transverse jet momentum due to finite detector resolution and fluctuations in the underlying event was considered. The fully corrected spectra and the corresponding nuclear modification factor were found to be compatible with the results of an analysis by Hadi Hassan, which employed an independent tagging approach.

After combining the results of both methods, it has been found that the obtained charged-particle b-jet cross section is consistent with the NLO pQCD calculations by the POWHEG dijet tune with the PYTHIA 8 fragmentation within experimental and theoretical uncertainties. In particular, the performed study shows that b jets constitute

only 2–4% of the inclusive non-tagged jet yield in the reported  $p_T$  range. The obtained fraction is compatible with the measurements by ATLAS [97] and CMS [11] and with the calculation by the POWHEG + PYTHIA model.

Furthermore, the nuclear modification factor of charged-particle b-jets is consistent with unity within the current experimental precision, suggesting that the b-jet production in minimum bias p–Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV is not affected by CNM effects or jet-quenching. The measured  $R_{pPb}$  was also found to be compatible with the NLO pQCD calculations by the POWHEG dijet tune with PYTHIA 8 fragmentation within the experimental and theoretical uncertainties. The simulation predicted a mild modification due to the anti-shadowing effect. Further comparison with the  $R_{pPb}$  of untagged charged-particle jets didn't reveal any modifications by mass-dependent effects, despite the expectations from the dead-cone effect. Thus, if these mass-effects are present, they are below the precision of the current measurements. The analysis presented in this thesis extends the reach of the previous measurements of b-jets in p–Pb by CMS [11] to lower transverse momenta, but in contrast to the CMS  $R_{pPb}$  it uses a pp reference from measured data. This significantly reduces the overall uncertainty of the measurement. The measurements of nuclear modification factor in minimum bias collisions of small systems benefit from the known analytic form of the scaling coefficient of the pp reference. Although a centrality bias could enhance medium-induced effects, the systematic uncertainty due to Glauber modeling could deteriorate precision of the final nuclear modification factor. The presented analysis also serves as a baseline for charged b-jet measurements in Pb–Pb and provides motivation for future higher precision measurements.

The b-jet tagging is based on selecting secondary vertices or tracks with a large displacement from a primary vertex. In Run 3 and Run 4, these methods will benefit from up to 5 times better pointing resolution of the upgraded ITS2 and up to 100 times larger amount of collected events w.r.t. Run 1 and Run 2 [14]. This will significantly help to determine the b-jet tagging purity and to reduce the corresponding systematic uncertainty. After four years of construction and commissioning, this new ITS detector is ready to perform data-taking. The first results from the ITS Quality Control software during the initial stage of Run 3 demonstrate that the performance of the new ITS is stable and the newly recorded data has good quality which opens opportunities for future heavy-flavor measurements.

# Appendix A

## List of publications

This section overviews my publications which are related to the topic of this thesis.

### Papers:

1. S. Acharya *et al.*, “Measurement of inclusive charged-particle b-jet production in pp and p–Pb collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV”, JHEP **01** (2022), 178.
2. S. Kushpil, F. Krizek and A. Isakov, “Recent Results From Beam Tests of the ALPIDE Pixel Chip for the Upgrade of the ALICE Inner Tracker”, IEEE Trans. Nucl. Sci. **66** (2019) no.11, 2319–2323.
3. F. Krizek, A. Isakov, *et al.*, “Irradiation setup at the U-120M cyclotron facility”, Nucl. Instr. Meth. Phys. Res. A, A 894 (2018) 87–95.

### Proceedings:

1. A. Isakov, “Analysis of b jets in p–Pb collisions at  $\sqrt{s_{\text{NN}}} = 5$  TeV with ALICE,” PoS **EPS-HEP2019** (2020), 319.
2. A. Isakov, “Heavy flavor jets in ALICE,” Int. J. Mod. Phys. E **29** (2020) no.11, 2040003.
3. A. Isakov, “ALICE ITS Upgrade“. In 20th Conference of Czech and Slovak Physicists Proceeding. Slovak Physical Society, Czech Physical Society, 2020. s. 57–58. ISBN: 978-80-89855-13-1.
4. A. Isakov, F. Krizek and T. Matlocha, “Irradiation setup at the U-120M cyclotron facility,” PoS **ICHEP2020** (2021), 986.
5. A. Isakov, “Analysis of b-jets production in p–Pb and pp collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV with ALICE,” PoS **LHCP2021** (2021), 187 [162].

6. A. Isakov, “ALICE measurements of inclusive untagged and heavy flavor-tagged jets in pp, p–Pb and Pb–Pb collisions,” PoS **PANIC2021** (2022), 228.

# Appendix B

## List of conferences

This section overviews my presentations and posters, which are related to the topic of this thesis.

1. *8th International Conference on New Frontiers in Physics (ICNFP 2019)*, Crete, Greece, 2019: “Heavy Flavor jets in ALICE” - **talk**.
2. *The European Physical Society Conference on High Energy Physics (EPS-HEP 2019)*, Ghent, Belgium, 2019: “Heavy Flavor jets in ALICE” - **talk**.
3. *20th Lomonosov Conference on Elementary Particle Physics*, Moscow, Russia, 2020: “Analysis of b jets production in p–Pb and pp at  $\sqrt{s_{NN}} = 5$  TeV with ALICE” - **talk**.
4. *LHCC poster session*, Geneva, Switzerland, 2021: “Analysis of b jets in p–Pb collisions at  $\sqrt{s_{NN}} = 5$  TeV with ALICE” - **poster**.
5. *40th International Conference on High Energy Physics (ICHEP 2020)*, Prague, Czech Republic, 2020: “Irradiation setup at the U-120M cyclotron facility” - **poster**.
6. *20th Conference of Czech and Slovak Physicists*, Prague, Czech Republic, 2020: “ALICE ITS Upgrade” - **talk**.
7. *9th LHC Physics Conference (LHCP 2021)*, Paris, France, 2021 (online): “Analysis of b-jets production in p–Pb and pp collisions at  $\sqrt{s_{NN}} = 5$  TeV with ALICE” - **talk**.
8. *22nd Particles and Nuclei International Conference (PANIC 2021)*, Portugal, Lisbon, 2021 (online): “Inclusive and heavy-flavour jets in ALICE” - **talk**.
9. *21st Zimányi School Winter Workshop on heavy ion physics*, Budapest, Hungary, 2021: “Status and prospect of the ALICE Inner Tracking System” - **talk**.



10. *XXIXth International Conference Quark Matter (QM 2022)*, Krakow, Poland, 2022: “ Analysis of b-jet production in p-Pb and pp collisions at  $\sqrt{s_{NN}} = 5$  TeV with ALICE” - **poster**.
11. *LHCC poster session*, Geneva, Switzerland, 2022: “ALICE Inner Tracking System performance in the initial stage of Run 3” - **poster**.

# Bibliography

- [1] D. de Florian and R. Sassot, “Nuclear parton distributions at next-to-leading order”, *Phys. Rev. D*, vol. 69, p. 074 028, 2004. arXiv: hep-ph/0311227.
- [2] R. Pasechnik and M. Šumbera, “Phenomenological review on quark–gluon plasma: Concepts vs. observations”, *Universe*, vol. 3, no. 1, 2017, ISSN: 2218-1997.
- [3] M. Gyulassy and M. Plümer, “Jet quenching in dense matter”, *Phys. Lett. B*, vol. 243, 1990.
- [4] B. Abelev, et al., “Constraints on jet quenching in p–Pb collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV measured by the event-activity dependence of semi-inclusive hadron-jet distributions”, *Phys. Lett. B*, vol. 783, pp. 95–113, 2018, ISSN: 0370-2693.
- [5] V. Khachatryan *et al.*, “Charged-particle nuclear modification factors in Pb–Pb and p–Pb collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV”, *JHEP*, vol. 04, p. 039, 2017. arXiv: 1611.01664 [nucl-ex].
- [6] J. Adam *et al.*, “Measurement of charged jet production cross sections and nuclear modification in p–Pb collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV”, *Phys. Lett. B*, vol. 749, pp. 68–81, 2015. arXiv: 1503.00681 [nucl-ex].
- [7] S. Tang, “Measurement of the jet-particle  $v_2$  in p–Pb and Pb–Pb collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV with ALICE at the LHC”, *PoS*, vol. EPS-HEP2021, p. 313, 2022. arXiv: 2110.15852 [nucl-ex].
- [8] J. Adam *et al.*, “Enhanced production of multi-strange hadrons in high-multiplicity proton-proton collisions”, *Nature Phys.*, vol. 13, pp. 535–539, 2017.
- [9] N. Armesto, C. A. Salgado, and U. A. Wiedemann, “Medium induced gluon radiation off massive quarks fills the dead cone”, *Phys. Rev. D*, vol. 69, p. 114 003, 2004. arXiv: hep-ph/0312106.
- [10] Y. Dokshitzer, V. Khoze, and S. Troyan, “On specific QCD properties of heavy quark fragmentation (‘dead cone’)”, *J. Phys. G*, vol. 17, p. 1602, Jan. 1999.
- [11] V. Khachatryan *et al.*, “Transverse momentum spectra of inclusive b jets in pPb collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV”, *Phys. Lett. B*, vol. 754, p. 59, 2016.
- [12] K. Aamodt *et al.*, “The ALICE experiment at the CERN LHC”, *JINST*, vol. 3, S08002, 2008.
- [13] B. Abelev *et al.*, “Performance of the ALICE Experiment at the CERN LHC”, *Int. J. Mod. Phys. A*, vol. 29, p. 1 430 044, 2014. arXiv: 1402.4476 [nucl-ex].
- [14] B. Abelev *et al.*, “Upgrade of the ALICE Experiment: Letter Of Intent”, *J. Phys. G*, vol. 41, p. 087 001, 2014.

- [15] B.-M. Blidaru, “Testbeam performance results of bent ALPIDE Monolithic Active Pixel Sensors in view of the ALICE Inner Tracking System 3”, in *12th International Conference on Position Sensitive Detectors*, Dec. 2021. arXiv: 2112.10414 [physics.ins-det].
- [16] G. Aglieri Rinella, “The ALPIDE pixel sensor chip for the upgrade of the ALICE Inner Tracking System”, *Nucl. Instr. Meth. Phys. Res. A*, vol. 845, G. Badurek, T. Bergauer, M. Dragicevic, *et al.*, Eds., pp. 583–587, 2017.
- [17] M. Richter, “Aliceo2group/aliceo2: First stable release”, 2018. DOI: <https://doi.org/10.5281/zenodo.1493334>.
- [18] S. Acharya *et al.*, “Measurement of inclusive charged-particle b-jet production in pp and p–Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”, *JHEP*, vol. 01, p. 178, 2022. arXiv: 2110.06104 [nucl-ex].
- [19] S. Kushpil, F. Krizek, and A. Isakov, “Recent Results From Beam Tests of the ALPIDE Pixel Chip for the Upgrade of the ALICE Inner Tracker”, *IEEE Trans. Nucl. Sci.*, vol. 66, no. 11, pp. 2319–2323, 2019.
- [20] F. Křížek, A. Isakov, J. Ferencei, *et al.*, “Irradiation setup at the U-120M cyclotron facility”, *Nucl. Instr. Meth. Phys. Res. A*, vol. 894, pp. 87–95, 2018, ISSN: 0168-9002.
- [21] G. A. Rinella *et al.*, “First demonstration of in-beam performance of bent Monolithic Active Pixel Sensors”, *Nucl. Instr. Meth. Phys. Res. A*, vol. 1028, p. 166 280, 2022. arXiv: 2105.13000 [physics.ins-det].
- [22] M. Tanabashi *et al.*, “Review of Particle Physics”, *Phys. Rev. D*, vol. 98, no. 3, p. 030 001, 2018.
- [23] R. Pasechnik and M. Šumbera, “Different faces of confinement”, *Universe*, vol. 7, no. 9, 2021, ISSN: 2218-1997.
- [24] Z. Fodor and C. Hoelbling, “Light Hadron Masses from Lattice QCD”, *Rev. Mod. Phys.*, vol. 84, p. 449, 2012. arXiv: 1203.4789 [hep-lat].
- [25] R. Gupta and A. Patel, “Calculation of the hadron spectrum in lattice QCD”, *Nucl. Phys. B*, vol. 226, no. 1, pp. 152–186, 1983, ISSN: 0550-3213.
- [26] S. Bethke, “The 2009 World Average of  $\alpha_s$ ”, *Eur. Phys. J. C*, vol. 64, D. H. Beck, D. Haidt, and J. W. Negele, Eds., pp. 689–703, 2009. arXiv: 0908.1135 [hep-ph].
- [27] E. Eichten, K. Gottfried, T. Kinoshita, J. Kogut, K. D. Lane, and T.-M. Yan, “Spectrum of charmed quark-antiquark bound states”, *Phys. Rev. Lett.*, vol. 34, no. 6, p. 369, 1975.
- [28] B. Andersson, *The Lund Model* (Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology). Cambridge University Press, 1998.
- [29] T. Sjöstrand, S. Ask, J. R. Christiansen, *et al.*, “An introduction to PYTHIA 8.2”, *Comput. Phys. Commun.*, vol. 191, pp. 159–177, 2015. arXiv: 1410.3012 [hep-ph].
- [30] G. P. Salam, “Towards Jetography”, *Eur. Phys. J. C*, vol. 67, pp. 637–686, 2010. arXiv: 0906.1833 [hep-ph].
- [31] “Event display of pp collisions”, eprint: <https://cms.cern/news/machining-jets>.

- [32] S. Acharya *et al.*, “Direct observation of the dead-cone effect in quantum chromodynamics”, *Nature*, vol. 605, no. 7910, pp. 440–446, 2022, [Erratum: *Nature* 607, E22 (2022)]. arXiv: 2106.05713 [nucl-ex].
- [33] S. Acharya *et al.*, “Exploration of jet substructure using iterative declustering in pp and Pb–Pb collisions at LHC energies”, *Phys. Lett. B*, vol. 802, p. 135 227, 2020. arXiv: 1905.02512 [nucl-ex].
- [34] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet User Manual”, *Eur. Phys. J. C*, vol. 72, p. 1896, 2012. arXiv: 1111.6097 [hep-ph].
- [35] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas”, *Phys. Lett. B*, vol. 659, pp. 119–126, 2008. arXiv: 0707.1378 [hep-ph].
- [36] J. C. Collins, D. E. Soper, and G. F. Sterman, “Factorization of Hard Processes in QCD”, *Adv. Ser. Direct. High Energy Phys.*, vol. 5, pp. 1–91, 1989. arXiv: hep-ph/0409313.
- [37] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, “Parton distributions for the LHC”, *Eur. Phys. J. C*, vol. 63, pp. 189–285, 2009. arXiv: 0901.0002 [hep-ph].
- [38] K. J. Eskola, P. Paakkinen, H. Paukkunen, and C. A. Salgado, “Towards EPPS21 nuclear PDFs”, in *28th International Workshop on Deep Inelastic Scattering and Related Subjects*, Jun. 2021. arXiv: 2106.13661 [hep-ph].
- [39] M. Hirai, S. Kumano, and T. H. Nagai, “Determination of nuclear parton distribution functions and their uncertainties in next-to-leading order”, *Phys. Rev. C*, vol. 76, p. 065 207, 2007. arXiv: 0709.3038 [hep-ph].
- [40] T.-J. Hou *et al.*, “New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC”, *Phys. Rev. D*, vol. 103, no. 1, p. 014 013, 2021. arXiv: 1912.10053 [hep-ph].
- [41] H. Paukkunen, “Status of nuclear PDFs after the first LHC p–Pb run”, *Nucl. Phys. A*, vol. 967, U. Heinz, O. Evdokimov, and P. Jacobs, Eds., pp. 241–248, 2017. arXiv: 1704.04036 [hep-ph].
- [42] M. Gockeler, R. Horsley, A. C. Irving, *et al.*, “A Determination of the Lambda parameter from full lattice QCD”, *Phys. Rev. D*, vol. 73, p. 014 513, 2006. arXiv: hep-ph/0502212.
- [43] R. D. Field and R. P. Feynman, “A Parametrization of the Properties of Quark Jets”, *Nucl. Phys. B*, vol. 136, L. M. Brown, Ed., p. 1, 1978.
- [44] R. Brüser, Z. L. Liu, and M. Stahlhofen, “Three-Loop Quark Jet Function”, *Phys. Rev. Lett.*, vol. 121, no. 7, p. 072 003, 2018. arXiv: 1804.09722 [hep-ph].
- [45] T. Becher and G. Bell, “The gluon jet function at two-loop order”, *Phys. Lett. B*, vol. 695, pp. 252–258, 2011. arXiv: 1008.1936 [hep-ph].
- [46] E. V. Shuryak, “Quark-Gluon Plasma and Hadronic Production of Leptons, Photons and Psions”, *Phys. Lett. B*, vol. 78, p. 150, 1978.
- [47] C. Schmidt and S. Sharma, “The phase structure of QCD”, *J. Phys. G*, vol. 44, no. 10, p. 104 002, 2017. arXiv: 1701.04707 [hep-lat].
- [48] A. Majumder and B. Muller, “Hadron Mass Spectrum from Lattice QCD”, *Phys. Rev. Lett.*, vol. 105, p. 252 002, 2010. arXiv: 1008.1747 [hep-ph].

- [49] T. Bhattacharya *et al.*, “QCD Phase Transition with Chiral Quarks and Physical Quark Masses”, *Phys. Rev. Lett.*, vol. 113, no. 8, p. 082 001, 2014. arXiv: 1402.5175 [hep-lat].
- [50] DOE/NSF Nuclear Science Advisory Committee, “The Frontiers of Nuclear Science, A Long Range Plan”, Sep. 2008. arXiv: 0809.3137 [nucl-ex].
- [51] E. W. Kolb and S. Wolfram, “Baryon Number Generation in the Early Universe”, *Nucl. Phys. B*, vol. 172, p. 224, 1980, [Erratum: Nucl.Phys.B 195, 542 (1982)].
- [52] U. W. Heinz and M. Jacob, “Evidence for a new state of matter: An Assessment of the results from the CERN lead beam program”, Jan. 2000. arXiv: nucl-th/0002042.
- [53] S. Cao *et al.*, “Determining the jet transport coefficient  $\hat{q}$  from inclusive hadron suppression measurements using Bayesian parameter estimation”, *Phys. Rev. C*, vol. 104, no. 2, p. 024 905, 2021. arXiv: 2102.11337 [nucl-th].
- [54] K. M. Burke *et al.*, “Extracting the jet transport coefficient from jet quenching in high-energy heavy-ion collisions”, *Phys. Rev. C*, vol. 90, no. 1, p. 014 909, 2014. arXiv: 1312.5003 [nucl-th].
- [55] R. Derradi de Souza, T. Koide, and T. Kodama, “Hydrodynamic Approaches in Relativistic Heavy Ion Reactions”, *Prog. Part. Nucl. Phys.*, vol. 86, pp. 35–85, 2016. arXiv: 1506.03863 [nucl-th].
- [56] J. Adams *et al.*, “Experimental and theoretical challenges in the search for the quark gluon plasma: The STAR Collaboration’s critical assessment of the evidence from RHIC collisions”, *Nucl. Phys. A*, vol. 757, pp. 102–183, 2005. arXiv: nucl-ex/0501009.
- [57] P. Kovtun, D. T. Son, and A. O. Starinets, “Viscosity in strongly interacting quantum field theories from black hole physics”, *Phys. Rev. Lett.*, vol. 94, p. 111 601, 2005. arXiv: hep-th/0405231.
- [58] F. Becattini and M. A. Lisa, “Polarization and Vorticity in the Quark Gluon Plasma”, *Ann. Rev. Nucl. Part. Sci.*, vol. 70, pp. 395–423, 2020. arXiv: 2003.03640 [nucl-ex].
- [59] G. Aad *et al.*, “Transverse momentum and process dependent azimuthal anisotropies in  $\sqrt{s_{NN}} = 8.16$  TeV p+Pb collisions with the ATLAS detector”, *Eur. Phys. J. C*, vol. 80, no. 1, p. 73, 2020. arXiv: 1910.13978 [nucl-ex].
- [60] B. Abelev *et al.*, “Transverse momentum distribution and nuclear modification factor of charged particles in p–Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”, *Phys. Rev. Lett.*, vol. 110, no. 8, p. 082 302, 2013. arXiv: 1210.4520 [nucl-ex].
- [61] E. Pereira de Oliveira Filho, “Measurements of electrons from heavy-flavour hadron decays in pp, p–Pb and Pb–Pb collisions with ALICE at the LHC”, *Nucl. Phys. A*, vol. 932, W. Horowitz, H. Weigert, T. Dietel, *et al.*, Eds., pp. 258–263, 2014. arXiv: 1404.3983 [hep-ex].
- [62] B. B. Abelev *et al.*, “Measurement of prompt D-meson production in p–Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”, *Phys. Rev. Lett.*, vol. 113, no. 23, p. 232 301, 2014. arXiv: 1405.3452 [nucl-ex].

- [63] P. Foka and M. A. Janik, “An overview of experimental results from ultra-relativistic heavy-ion collisions at the CERN LHC: Hard probes”, *Rev. Phys.*, vol. 1, pp. 172–194, 2016. arXiv: 1702.07231 [hep-ex].
- [64] G.-Y. Qin and X.-N. Wang, “Jet quenching in high-energy heavy-ion collisions”, *Int. J. Mod. Phys. E*, vol. 24, no. 11, X.-N. Wang, Ed., p. 1530014, 2015. arXiv: 1511.00790 [hep-ph].
- [65] R. Baier, Y. L. Dokshitzer, A. H. Mueller, S. Peigne, and D. Schiff, “Radiative energy loss of high-energy quarks and gluons in a finite volume quark - gluon plasma”, *Nucl. Phys. B*, vol. 483, pp. 291–320, 1997. arXiv: hep-ph/9607355.
- [66] K. Tywoniuk, “Is there jet quenching in pPb?”, *Nucl. Phys. A*, vol. 926, N. Armesto, L. McLerran, C. Pajares, and C. Salgado, Eds., pp. 85–91, 2014.
- [67] W.-T. Deng and X.-N. Wang, “Multiple parton scattering in nuclei: Modified Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution for fragmentation functions”, *Phys. Rev. C*, vol. 81, p. 024902, 2010.
- [68] J. Adam *et al.*, “Measurement of jet quenching with semi-inclusive hadron-jet distributions in central Pb–Pb collisions at  $\sqrt{s_{\text{NN}}} = 2.76$  TeV”, *JHEP*, vol. 09, p. 170, 2015. arXiv: 1506.03984 [nucl-ex].
- [69] Y. L. Dokshitzer and D. E. Kharzeev, “Heavy quark colorimetry of QCD matter”, *Phys. Lett. B*, vol. 519, pp. 199–206, 2001. arXiv: hep-ph/0106202.
- [70] A. Andronic, “An overview of the experimental study of quark-gluon matter in high-energy nucleus-nucleus collisions”, *Int. J. Mod. Phys. A*, vol. 29, p. 1430047, 2014. arXiv: 1407.5003 [nucl-ex].
- [71] S. Chatrchyan *et al.*, “Evidence of b-jet Quenching in Pb–Pb Collisions at  $\sqrt{s_{\text{NN}}} = 2.76$  TeV”, *Phys. Rev. Lett.*, vol. 113, no. 13, p. 132301, 2014. arXiv: 1312.4198 [nucl-ex].
- [72] M. L. Miller, K. Reygers, S. J. Sanders, and P. Steinberg, “Glauber modeling in high energy nuclear collisions”, *Ann. Rev. Nucl. Part. Sci.*, vol. 57, pp. 205–243, 2007. arXiv: nucl-ex/0701025.
- [73] A. Huss, A. Kurkela, A. Mazeliauskas, R. Paatelainen, W. van der Schee, and U. Wiedemann, “Discovering partonic rescattering in light nucleus collisions”, *Phys. Rev. Lett.*, vol. 126, May 2021.
- [74] J. Adam *et al.*, “Centrality dependence of particle production in p–Pb collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV”, *Phys. Rev. C*, vol. 91, no. 6, p. 064905, 2015. arXiv: 1412.6828 [nucl-ex].
- [75] A. Adare *et al.*, “Azimuthally anisotropic emission of low-momentum direct photons in Au+Au collisions at  $\sqrt{s_{\text{NN}}} = 200$  GeV”, *Phys. Rev. C*, vol. 94, no. 6, p. 064901, 2016. arXiv: 1509.07758 [nucl-ex].
- [76] G. Aad *et al.*, “Centrality and rapidity dependence of inclusive jet production in  $\sqrt{s_{\text{NN}}} = 5.02$  TeV proton-lead collisions with the ATLAS detector”, *Phys. Lett. B*, vol. 748, pp. 392–413, 2015. arXiv: 1412.4092 [hep-ex].
- [77] M. Kordell and A. Majumder, “Jets in d(p)-A Collisions: Color Transparency or Energy Conservation”, *Phys. Rev. C*, vol. 97, no. 5, p. 054904, 2018. arXiv: 1601.02595 [nucl-th].

- [78] A. Huss, A. Kurkela, A. Mazeliauskas, R. Paatelainen, W. van der Schee, and U. A. Wiedemann, “Predicting parton energy loss in small collision systems”, *Phys. Rev. C*, vol. 103, no. 5, p. 054903, 2021. arXiv: 2007.13758 [hep-ph].
- [79] A. Accardi, “Cronin effect in proton nucleus collisions: A Survey of theoretical models”, Dec. 2002. arXiv: hep-ph/0212148.
- [80] B. Abelev *et al.*, “Transverse momentum dependence of inclusive primary charged-particle production in p–Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”, *Eur. Phys. J. C*, vol. 74, no. 9, p. 3054, 2014. arXiv: 1405.2737 [nucl-ex].
- [81] J. Adams *et al.*, “Evidence from d + Au measurements for final state suppression of high  $p_T$  hadrons in Au+Au collisions at RHIC”, *Phys. Rev. Lett.*, vol. 91, p. 072304, 2003. arXiv: nucl-ex/0306024.
- [82] J. Adams *et al.*, “Identified hadron spectra at large transverse momentum in p+p and d+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV”, *Phys. Lett. B*, vol. 637, pp. 161–169, 2006. arXiv: nucl-ex/0601033.
- [83] G. Aad *et al.*, “Transverse Momentum Dependence of Charged Particle Production in p+Pb  $\sqrt{s_{NN}} = 5.02$  TeV collisions measured by ATLAS experiment at the LHC”, *ATLAS-CONF-2013-107*, Nov. 2013.
- [84] G. Aad *et al.*, “Measurement of the dependence of transverse energy production at large pseudorapidity on the hard-scattering kinematics of proton-proton collisions at  $\sqrt{s} = 2.76$  TeV with ATLAS”, *Phys. Lett. B*, vol. 756, pp. 10–28, 2016. arXiv: 1512.00197 [hep-ex].
- [85] J. A. Hanks, “PHENIX results on jets in d+Au”, *Nucl. Phys. A*, vol. 956, pp. 625–628, 2016, The XXV International Conference on Ultrarelativistic Nucleus-Nucleus Collisions: Quark Matter 2015, ISSN: 0375-9474.
- [86] J. Adam *et al.*, “Inclusive, prompt and non-prompt  $J/\psi$  production at mid-rapidity in Pb–Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV”, *JHEP*, vol. 07, p. 051, 2015. arXiv: 1504.07151 [nucl-ex].
- [87] S. Chatrchyan *et al.*, “ $J/\psi$  results from CMS in Pb–Pb collisions, with  $150 \mu\text{b}^{-1}$  data”, *CMS-PAS-HIN-12-014*, 2012.
- [88] A. Festanti, “Heavy-flavour production and nuclear modification factor in Pb–Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV with ALICE”, *Nucl. Phys. A*, vol. 931, P. Braun-Munzinger, B. Friman, and J. Stachel, Eds., pp. 514–519, 2014. arXiv: 1407.6541 [nucl-ex].
- [89] D. Kawana, “Measurement of Electrons from Charm and Beauty Hadron Decays in p–Pb Collisions at  $\sqrt{s_{NN}} = 8.16$  TeV with ALICE”, *JPS Conf. Proc.*, vol. 26, A. Dote, Y. Goto, A. Hosaka, S. Kumano, and A. Monnai, Eds., p. 024014, 2019. arXiv: 1901.03496 [hep-ex].
- [90] A. M. Sirunyan *et al.*, “Observation of Correlated Azimuthal Anisotropy Fourier Harmonics in pp and p–Pb Collisions at the LHC”, *Phys. Rev. Lett.*, vol. 120, no. 9, p. 092301, 2018. arXiv: 1709.09189 [nucl-ex].
- [91] S. Acharya *et al.*, “Charm-quark fragmentation fractions and production cross section at midrapidity in pp collisions at the LHC”, *Phys. Rev. D*, vol. 105, no. 1, p. L011103, 2022. arXiv: 2105.06335 [nucl-ex].

- [92] A. Morsch, “Multiple Parton Interactions with ALICE: from pp to p–Pb”, *J. Phys. Conf. Ser.*, vol. 535, p. 012 012, 2014. arXiv: 1407.3628 [nucl-ex].
- [93] T. Lang and M. Bleicher, “Possibility for  $J/\psi$  suppression in high-multiplicity proton-proton collisions at  $\sqrt{s} = 7$  TeV”, *Phys. Rev. C*, vol. 87, no. 2, p. 024 907, 2013. arXiv: 1302.0655 [hep-ph].
- [94] S. Acharya *et al.*, “Measurement of charged jet cross section in  $pp$  collisions at  $\sqrt{s} = 5.02$  TeV”, *Phys. Rev. D*, vol. 100, no. 9, p. 092 004, 2019. arXiv: 1905.02536 [nucl-ex].
- [95] S. Acharya *et al.*, “Measurement of the production of charm jets tagged with  $D^0$  mesons in pp collisions at  $\sqrt{s} = 7$  TeV”, *JHEP*, vol. 08, p. 133, 2019. arXiv: 1905.02510 [nucl-ex].
- [96] S. Acharya *et al.*, “Measurements of groomed-jet substructure of charm jets tagged by  $D^0$  mesons in proton-proton collisions at  $\sqrt{s} = 13$  TeV”, Aug. 2022. arXiv: 2208.04857 [nucl-ex].
- [97] M. Aaboud *et al.*, “Measurement of the nuclear modification factor of  $b$ -jets in 5.02 TeV Pb+Pb collisions with the ATLAS detector”, Apr. 2022. arXiv: 2204.13530 [nucl-ex].
- [98] A. Banfi, G. P. Salam, and G. Zanderighi, “Accurate QCD predictions for heavy-quark jets at the Tevatron and LHC”, *JHEP*, vol. 07, p. 026, 2007. arXiv: 0704.2999 [hep-ph].
- [99] A. Toia, “Bulk Properties of Pb–Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV measured by ALICE”, *J. Phys. G*, vol. 38, Y. Schutz and U. A. Wiedemann, Eds., p. 124 007, 2011. arXiv: 1107.1973 [nucl-ex].
- [100] S. Acharya *et al.*, “3D ALICE Schematic RUN2 ”, [Online]. Available: <https://alice-figure.web.cern.ch/node/11218>.
- [101] G. Aad *et al.*, “The ATLAS Experiment at the CERN Large Hadron Collider”, *JINST*, vol. 3, S08003, 2008.
- [102] S. Chatrchyan *et al.*, “The CMS Experiment at the CERN LHC”, *JINST*, vol. 3, S08004, 2008.
- [103] F. Reidt, “Studies for the ALICE Inner Tracking System Upgrade”, Ph.D. dissertation, Heidelberg U., Feb. 2016.
- [104] S. Acharya *et al.*, “Transverse impact parameter resolution vs  $p_T$ ”, eprint: <https://alice-figure.web.cern.ch/node/205>.
- [105] S. Acharya *et al.*, “Longitudinal impact parameter resolution vs  $p_T$ ”, eprint: <https://alice-figure.web.cern.ch/node/3644>.
- [106] J. Alme, Y. Andres, H. Appelshäuser, *et al.*, “The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events”, *Nucl. Instr. Meth. Phys. Res. A*, vol. 622, no. 1, pp. 316–367, 2010, ISSN: 0168-9002.
- [107] E. Abbas *et al.*, “Performance of the ALICE VZERO system”, *JINST*, vol. 8, P10016, 2013. arXiv: 1306.3130 [nucl-ex].
- [108] C. W. Fabjan, L. Jirdén, V Lindestruth, *et al.*, *ALICE trigger data-acquisition high-level trigger and control system: Technical Design Report*. Geneva: CERN, 2004. [Online]. Available: <https://cds.cern.ch/record/684651>.



- [109] *ALICE offline bible*. [Online]. Available: <https://alice-offline.web.cern.ch/sites/alice-offline.web.cern.ch/files/uploads/OfflineBible.pdf>.
- [110] S. Chatrchyan *et al.*, “Identification of b-quark Jets with the CMS Experiment”, *JINST*, vol. 8, P04013, 2013. arXiv: 1211.4462 [hep-ex].
- [111] L. Lyons, D. Gibaut, and P. Clifford, “How to Combine Correlated Estimates of a Single Physical Quantity”, *Nucl. Instr. Meth. Phys. Res. A*, vol. 270, p. 110, 1988.
- [112] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zancarini, “GEANT3”, Sep. 1987.
- [113] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, *JHEP*, vol. 05, p. 026, 2006. arXiv: hep-ph/0603175.
- [114] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner, “EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider”, *Phys. Rev. C*, vol. 92, no. 3, p. 034906, 2015. arXiv: 1306.0121 [hep-ph].
- [115] P. Skands, S. Carrazza, and J. Rojo, “Tuning PYTHIA 8.1: the Monash 2013 Tune”, *Eur. Phys. J. C*, vol. 74, no. 8, p. 3024, 2014. arXiv: 1404.5630 [hep-ph].
- [116] S. Chatrchyan *et al.*, “Shape, Transverse Size, and Charged Hadron Multiplicity of Jets in pp Collisions at  $\sqrt{s} = 7$  TeV”, *JHEP*, vol. 06, p. 160, 2012. arXiv: 1204.3170 [hep-ex].
- [117] S. Chatrchyan and other, “Measurement of the Underlying Event Activity in pp Collisions at  $\sqrt{s} = 0.9$  and 7 TeV with the Novel Jet-Area/Median Approach”, *JHEP*, vol. 08, p. 130, 2012. arXiv: 1207.2392 [hep-ex].
- [118] A. M. Sirunyan *et al.*, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”, *JINST*, vol. 13, no. 05, P05011, 2018. arXiv: 1712.07158 [physics.ins-det].
- [119] S. Frixione, P. Nason, and G. Ridolfi, “A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction”, *JHEP*, vol. 09, p. 126, 2007. arXiv: 0707.3088 [hep-ph].
- [120] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky, and W. K. Tung, “New generation of parton distributions with uncertainties from global QCD analysis”, *JHEP*, vol. 07, p. 012, 2002. arXiv: hep-ph/0201195.
- [121] K. J. Eskola, H. Paukkunen, and C. A. Salgado, “EPS09: A New Generation of NLO and LO Nuclear Parton Distribution Functions”, *JHEP*, vol. 04, p. 065, 2009. arXiv: 0902.4154 [hep-ph].
- [122] F. James, “MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1”, 1994.
- [123] A. Hocker and V. Kartvelishvili, “SVD approach to data unfolding”, *Nucl. Instr. Meth. Phys. Res. A*, vol. 372, pp. 469–481, 1996. arXiv: hep-ph/9509307.
- [124] T. Adye, “Unfolding algorithms and tests using RooUnfold”, in *PHYSTAT 2011*, Geneva: CERN, 2011, pp. 313–318. arXiv: 1105.1160 [physics.data-an].
- [125] G. D’Agostini, “A Multidimensional unfolding method based on Bayes’ theorem”, *Nucl. Instr. Meth. Phys. Res. A*, vol. 362, pp. 487–498, 1995.
- [126] B. Abelev *et al.*, “Measurement of charged jet suppression in Pb–Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV”, *JHEP*, vol. 03, p. 013, 2014. arXiv: 1311.0633 [nucl-ex].

- [127] S. Acharya *et al.*, “ALICE 2017 luminosity determination for pp collisions at  $\sqrt{s} = 5$  TeV”, *ALICE-PUBLIC-2018-014*, 2018.
- [128] B. Abelev *et al.*, “Measurement of visible cross sections in proton-lead collisions at  $\sqrt{s_{NN}} = 5.02$  TeV in van der Meer scans with the ALICE detector”, *JINST*, vol. 9, no. 11, P11003, 2014. arXiv: 1405.1849 [nucl-ex].
- [129] S. Acharya *et al.*, “ALICE luminosity determination for pp collisions at  $\sqrt{s} = 5$  TeV”, *ALICE-PUBLIC-2016-005*, 2016.
- [130] S. Acharya *et al.*, “Measurements of inclusive jet spectra in pp and central Pb–Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV”, *Phys. Rev. C*, vol. 101, no. 3, p. 034911, 2020. arXiv: 1909.09718 [nucl-ex].
- [131] I. Belikov, “Event reconstruction and particle identification in the ALICE experiment at the LHC”, *EPJ Web of Conferences*, vol. 70, p. 00 029, 2014. [Online]. Available: <https://doi.org/10.1051/epjconf/20147000029>.
- [132] S. Acharya *et al.*, “Transverse momentum spectra and nuclear modification factors of charged particles in pp, p–Pb and Pb–Pb collisions at the LHC”, *JHEP*, vol. 11, p. 013, 2018. arXiv: 1802.09145 [nucl-ex].
- [133] S. Acharya *et al.*, “The ALICE definition of primary particles”, *ALICE-PUBLIC-2017-005*, 2017.
- [134] S. Acharya *et al.*, “Production of light-flavor hadrons in pp collisions at  $\sqrt{s} = 7$  and  $\sqrt{s} = 13$  TeV”, *Eur. Phys. J. C*, vol. 81, no. 3, p. 256, 2021. arXiv: 2005.11120 [nucl-ex].
- [135] K. Aamodt *et al.*, “Production of pions, kaons and protons in pp collisions at  $\sqrt{s} = 900$  GeV with ALICE at the LHC”, *Eur. Phys. J. C*, vol. 71, p. 1655, 2011. arXiv: 1101.4110 [hep-ex].
- [136] S. Chatrchyan *et al.*, “Study of the Inclusive Production of Charged Pions, Kaons, and Protons in pp Collisions at  $\sqrt{s} = 0.9, 2.76,$  and  $7$  TeV”, *Eur. Phys. J. C*, vol. 72, p. 2164, 2012. arXiv: 1207.4724 [hep-ex].
- [137] A. Valassi, “Combining correlated measurements of several different physical quantities”, *Nucl. Instr. Meth. Phys. Res. A*, vol. 500, pp. 391–405, 2003.
- [138] K. J. Eskola, P. Paakkinen, H. Paukkunen, and C. A. Salgado, “EPPS16: Nuclear parton distributions with LHC data”, *Eur. Phys. J. C*, vol. 77, no. 3, p. 163, 2017. arXiv: 1612.05741 [hep-ph].
- [139] S. Alioli, K. Hamilton, P. Nason, C. Oleari, and E. Re, “Jet pair production in POWHEG”, *JHEP*, vol. 04, p. 081, 2011. arXiv: 1012.3380 [hep-ph].
- [140] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, *JHEP*, vol. 11, p. 070, 2007. arXiv: 0709.2092 [hep-ph].
- [141] B. Abelev *et al.*, “Upgrade of the ALICE Experiment: Letter Of Intent”, *J. Phys. G*, vol. 41, p. 087001, 2014.
- [142] P. Buncic, M. Krzewicki, and P. Vande Vyvre, “Technical Design Report for the Upgrade of the Online-Offline Computing System”, Apr. 2015.
- [143] B. Abelev *et al.*, “Technical Design Report for the Upgrade of the ALICE Inner Tracking System”, *J. Phys. G*, vol. 41, p. 087002, 2014.

- [144] C. Lippmann, “Upgrade of the ALICE Time Projection Chamber”, *CERN-LHCC-2013-020*, *ALICE-TDR-016*, Mar. 2014.
- [145] F. Reidt, “Upgrade of the ALICE ITS detector”, *Nucl. Instr. Meth. Phys. Res. A*, vol. 1032, p. 166 632, 2022. arXiv: 2111.08301 [physics.ins-det].
- [146] I. Ravasenga, “The upgrade of the ALICE Inner Tracking System”, *PoS*, vol. BOR-MIO2017, L. Fabbietti, C. Sfienti, and W. Kühn, Eds., p. 058, 2017.
- [147] G. Contin, “The STAR PXL detector”, *JINST*, vol. 11, no. 12, C. Gemme and L. Rossi, Eds., p. C12068, 2016.
- [148] M. Mager, “ALPIDE, the monolithic active pixel sensor for the ALICE ITS upgrade”, *Nucl. Instr. Meth. Phys. Res. A*, vol. 824, pp. 434–438, 2016, Frontier Detectors for Frontier Physics: Proceedings of the 13th Pisa Meeting on Advanced Detectors, ISSN: 0168-9002.
- [149] S. Senyukov, J. Baudot, A. Besson, *et al.*, “Charged particle detection performances of CMOS pixel sensors produced in a 0.18  $\mu\text{m}$  process with a high resistivity epitaxial layer”, *Nucl. Instr. Meth. Phys. Res. A*, vol. 730, pp. 115–118, 2013, Proceedings of the 9th International Conference on Radiation Effects on Semiconductor Materials Detectors and Devices, ISSN: 0168-9002.
- [150] H. Hillemanns *et al.*, “Radiation hardness and detector performance of new 180nm CMOS MAPS prototype test structures developed for the upgrade of the ALICE Inner Tracking System”, in *2013 IEEE Nuclear Science Symposium and Medical Imaging Conference and Workshop on Room-Temperature Semiconductor Detectors*, 2013.
- [151] G. Eulisse, P. Konopka, M. Krzewicki, M. Richter, D. Rohr, and S. Wenzel, “Evolution of the ALICE software framework for Run 3”, *EPJ Web of Conferences*, vol. 214, p. 05 010, Jan. 2019.
- [152] <https://restfulapi.net/>, “RESTful API”,
- [153] J. F. Grosse-Oetringhaus, C. Zampolli, A. Colla, and F. Carminati, “The ALICE online-offline framework for the extraction of conditions data”, *J. Phys. Conf. Ser.*, vol. 219, J. Gruntorad and M. Lokajicek, Eds., p. 022 010, 2010.
- [154] *QCAnalysis software for ITS post-processing QA*, <https://github.com/iravasen/QCanalysis>, Accessed: 2022-12-25.
- [155] D. Colella, *ALICE ITS Upgrade for LHC Run 3: Commissioning in the Laboratory*, 2020. [Online]. Available: <https://arxiv.org/abs/2012.01564>.
- [156] V. Raskina and F. Krížek, “Characterization of highly irradiated alpide silicon sensors”, *Universe*, vol. 5, p. 91, Apr. 2019.
- [157] R. Chandra, L. Dagum, D. Kohr, R. Menon, D. Maydan, and J. McDonald, *Parallel programming in OpenMP*. Morgan Kaufmann, 2001.
- [158] R. Brun and F. Rademakers, “ROOT: An object oriented data analysis framework”, *Nucl. Instr. Meth. Phys. Res. A*, vol. 389, M. Werlen and D. Perret-Gallix, Eds., pp. 81–86, 1997.
- [159] *TEfficiency Class Reference*, <https://root.cern.ch/doc/master/classTEfficiency.html>, Accessed: 2023-01-05.

- [160] N. Valle, *ITS – Run 3 data taking*, presented at ITS plenary meeting, Accessed: 2022-12-13.
- [161] M. Ferro-Luzzi, “Beam-Gas Interactions”, in *CAS - CERN Accelerator School: Vacuum for Particle Accelerators*, Jun. 2020. arXiv: 2006.06490 [physics.acc-ph].
- [162] A. Isakov, “Analysis of b-jets production in p–Pb and pp collisions at  $\sqrt{s_{NN}} = 5.02$  TeV with ALICE”, *PoS*, vol. LHCP2021, p. 187, 2021.