JYU DISSERTATIONS 1020

Laura Huhta

Probing QCD via Dijet Invariant Mass in pp and p–Pb Collisions and Performance of Forward Jets with ALICE FoCal

UNIVERSITY OF JYVÄSKYLÄ
FACULTY OF MATHEMATICS
AND SCIENCE

JYU DISSERTATIONS 1020

Laura Huhta

Probing QCD via Dijet Invariant Mass in pp and p–Pb Collisions and Performance of Forward Jets with ALICE FoCal

Esitetään Jyväskylän yliopiston matemaattis-luonnontieteellisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi Ylistönrinteen auditoriossa FYS1
joulukuun 5. päivänä 2025 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Mathematics and Science of the University of Jyväskylä,
in Ylistönrinne, auditorium FYS1, on December 5, 2025, at 12 o'clock.

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2025

Editors
Ilari Maasilta
Department of Physics, University of Jyväskylä
Päivi Vuorio
Open Science Centre, University of Jyväskylä

Copyright © 2025, by the author and University of Jyväskylä

ISBN 978-952-86-1117-2
ISSN 2489-9003

Permanent link to this publication: http://urn.fi/URN:ISBN:978-952-86-1117-2

ABSTRACT

Huhta, Laura
Probing QCD via dijet invariant mass in pp and p–Pb collisions and performance of forward jets with ALICE FoCal

Jets, collimated sprays of particles emerging from hard parton processes, are powerful probes of the strong interaction and of nuclear matter under extreme conditions at the CERN Large Hadron Collider (LHC). In this thesis, I present two studies of jets in the ALICE experiment at the LHC.

The first analysis reports an updated measurement of the invariant mass of charged dijets in proton-proton and proton-lead collisions at a nucleon-nucleon center-of-mass energy of $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV, using Run 2 data. I determine the dijet mass spectra and the nuclear modification factor R_{pA} in the mass range of $75 - 150$ GeV$/c^2$. The results, extending the earlier work presented in Ref. [1], are compared with predictions from Monte Carlo simulations of PYTHIA and POWHEG+PYTHIA. While the measured R_{pA} is found to be consistent with unity within experimental uncertainties, it lies systematically slightly above unity. Simulations suggest that this enhancement may result from dijets probing the anti-shadowing region of the nuclear PDFs of lead. While the current measurement uncertainties are too large for a definite conclusion, future studies with Run 3 data can provide greater sensitivity. This measurement also establishes a baseline for corresponding future analyses in Pb–Pb collisions.

The second analysis is the study of jet performance in the upcoming ALICE Forward Calorimeter (FoCal) upgrade. Using data and detector simulations, the achievable transverse momentum range as well as the jet energy scale and resolution are quantified. The feasibility of forward jet, dijet and photon-jet measurements is demonstrated. Such measurements will enable FoCal to probe the gluon saturation regime of QCD matter at low Bjorken-x down to $x \sim 10^{-6}$. In addition, results from test beam campaigns for the FoCal prototype are presented, validating the detector concept and performance in preparation for the construction and installation of FoCal for data taking in Run 4.

Keywords: ALICE, CERN, LHC, jet, dijet, heavy-ion, FoCal, pp, p–Pb, gluon saturation, QCD

TIIVISTELMÄ (ABSTRACT IN FINNISH)

Huhta, Laura
QCD-ilmiöiden tutkiminen dijettien invarianttimassan avulla pp ja p–Pb törmäyksissä sekä ALICEn FoCal-ilmaisimen suorituskyky etusuuntaisten jettien mittauksessa

Jetit, eli kollimoituneet hiukkassuihkut, syntyvät kovista partoniprosesseista. Jetit ovat tehokkaita työkaluja vahvan vuorovaikutuksen ja ydinaineen tutkimuksessa äärimmäisissä olosuhteissa, joita saavutetaan CERNin suuressa hadronitörmäyttimessä, LHCssa. Tässä työssä esitän kaksi jettitutkimusta CERNin ALICE-kokeessa.

Ensimmäinen tutkimus tuottaa päivitetyn mittauksen varattujen dijettien invarianttimassasta käyttäen protoni-protoni ja protoni-lyijytörmäysdataa joka on mitattu massakeskipiste-energialla $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV ALICEn Run 2 aikana. Määritän dijettimassaspektrin ja ydinmuutostekijän R_{pA} massa-alueella $75 - 150$ GeV$/c^2$. Tuloksia, jotka laajentavat aiempaa työtä [1], verrataan PYTHIA ja POWHEG + PYTHIA Monte Carlo -simulaatioiden ennustuksiin. Vaikka mitattu R_{pA} on mittausepävarmuuksien rajoissa yhteensopiva $R_{\mathrm{pA}} = 1$ kanssa, sijaitsevat mittauspisteet systemaattisesti $R_{\mathrm{pA}} = 1$ yläpuolella. Simulaatiot viittaavat siihen, että tämä lievä poikkeavuus voi johtua siitä, että dijetit osuvat lyijyn ydinpartonijakaumien anti-varjostusalueelle. Vaikka nykyiset mittausepävarmuudet ovat liian suuria varmoja johtopäätöksiä varten, tulevaisuuden tutkimukset uudella ALICEn Run 3-datalla voivat parantaa tulosten tarkkuutta. Tämä mittaus toimii myös vertailupohjana vastaaville tutkimuksille lyijy-lyijytörmäyksissä.

Työn toinen tutkimus käsittelee ALICEn tulevan Forward Calorimeter (FoCal) -ilmaisimen suorituskykyä etusuuntaisissa jettimittauksissa. Käyttämällä datan ja ilmaisimen simulaatioita määritän saavutettavissa olevan poikittaisliikemääräalueen sekä jettien energiaskaalan ja -resoluution. Osoitan, että etusuuntaisten jettien, dijettien ja fotoni-jettikorrelaatioiden mittaukset ovat toteutettavissa FoCal:lla. Nämä mittaukset mahdollistavat QCD-aineen gluonisaturaation tutkimuksen matalilla Bjorken-x:n arvoilla $x \sim 10^{-6}$ asti. Lisäksi esitän tuloksia FoCal-prototyypin testiajoista, jotka vahvistavat ilmaisimen konseptin ja suorituskyvyn, mahdollistaen FoCal:n rakentamisen ja asennuksen mittauksia varten ALICEn Run 4-ajossa.

Avainsanat: ALICE, CERN, LHC, jetti, dijetti, raskasioni, FoCal, pp, p–Pb, gluonisaturaatio, QCD

Author

Laura Huhta
Department of Physics
University of Jyväskylä
Finland

Supervisors

Associate Professor Sami Räsänen
Department of Physics
University of Jyväskylä
Finland

Adjunct Professor Dong Jo Kim
Department of Physics
University of Jyväskylä
Finland

Reviewers

Associate Professor Troels Petersen
Niels Bohr Institute
Denmark

Adjunct Professor Santeri Laurila
University of Helsinki
Finland

Opponent

Associate Professor Martin Spousta
Charles University
Czech Republic

PREFACE

It's turtles all the way down!

– Traditional anecdote

The work in this thesis has been carried out during the years 2021–2025. Before diving in, I would like to express my gratitude to those who have supported me in this work.

First, I would like to thank Associate Professor Troels Petersen and Adjunct Professor Santeri Laurila for reviewing my thesis and providing comments, and Associate Professor Martin Spousta for agreeing to be my opponent for the doctoral defense. I would like to acknowledge the Ellen and Artturi Nyyssönen fund, which funded a portion of my thesis work.

Thank you to the whole ALICE group at the department, first and foremost my supervisor, Sami Räsänen. It has been a pleasure to work under your guidance. Your support in the thesis work, as well as any other tasks and topics that arose during the past years, has been invaluable! I would also like to thank my co-supervisor, Dong Jo Kim, from whom I learned a lot about heavy-ion physics and coding. My thanks go to all the JYFL ALICE group members, former and current. Thank you to my office mates, especially Anna and Maxim, for the laughter, support, and lunchtime ski breaks. I wish to especially thank Oskari and Heidi, both of who generously helped me in my work, and all things related to coding and computers at the start of my PhD. I would also like to specifically thank Hadi for all the help and guidance in the FoCal jet analysis.

The ALICE group at CERN and worldwide has introduced me to the large community of experimental heavy-ion physics. I am grateful for my time spent at CERN, where I had the opportunity to work hands-on in data taking. To all the amazing people I met during my time at CERN, thank you for all the fun times! I also want to acknowledge the FoCal collaboration, whose hard work during the test beams was inspiring. In addition, interacting with the wider community of young researchers throughout Europe via ECR activities has been a wonderful, eye-opening experience.

The Department of Physics has been a great place to study and work. I have especially appreciated the opportunities to improve my teaching skills at the department. Thank you to all the friends we studied together with, especially Mirja, with whom I shared most of my courses and many wonderful memories.

The CoE in Quark Matter began when I had just started my PhD. The CoE has been a wonderful environment to learn more about a wide range of topics, be it through seminars, retreats or journal clubs. I wish to thank all of the members of the CoE, current and former, for the great working environment that everyone has played a part in. Special thanks go to the $\alpha_s cape$-team – who knew that creating a physics-themed escape room could be so much fun!

I am grateful for all the support from my friends and family. To my friends, I have appreciated all our tea times, sports activities, rants and hang-outs. I

couldn't have made it this far without you! To my family, thank you for all the guidance throughout the years, for supporting me through my studies and anything I have decided to do in life.

In Jyväskylä, November 2025

Laura Huhta

AUTHOR'S CONTRIBUTIONS

The author has contributed to the research discussed in this book as follows:

- The author has updated the analysis of the dijet invariant mass, first presented in [1], in pp and p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV measured in ALICE, to account for a new reconstruction pass of the ALICE Run 2 data.
- The author has studied and determined the jet reconstruction performance for the ALICE Forward Calorimeter (FoCal) through simulation studies. The results of the jet performance are published in Chapter 6 of Ref. [2]. The results were also discussed in the author's proceedings for the DIS2024 conference [3].
- As part of the author's service work in ALICE, the author worked on the simulation of a standalone detector prototype for the FoCal test beam setup and participated in the test beams, participating in data taking through setup, shifts and software work. The test beam results in 2021-2023 have been published in Ref. [4].

In addition to central data taking shifts in the ALICE control room at CERN, the author has completed a two-week-long Run Manager shift, being in charge of daily operations at the ALICE control room 24/7. This shift block also counts as service work for ALICE.

An Artificial Intelligence (AI)-based language tool was used in this thesis to check some grammar of text written by the author. All scientific text was produced by the author: no AI-generated scientific content, results or interpretation was used.

CONTENTS

ABSTRACT
TIIVISTELMÄ (ABSTRACT IN FINNISH)
PREFACE
AUTHOR'S CONTRIBUTIONS
CONTENTS

1 INTRODUCTION 11

2 ALICE: A LARGE ION COLLIDER EXPERIMENT 18
2.1 The Large Hadron Collider 18
2.2 ALICE in Run 2 19
2.2.1 Vertex and track reconstruction 22
2.2.2 Visible cross section 24
2.3 Upgrades for Run 3 25
2.4 Upgrades for Run 4 26
2.4.1 ALICE Forward Calorimeter 26

3 JETS AND GLUON SATURATION IN QCD 30
3.1 Jets 30
3.2 Dijets 34
3.3 Collinear factorization 35
3.4 Low-x regime 39

4 JET METHODS 45
4.1 Jet definitions 45
4.1.1 Jet areas 46
4.1.2 Jet reconstruction algorithms 46
4.2 Unfolding 51
4.3 Background subtraction methods 53
4.4 Monte Carlo simulations 55

5 DIJET INVARIANT MASS ANALYSIS 60
5.1 Dijet invariant mass 60
5.2 Update to new reconstruction pass 61
5.3 Underlying event fluctuations 65
5.4 Differential cross section 67
5.5 Nuclear modification factor 69
5.6 Unfolding 70
5.7 Systematic uncertainty 75
5.7.1 Tracking uncertainty 75
5.7.2 Unfolding uncertainty 77
5.7.3 Total systematic uncertainty 86
5.8 Results 88

6 JETS IN ALICE FORWARD CALORIMETER 92
6.1 FoCal test beams 92
6.2 Full detector simulations 96
6.2.1 Reconstruction and clustering 97
6.2.2 Calibration 98
6.3 Jet reconstruction in FoCal 99
6.4 Jet performance results 102

7 CONCLUSIONS AND OUTLOOK 112

REFERENCES 114

1 INTRODUCTION

The physics of our universe are governed by four fundamental forces. The Standard Model is capable of describing a wide range of experimental data on elementary particles and the way they interact via weak, electromagnetic and strong interactions. The Standard Model also classifies all the known elementary particles: the quarks, the leptons, the gauge bosons and the scalar Higgs boson. These are pictured in Fig. 1, along with their masses, charges and spins. There are still open questions on how to complete the Model in order to incorporate the fourth fundamental force, gravity. Also, parameters like masses and coupling constants are not given directly by the Standard Model, but have to be determined based on data from experiments [5]. In addition, while the Standard Model predicts a symmetry of matter and antimatter, an imbalance quantified as the baryon asymmetry is observed. Physics beyond the Standard Model is therefore required to explain the asymmetry [6]. As cosmological and astrophysical evidence suggests, ordinary matter constitutes only about 5% of the Universe, with the remainder consisting of dark matter and dark energy. A wide variety of Beyond the Standard Model (BSM) searches are actively performed at the LHC, with various models being tested and possible signals of new physics searched for in the experimental data [7, 8, 9].

The theory of quarks as building blocks of matter was formulated in the 1960's [11, 12], and experimentally confirmed in 1969 in electron-proton scattering at the Stanford Linear Accelerator Center (SLAC) [13, 14]. Quantum Chromodynamics (QCD) is the theory of the strong interaction, with fundamental fields of quarks and gluons described by matter and gauge fields, respectively. QCD is a gauge theory with the local symmetry group $SU(N_c)$, where $N_c = 3$ is the number of color degrees of freedom in the real world. If N_c is left as a free parameter, the theory can be studied with perturbative methods and simplified calculations using the large-N_c limit approximation [15, 16].

The strong force is not as simple as the electrodynamic force described by Quantum Electrodynamics (QED), which has been studied to a great extent since the 19th century, and for which the theory and values of parameters, such as the fine-structure constant $\alpha = e^2/4\pi\hbar c$ describing the strength of the interaction,

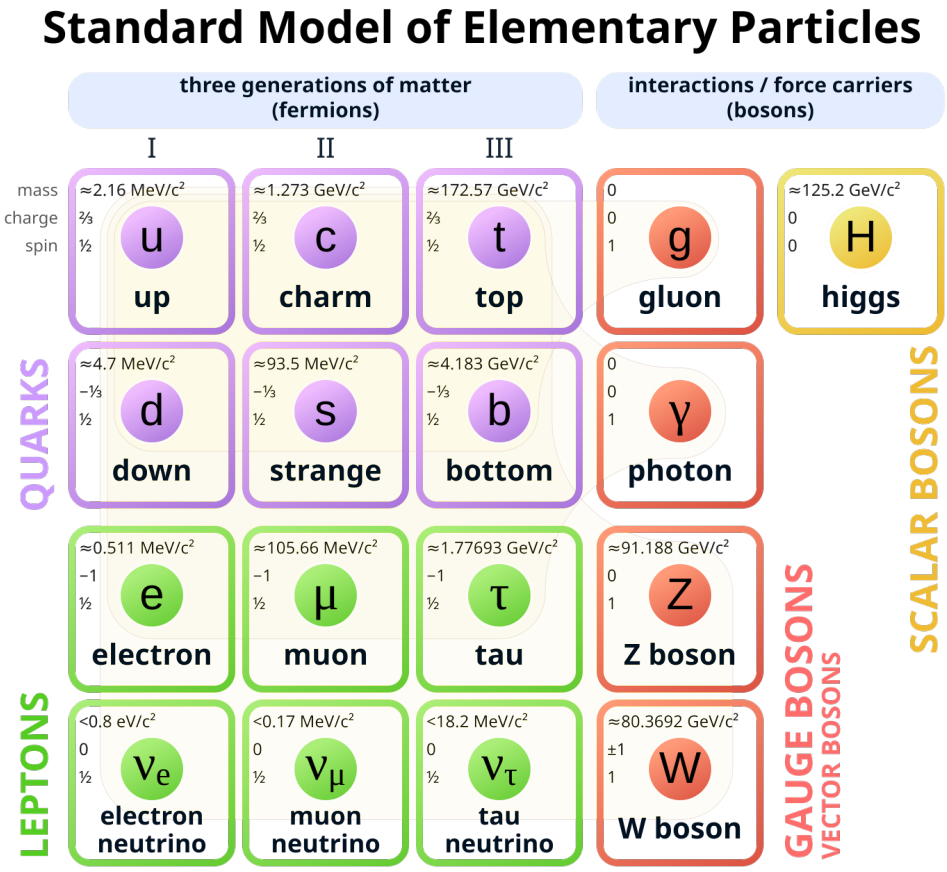


FIGURE 1 The Standard Model. Figure from Ref. [10], adapted under the license CC BY-SA 3.0.

are well-understood [5]. Just as QED involves the concept of electric charge transferred by photons, QCD contains the concept of color charge, which is transferred by gluons. What makes the QCD color charge different, however, is the fact that the gluon itself also carries a color charge, unlike the photon, which has a neutral electric charge. This allows gluons to interact with each other. Color confinement describes the phenomenon where color-charged particles cannot exist in isolation, but must combine into color-neutral states, hadrons. This means that, under normal circumstances, quarks and gluons cannot exist in isolation nor be directly observed. Searches for free quarks have been performed, yet no evidence for their existence have been found in experiments [17].

A major factor making the theory of QCD more complex than QED is asymptotic freedom. This arises because the strong coupling constant α_s, also called the running coupling constant, strengthens at long distances and weakens at short distances, a behavior that is a direct consequence of gluon self-interactions [18, 19]. Measurements of α_s as a function of the energy scale Q are presented in Fig. 2, where the weakening of the interaction strength at higher scales Q, corresponding to shorter distances $\sim 1/Q$, is apparent. In perturbative QCD (pQCD), the order of the calculation indicates the power of α_s. In leading order (LO), the calculations include the factor α_s, next-to-leading order (NLO) includes powers

α_s^2, next-to-next-to-leading order (NNLO) includes powers α_s^3, and so on. Increasing the order can quickly make the calculation increasingly complex and time-consuming.

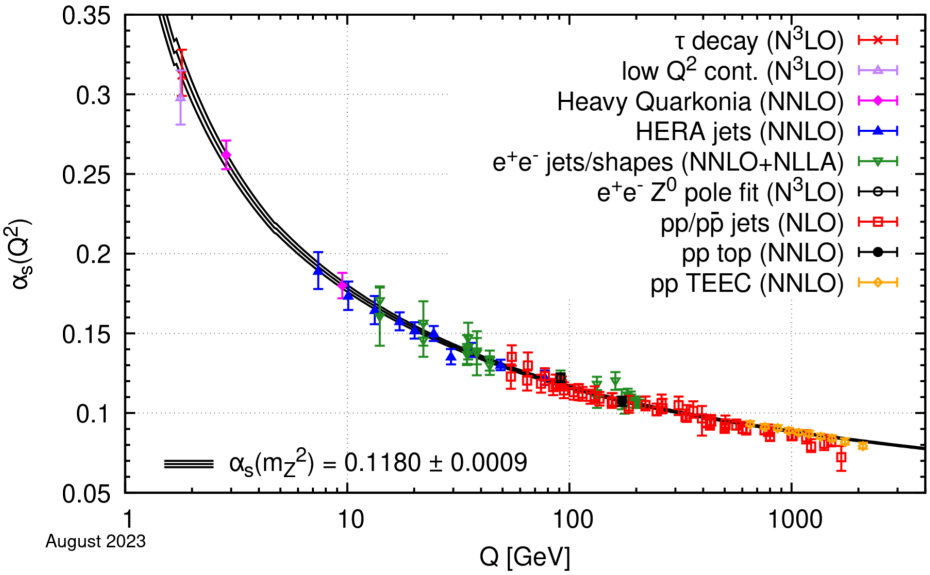


FIGURE 2 Summary of determinations of the running strong coupling constant α_s as a function of the energy scale Q. The black line represents α_s computed at five loops by taking as an input the current PDG average, $\alpha_s(m_Z^2) = 0.1180 \pm 0.0009$. Figure from Ref. [17], reprinted under the license CC BY 4.0.

In order for quarks and gluons to be able to exist as if they are free partons, we would need extreme conditions where temperatures and parton densities are extremely high. Our universe, created around 13.8 billion years ago, was extremely dense and hot in the time right after the Big Bang occurred – this is portrayed in Fig. 3. Neither atoms, nuclei nor nucleons could yet exist in the hot and dense state, but quarks and gluons formed a plasma where they could move freely, unbound to each other. This state of matter is called the quark-gluon plasma (QGP) [20]. The partons act as free partons because the strong interaction is weaker at short distances, which is why extreme conditions are required to create QGP. As the universe cooled down after the Big Bang, the QGP could not exist anymore. Quarks would start to stick together, forming hadrons, the state of matter in which quarks are bound together.

The phase diagram of QCD matter is presented in Fig. 4. In the figure, the usual nuclear matter at low temperatures is shown on the bottom of the plot, and experimental measurements [22, 23, 24, 25] of the chemical freeze-out parameters are shown as open points. Above the crossover line, shown in green and calculated with lattice QCD [26], QGP can exist. The region probed by LHC experiments is depicted in blue. At large baryochemical potential μ_B, a first-order phase transition is predicted, which would end at a critical point when moving toward lower μ_B. Experimental programs, such as the RHIC Beam Energy Scan (BES) [27], actively search for signatures of this critical point [28]. This region of

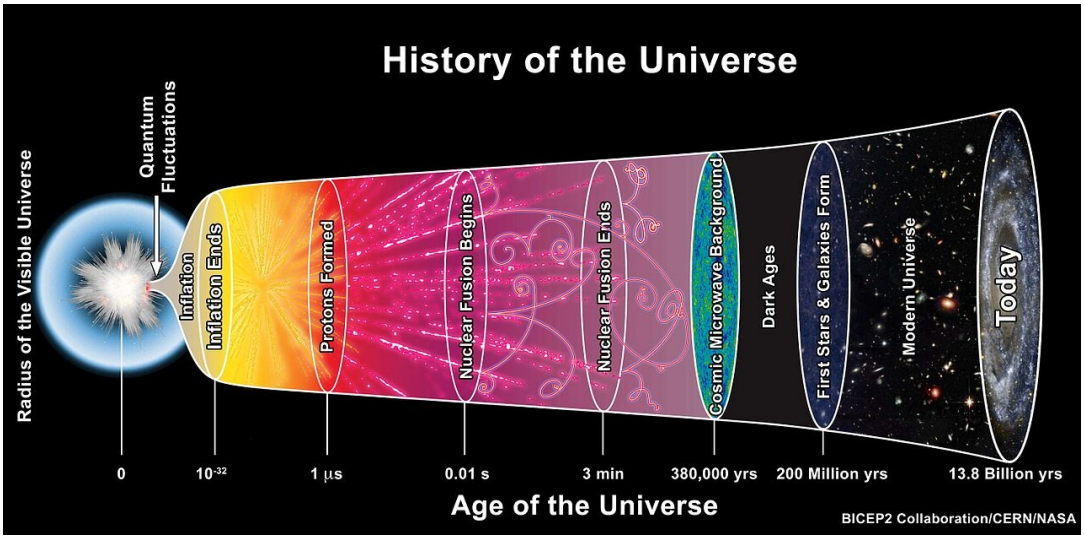


FIGURE 3 The timeline of our universe right after the Big Bang. © CERN/BICEP2 Collaboration/NASA. Reproduced for informational purposes from CERN website [21].

the phase diagram is relevant also for understanding astrophysical objects like dense neutron stars [29].

QCD matter can be studied in high-energy particle collisions, for which we need particle accelerators. Accelerators for high-energy nuclear physics were first developed in the 1930's, including the first circular accelerators, cyclotrons [5]. In Fig. 5, a summary of accelerators and colliders between 1950–2010 is shown, along with the center-of-mass (CM) energy reached with each machine. Hadron–hadron, electron–positron, and HERA electron–proton collision systems are shown with different marker styles. The Large Electron-Positron (LEP) collider collided electrons with positrons at the European Organization for Nuclear Research (CERN). The 27-kilometer-long tunnel of LEP now houses the Large Hadron Collider (LHC), which reaches the highest energies ever achieved with a particle accelerator.

The first lead-ion collisions were performed at CERN at the SPS, at energies of 158 GeV per nucleon [30]. In 2000, the first measurements of gold-ion collisions at the Relativistic Heavy Ion Collider (RHIC) at BNL were conducted [31]. To this day, energies up to $\sqrt{s_{\mathrm{NN}}} = 200$ GeV have been reached at RHIC.

In 1993, an experiment specialized in QCD physics to be built at the LHC was proposed. A Large Ion Collider Experiment (ALICE) [32] was envisioned to be a dedicated experiment to study the QGP, which was expected to be created in heavy-ion collisions. The vision laid out in the ALICE Letter of Intent [33] is to study strongly interacting matter at extreme energy densities, exploiting nucleus-nucleus collisions at LHC energies in order to form the expected new phase of matter, QGP. In 2010, the first heavy-ion data at the LHC and in ALICE was gathered.

The timeline of a heavy-ion collision is shown in Fig. 6. Jets, collimated showers of particles emerging from hard processes, are produced in the earliest phase of a heavy-ion collision. This means that they interact with the QGP medium throughout the whole time that the QGP exists. Through this interac-

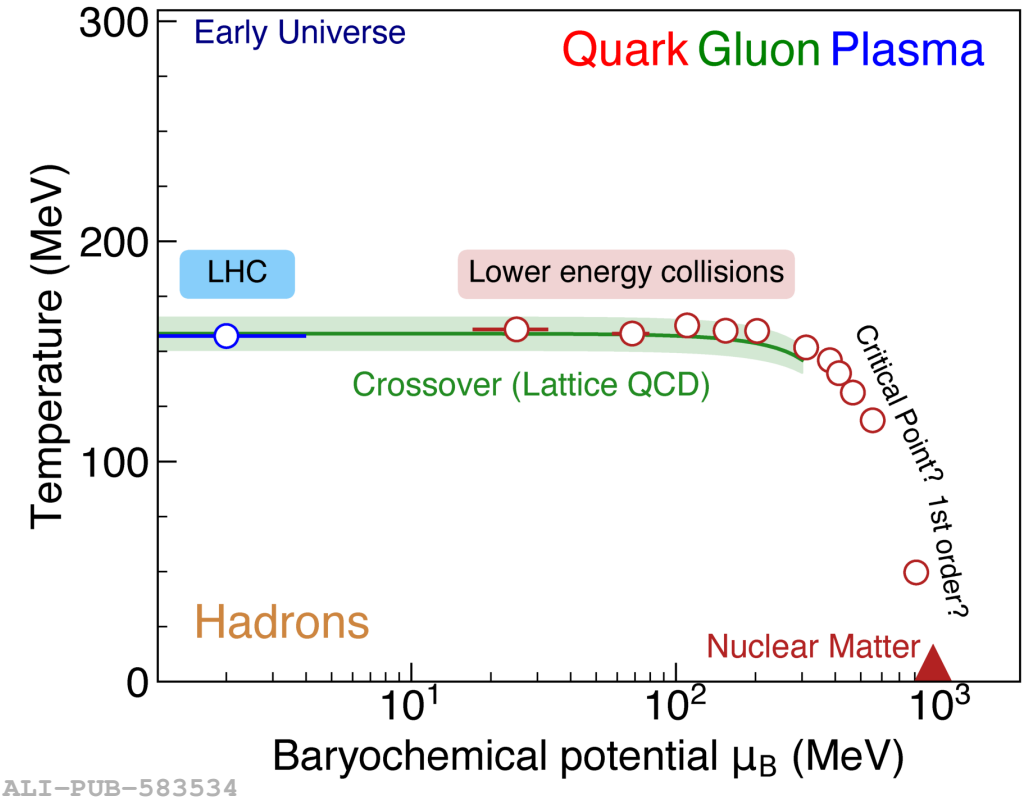


FIGURE 4 A schematic figure representing the phase diagram of strongly interacting matter. The open points show results from experiments. From Ref. [29], reprinted under the license CC BY 4.0.

tion, the jets are modified, as the jet can redistribute its energy with the medium itself [34]. By studying the ways the jets redistribute their energy when passing through the QGP, we can study the properties of the QGP [35].

Dijets are systems consisting of two jets created roughly back-to-back in the same hard parton collision. Dijets are important observables, as they can be used to constrain parton distribution functions (PDFs) [36], and to study jet quenching in heavy-ion collisions [37, 38, 39]. If we want to interpret dijet measurements in heavy-ion lead-lead collisions, we must first perform a reference measurement to study any possible effects arising from cold nuclear matter (CNM) instead of QGP. This can be done by studying the same observable in pp and p–Pb collisions, to create a baseline to compare future Pb–Pb measurements to. In addition, the cold nuclear matter effects are an interesting topic on their own. Many measurements have been performed to study CNM effects, with some signals of CNM effects of certain observables seen in p–Pb and d–Au collisions [40, 41, 42, 43].

Pseudorapidity is defined as

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right), \tag{1}$$

where θ denotes the scattering angle. At the high-energy limit, or when the mass of the particle can be approximated as zero, the definition of pseudorapidity converges to the definition of rapidity y. Larger absolute values of pseudorapidity

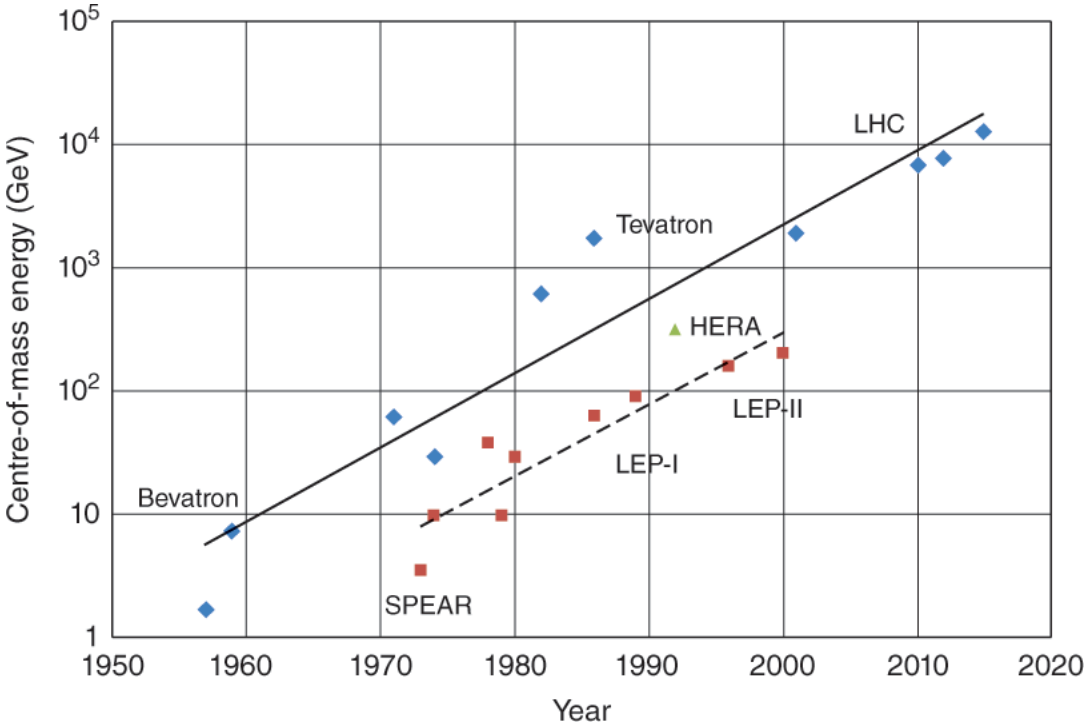


FIGURE 5 A summary of colliders between 1950-2010 colliding different particle types: hadron–hadron (diamonds), electron–positron (boxes), and HERA electron–proton (triangle). Figure from Ref. [5], reprinted under the license CC BY 4.0.

correspond to measurements closer to the beam axis, in the so-called forward region.

ALICE jet measurements have so far been performed at mid-rapidity, at $|\eta| < 0.9$. A new detector planned for data taking in Run 4, from the year 2030 onward, will enable more forward-rapidity measurements. This is the Forward Calorimeter (FoCal) [44], a detector upgrade that will be installed in the very forward region of ALICE, at $3.2 < \eta < 5.8$. Measurements at forward rapidity provide access to observables in kinematic regions that cannot be reached with current LHC experiments at mid-rapidity. In particular, FoCal will probe QCD in the region of gluon saturation, where the gluon density in hadrons is so large that gluon splitting is expected to be balanced by gluon recombination, taming the gluon density growth and reaching a gluonic saturation point. FoCal will also provide new, stringent constraints to nuclear PDFs [2].

In this thesis, I first present the experimental setup and methods in Sec. 2. The relevant background for jet measurements in ALICE is discussed in Sec. 3. In Sec. 4, jet methods and Monte Carlo simulations are described. In Sec. 5, I present an updated measurement of the dijet invariant mass of charged jets in the ALICE mid-rapidity measured during Run 2, continuing and updating the work presented in Ref. [1]. The results are also compared with Monte Carlo simulation results. Finally, in Sec. 6, I present a study on the jet performance of the upcoming FoCal detector upgrade. Using simulated data, I study the transverse momentum range and resolution at which these measurements can be performed. I also

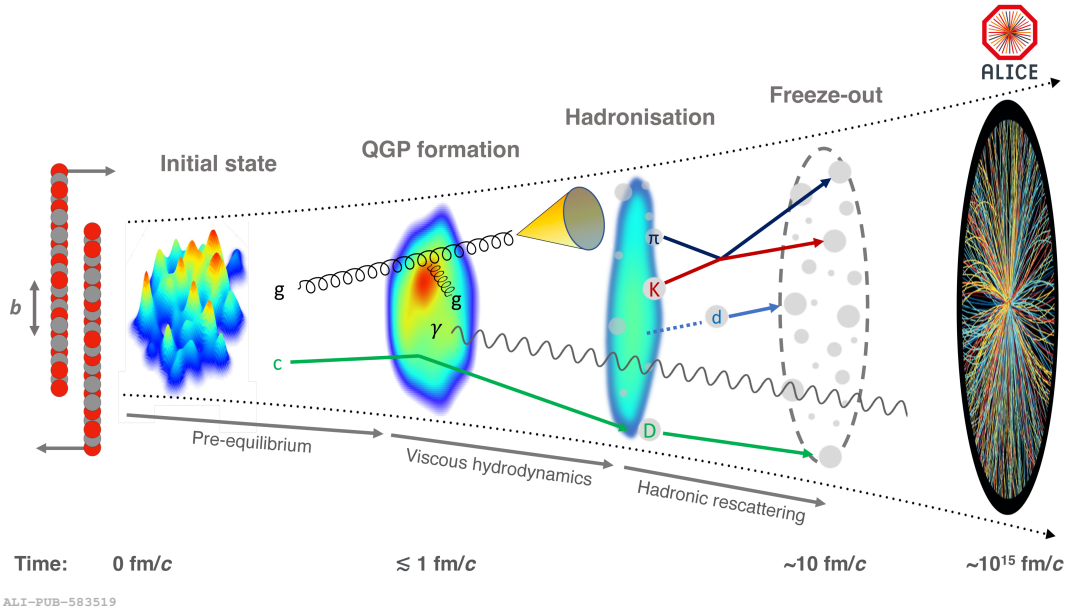


FIGURE 6 The timeline of a heavy-ion collision at LHC energies. A jet is portrayed by the yellow cone. From Ref. [29], reprinted under the license CC BY 4.0.

present work that has been done to test the prototypes of the FoCal in beam tests at CERN Test Beam facilities before its upcoming construction, assembly and installation in ALICE.

2 ALICE: A LARGE ION COLLIDER EXPERIMENT

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [45] is a 27-kilometer-long particle accelerator, situated in a tunnel around 100 meters underground at CERN, on the Franco-Swiss border nearby the city of Geneva, Switzerland. The LHC is the largest particle accelerator to date, with the capability to accelerate protons and ions to speeds near the speed of light. The LHC operations started in 2009 and, since then, CM energies of up to $\sqrt{s} = 13.6$ TeV have been reached for proton-proton (pp) collisions, and CM energies per colliding nucleon-nucleon pair $\sqrt{s_{\mathrm{NN}}} = 5.36$ TeV for lead-lead (Pb–Pb) collisions. The LHC has also collided other ion species, such as xenon [46], neon and oxygen [47].

The protons accelerated in the LHC start from negative hydrogen ions accelerated in the linear accelerator, LINAC4 [48], which strips the ions of their electrons, resulting in 160 MeV protons. The protons then enter the Proton Synchrotron Booster [49], where they are accelerated to 2 GeV and injected into the Proton Synchrotron (PS) [50]. In the case of ions, the Low Energy Ion Ring (LEIR) [51] receives pulses of lead ions from the Linear Accelerator 3 (LINAC3) [51], and accelerates them to 72 MeV to inject into the PS. The PS accelerates the proton beams to 26 GeV, after which they are sent to the Super Proton Synchrotron (SPS) [52], where they are further accelerated to 450 GeV. Finally, the particles are transferred to the two beam pipes of the LHC, circling clockwise and anticlockwise. The protons reach their maximum energy of 6.5 TeV in around 20 minutes of circulation in the LHC, after which they are ready to be brought to collision at the four experimental stations around the LHC.

In Fig. 7, the LHC schematic is presented, with the four large LHC experiments pictured around the LHC ring. Many other experiments also take place at CERN, and are situated in multiple positions around the accelerator complex. Of the four large experiments, ALICE [32], ATLAS [53], CMS [54] and LHCb [55], ALICE is the only one specifically optimized to study the collisions of heavy ions. With its precise particle identification and tracking capabilities, ALICE is

designed especially for the study of QGP in lead-ion collisions, but also conducts a wide range of measurements with all possible collision systems at the LHC.

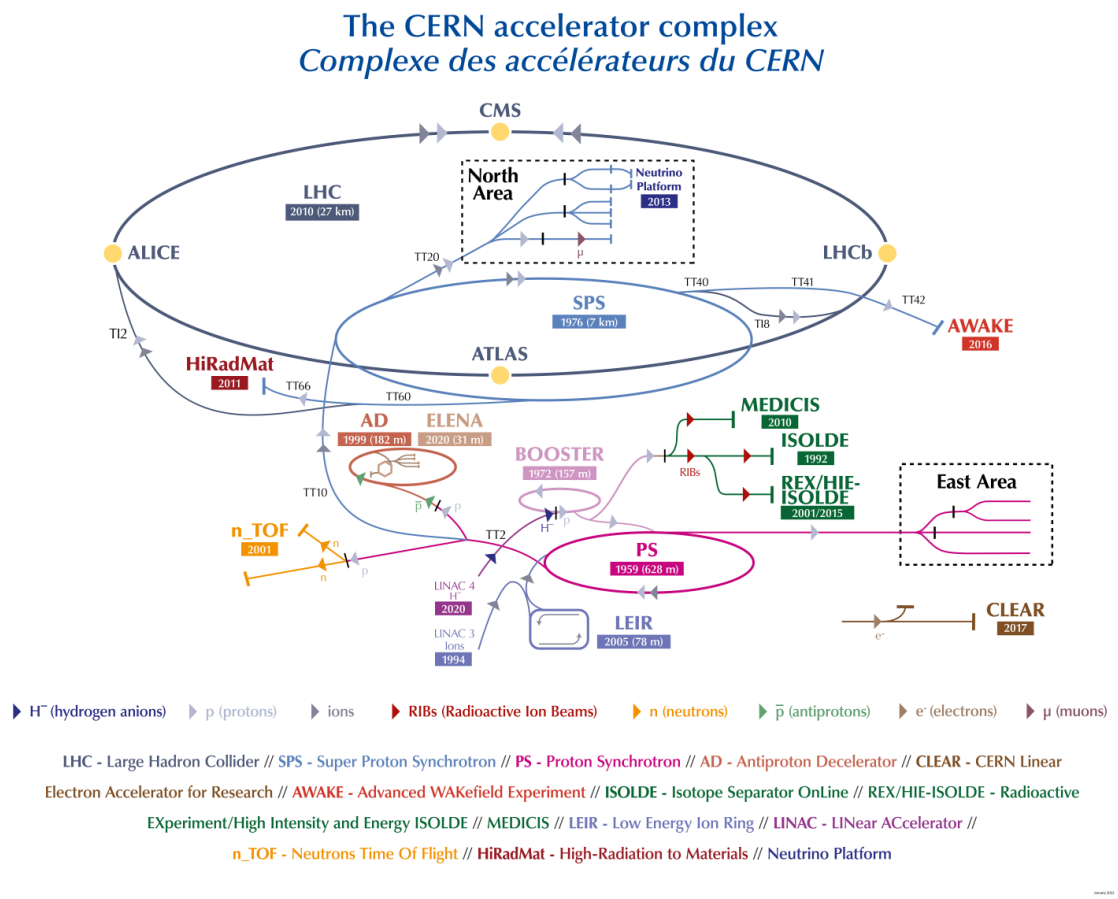


FIGURE 7 The CERN accelerator complex as it existed in 2022. Figure from Ref. [56], freely available for non-commercial use.

2.2 ALICE in Run 2

The dijet mass analysis described in Sec. 5 is performed on data measured during the LHC Run 2, during the years 2015–2018. A schematic of the ALICE detector during Run 2 is presented in Fig. 8. In between Run 1, 2010–2013, and Run 2, ALICE kept its general layout similar, but underwent some upgrades to improve performance.

The central part of ALICE [32] is enclosed within a solenoid, the L3 magnet, which has a radius of 5.75 m and length of 12.1 m. Inside the L3 magnet, the central detectors can be found: from the inside out, the central barrel contains the Inner Tracking System (ITS) [57], the Time-Projection Chamber (TPC) [58], the Transition Radiation (TRD) [59] detector, the Time-Of-Flight (TOF) [60] particle identification arrays, the High Momentum Particle Identification Detector

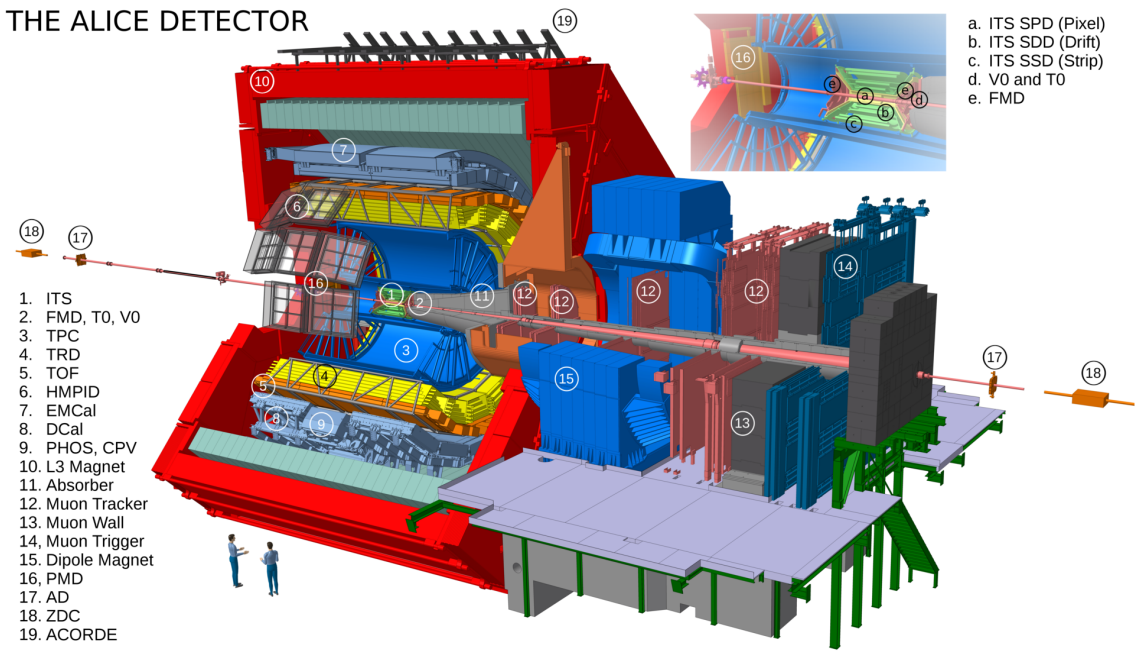


FIGURE 8 The ALICE detector in Run 2. Figure from Ref. [29], reprinted under the license CC BY 4.0.

(HMPID) [61], and two electromagnetic calorimeters (PHOS [62] and EMCal [63]). All but HMPID, PHOS and EMCal cover the full azimuth. The L3 magnet provides a moderate magnetic field of 0.5 T. Another magnet, the dipole magnet, is situated 7 m from the interaction point, and is an integral part of the muon arm.

ALICE concentrates on physics at or close to mid-rapidity, where the largest energy density is found in heavy-ion collisions. In the forward directions, there are a few more sub-detectors: the forward muon arm [64], the Forward Multiplicity Detector (FMD) [65], the Photon Multiplicity Detector (PMD) [66] and the Zero Degree Calorimeter (ZDC) [67]. Also, to provide triggering information, the T0 detector measures the event time, and the V0 detector is used as a minimum bias (MB) trigger and to reject beam-gas background [65]. Further information on sub-detectors relevant for our analysis is listed in Table 1 and described below. The two arrays of the T0 and V0 detectors are named according to the sides of ALICE they are located on: the A-side corresponds to positive pseudorapidity, on the left hand side of Fig. 8, and the C-side corresponds to the side of negative pseudorapidity, on the right hand side of Fig. 8.

The ITS [57] in Run 2 is formed of six cylindrical layers of silicon detectors. The innermost two layers are Silicon Pixel Detectors (SPD). The following two layers are Silicon Drift Detectors (SDD), and the two outer layers Silicon micro-Strip Detectors (SSD). The ITS allows to localize the primary vertex with a resolution better than 100 μm, to track and identify particles and to improve momentum and angle resolution for particles reconstructed with the TPC. The ITS can provide $\mathrm{d}E/\mathrm{d}x$ measurements for low-momentum particles. The silicon detectors have been designed to keep the amount of material in the active volume to a minimum, as multiple scattering effects in the detector material are dominant

TABLE 1 Summary of a selection of ALICE detector sub-systems, their pseudorapidity coverage, and distance from the interaction point to the detector face, either corresponding to the inner and outer radius of the barrel detectors, or position along the beam. Values are taken from Ref. [32].

Detector	η_{min}	η_{max}	**Position (m)**
ITS layer 1,2 (SPD)	-2.0	2.0	0.039, 0.076
ITS layer 3,4 (SDD)	-0.9	0.9	0.150, 0.239
ITS layer 5,6 (SSD)	-0.97	0.97	0.380, 0.440
TPC	-0.9	0.9	0.848, 2.466
TRD	-0.84	0.84	2.90, 3.68
TOF	-0.9	0.9	3.78
V0A	2.8	5.1	3.4
V0C	-3.7	-1.7	-0.897
T0A	4.61	4.92	3.75
T0C	-3.28	-2.97	-0.727

effects in the momentum resolution for low-momentum particles. The relative momentum resolution of the ITS is better than 2% for pions with transverse momentum p_{T} between 0.1 and 3 GeV/c.

The TPC [58] is the main tracking detector in the central barrel. The TPC is designed to provide measurements of charged-particle momentum with good particle identification, track separation and vertex determination [68]. The volume of the cylindrical barrel of the TPC is large, 5.6 m in diameter and 5 m in length. The detector is formed of a cylindrical field cage filled with a gas mixture of Ne-CO_2-N_2 with a ratio 90-10-5. The small amount of nitrogen was added after Run 1 to improve operation stability [69]. Charged particles passing through the TPC volume ionize the gas along their path, causing electrons to liberate and drift through the field cage to the end plates on either side of the central electrode. Each end plate is equipped with multi-wire proportional chambers mounted onto 18 trapezoidal sectors. The field cage is operated at high voltage gradients of 400 V/m with the central electrode at 100 kV, resulting in a maximum drift time of about 90 μs. The accurate measurement of both the arrival point and time of the ionization electrons allow the complete trajectory of charged particles in the TPC to be determined with good precision.

The main purpose of the TRD [59] is electron identification above transverse momenta of 1 GeV/c. In addition, it also is a part of the Level-1 trigger, providing fast triggers for high-p_{T} charged particles, enhancing, among others, recorded high-p_{T} jet yields. TOF [60], coupled with ITS and TPC for track and vertex reconstruction and $\mathrm{d}E/\mathrm{d}x$ measurements, provides particle identification in the intermediate momentum range.

The V0 [65] detector is a combination of two arrays of scintillator counters, V0A and V0C, placed on either side of the ALICE interaction point. The V0 detector provides MB triggers for the central barrel detectors, and indicates the centrality of the collision via the recorded event multiplicity. Also, the V0 detector

participates in luminosity measurements in pp collisions. The T0 [65] detector consists of two arrays of Cherenkov counters, T0A and T0C. The T0 generates a start time for the TOF detector, corresponding to ALICE's best estimate of the collision time. T0 also measures the vertex position and provides a trigger based on the position measurement, with the capability to generate an alternative MB and multiplicity trigger.

2.2.1 Vertex and track reconstruction

The track of a particle in the ALICE detector can be reconstructed by using multiple signals from a combination of detectors and by applying a tracking algorithm to the signals. For charged particles, which are used in our dijet mass analysis, the main detectors used in tracking are the TPC and ITS, with TRD and TOF used for particle identification and triggering.

The track reconstruction is performed in three stages [70], which are pictured in Fig. 9. In the first stage, the procedure starts with tracks found in the outer layer of the TPC. Track seeds are built with TPC clusters, and they are then propagated inwards, updated at each step with the nearest cluster corresponding to a proximity cut. Once the outermost ITS layer is reached, the reconstructed TPC tracks become seeds for the ITS track finding. They are then also propagated inwards, updated at each ITS layer with all clusters within the proximity cut, with each update saved as a new seed. Thus, for each track in the TPC, there is a tree of track hypotheses produced in the ITS. The highest quality candidate from each tree is added to the reconstructed event.

The second stage starts with all tracks from the previous stage being extrapolated to the point of closest approach to the preliminary interaction vertex. The tracks are refitted using the Kalman filter [72] in the outward direction, using clusters found in the first stage. The tracks are propagated outside the TPC to TRD, TOF and the other detectors, but they are not used to update the measured track kinematics.

In the third and final stage, all tracks are propagated inwards, starting from the outer layer of the TPC. The tracks are refitted in both the TPC and ITS with the previously found clusters. In this stage, the position, direction and curvature are determined. A vast majority of the tracks found in this way originate from the primary interaction vertex. In order to minimize tracks originating from secondary vertices, sets of cuts on the track distance of closest approach (DCA) can be imposed. The DCA to the primary vertex is required to be below 3.2 cm in z and 2.4 cm in the $x - y$ plane. In our analysis, tracks are required to have a p_{T} over 150 MeV/c, as the p_{T} resolution of tracks in ALICE is the most optimal between 150 MeV/c and 100 GeV/c [73].

The ITS provides information that can improve the track resolution significantly. Therefore, the ITS+TPC tracks are used when possible. However, the SPD layers of the ITS were partly inactive for some parts of the ALICE data taking during Run 2 [70]. This caused an uneven distribution of ITS+TPC tracks, and in order to balance the distribution, TPC-only tracks, reconstructed solely with TPC

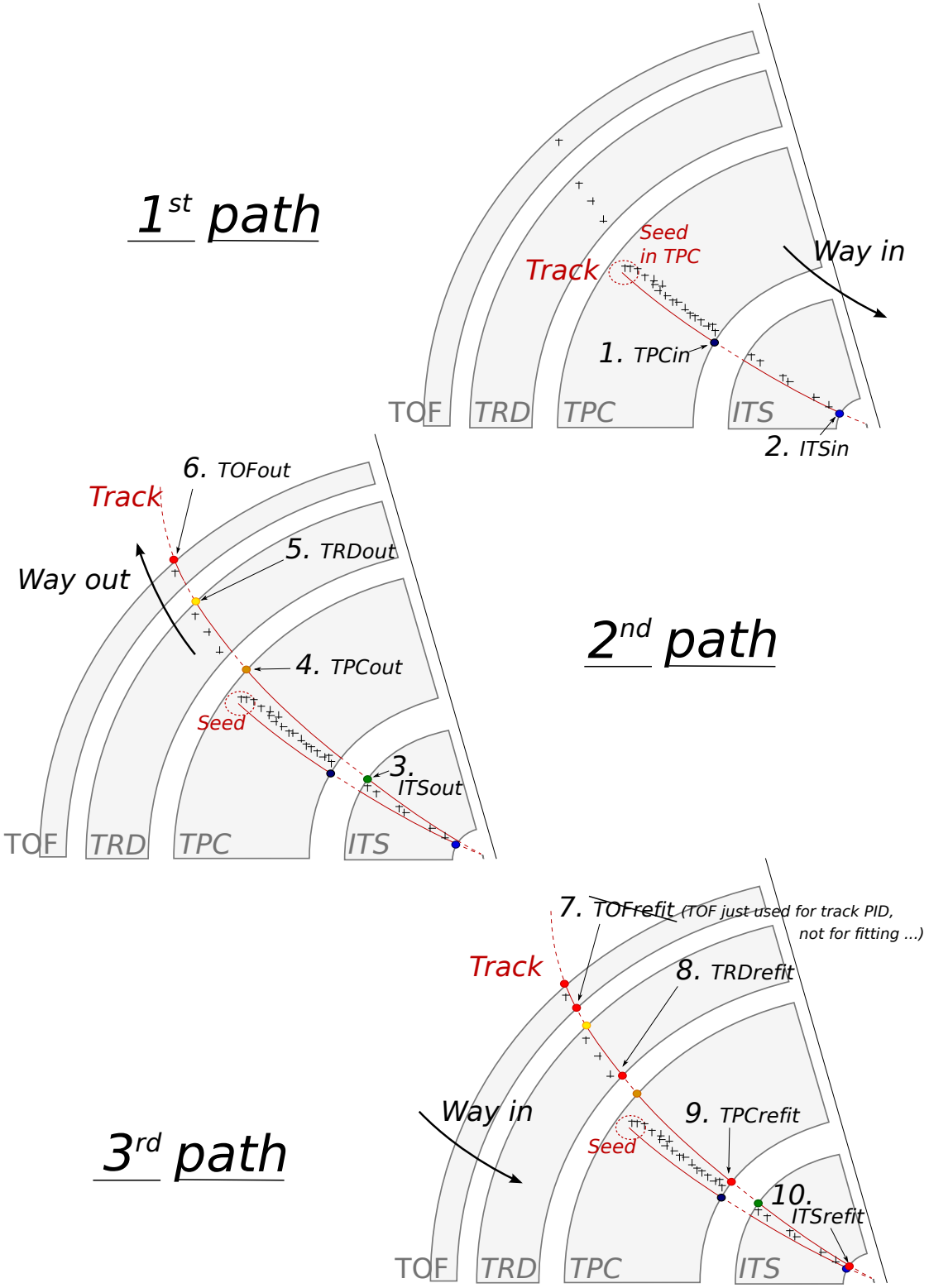


FIGURE 9 The tracking principles for an event in ALICE in Run 2, showing the three successive paths allowing to reconstruct a track and to refine its parameters. Figure from Ref. [71], © CERN, provided free of charge for educational and informational use.

information, are used whenever ITS is not fully usable. We call the combination of the two track classes, ITS+TPC and TPC-only, a hybrid track selection [74].

The reconstruction of the primary vertex is based on information provided by the innermost SPD layers of the ITS. For our analysis, the SPD vertex is required, even in the case of TPC-only tracks. The SPD vertex significantly improves track resolution when reconstructing with TPC-only track information. The SPD vertex is also required to be found within 0.5 cm of the reconstructed primary vertex. We also restrict the primary collision vertex $|z_{\mathrm{vtx}}| < 10$ cm, a default cut applied in most ALICE analyses in order to ensure high-quality events.

As mentioned, the events selected for the analysis are required to have a reconstructed vertex. However, the MB trigger sometimes triggers on events with too few tracks for the vertex to be reconstructed. This means we have to implement a correction factor to match the number of accepted events of the analysis to the minimum bias events. As the vertex position is required to be a 10 cm distance away from the interaction point in z, this allows us to write the total number of events with a vertex within the z acceptance as

$$N^{\mathrm{tot}}_{<10\ \mathrm{cm}} = N^{\mathrm{vtx}}_{<10\ \mathrm{cm}} \frac{N^{\mathrm{tot}}_{<10\ \mathrm{cm}}}{N^{\mathrm{vtx}}_{<10\ \mathrm{cm}}} \approx N^{\mathrm{vtx}}_{<10\ \mathrm{cm}} \frac{N^{\mathrm{tot}}}{N^{\mathrm{vtx}}}, \tag{2}$$

where N^{tot} is the number of all the MB triggered events, N^{vtx} is the number of triggered events with a reconstructed vertex, and the subscript < 10 cm signifies the events with vertices within the z acceptance.

The ratio of events with a reconstructed vertex divided by the total number of triggered events is then defined as the vertex reconstruction efficiency:

$$\epsilon_{\mathrm{vtx}} \equiv \frac{N^{\mathrm{vtx}}}{N^{\mathrm{tot}}}. \tag{3}$$

This correction factor can then be used to correct the total number of events considered in the analysis, and we rewrite Eq. (2) as

$$N^{\mathrm{tot}}_{<10\ \mathrm{cm}} \approx \frac{N^{\mathrm{vtx}}_{<10\ \mathrm{cm}}}{\epsilon_{\mathrm{vtx}}}. \tag{4}$$

2.2.2 Visible cross section

For the dijet mass analysis, we use a MB triggered data set called V0AND. The name refers to a logical AND condition requiring hits in the forward and backward direction of the V0 detector [70]. ALICE has measured the corresponding visible cross section, σ_{V0AND}, using a van der Meer scan [75]. The visible cross section, determined using the V0AND trigger condition, represents a fraction of the total inelastic interaction cross section, the two being related by an efficiency factor. To calculate this efficiency, ALICE performed event-generator studies, simulating the visible cross section by applying the V0AND trigger condition at the generator level [76]. From these studies, the efficiency, corresponding to the ratio of the visible cross section to the total inelastic cross section, was found to be $74.2^{+5.0}_{-2.0}\%$ at $\sqrt{s} = 7$ TeV.

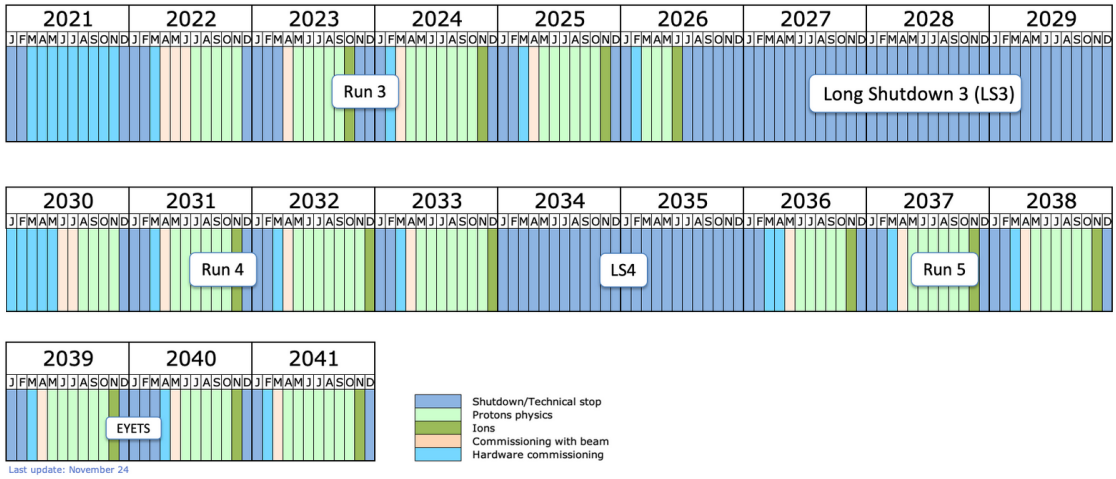


FIGURE 10 The long term schedule of the LHC, as foreseen in 2025. Obtained from the LHC Commissioning website (CERN) [84]. © CERN, used for informational purposes only.

For V0AND triggered data, ALICE has measured the visible cross section for pp collisions at $\sqrt{s} = 5$ TeV to be $50.87 \pm 0.04(\text{stat.}) \pm 0.92(\text{syst.})$ mb [77] and for p–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV to be $2090 \pm 70(\text{syst.})$ mb [78].

2.3 Upgrades for Run 3

After Run 2, the LHC went into a period called the Long Shutdown 2 (LS2) in years 2019–2022, during which upgrades to the detectors and the LHC could take place [79]. Figure 10 shows the long term schedule of the LHC at the time of writing. For ALICE, the major upgrades were the TPC [80], ITS [81], the new fast interaction trigger (FIT) [82], and the new Muon Forward Tracker (MFT) [83]. Also, the beam pipe was replaced with a new pipe, made of beryllium, which is thinner and smaller in diameter than before. With these upgrades, the performance of tracking and secondary vertex reconstruction, especially at low momenta, was enhanced, and the material budget of the tracking system was decreased.

The main physics topics addressed by the ALICE upgrade for Run 3 are the precise measurements of heavy-flavor hadrons, low-momentum quarkonia, and low-mass di-leptons. As these measurements have a very small signal-to-background ratio, very large statistics are necessary for their measurement. Also, the large background makes triggering techniques inefficient or even impossible. Due to this, the whole data taking system of ALICE was upgraded. Before, a dedicated trigger was used to select events of interest to be saved. In Run 3 and beyond, ALICE has moved to a continuous common readout system, called Online/Offline processing (O^2) [85].

The increase in data flow with the readout system update is roughly two orders of magnitude compared to Run 1. The raw data input reaches values of 3.5 TB/s for Pb–Pb collisions at 50 kHz [86]. A maximal reduction of the data

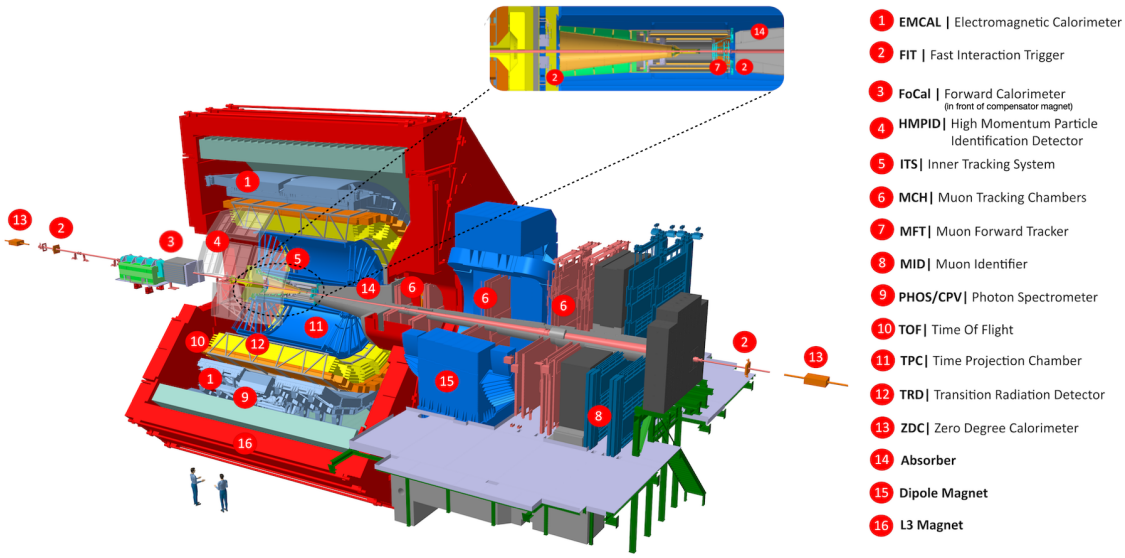


FIGURE 11 The ALICE detector in Run 4. Figure from Ref. [89], reprinted under the license CC BY 4.0.

read-out from the detector is performed as early as possible within the data processing. First, the non-suppressed data is shipped to the O^2 facility at a rate of up to 50 kHz for Pb–Pb and 200 kHz for pp and p–Pb collisions. This reduces the aggregated detector data rate for Pb–Pb collisions from 3.5 TB/s to 900 GB/s. The data is reconstructed synchronously with the data taking in multiple steps, for example using online cluster-finding for the TPC raw data. During the synchronous reconstruction pass, the Pb–Pb data rate is further reduced to 130 GB/s before writing the data to storage [86]. This data is stored temporarily, and the second reconstruction stage is performed asynchronously [85].

2.4 Upgrades for Run 4

The analysis described in Sec. 6 is performed on simulated data in the new Forward Calorimeter (FoCal) to be installed in ALICE during the Long Shutdown 3, 2026-2030, for data taking during Run 4, 2030–2033. For Run 4, also the ITS will be upgraded, with a new ITS3 detector [87, 88] replacing the ITS2 from Run 3. In Fig. 11 is a schematic of the ALICE detector as it is planned for Run 4, with FoCal installed in the forward region.

2.4.1 ALICE Forward Calorimeter

The ALICE Forward Calorimeter (FoCal) [44, 89] consists of a highly granular silicon-tungsten (Si+W) electromagnetic calorimeter (FoCal-E) combined with a copper (Cu) and scintillating fiber hadronic calorimeter (FoCal-H). The detector is 0.9 m long on its sides, shaped nearly square in the transverse plane. There is a square opening in the middle, of dimension $5 \times 5\,\mathrm{cm}^2$, through which the

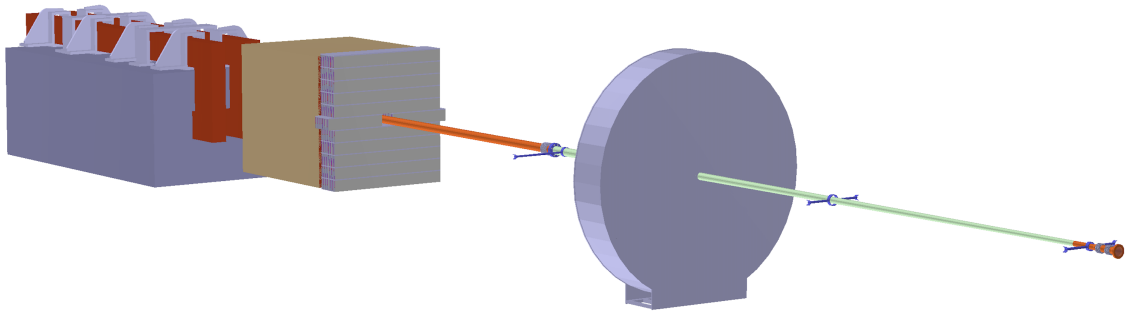


FIGURE 12 The geometry of FoCal as implemented in the simulation of the ALICE detector. FoCal is on the left, at 7 m from the interaction point, placed in front of the compensator magnet on the far left. The figure also shows the details of the beam pipe and the FIT-A detector, with other details omitted for visibility. Figure from Ref. [2], reprinted under the license CC BY 4.0.

beam pipe passes. This results in a pseudorapidity coverage of approximately $3.2 < \eta < 5.8$ of which the range $3.4 < \eta < 5.5$ is full coverage in the azimuth. The geometry of FoCal in a simulation of the ALICE Run 4 detector is shown in Fig. 12.

As the most challenging measurements within the FoCal physics program are those of direct photons and high p_{T} neutral pions, the detector must be designed to fulfill these requirements [44]. The detector must be placed close to the beam pipe to ensure large rapidity. Placing the detector a large distance away from the primary interaction vertex is also necessary, since at forward rapidities, the kinematic boost of particles creates very high particle densities. Thus, FoCal is planned to be installed 7 m away from the interaction point, on the A-side of the ALICE experiment, in front of the compensator magnet.

The necessity to discriminate between direct photons and decay photons arising from neutral pion decays dictates that the spatial separation power of the detector must be very good, ideally as low as 1 mm. To achieve this, the readout cells of the FoCal-E need to be small. The FoCal-E is designed as a sampling calorimeter, the schematic of which is pictured in Fig. 13. In the transverse plane, FoCal-E consists of 2×11 modules. With each module having a transverse size of $45 \times 8\,\mathrm{cm}^2$, the whole FoCal-E spans $90 \times 90\,\mathrm{cm}^2$ in the transverse plane.

FoCal-E is formed of 20 layers, each layer consisting of a W absorber sheet followed by an active Si sensor layer. Tungsten has been chosen for the absorber as it has a small Molière radius $R_M = 9$ mm and radiation length $X_0 = 3.5$ mm. Each W sheet is of thickness 3.5 mm, the size of the radiation length X_0. The total thickness of each layer, including the tungsten sheet, silicon, supportive materials and air, is 8.5 mm. Two different Si readout technologies are used in the sensor layers: pad layers, with transverse cell size of $\approx 1\ \mathrm{cm}^2 \approx R_M^2$, and pixel layers, with cell size $\approx 30 \times 30\ \mu\mathrm{m}^2$. In the schematic in Fig. 13, the 18 pad layers and two pixel layers, positioned in the 5th and 10th layer, are shown. The positioning of the pixel layers has been chosen in order to balance the spatial separation of two showers and the energy resolution of the individual showers in the shower pair. The two pixel layers use a monolithic active pixel sensor (MAPS) design

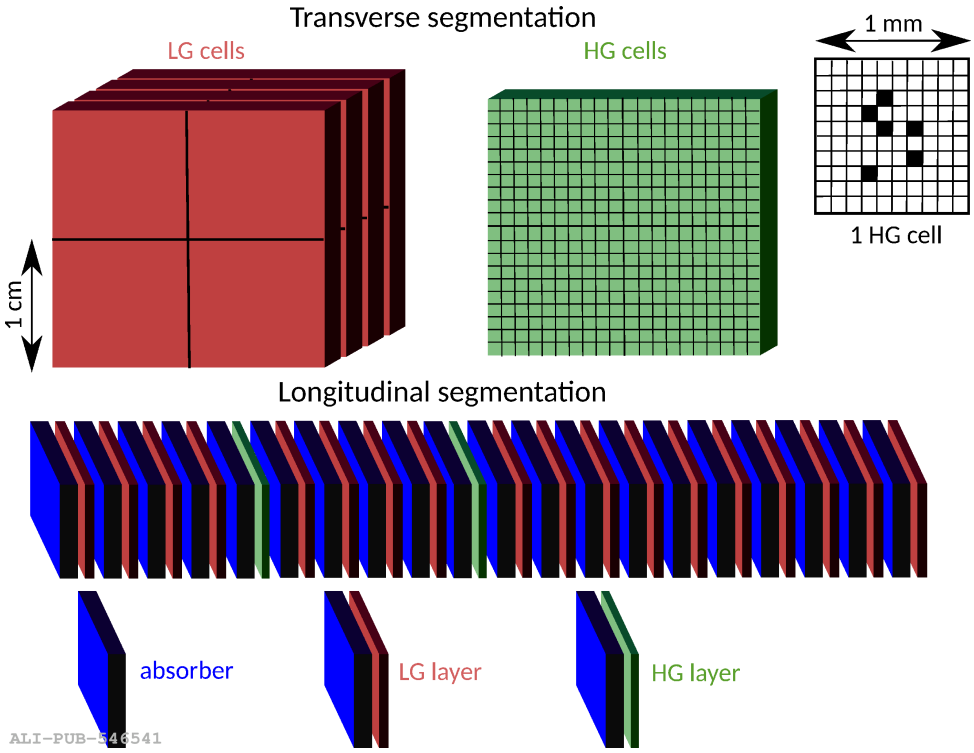


FIGURE 13 A schematic picture of the FoCal-E detector, where low-granularity (LG) pad layers are shown in red, and high-granularity (HG) pixel layers are in green. The blue layers represent the W absorber sheets. Figure from Ref. [44], reprinted under the license CC BY 4.0.

based on the ALPIDE chip [90], which was developed for use in the ALICE ITS and MFT detectors. The rest of the 18 layers of FoCal-E are composed of active layers of low-granularity silicon pad sensors [91].

The FoCal-H, the hadronic calorimeter of FoCal, is necessary for measurements of photon isolation and jets. The placement of FoCal-H is close behind FoCal-E, as shown in Fig. 14. Unlike FoCal-E, FoCal-H is not segmented longitudinally. Different implementations of the detector design have been considered and studied, such as a sandwich-type with scintillator plates, and a spaghetti-type with scintillator fibers within copper tubes. An implementation of the FoCal-H which uses copper plates with grooves enclosing scintillating fibers has also been studied at the time of writing, and is to be tested during 2025.

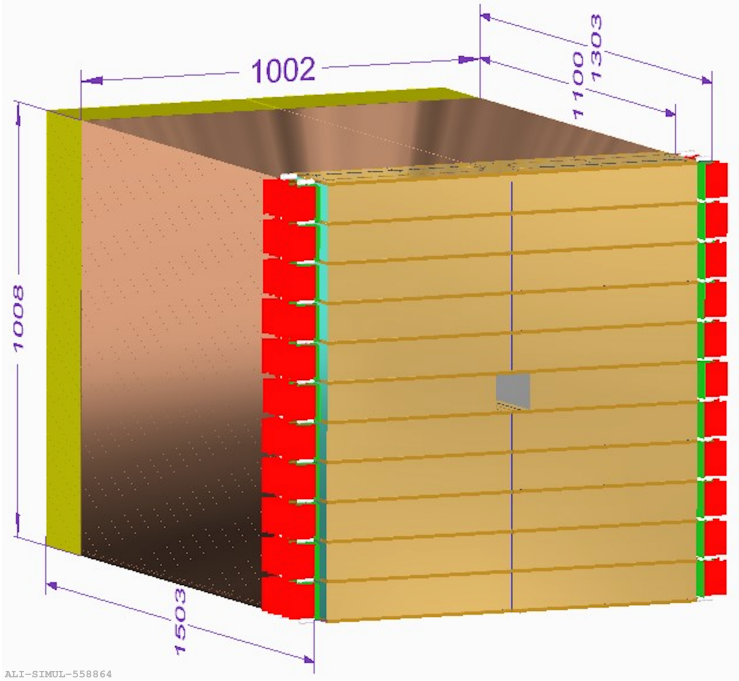


FIGURE 14 FoCal-E, in the forefront of the figure, and FoCal-H in the back. The rectangular opening in the middle is the location of the beam pipe. The readout for FoCal-E is on the sides of the detector, for FoCal-H the readout is on the back of the detector. Dimensions are in mm. Figure from Ref. [89], reprinted under the license CC BY 4.0.

3 JETS AND GLUON SATURATION IN QCD

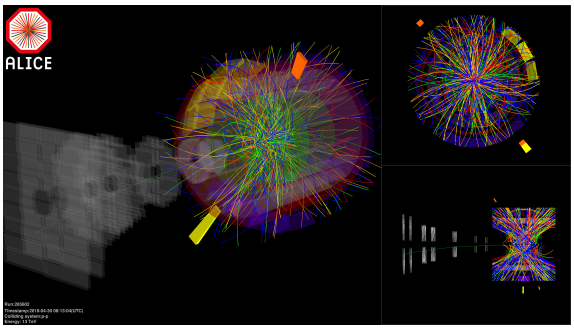
3.1 Jets

Fragmentation processes of hard partons can typically be characterized by a large momentum transfer Q, and can thus be calculated using methods of perturbative QCD (pQCD) [92]. Jets are experimental signatures of quarks and gluons, collimated sprays of particles that emerge from these hard parton fragmentation processes. Experimentally, they can be visually obvious from event displays, as shown in Fig. 15a in ALICE and Fig. 15b in ATLAS.

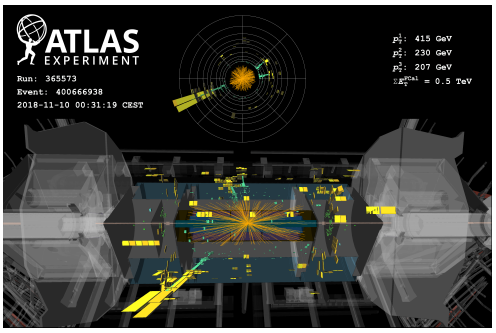
An important distinction between jet measurements in pp and heavy-ion collisions is the surrounding environment. In pp collisions, as shown in Fig. 15a, even in a high multiplicity event the jet signals appear localized, as there is not that much underlying activity outside of the two signals. In contrast, when looking at the event display of a Pb–Pb collision, pictured in Fig. 16, the large multiplicity of soft background particles is clearly visible. This background has to be reliably understood and subtracted from the event, so that the signals can be found and studied. This will be further discussed in Sec. 4.3.

In order to interpret the results of jet measurements and differentiate the effects arising from the existence of a QGP medium in Pb–Pb collisions, we must first perform a baseline measurement in pp collisions. Measurements of jets in pp collisions performed with ALICE and other experiments have demonstrated that jets are well-understood and consistent with pQCD expectations [96].

In heavy-ion collisions, jets are produced in the earliest phase of the collision. The high-energy partons scattered early on in the collision have to pass through a significant amount of QGP as they travel from inside the QGP out and toward the detectors. The partons within the jet interact with the QGP medium, causing parton energy loss [97], more commonly called jet quenching. This phenomenon was first observed at the BNL RHIC [98, 99] and afterwards at the CERN LHC [100, 101, 39, 102]. By comparing high-p_{T} charged particle and jet yields in heavy-ion collisions to those in proton-proton reference collisions, a suppression in the heavy-ion collision yields is seen, which is attributed to jet



(a) A pp collision at 13 TeV center-of-mass energy as recorded by ALICE in Run 2. Reconstructed trajectories of charged particles produced in the collision are visible, as well as energy deposited in the photon spectrometer and calorimeter. Many simultaneous collisions, known as pile-up, cause the high multiplicity. Figure from Ref. [93], © CERN, used for informational purposes only.



(b) A high-p_T, three-jet event is recorded by ATLAS during the Run 2 Pb–Pb run at 5.02 TeV center-of-mass energy. The leading 415 GeV jet is azimuthally balanced by the two smaller sub-leading jets (230 and 207 GeV). Figure from Ref. [94], © CERN, used for informational purposes only.

FIGURE 15 Event displays portraying a pp collision in ALICE (left) and a Pb–Pb collision in ATLAS (right).

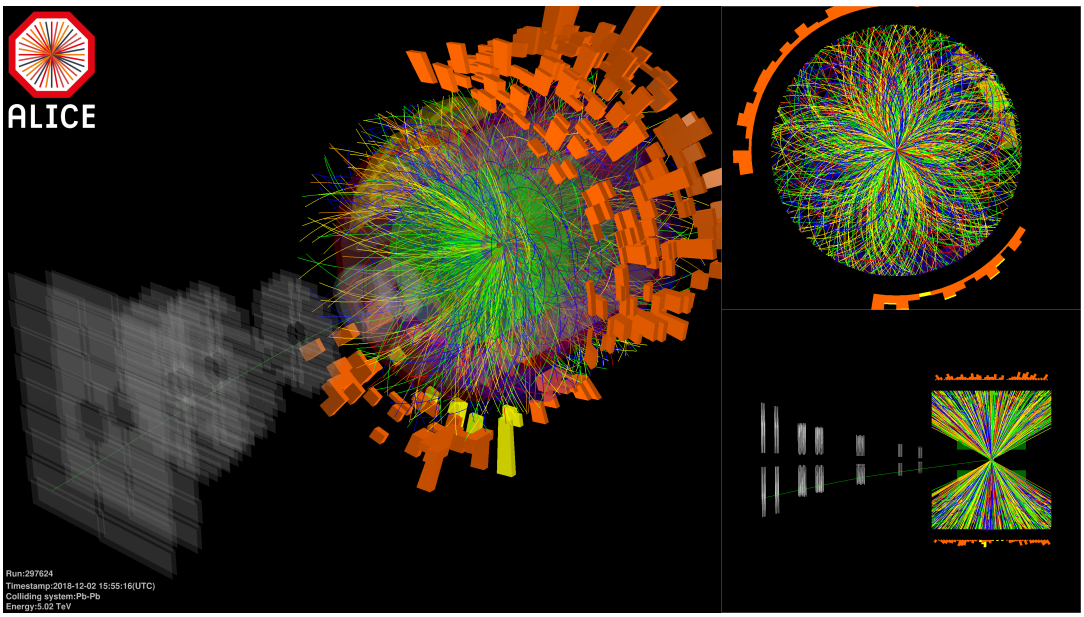


FIGURE 16 An event display showing the collision of lead ions with energy 5.02 TeV per nucleon pair recorded by ALICE in Run 2. The colored lines represent the charged tracks, and the orange and yellow towers represent energy measured by the electromagnetic calorimeters and the photon spectrometer. The green line in the forward region corresponds to a muon traversing the muon spectrometer. Figure from Ref. [95], © CERN, used for informational purposes only.

quenching.

While traversing the QGP medium, the structure of the jet is also modified. There have been a wide variety of measurements performed to study the QGP via jet structure. The modification of the hard splitting inside the jets has been studied by grooming away soft, wide-angle radiation from the edges of the jet cone using a soft drop grooming algorithm [103], isolating the hard core of the jet. This offers a way to compare narrow and broad jets modified by the medium. It has been found [104, 105] that wider jets are more suppressed in heavy-ion collisions, showcased by an enriched distribution of narrower jets in a fixed p_{T} range in Pb–Pb compared to pp collisions.

The quark/gluon content fraction of jets can also play a role in their interaction with the medium. Gluon-enriched jets with a color charge of $C_g = 3$ are expected to lose more energy than quark-enriched jets with a color charge of $C_q = 4/3$, as the larger color charge implies more interaction with the QCD medium. A wide range of tools exist [106] to discriminate between the two jet types. One category of these tools focuses on the jet shape, namely on angularities [107], which are a measure of QCD radiation around the jet axis, and energy-correlation functions [108], which identify the jet substructure. Also, heavy-quark jets are expected to lose energy differently in the medium, since their large masses imply that they are produced predominantly in the initial hard-scattering process, and therefore experience the full QGP evolution. This makes heavy-quark jets effective probes of the QGP, and an interesting observable studied in ALICE [109].

While Pb–Pb collisions are shown to exhibit jet quenching, the question remains whether such modifications can be seen in collisions of protons and lead ions, p–Pb collisions. An important observable in the study of jet energy loss in the dense nuclear medium in nucleus-nucleus collisions is the nuclear modification factor, R_{AA}. For p–Pb collisions, a similar factor, R_{pPb}, can be defined. Experimentally, the factor R_{pPb} is defined as the ratio between yields in collisions of lead nuclei with protons and corresponding yields in pp collisions:

$$R_{\mathrm{pPb}} = \frac{1}{\langle N_{\mathrm{coll}} \rangle} \frac{\mathrm{d}^2 N_{\mathrm{pPb}} / \mathrm{d}p_{\mathrm{T}} \mathrm{d}y}{\mathrm{d}^2 N_{\mathrm{pp}} / \mathrm{d}p_{\mathrm{T}} \mathrm{d}y}. \tag{5}$$

Here, N_{pPb} and N_{pp} represent the yields in p–Pb and pp collisions, and $\langle N_{\mathrm{coll}} \rangle$ is the average number of binary nucleon-nucleon collisions.

Cold nuclear matter (CNM) effects refer to any observed differences between pp and pA, or deuteron-A (dA), collisions in which a hot QCD medium is not expected. The presence of a nucleus in the initial state of the collision can still influence the production of a final state observable, even without a QGP medium. Possible reasons for CNM effects include coherent multiple scattering within the nucleus [110], gluon shadowing [111], and parton energy loss inside the nucleus [112]. The study of these effects is an interesting study in its own right, while also providing a basis for the interpretation of heavy-ion collisions. It has been indicated by studies at forward rapidities that, for instance, heavy flavor [41, 42] and J/ψ [43] are both suppressed in CNM.

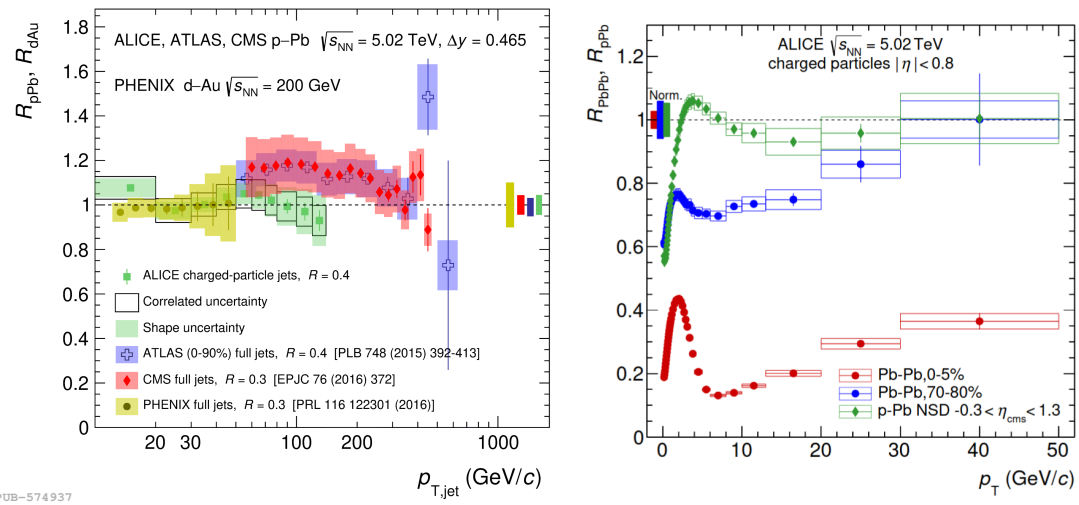


FIGURE 17 Left: A comparison of the nuclear modification factors of jets in p–Pb and d–Au measurements at the LHC and RHIC. From Ref. [113], reprinted under the license CC BY 4.0. Right: Nuclear modification factors measured by ALICE in central and peripheral Pb–Pb collisions, and in p–Pb collisions. From Ref. [40], reprinted under the license CC BY 4.0.

Studies of jet production in p–Pb collisions are needed in order to quantify the effects from the CNM and decouple them from effects observed in Pb–Pb collision data arising from the QGP medium interactions. The R_{pPb} for jet production in pp and p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV [113] has been recently measured in ALICE. The result of this calculation is shown in Fig. 17 (left), along with measurements of R_{pPb} and R_{dAu} by other experiments. With the precision of the measurement in Ref. [113], the R_{pPb} is observed to be consistent with one, implying that any nuclear effects on single charged jet production in p–Pb collisions are below the resolution of the measurement.

For charged hadrons, a similar measurement has been performed at the same energy for pp, p–Pb and Pb–Pb collisions [40] and the resulting nuclear modification factors are shown in Fig. 17 (right). As is visible from the red markers signifying the most central Pb–Pb collisions, there is a notable suppression of the charged hadron yield in central heavy-ion collisions. The suppression becomes weaker in less-central Pb–Pb collisions, shown in blue markers, and even weaker in p–Pb collisions, shown in green. For the intermediate p_{T} range below 10 GeV/c, there is a slight enhancement of the p–Pb ratio. This feature is known as the Cronin peak [114].

It has been noted in Ref. [35] that, based on current evidence from p–Pb and d–Au collisions, pp collisions can function as an appropriate reference for heavy-ion jets. However, as there have been some CNM effects found for jets, the suggestion is for each heavy-ion collision observable to be measured also in cold nuclear matter, in order to properly interpret the heavy-ion event with hot nuclear matter. Therefore, for hard processes in Pb–Pb collisions, p–Pb collisions can function as appropriate reference systems [35]. Some care must, however, be taken when working at large rapidities and multiplicities.

As the ALICE detector does not have a hadronic calorimeter to provide full jet measurements around the full acceptance, the jet analysis in ALICE is largely focused on jets formed from charged tracks. There have been some studies done on full jets in ALICE, using calorimeter cluster information from the EMCal [63]. In Run 4, once FoCal has been installed in ALICE, the measurement of full jets in the forward direction will be possible with the combination of the FoCal electromagnetic and hadronic calorimeters.

3.2 Dijets

In leading order of pQCD, dijets originate from $2 \rightarrow 2$ processes. Due to momentum conservation, the two outgoing partons from the hard interaction of this type of process have transverse momenta that are opposite and equal in magnitude. Higher order processes break this momentum balance, including hard $2 \rightarrow 3$ scatterings [115], the QCD evolution of the incoming partons [116, 117, 118, 119] and multiple interactions [120]. Longitudinal momenta, however, are not necessarily balanced, due to the longitudinal momentum fractions $x_{1,2}$ that the incoming partons carry from the colliding protons not generally being equal to each other. Experimentally, a cut is usually applied for the magnitude of the transverse momenta of the outgoing jets, while simultaneously requiring that the two observed jets are in opposite hemispheres of the transverse plane.

As in single jet measurements, in heavy-ion collisions the hard-scattered partons within a dijet system of two jets interact with the QGP medium. The energy loss of the two jets thus contains information about the medium. When we observe a single-jet observable, it can have been created anywhere in the medium. We cannot know exactly whether the jet has been created close the QGP fireball edge at which the measurement is performed, or on the opposite edge, or somewhere in between. If the jet has had to pass through the whole fireball, it has been quenched more by the QGP medium than jets closer to the edge of the measurement-side of the fireball, possibly quenched to the extent that it cannot be identified. The consequence of this effect is called a surface bias, describing how the hard partons captured by jet measurements are more likely to have been created near the surface of the QGP [121, 122]. There are some model indications [123] that a dijet measurement can better probe the center of the fireball than the single-jet measurement. By requiring two jets to be measured simultaneously, we include events where either both jets have been created close to the center of the fireball and thus travel a similar length of the medium, or where the other jet has enough energy to pass through the medium from the other edge and still be identified.

A well-known CNM effect is the suppression of partons at small momentum fractions, x, described through parton distribution functions (PDFs). Nuclear PDFs (nPDFs) can be constrained using dijet production in p–Pb collisions [36]. Also, jet imbalance has been studied in pp and p–Pb collisions in order to search

for the onset of the creation of QGP. Previous results from ATLAS [37, 38] and CMS [39] show that, in non-central lead-lead collisions, observables measuring the energy imbalance between the two jets forming a dijet system converge towards measurements from proton-proton collisions. This means that no collective effects are visible in the non-central collisions via these observables. However, when the same observables were studied in central lead-lead collisions, significant changes attributed to the QGP medium are visible. Based on these measurements, there is a clear increase of momentum imbalance in dijet systems in central heavy-ion collisions compared to pp collisions.

The invariant mass of the two jets which form the dijet system, here referred to as the dijet mass, provides a Lorentz-invariant measure of the hardness, or the scale Q, of the hard interaction. Dijet mass has been studied at very high masses in ATLAS [124] and CMS [125], motivated by looking for signatures of physics beyond the Standard Model, which an unexpected resonance peak would signify [126, 127]. ALICE can push dijet mass measurements to a lower mass range, corresponding to a low transverse momenta of the jets. At lower jet momenta in heavy-ion collisions, the medium modification can be expected to be more prominent.

In order for a dijet mass measurement in Pb–Pb to be possible, we first need to analyze and compare the dijet mass distribution in pp and p–Pb collisions. This allows us to distinguish any effects from the cold nuclear matter on the Pb–Pb dijet mass distribution.

3.3 Collinear factorization

When the momentum transfer Q^2 is large, the strong coupling $\alpha_s(Q^2)$ becomes small. In this regime, pQCD can be applied, and hadron–hadron collisions are described within the quark–parton model. For hard processes, where Q^2 is sufficiently high, reliable QCD predictions of cross sections can be obtained. At the parton level, a hadron-hadron collision can be viewed as an interaction between a parton with momentum fraction x_a and another parton with a momentum fraction x_b. The probability of finding a parton with a given momentum fraction x inside the hadron is described by a PDF, defined separately for the quark and the gluon.

If we want to study inclusive hadron production with the collinear framework and pQCD, we can convolute the production of a parton c with the corresponding fragmentation function of that parton into a hadron h. In this manner, we obtain the inclusive cross section for hadron production in a pp collision [128]:

$$\begin{aligned} \mathrm{d}\sigma^{\mathrm{pp}\to hX} = \sum_{a,b,c} f_a(x_a, Q^2) \otimes f_b(x_b, Q^2) \otimes \mathrm{d}\hat{\sigma}^{ab\to cX}(x_a, x_b, Q^2, \alpha_s(\mu^2)) \\ \otimes D_{c\to h}(z, \mu_F^2). \end{aligned} \tag{6}$$

In Eq. (6), the factorization scale is denoted by Q, μ denotes the renormalization

scale, and μ_F denotes the fragmentation scale. The partonic cross section $\hat{\sigma}^{ab\to cX}$ is expressed as a power expansion in the strong coupling constant α_s. $D_{c\to h}(z, \mu_F^2)$ denotes the fragmentation function of a hadron h from the parton c with momentum fraction $z = p_h/p_c$.

For these types of processes, Fig. 18 shows a general hadron-hadron collision, with labels a, b referring to the incoming particles participating in the reaction, c an outgoing particle, h the outgoing hadron, and X referring to any final state of particles.

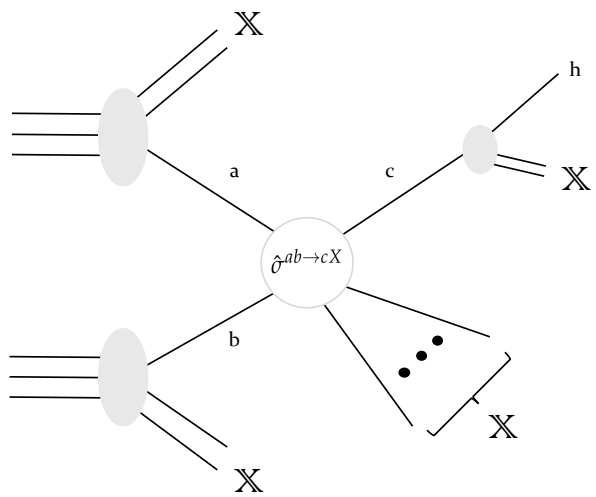


FIGURE 18 The quark-parton model for a hadron-hadron collision into a hadron h and any other final state X.

The collinear factorization presented in Eq. (6) is defined for pp collisions. We can modify the equation to include nuclear effects, in order to calculate cross sections for nucleus collisions. The equation for collisions of nuclei A and B is [34]

$$\begin{aligned} \mathrm{d}\sigma^{AB\to hX} = \sum_{a,b,c} f_{a/A}(x_a, Q^2) \otimes f_{b/B}(x_b, Q^2) \otimes \mathrm{d}\hat{\sigma}^{ab\to cX}(x_a, x_b, Q^2, \alpha_s(\mu^2)) \\ \otimes \tilde{D}_{c\to h}(z, \mu_F^2). \end{aligned} \tag{7}$$

In Eq. (7), the PDFs have been altered to nuclear PDFs (nPDFs) which account for the effects of the nucleus on the PDF. Also, the fragmentation function has been modified to take into account the fact that medium interactions are able to alter the properties of the hard parton c. The medium-modified fragmentation function is approximated by convoluting the vacuum fragmentation function with the parton-medium interaction function $P_{j\to j'}$ as [34]

$$\tilde{D}_{c\to h}(z, \mu_F^2) \approx \sum_{j'} P_{j\to j'}(p_{j'}|p_j) \otimes D_{c\to h}(z_{j'}, \mu_F^2). \tag{8}$$

The partonic cross sections in the above equations involve large momentum transfers, which makes them calculable with pQCD. In contrast, the fragmentation functions describe the hadronization of partons at low momentum scales

where perturbation theory breaks down, and therefore they cannot be computed from first principles in pQCD. Thus, they are determined phenomenologically, most commonly by fitting to experimental data. This is done by using a global analysis of data from e^+e^- collisions [129, 130]. According to QCD factorization, the fragmentation functions are process-independent, thus, they can be used to predict inclusive single hadron cross sections in a variety of other processes beyond e^+e^- collisions. In order to obtain the medium-modified fragmentation functions necessary for Eq. (8), several models have been developed to model the parton-medium interaction, such as solving medium evolution equations or via Monte Carlo studies [131].

The theoretical nuclear modification factor of the full nuclear PDFs is defined as the ratio of the full nuclear nPDF, $f_i^A(x, Q^2)$, to the free proton PDF, $f_i^p(x, Q^2)$ [132]. Both can be written as a linear combination depending on the number of protons Z and neutrons N, which results in the ratio

$$R_i^A(x, Q^2) = \frac{Z f_i^{\mathrm{p}/A}(x, Q^2) + N f_i^{\mathrm{n}/A}(x, Q^2)}{Z f_i^{\mathrm{p}}(x, Q^2) + N f_i^{\mathrm{n}}(x, Q^2)}. \tag{9}$$

Here, i denotes the parton flavor, and A the mass number of the nucleus.

PDFs and nPDFs are obtained from fits to experimental data, as they are not calculable from pQCD directly. A wide set of available data will provide a better fit. PDF sets provide the proton baseline. The main constraints for proton PDFs come from fits to deep inelastic scattering (DIS) data, most notably data measured in HERA [133]. In Fig. 19, the PDFs for the valence (u_v and d_v) quarks, sea (S) quarks and gluons (g) from the HERAPDF2.0 NLO [133] are shown at the scale of 10 GeV/c^2. The experimental, model and parametrization uncertainties are shown separately. The valence quark distributions peak around $x = 0.2$, but as we head toward lower values of x, the gluon and sea quark distributions begin to rise rapidly. This means that when we probe the proton at a lower value of x, we are mostly probing gluons. There are multiple other parameterizations for the free proton PDF, for instance NNPDF [134] and the CTEQ-TEA [135] PDFs.

For the calculation of nPDFs, a widely used fit is EPPS21 [132], which includes more data from p–Pb collisions at the LHC than the previous EPPS16 [136] fit. The new data has led to better-constrained gluon distributions at lower values of momentum fraction x. This provides access to study the nuclear modification factor at varying values of x. The average-nucleon nuclear modification for gluons in lead, R_g^{Pb}, is shown in Fig. 20 (left). In the figure, a suppression is seen at lower values of x. This effect is called nuclear shadowing. The slight enhancement seen around $x = 0.1$ is called anti-shadowing. The proton baseline for EPPS21 is the set CT18ANLO [135]. There are also other models that fit nPDFs, for instance, nCTEQ15 [137] and nNNPDF [138]. The comparison of the R_g^{Pb} calculated from these three fits is shown in Fig. 20 (right).

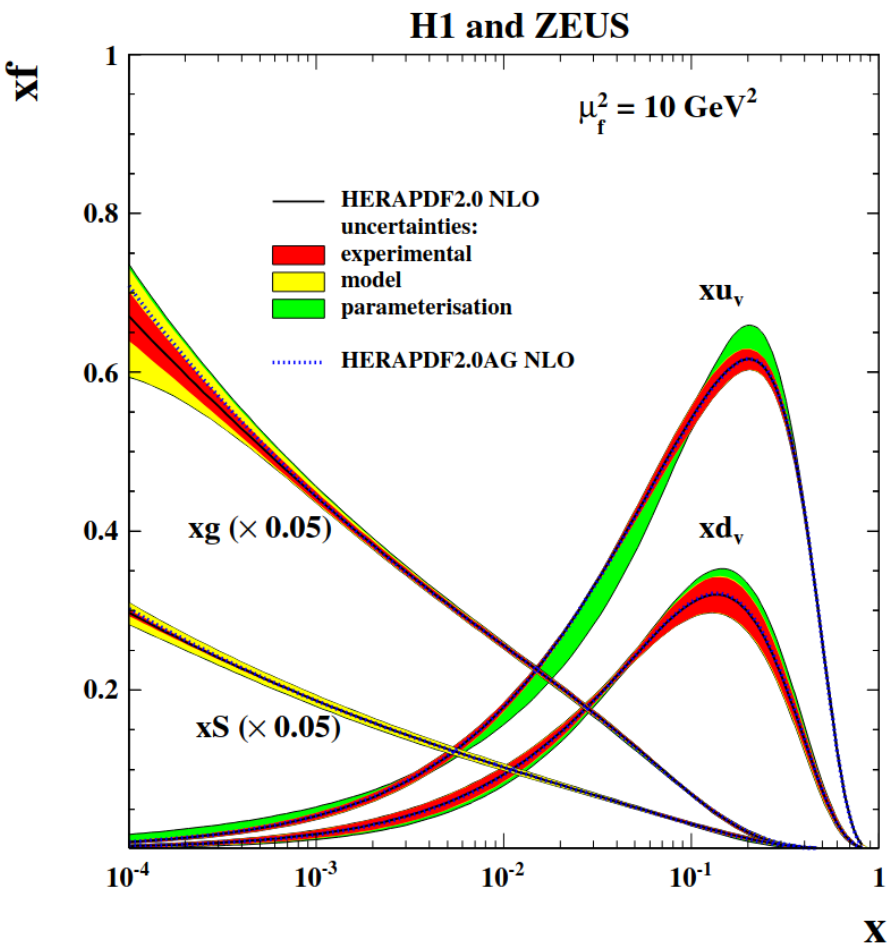


FIGURE 19 The PDFs xu_v, xd_v, xS and xg of HERAPDF2.0 NLO at $\mu_f^2 = 10$ GeV^2. The gluon and sea distributions are scaled down by a factor 20. From Ref. [133], reprinted under the license CC BY 4.0.

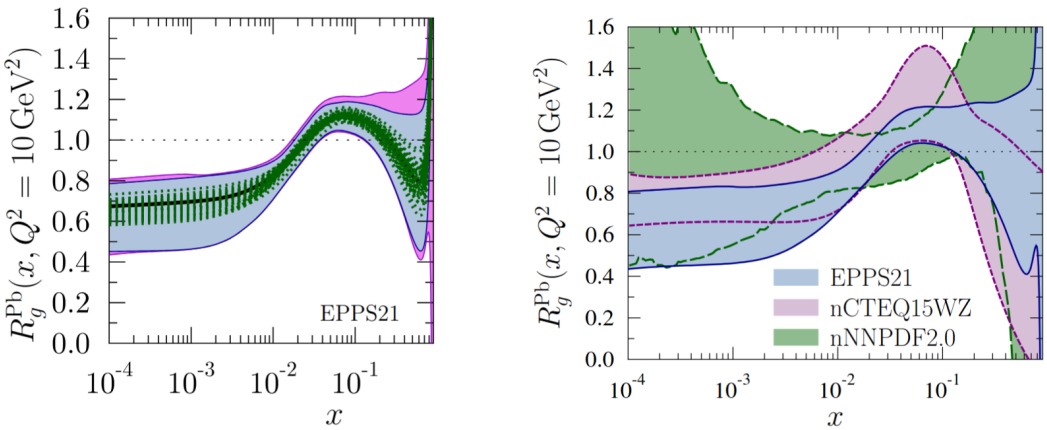


FIGURE 20 Left: The average-nucleon nuclear modification for lead at the parametrization scale $Q^2 = 10$ GeV^2. The central results are denoted by black curves, and the nuclear error sets by green curves. The blue bands denote nuclear errors, the purple bands the full uncertainty including uncertainties from free proton PDFs. Right: A comparison of the EPPS21 [132] (blue), nCTEQ15WZ [139] (purple) and nNNPDF2.0 [138] (green) R_g^{Pb}. Figures from Ref. [132], reprinted under the license CC BY 4.0.

3.4 Low-x regime

While the perturbative regime of QCD is well understood, describing the interactions at short distances and high momentum transfer Q successfully, the non-perturbative behavior of QCD at long-range and small momentum-transfer is not as well understood [44]. The non-perturbative physics is absorbed into the PDFs, which are determined from experimental data, as described in the previous section. The PDFs are commonly determined from global fits to data that cover a wide range of momentum fraction x and momentum transfer scale Q.

In comparison to DIS, hadronic collisions do not provide direct access to the kinematics of scattered partons. Therefore, for these measurements, the momentum fraction x probed is approximated for LO $2 \to 2$ processes as

$$x_{1,2} \approx \frac{2p_{\mathrm{T}}}{\sqrt{s}} e^{\pm y}, \tag{10}$$

where $x_{1,2}$ depends on the transverse momentum p_{T} of the emitted parton, rapidity y, and the center-of-mass energy of the collision, $\sqrt{s}$. We denote the negative-rapidity case x_2 by x. We can reach lower values of x at large rapidities and low p_{T}.

In Fig. 21, the kinematic coverage of various experiments probing hadronic structure are presented. The coverage is shown separately for electromagnetic (EM) probes and DIS measurements on the left, and hadronic and ultra-peripheral collisions (UPCs) on the right. As seen, the FoCal is expected to extend measurements down to lower values of x than before.

As can be seen from Fig. 21, the full range of (x, Q^2) has not yet been probed, so the areas not covered by measurements must be extrapolated into based on the

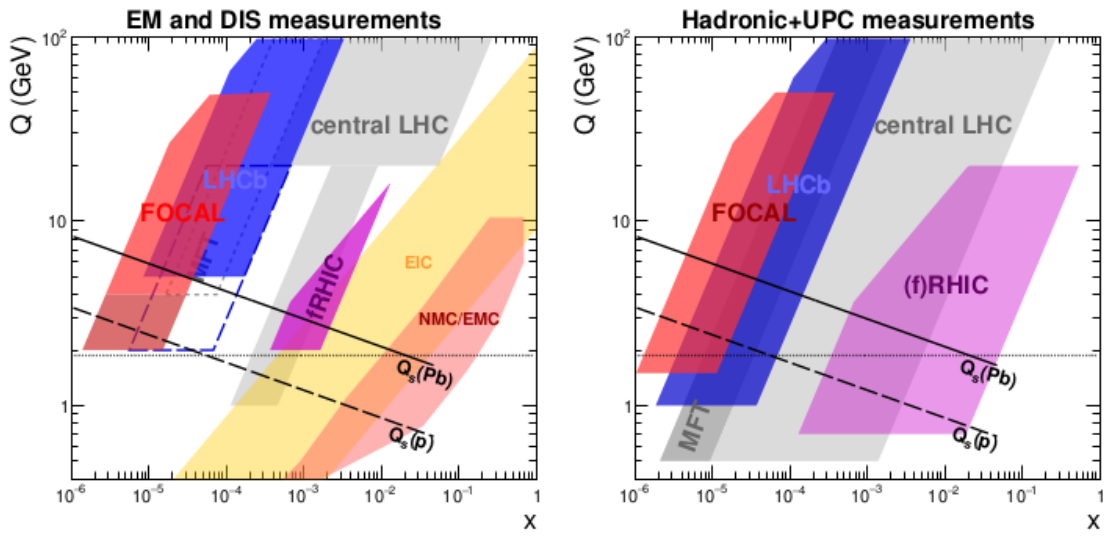


FIGURE 21 The momentum fraction x and momentum transfer scale Q of partons probed with various measurements within the acceptance of current and proposed experiments. Left: EM probes and DIS measurements. Right: hadronic and UPC measurements. The horizontal line describes the cut above which data is commonly included in PDF determination. Figures from Ref. [44], reprinted under the license CC BY 4.0.

available measurements. This is done using evolution equations, which are visualized in Fig. 22. The DGLAP [140, 141, 142] evolution equations are valid in the region where Q^2 and x are moderate to large, where parton densities are small enough. At moderate Q^2, the gluon density increases as x becomes smaller. The BFKL equations [143] describe this stage of evolution. At small enough x, the abundance of soft gluons resulting from increased gluon splitting leads to very high parton densities. This in turn increases the occurrence of parton recombination, notably gluon recombination. Once the saturation region is approached, the non-linear effects arising from parton recombination start to become significant, and the linear DGLAP and BFKL equations do not describe the evolution anymore. The non-linear evolution in this region can be described by the JIMWLK [144] equations, or approximated by the BK [145] equations. Once the non-linear effects limit the growth of the PDFs, we reach a sort of equilibrium, which is called gluon saturation.

The dashed and full line in Fig. 21 mark the estimated saturation lines for the proton and lead. At a given x, the saturation scale is approximated by

$$Q^2_{\text{sat}} \approx \frac{xg(x, Q^2)A}{\pi R_A^2} \propto A^{1/3} x^{-\lambda}, \tag{11}$$

where $g(x, Q^2)$ is the gluon PDF in a proton, R_A is the nuclear radius, A the mass number of the nucleus, and the exponent $\lambda \approx 0.3$ [44, 147]. At small values of x and for large nuclei like Pb, the saturation scale increases, corresponding to an increase in gluon density. Various models describe the gluon saturation regime. A prominent model is the Color Glass Condensate (CGC) [148, 149], in which the

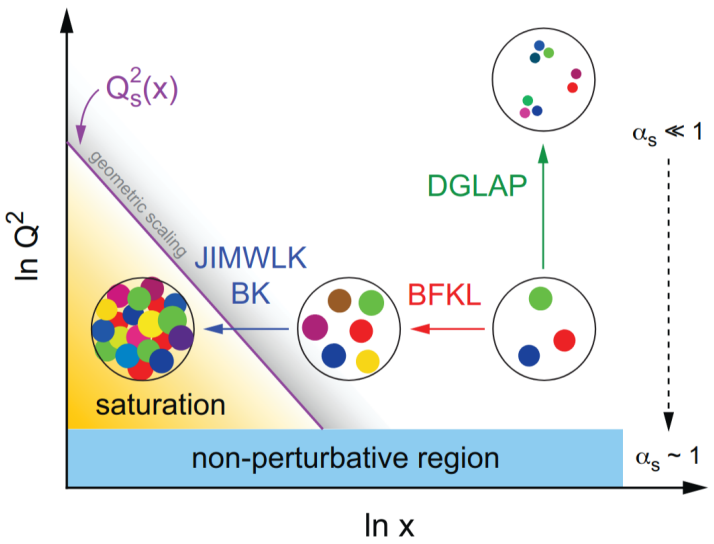


FIGURE 22 A schematic of the resolution probed with varying x and Q^2, and the relevant evolution equations in their regions of validity. Figure from Ref. [146], reprinted under the license CC BY 4.0.

target hadron can be described with classical color fields, due to the high density of low-x, soft gluons.

Currently, the most precise measurement at forward rapidity to probe the small-x region at the LHC is the prompt D-meson measurement by LHCb, at rapidity $2.5 < y < 4.0$ [150]. Including the D-meson data in nuclear PDF determination reduces the uncertainties by a factor of two [151]. However, hadronic final state effects, mainly rescattering, can affect the observations, complicating the interpretation of the measurements and the determination of the amount of gluon shadowing.

With the FoCal, measurements of isolated photons can provide excellent sensitivity to gluon distributions at small x. Isolated prompt photons, born at LO in the parton interaction vertex, are not affected by final state effects. It has been shown in Ref. [2] that, with the use of isolation cuts, the isolated photons arising from quark-gluon Compton scattering dominate the measured photon population. Therefore, the inclusion of FoCal measurements of isolated prompt photons in global PDF fits will provide a probe of the low-x region without uncertainties arising from final state hadronization.

Figure 23 shows the nuclear modification factor R_{pPb} for inclusive prompt photons at energy $\sqrt{s_{\mathrm{NN}}} = 8.8$ TeV [2]. FoCal pseudo-data, in black markers, has been simulated in pp and p–Pb collisions using INCNLO [152, 153, 154, 155, 156] calculations at NLO. The distribution of the pseudo-data is compared with NLO calculations using the nNNPDF3.0 [157] PDF set. The INCNLO calculation is performed incorporating nNNPDF3.0 without a constraint from the LHCb D-mesons. The distribution and corresponding uncertainty prior to any reweighting are shown by the gray line and band. The prediction is then reweighted using Bayesian inference by either incorporating the LHCb D-mesons or the FoCal pseudo-data. The reweighted results with LHCb D-mesons are shown in blue,

and the results reweighted with the FoCal pseudo-data in red. The inclusion of data, both FoCal and LHCb, decreases the nPDF uncertainties by about or more than 50%. Including both measurements will provide a global and complementary constraint to the PDFs.

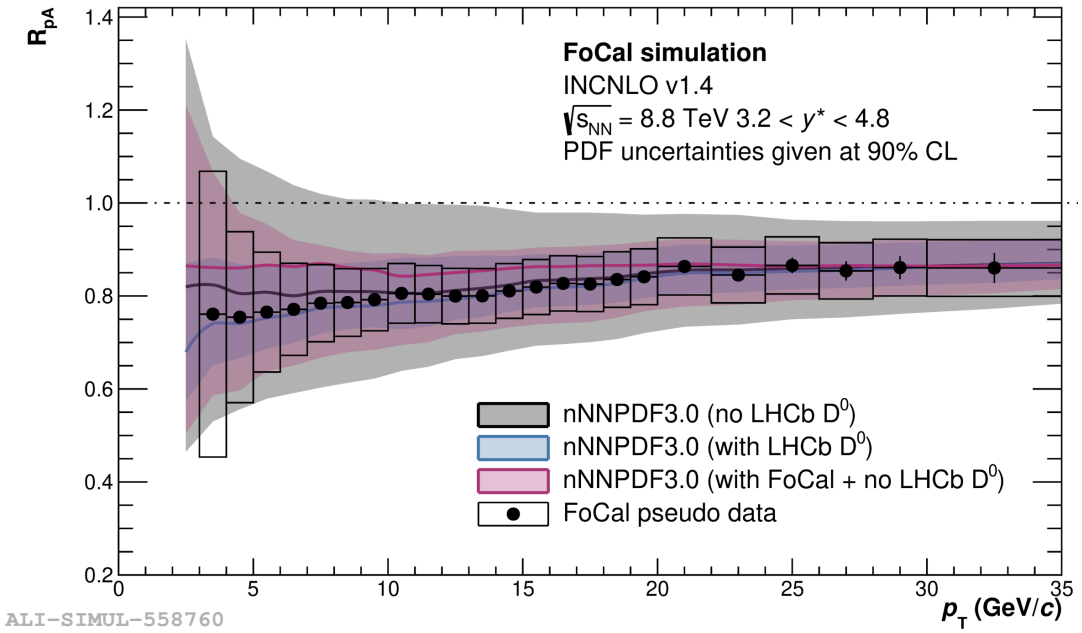


FIGURE 23 Distribution of R_{pA} for inclusive prompt photons at $\sqrt{s_{NN}} = 8.8$ TeV for FoCal pseudo-data (black markers) based on NLO calculations using the INCNLO program. The prediction for nNNPDF3.0 before any reweighting is shown in gray. In red is the resulting posterior distribution and its uncertainty after reweighting with the FoCal pseudo-data. The distribution constrained by LHCb D-meson measurements are shown in blue. From Ref. [2], reprinted under the license CC BY 4.0.

The measurement of isolated photons is one of the major goals of the FoCal upgrade. In addition, photon and hadron-triggered correlations, photon-jet and dijet correlations, and inclusive jet measurements are observables to be measured by FoCal. Inclusive jet production in the forward region in p–Pb collisions has been calculated to NLO with the CGC framework [158, 159, 160]. It is noted [158], however, that while single forward jet production in pA collisions is an ideal process for observing saturation phenomena, the saturation effects are expected to be small for high p_T jets. Therefore, the experimental challenge is to reliably measure jets with a low p_T of the order of a few times Q_{sat}. Also, production of J/ψ and $\psi(2S)$ in UPCs are to be studied with the FoCal.

Dijets at forward rapidity may have an enhanced sensitivity to low-x saturation effects compared to inclusive jet measurements. This is due to the fact that there exist three momentum scales in the forward dijet process: Q_{sat} characterizing the saturation scale of the target at small x, the individual jet p_T^{jet}, and the momentum imbalance k_T of the dijet pair. The individual jet momenta are much larger than Q_{sat}, while the magnitude of k_T ranges from order of Q_{sat} to values much higher. When $Q_{sat} \ll k_T$, the small-x QCD dynamics in the process are

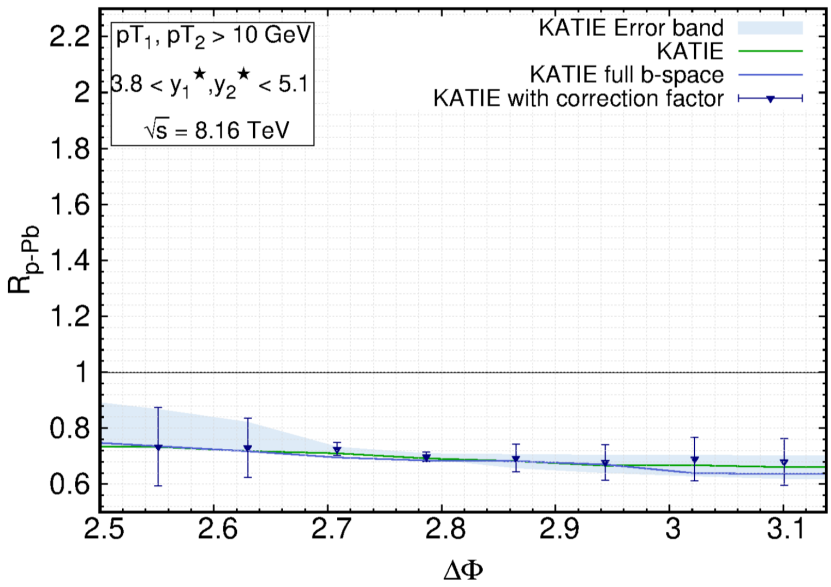


FIGURE 24 Nuclear modification ratio R_{pPb} for dijets as a function of the azimuthal angle $\Delta\phi$ in the FoCal kinematics, obtained from the KaTie MC [163] multiplied by the non-perturbative correction factors from PYTHIA. The error bars show the cumulative uncertainty from the factorization scale and correction factors. The blue error band represents the uncertainty due to the variation of the factorization scale. Figure from Ref. [163], reprinted under the license CC BY 4.0.

linear, while when $Q_{\mathrm{sat}} \sim k_{\mathrm{T}}$, non-linear effects are involved. By probing varying values of k_{T}, the region between the small (x, Q^2) limit of a transverse momentum dependent (TMD) factorization framework and the high-Q limit of high-energy factorization can be studied [161].

In Fig. 24, a calculation of the nuclear modification ratio R_{pPb} for dijet events in the FoCal kinematic range at collision energies of $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV is shown as a function of the azimuthal angle $\Delta\phi$ between the jets. The calculation has been performed using a KaTie MC [162, 163, 164] code. KaTie is based on the improved TMD (iTMD) factorization, together with Sudakov resummation and non-perturbative corrections due to hadronization and showers taken into account using PYTHIA. In Ref. [163], R_{pPb} was also studied in the kinematic range of the FCal calorimeter [165] in ATLAS, where the rapidity range for jets is $2.7 < y < 4.0$. FoCal will expand these measurements to a more forward rapidity and to lower values of jet p_{T}.

Figure 25 shows the R_{pPb} for two channels: photon-jet coincidence events and dijet events. A suppression of 30% is seen in the back-to-back configuration, corresponding to low k_{T}, for both channels when $p_{\mathrm{T}} > 10$ GeV$/c$. The suppression weakens when moving to larger k_{T}, away from the back-to-back configuration.

A CMS measurement [167] of the momentum imbalance and η-dependence of the yield in forward dijets in p–Pb collisions has been incorporated in the EPPS21 [132] nPDF fit, and has been found to have significant constraining power

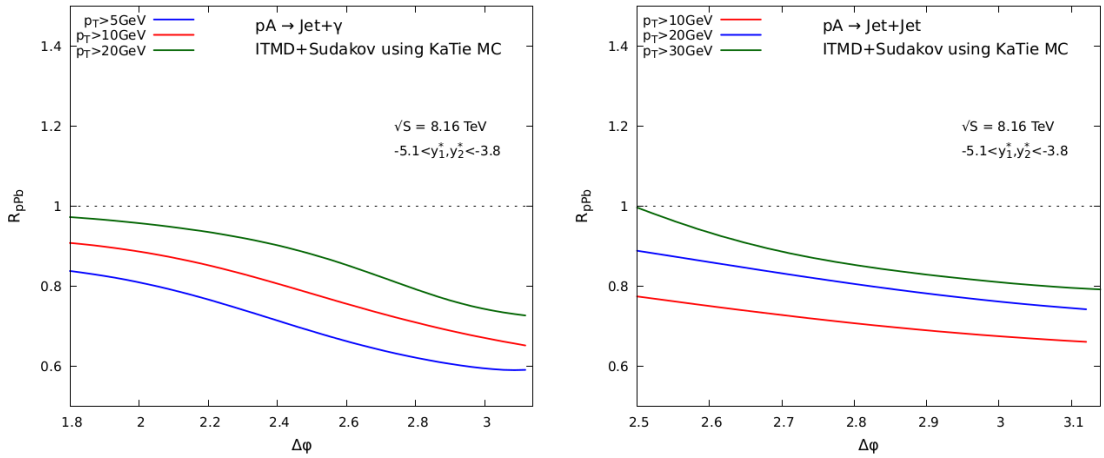


FIGURE 25 The $R_{\mathrm{p-Pb}}$ as a function of the azimuthal angle $\Delta\phi$ for various p_{T} thresholds. Left: γ+jet channel. Right: dijet channel. Figure from Ref. [166], reprinted under the license CC BY 4.0.

in the fits. While the CMS measurement overlaps with the FoCal in acceptance ($4 < \eta_{\mathrm{jet}} < 5.2$), the jet p_{T} values probed are higher than that expected of FoCal, with leading jet $p_{\mathrm{T}} > 120$ GeV/c, and subleading jet $p_{\mathrm{T}} > 30$ GeV/c. The high p_{T} values make a measurement of the dijet imbalance at the lower range of $Q_{\mathrm{sat}} \sim$ few GeV/c challenging. The FoCal measurements of forward jet and dijet measurements at lower values of p_{T} will complement and extend these measurements in the future.

To investigate low-p_{T} jets, dijets and their correlations with FoCal, it is first necessary to characterize the jet performance of the detector, in particular, the p_{T} range that can be reliably probed. This work, carried out by the author, investigates the FoCal jet measurement performance, focusing on the jet energy resolution and scale for jets down to $p_{\mathrm{T}} = 5$ GeV. The results of this work are presented in Sec. 6.

4 JET METHODS

4.1 Jet definitions

In order to be able to compare jet measurements between different experiments, simulations and theory calculations, a standardization of jet definitions and their reconstruction algorithms is necessary. A jet reconstruction algorithm starts by taking in a list of particles, each associated with a four-momentum p_i^μ. To create jets, we need to construct a subset of the input list of particles that are close to each other in phase space, and combine their momenta in order to obtain the jet momentum. The chosen jet algorithm defines how the particles are selected from the list, and the chosen recombination scheme defines how the momenta of the particles are combined.

There exist several recombination schemes for merging of the four-momenta of particles during the clustering procedure. The simplest procedure simply adds the four-vectors together – this is called the E-scheme, and it has been advocated in Ref. [168] as a standard procedure, and is the default option in FastJet [169]. The E-scheme retains the masses of the particles and conserves energy and momentum.

Calculating the energy of a measured particle requires knowledge of its mass. In experiments like ALICE, particle identification is limited to a finite p_T range and cannot be perfectly accurate. In LHC experiments, a common approach [170] is to assign the mass of all charged tracks to that of a charged pion, $m_{\pi^\pm} \approx 140$ MeV, since pions are the most abundant hadrons produced in particle collisions. The effect of the pion mass assumption is greater in low-p_T measurements, where uncertainties in the masses of measured observables can become large. Monte Carlo (MC) methods can be used to take into account uncertainties arising from the pion mass assumption [170], or to correct for the energy loss of heavier particles [171, 172]. As such, since the masses of single jets are not exactly known, the chosen recombination scheme preserving the particle mass is not an important requirement. Jet mass is also often not included in theoretical calculations to which the experimental results are compared.

There exist many jet recombination schemes which set the particle masses to zero, and which define the recombination only in terms of particle transverse momentum p_{T}, pseudorapidity η and azimuthal angle ϕ. By using this kind of scheme in our dijet mass analysis, we can get rid of the uncertainty related to the determination of the mass of the single jets, which would also affect the uncertainty in the dijet invariant mass measurement. We also do not need to consider single jet mass uncertainty in the unfolding.

The p_{T} scheme [173], which is used in our analysis of the dijet mass, is one of the schemes where jets are treated as massless. In the p_{T} scheme, a preprocessing stage is incorporated to make the initial momenta of particles massless. This is done by rescaling the four-momenta of the particles so that their energy is equal to the amplitude of their three-momentum. This makes the four-momentum massless. After the preprocessing, the combination, p_r, of two four-momenta, p_i and p_j, which satisfies

$$p_{\mathrm{T},r} = p_{\mathrm{T},i} + p_{\mathrm{T},j} \tag{12}$$

$$\phi_r = \frac{p_{\mathrm{T},i}\phi_i + p_{\mathrm{T},j}\phi_j}{p_{\mathrm{T},i} + p_{\mathrm{T},j}} \tag{13}$$

$$\eta_r = \frac{p_{\mathrm{T},i}\eta_i + p_{\mathrm{T},j}\eta_j}{p_{\mathrm{T},i} + p_{\mathrm{T},j}}, \tag{14}$$

is a massless four-vector.

4.1.1 Jet areas

Jet areas measure how a jet extends over a surface in the $\eta - \phi$ plane. As a jet is not a set of evenly distributed partons but consists of a finite number of particles, the jet area must be specifically defined in order to avoid ambiguities in the area definition. The jet area functions as a measure of the jet's susceptibility to radiation, including the underlying event. Three definitions of the jet area have been proposed in Ref. [174]. These have been implemented in FastJet [169].

The active jet area is determined by first adding a uniform background of ghost particles to the event and then including them in the clustering. Ghost particles are infinitely soft, near-zero-p_{T} particles injected into the simulated event. The active area of a jet is taken to be proportional to the number of ghosts it contains. In contrast, the passive jet area is evaluated by adding a single ghost particle at a randomly chosen position to the event and checking whether the ghost ends up within a given jet. The procedure is repeated N times, and the passive area of the jet is defined to be proportional to the probability that it contains the ghost particle added to the event. A third definition of jet area is the Voronoi jet area, described in Ref. [174].

4.1.2 Jet reconstruction algorithms

Only a small fraction of the final-state hadrons in a hadron-hadron collision are associated with high-p_{T} jets, while the rest are part of the underlying event. To

reduce the influence of the underlying event, jets were historically defined using cone algorithms. Such algorithms were proposed as a standard in 1990 [175]. In these algorithms, a jet is defined by a cone of radius R_{jet} in the $\eta - \phi$ phase space. For any particle, its distance from the center of the jet is given by

$$\Delta R_{i,0} \equiv \sqrt{(\phi_i - \phi_0)^2 + (\eta_i - \eta_0)^2}, \tag{15}$$

where ϕ_0 and η_0 denote the center of the cone and (ϕ_i, η_i) are the coordinates of the particle. A particle is associated with the jet if it lies within the cone, i.e. if $\Delta R_{i,0} < R_{\text{jet}}$. Cone algorithms typically use some or all of the particles in the event as seeds. For each seed, a trial cone is created, and the list of particles within the trial cone is evaluated. The sum of the four-momenta of these particles acts as the four-momentum of a new seed. This process is iterated until the cone becomes stable, meaning the seed no longer changes [176].

The basic cone algorithm provides a simple method to form jets from all particles whose momentum vectors lie within a well-defined cone. However, the algorithm can result in ambiguous cases when jet cones overlap, resulting in two jets containing the same particles. Cone algorithms have largely been replaced by sequential recombination algorithms. These algorithms ensure an infrared and collinear (IRC) safe procedure for jet reconstruction, which was not the case for the simplest cone algorithms. IRC safety in jet algorithms is required for a comparison between experimental measurements of jets and pQCD calculations to be meaningful.

Using an infrared safe algorithm, if we add a very soft parton into the jet, the properties of the jet and the output of the jet algorithm do not change. An example case in which infrared problems can arise is portrayed in Fig. 26. On the left, two partons, portrayed by arrows, have been assigned to two separate jet cones. The distance between the two partons is greater than the cone radius R, which separates them into two jets, but still less than $2R$. If we then add a soft, third particle in between these two partons, with a distance less than than the cone radius to both original partons, the partons are then considered to be contained within the same cone. Thus, two jets are now one, meaning that the soft particle has changed the jet definition. This is an example of an infrared unsafe jet algorithm.

When using a collinearly safe algorithm, replacing a parton in the jet with two collinear particles carrying the same total momentum does not affect the jet properties or definition. In Fig. 27, on the left, three partons are contained within a jet cone of radius R. If we have an algorithm that takes the highest energy parton as the one to draw a cone around, this can result in issues with collinear safety: if the central parton on the left, with the largest energy, is split collinearly, the right hand side parton becomes the hardest parton, as shown on the right hand side of Fig. 27. Then, the left hand side parton is left out of the jet, and will join another jet. This results in the number of jets being affected by a collinear splitting, showcasing a collinearly unsafe algorithm.

The basic cone algorithms are not IRC-safe due to the seeds used in the iterative search for stable cones [176]. If only those particles whose p_{T} is higher

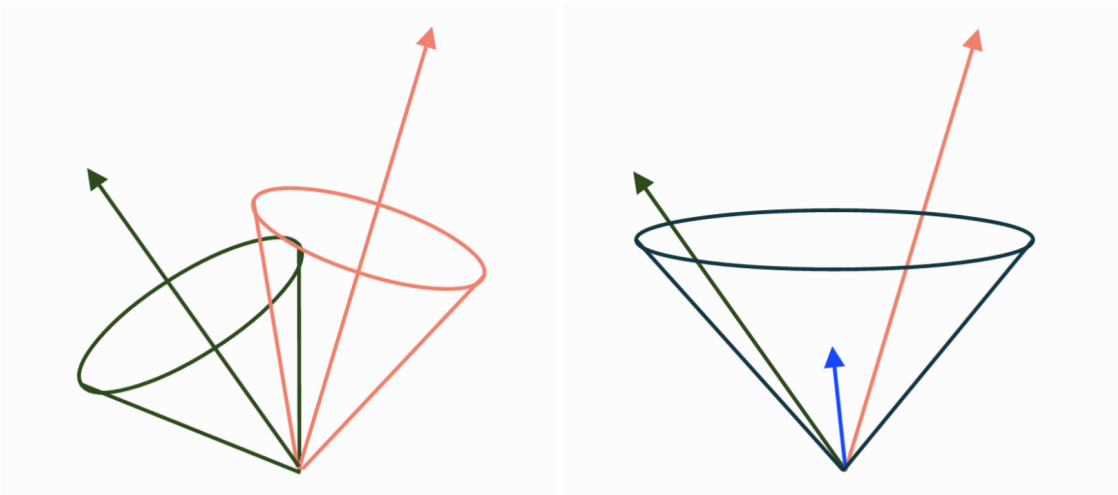


FIGURE 26 On the left, two jets are visualized with their cones of radius R , and the arrows represent particles within the jets, within a distance $< 2R$ between them. On the right, a soft particle has been added, which has affected the jet definition, resulting in all three particles being contained in one jet.

than a certain threshold are used as seeds, the procedure is collinearly unsafe. If all particles are considered for seeding, the process is IR unsafe if the addition of an infinitely soft particle causes the stable cones to be redefined. The SISCone (Seedless Infrared Safe Cone) algorithm [176] is an IRC-safe jet algorithm based on searching for stable cones. A split–merge step disentangles overlapping stable cones, making sure that there is no particle that ends up in more than one jet.

Many other IRC-safe jet algorithms have also been implemented in the FastJet software package [169]. Of these, the most commonly used are the k_T [177], anti-k_T [178] and Cambridge/Aachen [179, 180] clustering algorithms. Jet clusters formed from these algorithms, as well as from the SISCone algorithm, are visualized in Fig. 28. To create the shapes in the figure, which illustrate the active catchment areas of the resulting hard jets, a parton level event has been injected with around 10^4 ghost particles and the jets have been clustered using the four jet algorithms. For each parton level jet, the region within which the ghosts have been clustered within the jet are shown [178]. The algorithms are discussed in more detail below.

The k_T algorithm starts with a list of particles, called protojets, and their momenta, p_i^μ. The masses of the protojets are small compared to their transverse momenta, so the momenta are considered light-like. The algorithm proceeds recursively as follows [177]:

1. For each protojet i we define the beam distance

$$d_{iB} = p_{\mathrm{T},i}^2, \tag{16}$$

and for each pair of protojets, i and j, we define a distance

$$d_{ij} = \min(p_{\mathrm{T},i}^2, p_{\mathrm{T},j}^2)\frac{\Delta R_{ij}^2}{R^2}, \tag{17}$$

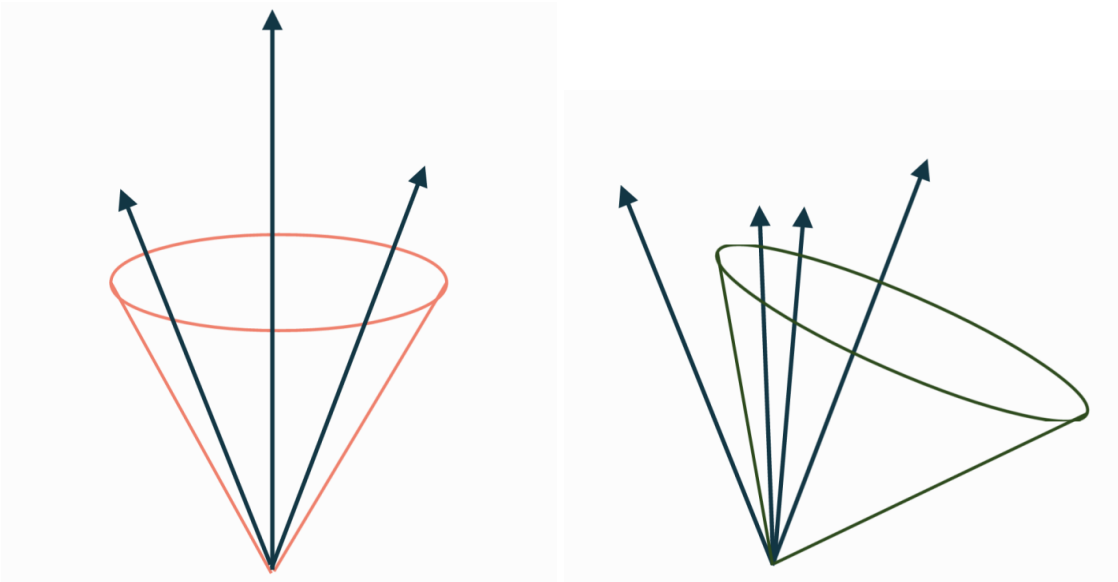


FIGURE 27 On the left, a jet with three parton constituents is visualized, with the central parton being the hardest one. On the right, the hardest parton has split collinearly, making the right hand side parton the hardest. When using a jet algorithm that creates the cone around the hardest parton, the jet definition can change when partons split, showcasing collinear unsafety.

where R is a chosen resolution parameter, of the order of 1, and $\Delta R_{ij}^2 = \Delta\phi_{i,j}^2 + \Delta\eta_{i,j}^2$ is the distance between the protojets in the $\phi - \eta$ plane, similarly to Eq. (15).

2. We find the smallest of the d_{iB} and d_{ij}, and label the smallest value found $d_{\min}$.
3. If the smallest value $d_{\min}$ is a distance d_{ij}, we merge the protojets i and j into a new protojet, k. This can be done for instance with the E-scheme as

$$p_{\mathrm{T},k} = p_{\mathrm{T},i} + p_{\mathrm{T},j}. \tag{18}$$

4. If the smallest value $d_{\min}$ is a d_{iB}, we remove the corresponding protojet i from the list of protojets, and add it to a list of completed jets.
5. We continue from step 1 until the list of protojets is empty.

At the end of the algorithm, we are left with a list of jets with successively larger values of d_{iB} for each collision event. The last jets added to the list are those of interest, as they have large energies, and the jets with smaller energy, sometimes called minijets, or sometimes resulting purely from debris from the beam, are often not considered for analysis.

The Cambridge/Aachen (C/A) algorithm is otherwise identical to the k_{T} algorithm, but the distance measure is defined as

$$d_{ij} = \frac{\Delta R_{ij}^2}{R^2}, \tag{19}$$

and the beam distance

$$d_{iB} = 1. \tag{20}$$

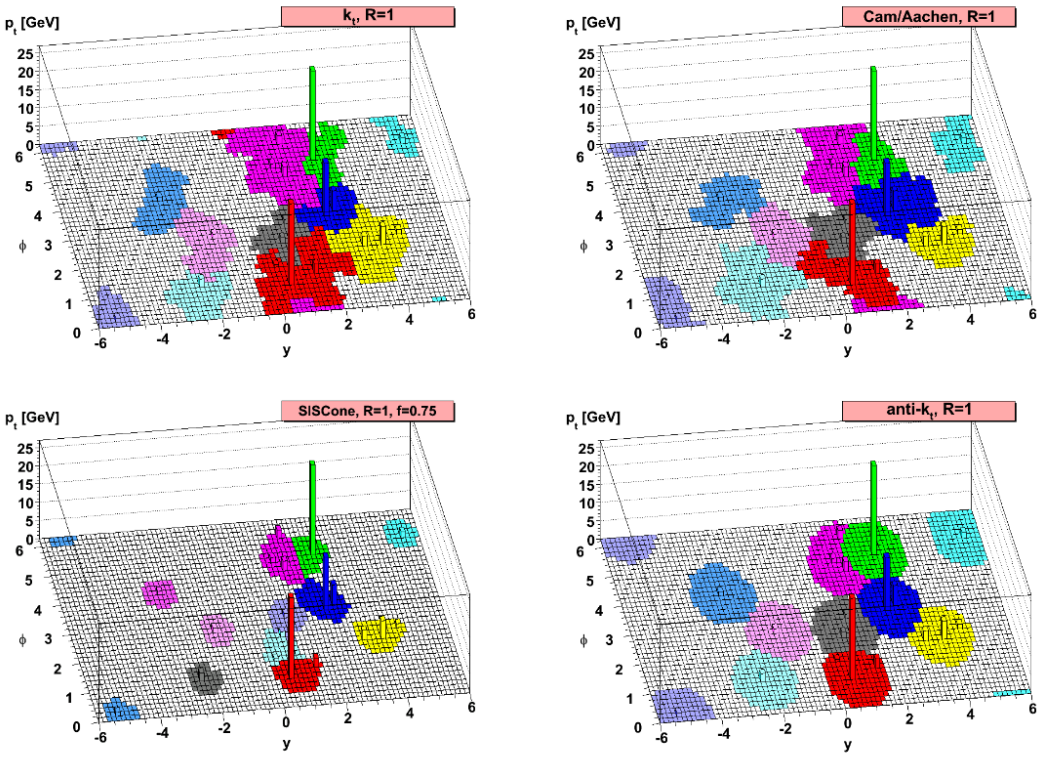


FIGURE 28 A parton-level event has been generated with Herwig along with around 10^4 ghost particles. The jets have been clustered with four different jet algorithms, resulting in active catchment areas of the resulting jets. Figure from Ref. [178]. Copyright © 2008, SISSA; permission to reprint granted by the copyright holder.

For the k_T and C/A algorithms, the regions in Fig. 28 are jagged and irregular. This is a consequence of the randomness of the ghosts, as the shapes of the regions are affected by the specific sets of ghosts used in the event. For the SISCone algorithm, the regions are more regular, with the more-jagged regions associated with composite jets consisting of more than one particle.

The anti-k_T algorithm is defined exactly like the k_T algorithm, except the distance measure is defined as

$$d_{ij} = \min(1/p_{\mathrm{T},i}^2, 1/p_{\mathrm{T},j}^2)\frac{\Delta R_{ij}^2}{R^2}, \tag{21}$$

and the beam distance as

$$d_{iB} = 1/p_{\mathrm{T},i}^2. \tag{22}$$

While the anti-k_T algorithm is a sequential recombination algorithm like the k_T and C/A algorithms, it also behaves to a certain degree like a cone algorithm, since the hard jets are circular in the $y - \phi$ plane, as seen in Fig. 28. The softer jets create a slightly more jagged region, clipped by the harder, circular, jet regions.

4.2 Unfolding

Distributions of measured observables in high-energy physics experiments are affected by the limited acceptance, finite resolution and efficiency of the detectors. These effects distort the measurements of observables in comparison to Monte-Carlo (MC) generated particle-level information. Also, physical effects like radiation, parton fragmentation, and particle decay cause the measured distribution to differ from the parton level information given by many theory calculations. When we want to compare MC simulations to data, we must correct for detector inefficiencies in the measured data in order for the sets to be comparable. The task of transforming the measured distribution, $g(y)$, to a true distribution, $f(x)$, of a kinematical quantity x is a complex task called unfolding.

In the measurement of just one physical observable, the simplest unfolding method would be a so-called bin-by-bin correction. This correction can be obtained from the MC by calculating the ratio of the number of events which fall into a certain bin of the reconstructed variable y, and the number of events in the same bin of true variable x. The ratio is then used as a generalized efficiency to estimate the number of true events based on the number of measured events in that specific bin. This method is often not used, however, as it requires the same subdivision of bins in the true and experimental variable, neglects correlations between bins, and is only valid if the migration between bins is negligible and if the standard deviation of smearing is not the size of the bin width [181].

In mathematical terms, the problem of unfolding can be formulated as an integral equation of the form [182]

$$g(y) = \int A(y,x) f(x) \mathrm{d}x, \tag{23}$$

where, given the observation of the distribution $g(y)$ of a quantity y, related to x, and the resolution function $A(y,x)$ representing the detector effects, the distribution $f(x)$ of the true variable x can be calculated. The unfolding is then the inverse problem of determining $f(x)$ from $g(y)$.

In reality, we often study a finite set of parameters, representing the distributions using histograms with a finite number of bins, and the resolution function as a matrix. Equation (23) can then be represented instead by a matrix equation:

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \tag{24}$$

where the vectors $\mathbf{x}$ and $\mathbf{y}$ of dimensions m and n represent the histograms of the distributions of quantities x and y, respectively, and the transition from $\mathbf{x}$ to $\mathbf{y}$ is described by the n-by-m response matrix $\mathbf{A}$. To understand this more intuitively, the elements A_{ij} of the response matrix $\mathbf{A}$ are related to the probability to find an entry in the histogram bin i of the histogram $\mathbf{y}$ in the case that the true value x is produced from the bin j in the histogram $\mathbf{x}$. As neither the resolution function $A(y,x)$ nor the response matrix $\mathbf{A}$ can be known analytically, often $\mathbf{A}$ is estimated using Monte Carlo techniques simulating the detector effects.

If the dominant effect of the response matrix $\mathbf{A}$ is limited acceptance, all non-diagonal elements of $\mathbf{A}$ are zero, and the unfolding of $\mathbf{x}$ by inverting Eq. (24) is not a difficult procedure. However, solving the solution of

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}, \tag{25}$$

where $\mathbf{A}^{-1}$ is the inverse matrix of $\mathbf{A}$, becomes difficult with non-diagonal terms of $\mathbf{A}$. The problem becomes even more difficult when dealing with singular matrices, which do not have an inverse. As the response matrix is often filled based on information from a MC simulation, the statistics of the matrix are not infinite. This can result in fluctuations in the bins, especially in those near the edges of the matrix. When directly inverting the matrix, contributions from the small fluctuations can become very large, affecting the sensibility of the inverted matrix and the results unfolded with it. Therefore, in experimental physics we often use specific unfolding algorithms that are equipped to deal with such fluctuations.

One unfolding method which can handle such fluctuations is a commonly used method based on Bayes' Theorem [181]. Bayes' Theorem is a mathematical theorem that tells the probability of a cause given its effect. Let's say we have several independent causes, $C_i, i = 1, 2, ..., n_C$, which can produce a single effect, E. We assume we know the initial probability of the causes, $P(C_i)$, as well as the conditional probability of each cause to produce the effect, $P(E|C_i)$. Then, Bayes' formula tells us the probability of the observed effect having been caused by the i-th cause:

$$P(C_i|E) = \frac{P(E|C_i)P(C_i)}{\sum_{j=1}^{n_C} P(E|C_j)P(C_j)}. \tag{26}$$

In experimental particle physics, Eq. (26) can be used to estimate the true, original result, without detector effects, given a certain measurement, which has been affected by the detector and other effects in the measurement process. Bayes' Theorem can thus be used as a basis for the multidimensional unfolding of experimental distributions, following the D'Agostini iterative method [181, 183].

Experimentally, the effect E can be considered as the measured event, for instance a measured or detector level dijet, and the causes C_i as the true event, like truth or particle level dijets. The Bayesian method allows the use of different sized cells for the distributions of the true and detector level values. Matrix inversion is not necessary, so the method works for any kind of response matrix.

In Eq. (26), the probability $P(C_i|E)$ depends on the initial probability of the causes, $P(C_i)$. The distribution of these probabilities is called the prior, which represents a prior guess of the distribution of the true values. A realistic guess of the true value distribution results in the best outcome, but even in the case of a totally ignorant guess, a uniform distribution, results can still be satisfactory. This is because, in the iterative unfolding method, the prior is updated after each iteration, using the result from the previous step. The probability $P(C_i|E)$ also depends on the distribution of the probabilities $P(E|C_i)$. This corresponds to the unfolding matrix, and has to be calculated or estimated using Monte Carlo methods.

Let's say we have measured, after N_{obs} experimental observations, frequencies of dijets of a certain type j, $n(E_j)$. To get the expected number of true dijets, we calculate

$$\hat{n}(C_i) = \sum_j n(E_j) P(C_i|E_j), \tag{27}$$

and the sum of all $\hat{n}(C_i)$s is the estimate of the true total number of events, $\hat{N}_{\text{true}}$. In order to calculate the $P(C_i|E_j)$, we first need to choose a prior distribution to plug into Eq. (26), $P_0(C_i)$. The initial expected number of events is thus $n_0(C_i) = P_0(C_i)N_{\text{obs}}$. After the first iteration, we calculate the values $\hat{n}(C_i)$ and $\hat{P}(C_i) = \hat{n}(C_i)/\hat{N}_{\text{true}}$. Then, the $\hat{P}(C_i)$ values replace the priors $P_0(C_i)$ and the $\hat{n}(C_i)$ replace $n_0(C_i)$ values, and the algorithm starts again. The χ^2 of all $\hat{n}(C_i)$ and $n_0(C_i)$ are compared after the second iteration. If the χ^2 is not small enough, we continue the procedure until some cutoff number of iterations.

Another commonly used unfolding method is the Singular Value Decomposition (SVD) method [184], which is a regularization method for unfolding measured distributions using the SVD of the response matrix. Using a regularization condition [182] aids in overcoming the instability of unfolding, resulting from small variations in the system, statistical fluctuations and consequent issues in creating the inverse of the unfolding matrix. The regularization of the SVD unfolding is set with the k-value. Too small values of k can result in over-regularization, biasing the unfolded spectrum towards the input MC, while under-regularization with too large values of k leads to large fluctuations in the unfolded spectrum [185]. The optimal regularization is determined according to guidelines in Ref. [184]

The Bayesian unfolding method is widely used, in part because it can be generalized to multi-dimensional distributions. In our dijet invariant mass analysis, we perform the unfolding with both methods, taking the Bayesian method as the default, and the SVD method as a comparison to evaluate the reliability of results and to study systematic errors arising from the choice in the unfolding method.

4.3 Background subtraction methods

In heavy-ion collisions, there is a significant background impacting the reconstructed jet four-momentum, caused by particles from both soft scattering processes as well as other hard scatterings than the jet under study [186, 187]. FastJet [169] provides a standard method to measure and remove background on an event-by-event basis. In the method, all particles are first clustered into jets using the k_{T} algorithm. The background density, ρ, is estimated as the median of the transverse momentum density in the jets in the event,

$$\rho = \text{median} \left\{ \frac{p_{\text{T,jet}}^{\text{raw}}}{A_{\text{jet}}} \right\}, \tag{28}$$

where $p_{\mathrm{T,jet}}^{\mathrm{raw}}$ is the jet transverse momentum without any background subtraction, and A_{jet} is the area of the jet. In the calculation of the median, the two leading p_{T} jets are excluded. The transverse momenta of all jets in the analysis are then corrected using the median ρ to obtain the background-subtracted jet transverse momentum

$$p_{\mathrm{T,jet}} = p_{\mathrm{T,jet}}^{\mathrm{raw}} - \rho A_{\mathrm{jet}}. \tag{29}$$

In p–Pb collisions, the charged-particle multiplicity density [188] is much smaller than in central Pb–Pb collisions [189], by two orders of magnitude. Therefore we assume the jet background to be reduced correspondingly. The more sparse p–Pb collision event necessitates a more reliable estimate. For this purpose, an approach developed for pp collisions, the sparse event method [190], is used. The method is similar to that used for Pb–Pb collisions, but with an additional correction factor accounting for more sparse events having regions without particles.

The transverse momentum density ρ of the event is calculated for k_{T} jets for both pp and p–Pb events,

$$\rho = \underset{\text{real } k_{\mathrm{T}} \text{ jets}}{\text{median}} \left\{ \frac{p_{T,\mathrm{jet}}}{A_{\mathrm{jet}}} \right\} \times C, \tag{30}$$

where A_{jet} is the jet area in the ϕ–η plane calculated using the explicit ghost method [169]. Ghost particles, with a vanishing momentum, do not influence the jet finding procedure, and they are distributed into the area that is to be clustered. As such, the number of ghost particles in a jet is a direct measure of the area of that jet. The term "real k_{T} jets" refers to k_{T} jets with at least one track, thus discarding pure ghost jets. Also, the leading and subleading jets are dropped from the determination of ρ.

The coefficient C estimates the empty ϕ–η area of the event using jets clustered with the k_{T} algorithm. This is done by dividing the sum of the area of real jets by the sum of the area of all jets, including empty ghost jets [191]:

$$C = \frac{\sum_{\text{real } k_{\mathrm{T}} \text{ jets}} A_{\mathrm{jet}}}{\sum_{\text{all } k_{\mathrm{T}} \text{ jets}} A_{\mathrm{jet}}}. \tag{31}$$

The soft background is subtracted from the underlying event using the four-momentum background subtraction [192, 193]

$$p_{\mathrm{corr}}^{\mu} = p^{\mu} - \left[(\rho + \rho_m) A_{\mathrm{jet}}^{E}, \rho A_{\mathrm{jet}}^{x}, \rho A_{\mathrm{jet}}^{y}, (\rho + \rho_m) A_{\mathrm{jet}}^{z} \right], \tag{32}$$

with p^{μ} corresponding to the uncorrected jet four-momentum, ρ_{m} the mass density calculated in a similar manner as ρ, and A^{μ} the jet-area four-vector, as calculated by the FastJet package [169].

The issue with using only the soft background subtraction method defined above is that it neglects to account for background that is not uniformly distributed in the ϕ-η plane, but instead fluctuates between different regions. These fluctuations are mostly Poissonian in nature, but also include some correlated

variations of the particle multiplicity and mean p_{T} by region, resulting from both detector effects and the collision itself [186]. The density fluctuations around the soft background density estimate ρ can be evaluated using a Random Cone (RC) approach, as described in [186].

Cones of a radius corresponding to the value of R used in the jet finding algorithm are randomly placed into the ϕ-η acceptance. Then, the transverse momenta of all of the charged tracks that coincide with the cone are summed, and the background estimate is subtracted from this sum:

$$\delta p_{\mathrm{T,ch}} = \sum_i p_{\mathrm{T,ch}} - \rho_{\mathrm{ch}} A, \;\; A = \pi R^2. \tag{33}$$

The distribution of $\delta p_{\mathrm{T,ch}}$ directly measures the fluctuations of the background and can be directly used in the unfolding procedure. In the dijet mass analysis presented in this thesis, inspired by two-particle correlation analyses, we determine the background using cones rotated by 90 degrees in ϕ instead of randomly rotated cones [194]. With the non-random rotation in ϕ, the rotated cone better avoids the signal from the correlated particle or jet ending up within the rotated cone, which would distort the background evaluation.

4.4 Monte Carlo simulations

Monte Carlo (MC) event generators are a central tool in simulating physical processes, notably collisions of particles and heavy ions. These event generators are a tool to obtain a prediction to which data can be compared. They utilize predictions for QED and QCD in collision systems, and by simulating a large variety of outcomes over a large amount of collisions, give out a final-state particle level output. Over the past few decades, techniques to describe most aspects of hadronic collisions have been gathered, refined and implemented into general-purpose event generators, capable of simulating a wide range of processes important for particle physics experiments. For instance, HERWIG [195], PYTHIA [196] and SHERPA [197] are some widely used general-purpose generators. In addition, a large number of generators specific to certain processes have been developed.

The event generators applicable for the study of jets in heavy-ion collisions can be coarsely split into two categories. The first are generators that modify simulated pp collisions according to assumed changes to the jet evolution due to the QGP medium, namely jet quenching. Generators like JEWEL [198] and the JETSCAPE [199] framework fit into this category. Another category of MC event generators are ones that attempt to describe the full collision event. This is the case for generators like HIJING [200], EPOS [201] and the Angantyr [202] addition to PYTHIA. These generators are useful for studying the influence of the underlying event on the jet signal [203].

Most MC event generators available for jet physics are at leading order of calculations. Some frameworks, like MC@NLO [115] and POWHEG [204],

TABLE 2 Summary of commonly used Monte Carlo event generators and frameworks.

Generator	Category	Main Features
PYTHIA [196]	General-purpose	pp generator; hard processes, parton showers, MPI, hadronization.
SHERPA [197]	General-purpose	Lepton-lepton, lepton-photon, photon-photon, lepton-hadron and hadron-hadron collisions.
HERWIG [195]	General-purpose	Hard lepton-lepton, lepton-hadron and hadron-hadron collisions. NLO accuracy.
JETSCAPE [199]	General purpose heavy-ion	Modular framework; combines hard scattering (PYTHIA) with in-medium shower models (e.g. MATTER, LBT, MARTINI) and bulk evolution.
JEWEL [198]	Jet quenching	QCD jet evolution in heavy-ion collisions.
HIJING [200]	Full heavy-ion event	Jet production in pp, pA and AA collisions.
ANGANTYR [202]	Full heavy-ion event (PYTHIA8 extension)	High-energy nucleus collisions.
EPOS [201]	Full heavy-ion event	Heavy-ion interactions and cosmic ray air shower simulations.
POWHEG [204]	Framework (NLO matching)	NLO calculations interfaced with parton shower.
MC@NLO [115]	Framework (NLO matching)	NLO QCD matrix elements interfaced with parton shower.

however, are based on next-to-leading-order (NLO) calculations, interfaced with showers from generators like PYTHIA or HERWIG. Sherpa can be run with NLO elements, using packages like OpenLoops [205] to generate NLO loop corrections. In Table 2 are listed all the mentioned generators and frameworks and their basic properties.

A general-purpose MC event generator simulates a process in several phases. A schematic of a PYTHIA generated pp event is shown in Fig. 29. In the figure, different components of the simulated high-energy collision are shown. The different momentum regimes in a particle collision can be factorized, and we end up with regimes of higher and lower momentum particles. MC techniques can be used for both hard and soft regimes, connected by an evolutionary process calculable from pQCD [206]. The large phase space with dimensions $3n - 4$, resulting from the three components of momentum for each of the n particles while accounting for energy-momentum conservation, works well with MC integration methods, for which accuracy increases inversely with the square root of the number of integration points.

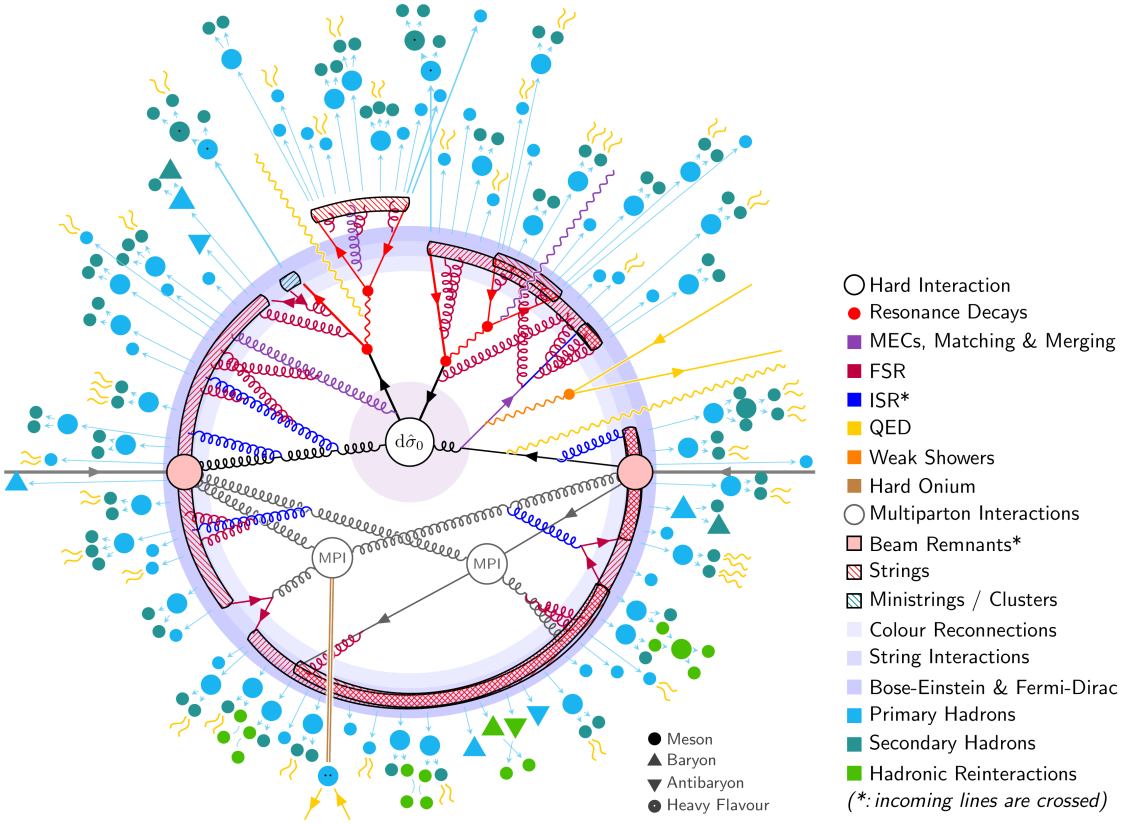


FIGURE 29 A schematic of a PYTHIA pp→ $t\bar{t}$ event. Figure from Ref. [207], reprinted under the license CC BY 4.0.

First, there is a primary hard sub-process. The hard sub-process can oftentimes be selected by the user, in order to obtain events of a certain type. Gluons radiating from partons from the hard sub-process generate showers of outgoing partons. This is simulated with a Markov chain, called a parton shower algorithm. The parton shower includes initial-state radiation (ISR) and final-state radiation (FSR). Further scattering processes between additional partons from the incoming beams are also possible, called multi-parton interactions (MPI).

Hadronization happens at the end of a parton shower. Once the event has been evolved down in momentum to a region where perturbative QCD breaks down, the evolution must be replaced by a non-perturbative hadronization model, confining the colored partons into hadrons. This results in hadronization models not being directly derivable from QCD, but tunable with an assortment of parameters. While most interesting signal processes studied in high-energy collisions are from hard interactions, there is also a large contribution from soft collisions. Also non-perturbative in nature, these collisions have to be modeled with tunable parameters, in order to describe data.

To incorporate PDFs and nPDFs into MC simulations of pp and p–Pb collisions, each incoming beam is assigned a PDF set consistent with the perturbative order of the hard process. In pp collisions, both beams can use a free-proton PDF set. In p–Pb collisions, the proton beam is assigned a free-proton PDF, while the lead beam is described using a nPDF for bound protons or neutrons in the lead nucleus. The LHAPDF 6 library [208] provides standardized access to a wide range of PDFs and nPDFs, and is supported by many MC generators and other physics programs. Running p–Pb simulations also requires accounting for the energy asymmetry between the proton and nucleons in the lead ion. A lead nucleus contains 82 protons and 126 neutrons for a total of $A = 208$ nucleons, giving a charge-to-mass ratio $Z/A = 82/208$. While a proton in the LHC can be accelerated up to an energy E_{p}, the energy per nucleon of the lead ion is limited to $E_{\mathrm{Pb}} = \frac{Z}{A} E_{\mathrm{p}}$. As a result, the nucleon–nucleon center-of-mass system in p–Pb collisions at 5.02 TeV is boosted in the proton beam direction with a rapidity shift of $\Delta y = -0.465$ [209]. The asymmetry can be taken into account in the MC simulation either by setting asymmetric beam energies or by applying the rapidity shift into the center-of-mass system.

To study jets in heavy-ion collisions, a commonly used approach is embedding, in which hard jets generated, for instance, with PYTHIA are added to simulated heavy-ion events produced with a generator such as HIJING or EPOS. Embedding allows testing background subtraction methods and jet-finding algorithms: by comparing the truth-level jets to the reconstructed jet after background subtraction, the efficiency and performance of the algorithms can be evaluated.

Running PYTHIA as a minimum bias (MB) production results in a steeply falling jet spectrum. To ensure sufficient statistics for high-p_{T} jets, the simulations are enhanced by over-sampling the hard-scattered process. In order to achieve this in our dijet mass analysis, PYTHIA simulations were run in eight different bins of the transverse momentum of the hard $2 \to 2$ parton scattering, $\hat{p}_{\mathrm{T}}$, with upper and lower limits set for each $\hat{p}_{\mathrm{T}}$ bin. To merge the $\hat{p}_{\mathrm{T}}$ bins, the result-

ing spectra have to be scaled with a weight, which is determined according to the corresponding cross section provided by PYTHIA divided by the number of events in that bin. After weighting, the merged spectra are properly normalized and provide sufficient statistics across the full p_{T} range.

All MC event generators contain numerous free parameters that can be varied independently of each other. These parameters control aspects such as hadronization and multi-parton interactions, both containing pQCD physics. To produce physically sensible simulations, these parameters are tuned simultaneously to describe a broad set of experimental data. In PYTHIA, such prepackaged sets of parameters are called tunes. For the dijet mass analysis in pp collisions, we use the Monash 2013 tune [210], which is optimized using a wide range of LHC measurements.

The MC simulations used for the unfolding of the dijet mass spectrum in the dijet mass analysis, described in Sec. 5, are generated centrally in ALICE. The simulated signals should reproduce the real data conditions as closely as possible, which means that the detector geometry, data collection efficiency, as well as effects from the LHC conditions present in the data have to be included within the simulation procedure. Anchoring a MC simulation means that the conditions are matched with those during a real data-taking run, including information from the LHC filling scheme, the ALICE detector states, and the interaction rate.

The ALICE detector and other material, like the beam pipe and magnet, are simulated with GEANT [211, 212]. The MC generated particles pass through this material, interacting with it as in real life. The detector response is simulated through the digitization of the detector signals, after which the reconstruction procedure begins [32]. During the MC reconstruction phase, the same version of the reconstruction software should be used as in the processing of the real data, to ensure consistent tracking performance and efficiency, and overall data quality. If significant changes are made to the data reconstruction procedure and a new reconstruction pass of the data is used, the MC simulation must also be updated to match the reconstruction of the simulated events to the data reconstruction conditions. This is the case in our dijet analysis where the reconstruction was updated from pass1 to pass2, the details of which are described in Sec. 5.

5 DIJET INVARIANT MASS ANALYSIS

The goal of this analysis is to measure the invariant mass spectrum of charged dijets measured by ALICE in proton–proton and proton–lead collisions, in the years 2017 and 2016, respectively, at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV. We calculate the nuclear modification factor R_{pA} in the dijet invariant mass range of 75–150 GeV/c^2 and compare the results with different Monte Carlo model calculations. This works continues that described in Ref. [1], in which the analysis was first performed and described. In this thesis, the analysis has been continued in order to update the results to a new ALICE data reconstruction pass. The analysis is described in this section, with emphasis on the steps required to update to the new data pass.

5.1 Dijet invariant mass

Jets are reconstructed using the FastJet package [169] with the anti-k_{T} algorithm [178]. The jet radius of $R = 0.4$ is used, and the jets are required to be within the ALICE central tracking acceptance, restricting the jet pseudorapidity to $|\eta_{\mathrm{jet}}| < 0.9 - R = 0.5$. As incoming tracks are clustered with the p_T scheme, they are scaled to be massless, as described in Sec. 4.1. The p_{T} scheme also performs the recombination of these tracks such that the final jet is massless.

In this analysis, we consider dijets only in those events where the transverse momenta of both jets in the dijet system exceed 20 GeV/c. The jets are required to be measured in opposite hemispheres of the transverse plane, which means that their azimuthal difference $|\Delta\phi| > \pi/2$. This cut allows us to avoid contributions from $2 \to 3$ processes, where one of the hard jets is outside the detector acceptance, and the other jet splits into two, forming a low-mass dijet system with $|\Delta\phi| < \pi/2$. When searching for the leading and subleading jets in an event, if the azimuthal angle of the subleading jet does not fit this criterion, the subleading jet is discarded from the search. Then, the next hardest jet is taken as the subleading jet candidate, and tested for the azimuthal and p_{T} criteria. Once a subleading jet within the conditions is found, the dijet is formed.

Most events in this study contain a maximum of one dijet which fulfills the cuts defined. However, in some rare cases, the collision event under study contains more than one acceptable dijet. In this case, the leading pair of jets fulfilling the cuts is selected based on the momentum of the leading jet, for which the subleading jet is further selected. Jets containing constituents with p_{T} above over 100 GeV/c are removed from the analysis, due to the lowered p_{T} resolution of tracks with $p_{\mathrm{T}} > 100\,\mathrm{GeV}/c$.

The dijet invariant mass, also called dijet mass here, is defined as

$$M_{\mathrm{jj}}^2 = (p_1 + p_2)^2 \tag{34}$$

$$= m_1^2 + m_2^2 + 2\left(m_{\mathrm{T},1} m_{\mathrm{T},2} \cosh(\Delta y) - p_{\mathrm{T},1} p_{\mathrm{T},2} \cos(\Delta\phi)\right) \tag{35}$$

$$\approx 2 p_{\mathrm{T},1} p_{\mathrm{T},2} \left(\cosh(\Delta\eta) - \cos(\Delta\phi)\right), \tag{36}$$

where p_i is the four-momentum, $p_{\mathrm{T},i}$ the transverse momentum, m_i the mass and $m_{\mathrm{T,i}}$ the transverse mass of the leading ($i = 1$) or subleading ($i = 2$) jet. The last approximation holds when $m/p_{\mathrm{T}} \ll 1$, and as such, this approximation is exact for massless jets.

5.2 Update to new reconstruction pass

The ALICE datasets undergo many reconstruction passes before the data is ready to be analyzed [70]. The first two reconstruction passes are used for calibration purposes, performed on a sample of events from each run. The first of these is cpass0, which provides input for the calibration of the TPC, TRD, TOF and T0, as well as for the centrality. The second pass, cpass1, applies the calibration, and the calibrated, reconstructed events are then used as input for the data quality assurance and updated calibration of the SDD, TPC and EMCal. For each data taking period, normally spanning a period of 4–6 weeks, a manual calibration is performed. This calibration is then verified with a validation pass, vpass. Finally, a physics reconstruction pass is performed an all events, providing the input for physics analysis. Passes are updated when changes are made to the reconstruction process, such as when new reconstruction and calibration methods are implemented or detector alignment is improved [70]. The original analysis of dijet mass was developed and performed in Ref. [1] for the first reconstruction pass of the datasets LHC17pq (pp) and LHC16qt (p–Pb). I updated the analysis to run on the new, second reconstruction passes of both datasets.

The main improvements of pass2 with respect to pass1 include the following for pp and p–Pb datasets, as reported by the ALICE Data Preparation Group. The TRD was used in refitting the tracks, the alignment of some SPD sectors was improved, the TPC reconstruction and error parametrization as well as distortion map binning and material alignment were improved, and the treatment of the difference between the ALTRO (ALICE TPC Read Out chip [213]) clock and the LHC clock was improved. In addition, for p–Pb pass2 there was also an improved TOF calibration and a fix for the innermost SPD layer Lorentz angle. For

the pass2 datasets used in our dijet mass analysis, the vertex cuts together with trigger and event selection, which were described in Sec. 2.2.1, select around 890 million events out of a total of 1040 pp events, and around 620 million events out of 730 million p–Pb events. These can be compared to the pass1 datasets, where around 705 million pp events were selected out of 820 million events, and 530 million p–Pb events were selected out of 625 million.

The update to the new reconstruction pass required running a new Monte Carlo production anchored to the pp and p–Pb pass2 datasets to use in the unfolding procedure. To create the pp MC datasets for both pass1 and pass2, PYTHIA was used to generate jet events. The $\hat{p}_{\mathrm{T}}$ bins used for this production were 5-7, 7-9, 9-12, 12-16, 16-21, 21-28, 28-36, 36-45, 45-57, 57-70, 70-85, 85-99, 99-115, 115-132, 132-150, 150-169, 169-190, 190-212, 212-235, >235 GeV/c^2.

For p–Pb, the MC production anchored to the pass1 dataset was generated using PYTHIA and EPOS. EPOS was used to generate an underlying event of the p–Pb collisions, and PYTHIA was used to generate hard QCD jets. These jets were then embedded on top of the underlying event. Due to technical limitations in the centralized ALICE MC generation process, it was not possible to use EPOS for the new pass2 MC. Therefore, the pass2 anchored MC is a PYTHIA+HIJING production, in which PYTHIA jets are embedded into a HIJING underlying event background. Both productions were performed in $\hat{p}_{\mathrm{T}}$ bins 5-11, 11-21, 21-36, 36-57, 57-84, >84 GeV/c^2.

In the EPOS production, the $\hat{p}_{\mathrm{T}}$ restriction of the MC production bins did not appear to be taken into account by the EPOS-generated particles, which resulted in large statistical fluctuations especially in low $\hat{p}_{\mathrm{T}}$ bins. All particles originating from the EPOS generator were therefore filtered out with the help of *AliAnalysisTaskSVtaskMCFilter* in AliPhysics [214]. This task reads the original track and particle lists, filters them and feeds new lists of tracks and particles with filtered information back to the original analysis task. I updated the filter task for the pass2 p–Pb MC to filter HIJING particles similarly as the EPOS particles were removed by the original task. To study the HIJING production and its filtering, I tested how the analysis performs with and without the filtering included. I found that fluctuations in the raw, truth-level track p_{T} are removed with the filtering of HIJING particles in the pass2 MC, and these fluctuations were more pronounced than in the EPOS production anchored to pass1 MC. This suggests that the filtering is necessary, even more so for the new HIJING production than for the EPOS production.

For MC runs in the analysis, if we find a jet with transverse momentum p_{T} over four times as large as the $\hat{p}_{\mathrm{T}}$ of the event, we reject the event. This is a standard procedure in ALICE analysis of jets in $\hat{p}_{\mathrm{T}}$-binned MC simulations. If the jet has a much larger p_{T} than $\hat{p}_{\mathrm{T}}$, it can result in an outlier when merging the bins. As merging the bins involves scaling with the cross section, and as this scaling factor is largest in the smallest $\hat{p}_{\mathrm{T}}$ bins, if there is an outlier in a smaller $\hat{p}_{\mathrm{T}}$ bin, its effect can become relevant when it is scaled and merged. The factor 4 has been found to minimize the outlier contributions, while remaining a high enough value to create a sensible p_{T} distribution.

In order to visualize any possible differences in jet and track p_{T} distributions in the passes of the data and the correspondingly anchored Monte Carlo simulations, I plot spectra from the pp and p–Pb datasets and Monte Carlo simulations. In Fig. 30, the comparison of jet p_{T} spectra in the raw data is shown. In the figure, it can be seen that there is some fluctuation in the final jet p_{T} bins above 100 GeV/c in the pp data, shown in the left-hand side plot. In the p–Pb data, on the right-hand side, this fluctuation is also seen but, more strikingly, there is also a 10% increase in the spectrum of pass2 as compared to pass1. This can be due to improvements in the tracking, which can increase the amount of tracks found in the data set.

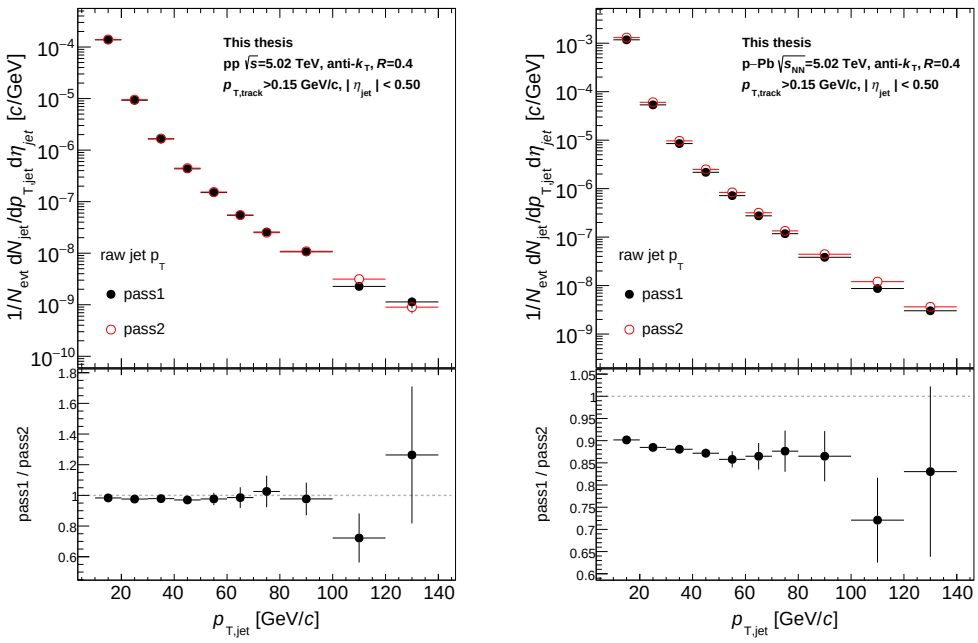


FIGURE 30 Jet transverse momentum spectra in the pp (left) and p–Pb (right) raw data.

In Fig. 31, the detector level jet p_{T} spectra in the MC is presented. As is expected based on the data, the detector level jets in the p–Pb MC also show a significant difference between passes, although here the difference is around 6-8% instead of the 10% seen in the data. For the pp MC, the spectrum shows a difference of around 1% in the higher bins of jet p_{T}. In Fig. 32, the truth level MC jet p_{T} spectra are plotted. In the truth level MC, the two passes give similar spectra for p–Pb. From the pp spectrum, it is seen that the 1% difference in the detector level MC is visible also on the truth level. Similar differences are reflected in charged particle distributions.

In Eq. (36), we can see that the dijet mass is formed from the transverse momenta and the rapidity and azimuthal differences of the jets in the dijet system. The respective contributions to the dijet mass of the two components of Eq. (36) were studied in detail in [1] by studying the kinematic and geometrical components of the jets before any background subtraction. I performed the same studies

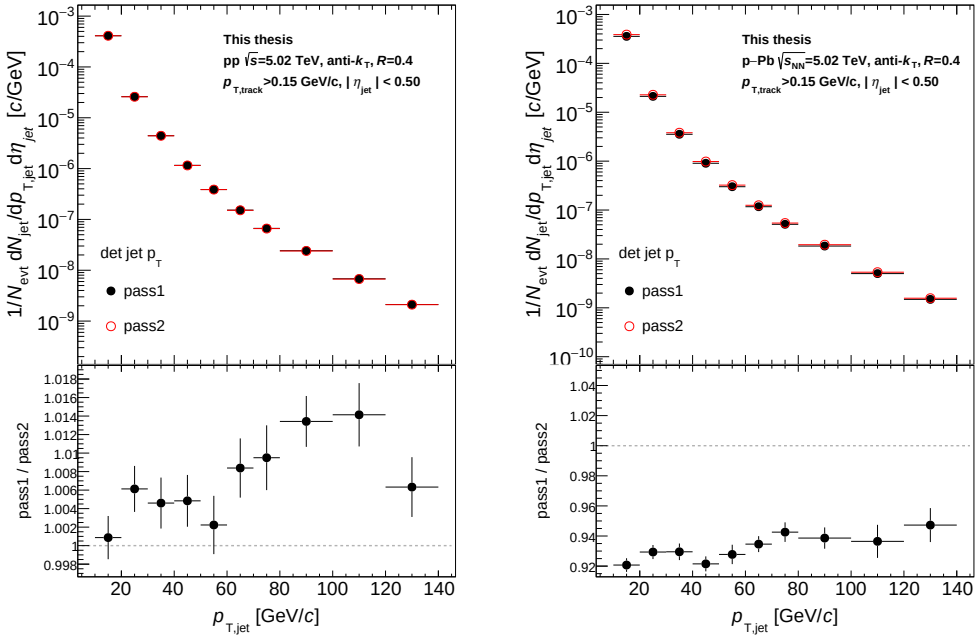


FIGURE 31 Jet transverse momentum spectra in the detector level pp (left) and p–Pb (right) MC.

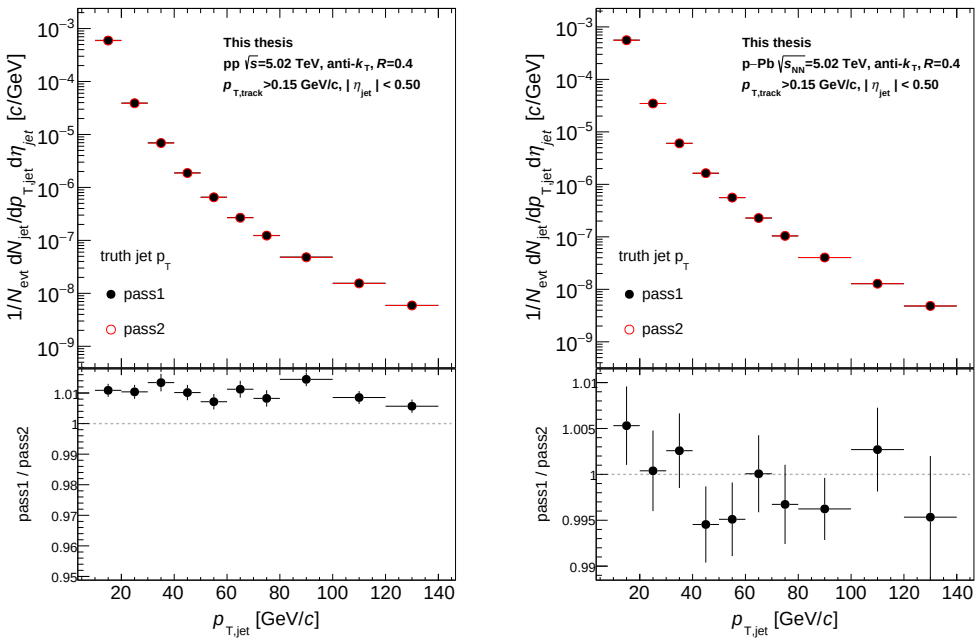


FIGURE 32 Jet transverse momentum spectra in the truth level pp (left) and p–Pb (right) MC.

for the new data pass, and no significant differences were seen in pass2 compared to pass1.

The raw dijet invariant mass spectrum is presented in Fig. 33. When looking at the ratio of p–Pb to pp, we note that the value of the ratio is of the same order as the number of nucleon-nucleon binary collisions $\langle N_{\mathrm{coll}} \rangle = 6.87$ [215] in the mass range under study in this analysis, between 65 and 150 GeV/c^2.

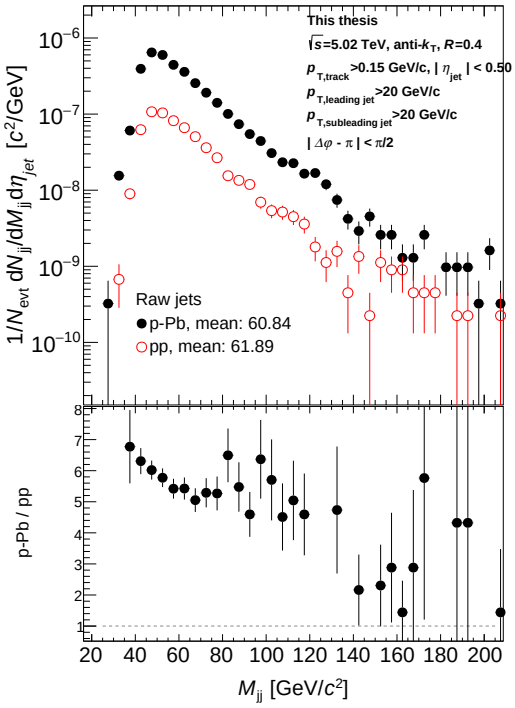


FIGURE 33 The raw dijet invariant mass spectrum formed from the raw jets in pass2 data, with no background subtraction.

5.3 Underlying event fluctuations

The dijet mass is calculated from jets which have had their background already subtracted from them, by using the background-subtracted four-momentum of the jets p^{μ}_{corr} from Eq. (32) in Eq. (36). This results in a M_{jj} distribution with background removed. However, the exact background at the location of an individual jet can be different from the overall event density determined by ρ. Hence, the dijet mass distributions calculated using this method are affected by background fluctuations. To study these fluctuations, we perform the background density measurement in two different ways, and form a distribution from their difference. The method we use is a generalization, developed in [1], of the random cone method for the p_{T} of single jets which was described in Sec. 4.3.

We measure a local background density ρ' for the dijet system using rotated

cones with a radius $R_{\text{cone}} = 0.4$. Two rotated cones are defined such that the first cone is set perpendicular to the leading jet in the azimuthal plane, and the second cone is set perpendicular to the subleading jet. Both cones are rotated in the same direction so that the cones have the same $\Delta\phi$ as the original dijet. The pseudorapidity for each cone is the same as for the jets.

The local background p_{T} density ρ' is estimated separately for each cone as

$$\rho' \equiv \sum_{\text{cone}} p_{\text{T},i}/(\pi R^2_{\text{cone}}), \tag{37}$$

where the summed $p_{\text{T},i}$ are the transverse momenta of all charged tracks within the cone. This density is used in Eq. (32) to obtain a four-momentum for each jet separately with the local background, including its fluctuations, subtracted. These momenta are then used to calculate a dijet mass, notated as M'_{jj}, which now has the background and fluctuations subtracted. The size of the fluctuations are then estimated with the distribution

$$\delta M_{\text{jj}} \equiv M_{\text{jj}} - M'_{\text{jj}}, \tag{38}$$

where M_{jj} is calculated using the default background density ρ.

There are rare cases where the rotated cone overlaps with a hard jet in the event. These cases result in a significantly long tail at values of $\delta M_{\text{jj}} > 0$, when M'_{jj} over-corrects the background, as there is another jet inside the cone increasing the value of ρ' significantly. All events which have a negative M'_{jj} are rejected from the δM_{jj} distribution, as these cases do not reflect true background fluctuations. Also events in which a cone overlaps with either the leading or subleading jet are rejected from the δM_{jj} distribution. In Fig. 34, the distributions of δM_{jj} are shown in two different bins: $\delta M_{\text{jj}} > 85$ GeV$/c^2$ and $\delta M_{\text{jj}} < 85$ GeV$/c^2$. In the furthest δM_{jj} bins of the histogram, large fluctuations are enhanced in the high M_{jj} bin.

The full distributions for the underlying event fluctuations δM_{jj} of the dijet mass in pp and p–Pb are shown in Fig. 35. The δM_{jj} distribution peaks near zero, and is asymmetric around the mean μ. Thus, we fit the distribution with an asymmetric generalized Gaussian distribution:

$$f(\delta M_{\text{jj}}) = C \times \begin{cases} \exp\left\{\left[-\left(\frac{|x-\mu|}{\sigma^+}\right)^{p^+}\right]\right\}, & x-\mu \geq 0 \\ \exp\left\{\left[-\left(\frac{|x-\mu|}{\sigma^-}\right)^{p^-}\right]\right\}, & x-\mu < 0 \end{cases}. \tag{39}$$

This allows us to have different width parameters ($\sigma^{\pm}$) and powers ($p^{\pm}$) for positive ($x-\mu \geq 0$) and negative ($x-\mu < 0$) sides of the mean. The fit parameters from the asymmetric Gaussian distribution are presented in Table 3. For the p–Pb fit, we wanted to limit the behavior of the tail at very large δM_{jj}, in order for the pp and p–Pb fit tails to never overlap. To achieve this, the parameter p^+ was set to the range $0.5 - 0.6$. The fit parameter ends up being at the upper limit of this range.

Later, in the unfolding procedure, the fit of the δM_{jj} distribution is used instead of the measured distribution, in order to avoid statistical fluctuation effects in the unfolding procedure.

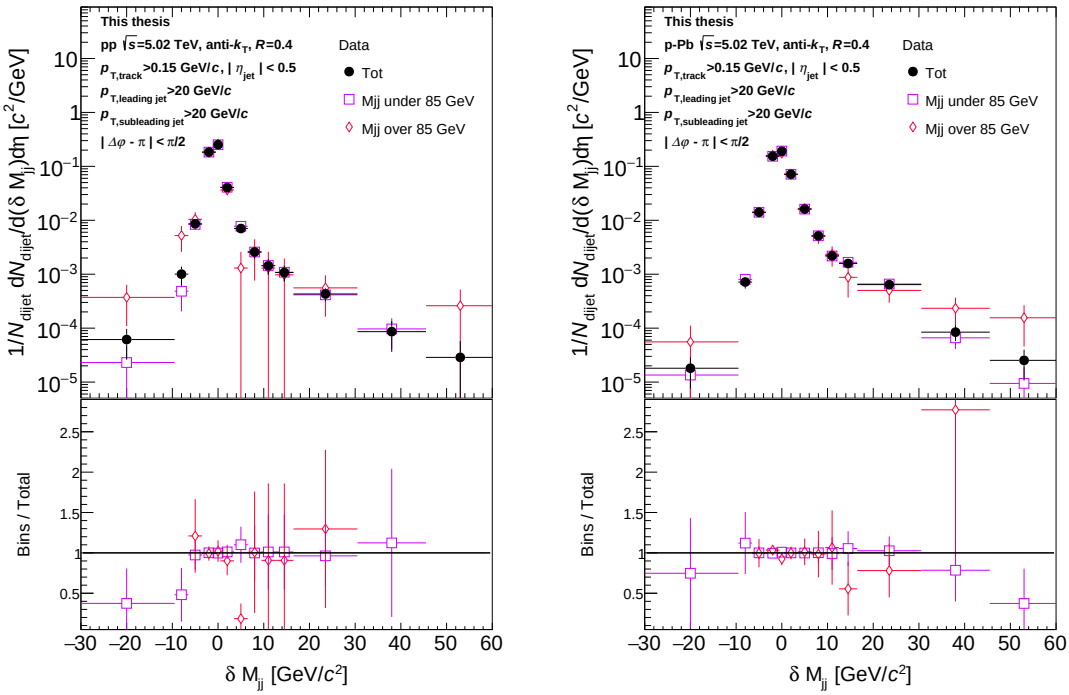


FIGURE 34 The underlying event fluctuations δM_{jj} of the dijet mass for pp (left) and p–Pb (right) in bins of dijet mass.

TABLE 3 The fit parameters for the background fluctuations shown in Fig. 35, see text for details. Parameter errors are as given by the ROOT fit.

	C	μ [GeV/c^2]	σ^- [GeV/c^2]	σ^+ [GeV/c^2]	p^-	p^+	χ^2/ndf
pp	0.67 ± 0.10	-0.26 ± 0.07	1.17 ± 0.19	0.28 ± 0.09	1.09 ± 0.12	0.53 ± 0.07	0.63
p–Pb	0.347 ± 0.013	-0.26 ± 0.03	1.92 ± 0.10	0.77 ± 0.02	1.36 ± 0.07	0.60(at limit)	2.22

5.4 Differential cross section

The differential cross section of the dijet invariant mass is calculated as

$$\frac{\mathrm{d}^2\sigma_{\mathrm{dijet}}}{\mathrm{d}M_{\mathrm{jj}}} \equiv \frac{\sigma_{\mathrm{V0AND}}}{N_{\mathrm{V0AND}}^{\mathrm{tot\ evt}}} \frac{\mathrm{d}N_{\mathrm{dijet}}}{\mathrm{d}M_{\mathrm{jj}}\mathrm{d}\eta_{\mathrm{jet}}}. \tag{40}$$

Here, $\mathrm{d}N_{\mathrm{dijet}}$ is the number of dijets in a mass bin defined by a bin width of $\mathrm{d}M_{\mathrm{jj}}$. The visible cross section σ_{V0AND} is as described in Sec. 2.2.2, $N_{\mathrm{V0AND}}^{\mathrm{tot\ evt}}$ is the number of all events measured by ALICE with the V0-AND trigger, and $\mathrm{d}\eta_{\mathrm{jet}}$ represents the jet acceptance in pseudorapidity, equal to 1.0 in this analysis.

Some events are not accounted for in the visible cross section due to the primary collision vertex not being reconstructed for these events. Therefore, the dijet invariant mass cross section has to be calculated using the number of events measured with the vertex reconstructed, $N_{\mathrm{V0AND}}^{\mathrm{has\ vtx}}$, and correct for this difference

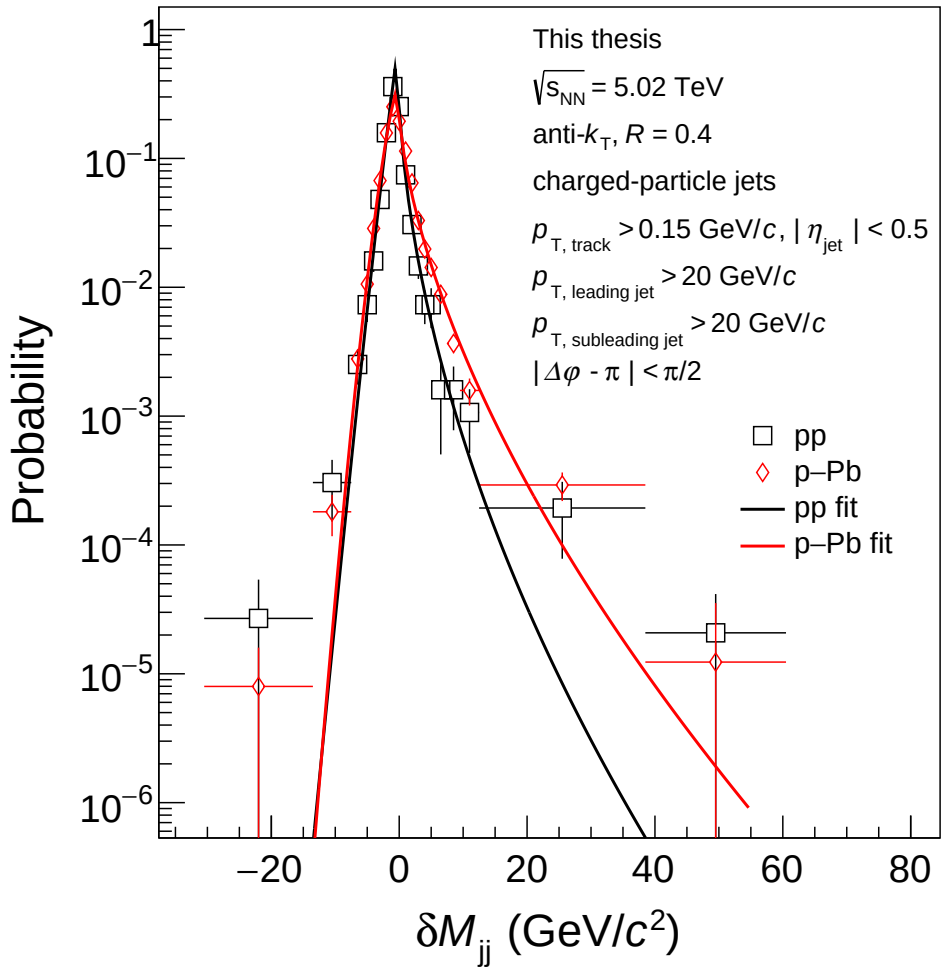


FIGURE 35 The underlying event fluctuations δM_{jj} of the dijet mass for both pp (black) and p–Pb (red). The parameters of the asymmetric generalized Gaussian fit can be found in Table 3. This fit is used to generate the unfolding matrix corresponding to the fluctuations.

using the vertex correction factor ϵ_{vtx}, as defined in Sec. 2.2.1. This results in the equation

$$\frac{\mathrm{d}^2\sigma_{\mathrm{dijet}}}{\mathrm{d}M_{\mathrm{jj}}} \equiv \epsilon_{\mathrm{vtx}} \frac{\sigma_{\mathrm{V0AND}}}{N_{\mathrm{V0AND}}^{\mathrm{has\ vtx}}} \frac{\mathrm{d}N_{\mathrm{dijet}}}{\mathrm{d}M_{\mathrm{jj}}\mathrm{d}\eta}. \tag{41}$$

The visible cross section divided by the number of V0AND triggered events in Eq. (41) is a close approximation to the inelastic differential cross section divided by the number of all events. This is due to the fact that events not triggered by the minimum bias trigger seldom contain a jet, as having a jet in the event is correlated with a large number of tracks. Therefore, the number of dijets, $\mathrm{d}N_{\mathrm{dijet}}$, is nearly the same in both inelastic and visible cross sections. An exception to this highlighted in [1] is central diffractive events, in which a dijet is created diffractively with both projectile particles remaining intact. There have been Monte Carlo studies [216] performed by CMS-TOTEM [217] where a visible cross section of 26.1 pb was estimated for dijets with a mass of over 100 GeV/c^2 or over in central diffractive events. This implies a small contribution to dijets not measured with the V0-AND trigger.

5.5 Nuclear modification factor

In order to study the parton energy loss in a heavy-ion system, we generally use Eq. 5 to define the nuclear modification factor. Experimentally, and for the purposes of this analysis, the nuclear modification factor for p–Pb collisions as a function of dijet invariant mass becomes

$$R_{\mathrm{pA}}(M_{\mathrm{jj}}, \eta_{\mathrm{jet}}) = \frac{\frac{1}{N^{\mathrm{pA}}} \frac{\mathrm{d}N^{\mathrm{pA}}}{\mathrm{d}M_{\mathrm{jj}}\mathrm{d}\eta}}{\left\langle N_{\mathrm{coll}}^{\mathrm{pA}} \right\rangle \frac{1}{N^{\mathrm{pp}}} \frac{\mathrm{d}N^{\mathrm{pp}}}{\mathrm{d}M_{\mathrm{jj}}\mathrm{d}\eta_{\mathrm{jet}}}}, \tag{42}$$

where $N^{\mathrm{pp/pA}}$ are the number of events for pp and p–Pb collisions and $\left\langle N_{\mathrm{coll}}^{\mathrm{pA}} \right\rangle$ is the average number of inelastic nucleon-nucleon collisions in the p–Pb events. In minimum bias p–Pb collisions, with no selection on the multiplicity of the event, the average number of inelastic collisions can be rewritten to be [218]

$$\left\langle N_{\mathrm{coll}}^{\mathrm{pA}} \right\rangle = A \frac{\sigma_{\mathrm{inel}}^{\mathrm{pp}}}{\sigma_{\mathrm{inel}}^{\mathrm{pA}}}, \tag{43}$$

where A is the mass number of the nucleus, in this case $A = 208$ for the lead nucleus. Combining Eq. (41) and Eq. (43), we finally write our nuclear modification factor as

$$R_{\mathrm{pA}}(M_{\mathrm{jj}}, \eta_{\mathrm{jet}}) = \frac{\frac{\mathrm{d}^2\sigma^{\mathrm{pA}}}{\mathrm{d}M_{\mathrm{jj}}\mathrm{d}\eta_{\mathrm{jet}}}}{\frac{\mathrm{d}^2\sigma^{\mathrm{pp}}}{\mathrm{d}M_{\mathrm{jj}}\mathrm{d}\eta_{\mathrm{jet}}}}. \tag{44}$$

5.6 Unfolding

An unfolding procedure is required to account for background fluctuations and instrumental effects in the measurement, as described in Sec. 4.2. The procedure translates the measured dijet mass distribution into the physical cross section, allowing us to compare with theory. A response matrix mapping the true and detector level distributions is constructed using a MC simulation.

Simulated events were generated with PYTHIA [196], using the tune Monash 2013 for pp and tune Perugia 2011 for p–Pb, in which only hard QCD interactions were selected. These simulations were run in separate bins of $\hat{p}_{\mathrm{T}}$ to enhance the production of dijets. In the case of truth level particles, the dijet mass distribution is constructed from all final charged hadrons given by PYTHIA in the fiducial acceptance. For the detector level tracks, the final PYTHIA particles were transported through ALICE using GEANT [211, 212], and the full simulation of detectors and track reconstruction was performed.

The dijet mass analysis was performed on these detector level tracks in the same manner as for the data. A dijet mass unfolding matrix, $\hat{M}_{\mathrm{det}}$, was filled with the masses of dijets that were found for both the truth and detector level dijets matched with the criteria $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} < 0.3$, with a match required for both the leading and subleading jet of the event. If the leading and subleading jet switch roles in the matching process, the dijet event is still considered a match. Fakes and misses have been taken into account in the unfolding. Fakes describe the cases where a dijet appears in the detector level MC but not in the truth level. Misses on the other hand describe cases in which truth level dijets are not seen in the detector level. The fractions of fakes and misses have been calculated by taking the projections of the detector response to its detector or truth level axis, and comparing them to the detector or truth level spectra, respectively. This has been described in detail in Ref. [1].

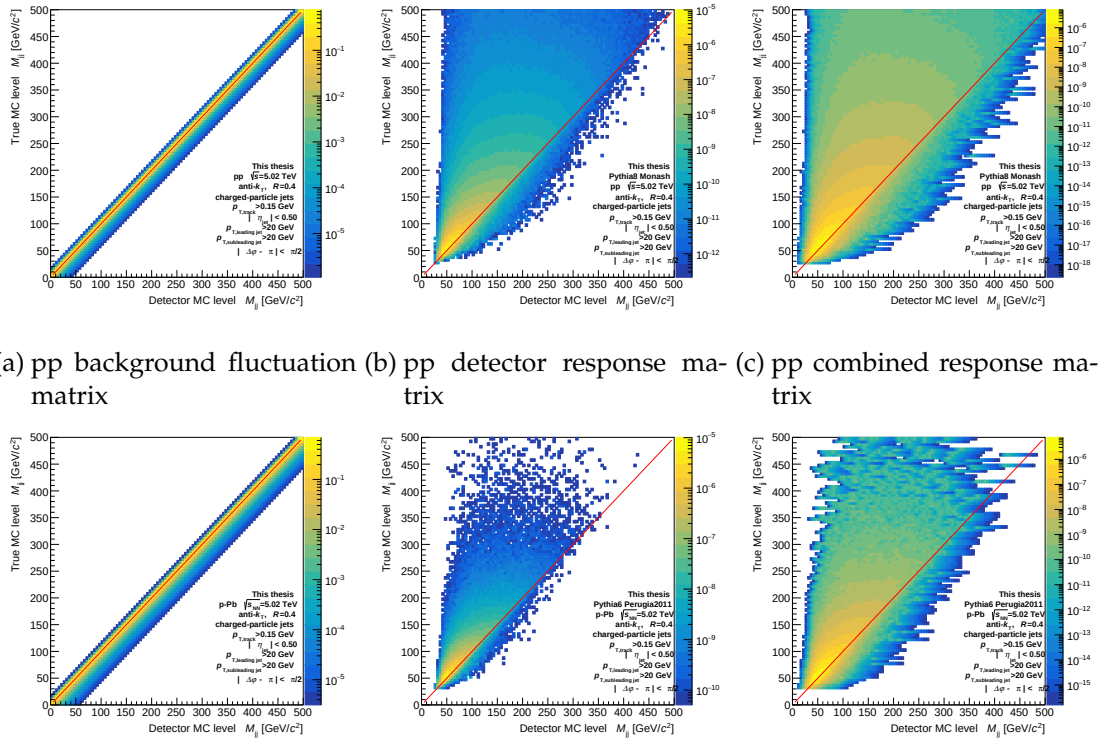
The unfolding matrix $\hat{M}_{\mathrm{det}}$ contains only information on instrumental detector effects. We want to include the background fluctuations, as described in Sec. 5.3, in the unfolding procedure. Following methods used in previous ALICE jet analyses [191, 209, 219], we construct a matrix, $\hat{M}_{\mathrm{fluct}}$, to describe the fluctuations measured at the detector level. The 2-dimensional $\hat{M}_{\mathrm{fluct}}$ is constructed from the 1-dimensional distribution δM_{jj}. For each y-axis bin of $\hat{M}_{\mathrm{fluct}}$, corresponding to the truth level M_{jj}, we copy the δM_{jj} histogram bin-by-bin. The bin of M_{jj} that is centered around zero is assigned to the diagonal of the $\hat{M}_{\mathrm{fluct}}$ histogram. After this, the $\hat{M}_{\mathrm{fluct}}$ histogram is re-binned to the binning used in the unfolding procedure. Then, each $\hat{M}_{\mathrm{fluct}}$ y-axis bin is scaled separately to represent a probability distribution.

With these two unfolding matrices now constructed, we can multiply them to obtain the total unfolding matrix

$$\hat{M}_{\mathrm{tot}} = \hat{M}_{\mathrm{fluct}}\hat{M}_{\mathrm{det}}. \tag{45}$$

The unfolding matrices corresponding to the background fluctuations for pp and

p–Pb are shown in Fig. 36 (left). In the middle of the figure are the unfolding matrices corresponding to the detector response. The unfolding matrices used for the final unfolding, obtained from Eq. (45), are shown on the right-hand side of the same figure. These final unfolding matrices are also pictured in Fig. 37, with the asymmetric binning used in the final unfolding procedure.



(a) pp background fluctuation matrix (b) pp detector response matrix (c) pp combined response matrix

(d) p–Pb background fluctuation matrix (e) p–Pb detector response matrix (f) p–Pb combined response matrix

FIGURE 36 The unfolding matrices of pp and p–Pb.

As described in Sec. 4.2, we choose the Bayesian unfolding algorithm [181, 183] as our default unfolding algorithm, while the SVD unfolding algorithm [184] is used to evaluate the systematic uncertainty arising from the choice of the unfolding procedure. The unfolding parameter for the Bayesian unfolding algorithm, *nIter*, describes the number of iterations used in the unfolding process. The parameter for pp is selected to be 5, and for p–Pb , the parameter is 3. These values have been selected by studying the unfolding performance at varying values of *nIter*, and selecting the smallest value for which the results do not differ significantly when compared to results obtained with $nIter - 1$ and $nIter + 1$. The SVD unfolding algorithm requires setting a regularization value, the k-value, which is selected to be 9 and 5 for pp and p–Pb , respectively. It is possible that the optimum unfolding parameters can change when the data reconstruction pass is updated. However, we found that in this analysis, the optimum choice of unfolding parameters for the pass2 datasets corresponds to the choice for the pass1 datasets [1].

In order to study whether the unfolding works as expected, the unfolded

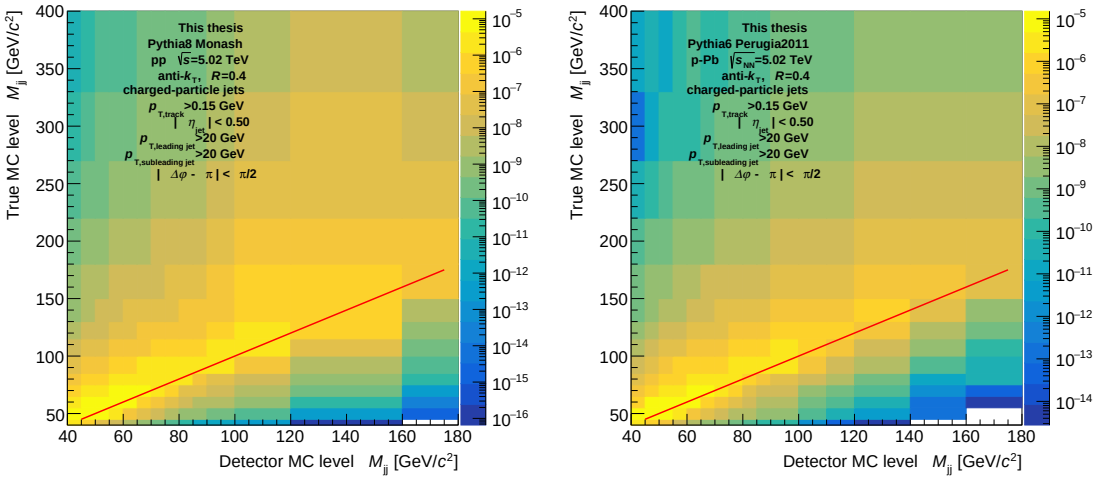


FIGURE 37 The unfolding matrices of pp and p–Pb with the binning used in the final unfolding procedure. The binning used is asymmetric.

detector level MC spectra were compared to the truth level MC spectra. Figure 38 shows the comparison of unfolded detector level MC spectra to truth level spectra. The test is performed using a set of run numbers from each of the $\hat{p}_{\mathrm{T}}$ bins set aside to only be used as a closure test, not in the creation of the unfolding matrix. The closure test is a success, as the ratio is to a great extent equal to unity within uncertainties. A trivial closure test was also conducted using the same dataset as the unfolding matrix was generated with, therefore called trivial. This test also resulted in a ratio agreeing with unity.

Another validation check of the unfolding is to refold the already unfolded spectrum. We want to check whether the unfolded detector level MC spectrum can be refolded back to the detector level result, which we can then compare to the original detector level spectrum. As unfolding requires a matrix inversion, folding is a much simpler process, requiring just a matrix multiplication. The result of such a refolding test is shown in Fig. 39. As can be seen from the figure, the refolded detector level spectrum is consistent with the original detector level spectrum, therefore validating our unfolding procedure.

Even with the $p_{\mathrm{T}} < 4\hat{p}_{\mathrm{T}}$ cut implemented, an outlier was found in the p–Pb MC dataset in the dijet mass bin 115–130 GeV/c^2. The outlier was found when studying the closure test of the p–Pb MC, Fig. 40 (left), and seen in Fig. 40 (right), where I compare the raw dijet spectrum as a function of dijet mass. A single outlier-dijet was then found in the $\hat{p}_{\mathrm{T}}$ bin 11–21 GeV/c of the p–Pb MC production. Out of just five dijets found in this $\hat{p}_{\mathrm{T}}$ bin of the MC production, one was at a much higher dijet mass $M_{\mathrm{jj}} = 118$ GeV/c^2 than the others, which were around 50 GeV/c^2.

If we assume, naively, that the dijet mass is of the order $2p_{\mathrm{T,jet}}$, the $4\hat{p}_{\mathrm{T}}$-cut would correspond to cutting on the dijet mass as $M_{\mathrm{jj}} < 8\hat{p}_{\mathrm{T}}$. I took the mean of the $\hat{p}_{\mathrm{T}}$ distribution of this bin, 13.74 GeV/c^2, as the cut, as the event-by-event information was not available at this stage of the analysis. I applied the dijet mass cut on the $\hat{p}_{\mathrm{T}}$ bin 11–21 GeV/c only, and, out of the five dijets in the bin,

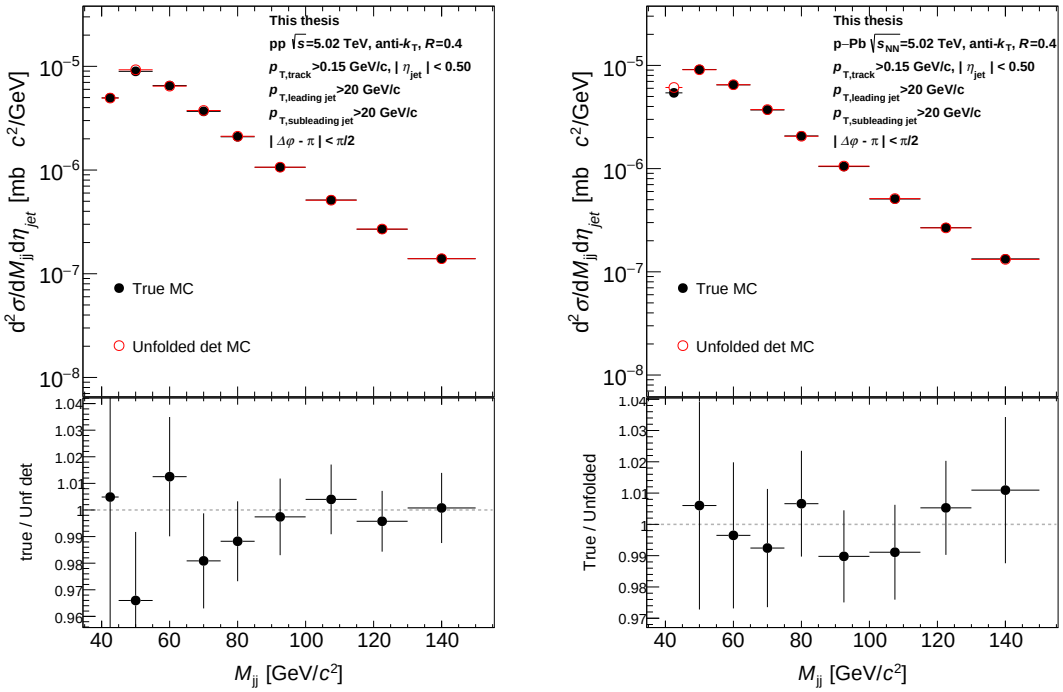


FIGURE 38 Unfolded detector level MC spectra are compared to the truth level MC spectra for pp (left) and p–Pb(right) simulations. A statistically separate dataset not used in the creation of the unfolding matrix was used for this test. The default number of unfolding iterations are used.

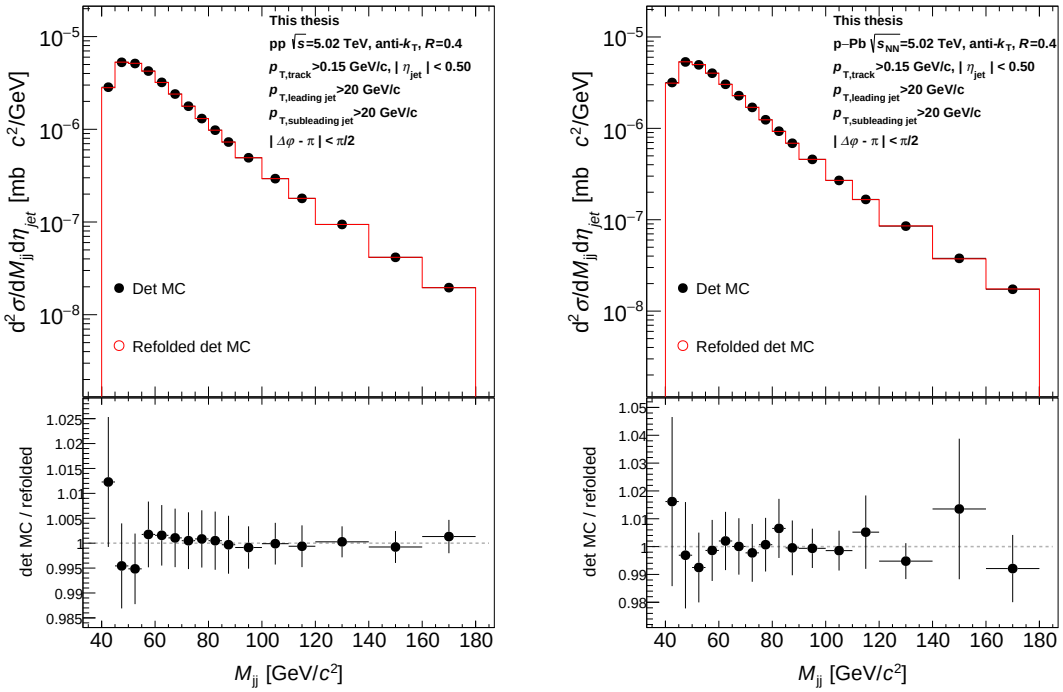


FIGURE 39 A refolding check performed for the dijet invariant mass in pp (left) and p–Pb (right). This check folds the already unfolded detector-level MC spectrum, and then compares the refolded spectrum with the original detector-level MC spectrum.

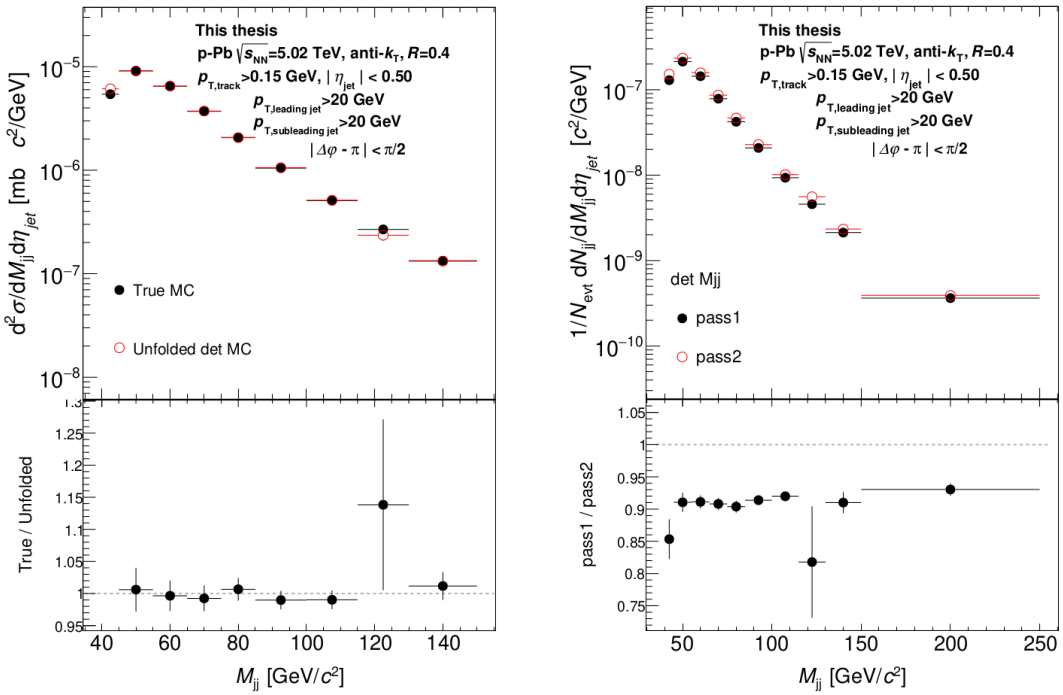


FIGURE 40 Left: the closure test of the p–Pb before removing the outlier. Right: comparing the detector level dijet mass spectra in the Monte Carlo passes before outlier removal.

the outlier was discarded from the analysis. The results presented use this outlier cut.

5.7 Systematic uncertainty

Sources of systematic uncertainty can be gauged by systematically altering parameters, methods or the data, and studying the impact of the change on the analysis results. If the effect presents a systematic difference as compared to the default results, we can assign the change of method or parameter as a systematic uncertainty source. In the context of this dijet mass analysis, we divide the sources of systematic uncertainty into two main sources: tracking efficiency and unfolding. The default settings are first selected. Then, each source of uncertainty is varied systematically, after which they are compared to the default case, separately for pp, p–Pb and their ratio R_{pPb}.

5.7.1 Tracking uncertainty

The performance of the tracking has been studied and estimated within ALICE, using a full MC simulation accounting for realistic experimental tracking and quality selections. The variations of the tracking quality with changes in the

track quality cuts have been studied and determined by the ALICE data processing group. We use these recommended variations to calculate the tracking uncertainty for this analysis. The tracking uncertainty for the data sets used in this analysis have been estimated to be 3% for pp collision data, and a varying value between 1% and 2.5% for p–Pb, dependent on the track p_{T}. The precise values used in this analysis can be found from the publicly available AliPhysics code repository [214] in file *TrackEfficiencyConfiguration.yaml*.

In order to study the effect of this uncertainty on the dijet mass analysis, we remove a percentage of particles in the detector level MC simulation output corresponding to the uncertainty approximated for the data. The unfolding matrix that is formed from this MC simulation is used to unfold the data, and the result is then compared to the data unfolded with the original unfolding matrix. The result of this comparison is seen in Fig. 41, in which we can see a difference in the dijet mass spectrum between the two dijet mass spectra for both the pp and p–Pb datasets.

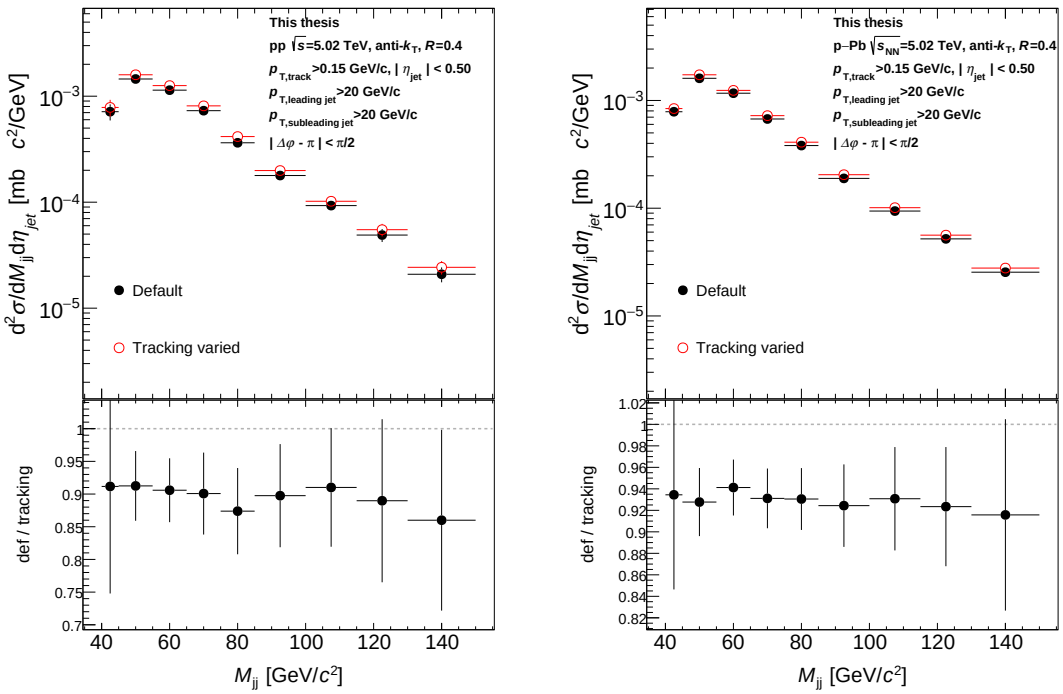


FIGURE 41 Tracking efficiency systematic uncertainty for pp and p–Pb. On the left, the pp spectrum has been varied, and on the right, the p–Pb has been varied.

The ratio from Fig. 41 tells us the effect of the extra track loss for any observable, in our case, the dijet invariant mass, ranging from 10% to 6%. It is not fully understood whether the tracking efficiency uncertainty is correlated between the pp and p–Pb datasets. Therefore, we do not assume that they would cancel in the ratio of the spectra. This means that our tracking-related systematic uncertainty is calculated in a conservative manner.

The systematic behavior of R_{pA} is studied by calculating a value first by

varying the tracking in pp as described above while keeping the default spectrum for p–Pb, and then vice versa. The results of these studies are shown in Fig. 42, where the systematic uncertainty results are compared and the larger of the two values is selected as the systematic tracking uncertainty for R_{pA}. The tracking systematic uncertainty for R_{pA} is shown in Fig. 43. As has been described, the tracking uncertainty has been evaluated by studying the loss of tracks, not by adding extra tracks. Therefore, as we assume the effect to be symmetrized in the case of extra tracks being included, we mirror the tracking uncertainty for positive contributions, for both pp and p–Pb spectra as well as the R_{pA}.

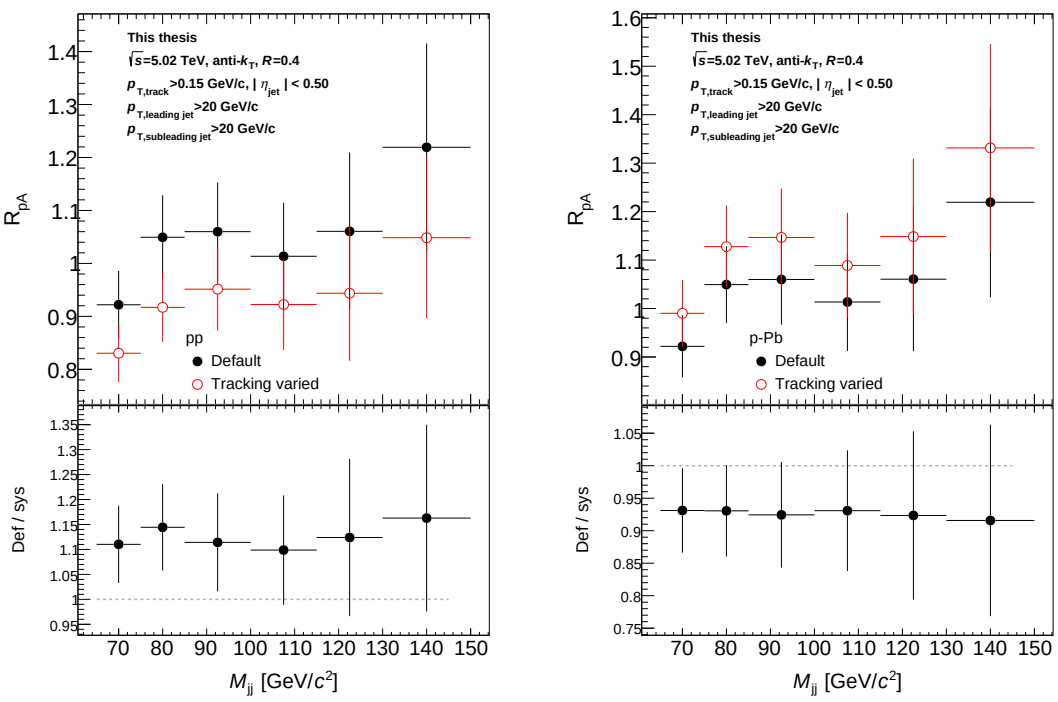


FIGURE 42 Tracking efficiency systematic uncertainty determination for the R_{pA}. On the left, the pp spectrum has been varied, and on the right, the p–Pb has been varied.

5.7.2 Unfolding uncertainty

The unfolding procedure described in previous sections includes choices on, for instance, the unfolding algorithm, number of iterations, and other parameters. These choices can all be sources of uncertainty, and therefore we perform a systematic check on all of these choices. A list of all the sources of systematic uncertainty related to the unfolding procedure are listed below, after which I explain all of them in some detail.

- Data truncation
- Closure
- Unfolding binning choice

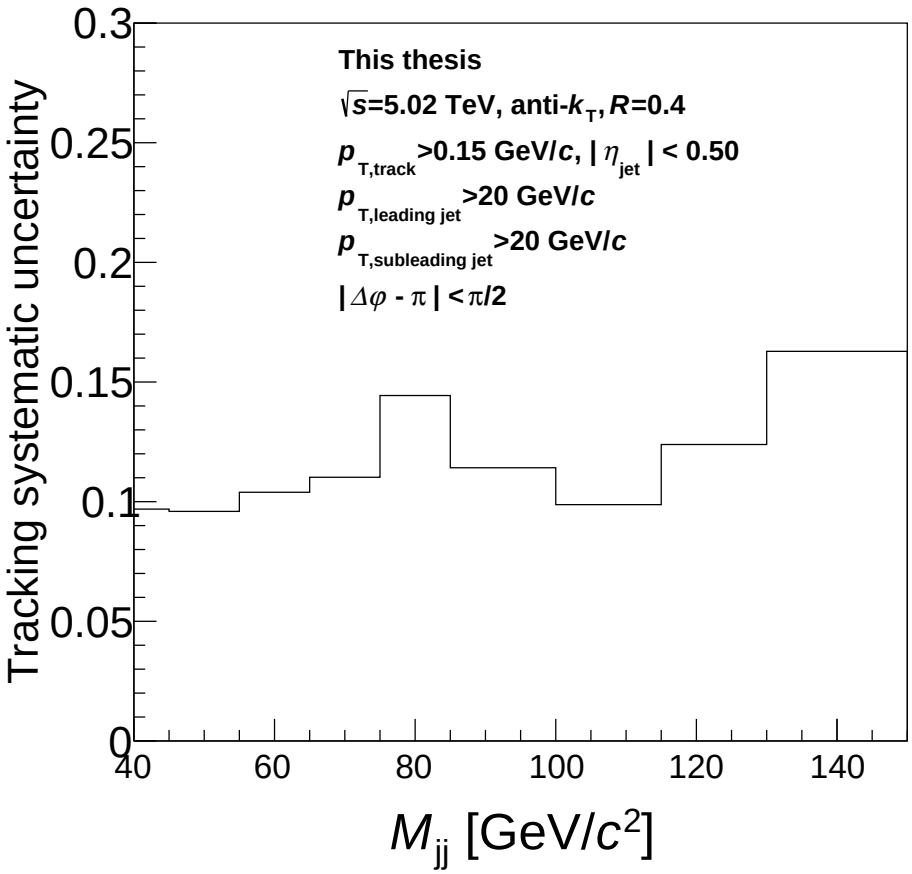


FIGURE 43 The final tracking efficiency systematic uncertainty for R_{pA}.

- Prior choice
- Unfolding algorithm
- Number of iterations

Data truncation The unfolding algorithms perform best when there are enough events in the unfolding bins. Therefore, we want to find a good cut-off for the mass bins, as there is a limited number of entries in the highest mass bins. We use a default value of 180 GeV/c^2, and in order to study the effect of this cut, we compare the resulting spectrum to one in which we instead use a cut-off of 160 GeV/c^2. The effect of this is shown in Fig. 44. As is seen, only the very last bins are affected by the cut. The default cut results in a value of R_{pA} that is higher than with the tighter cut, however, these results are within the statistical uncertainties. As is seen in the top left plot, the truncation affects the final bin of the pp spectrum the most, which then affects the fluctuation of the final bin of R_{pA}.

Closure As was described in Sec 5.6, a closure test was performed in order to validate the unfolding procedure. The closure test result using a statistically separated dataset to unfold the detector level MC, shown in Fig. 38 for the pp and p–Pb spectra, is used to calculate the R_{pA} for the truth level and unfolded detector level MC separately. The comparison of these is shown in Fig. 45, and the difference in the R_{pA} is taken as the systematic error related to the closure test of the unfolding procedure.

Unfolding binning choice In the analysis of the pass1 data, we had to find a suitable binning for the pass1 datasets. With the update of the reconstruction pass, I studied different sets of detector and truth level binning in the unfolding procedure to select a suitable combination to minimize systematic errors. I found that merging some bins together in the detector level binning improved the unfolding of the pp data, while the truth level bins were kept as in pass1. The binning choice used for unfolding the pp spectra was the following:

$$\text{Detector level} : \{40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 120, 160, 180\}$$
$$\text{Truth level} : \{40, 45, 55, 65, 75, 85, 100, 115, 130, 150, 180, 220, 270, 330, 400\}$$

For the p–Pb data, the same binning as was used for the pass1 dataset resulted in a successful unfolding of the pass2 dataset. This binning includes two more bins in the detector level than for the pp data:

$$\text{Detector level} : \{40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 110, 120, 140, 160, 180\}$$
$$\text{Truth level} : \{40, 45, 55, 65, 75, 85, 100, 115, 130, 150, 180, 220, 270, 330, 400\}$$

Systematic uncertainty related to changes in the binning choice are evaluated by selecting an altered binning which is a denser binning than the default described above. The altered binning is selected to be the following:

$$\begin{aligned}\text{Detector level} : \{&40, 43, 45, 50, 55, 60, 63, 65, 70, 75, 80, 85, 90, 95, 100,\\ &105, 110, 120, 130, 140, 150, 170, 180, 200\}\\ \text{Truth level} : \{&40, 43, 45, 50, 55, 65, 70, 75, 85, 92, 100, 108, 115, 122,\\ &130, 140, 150, 200, 250\}.\end{aligned}$$

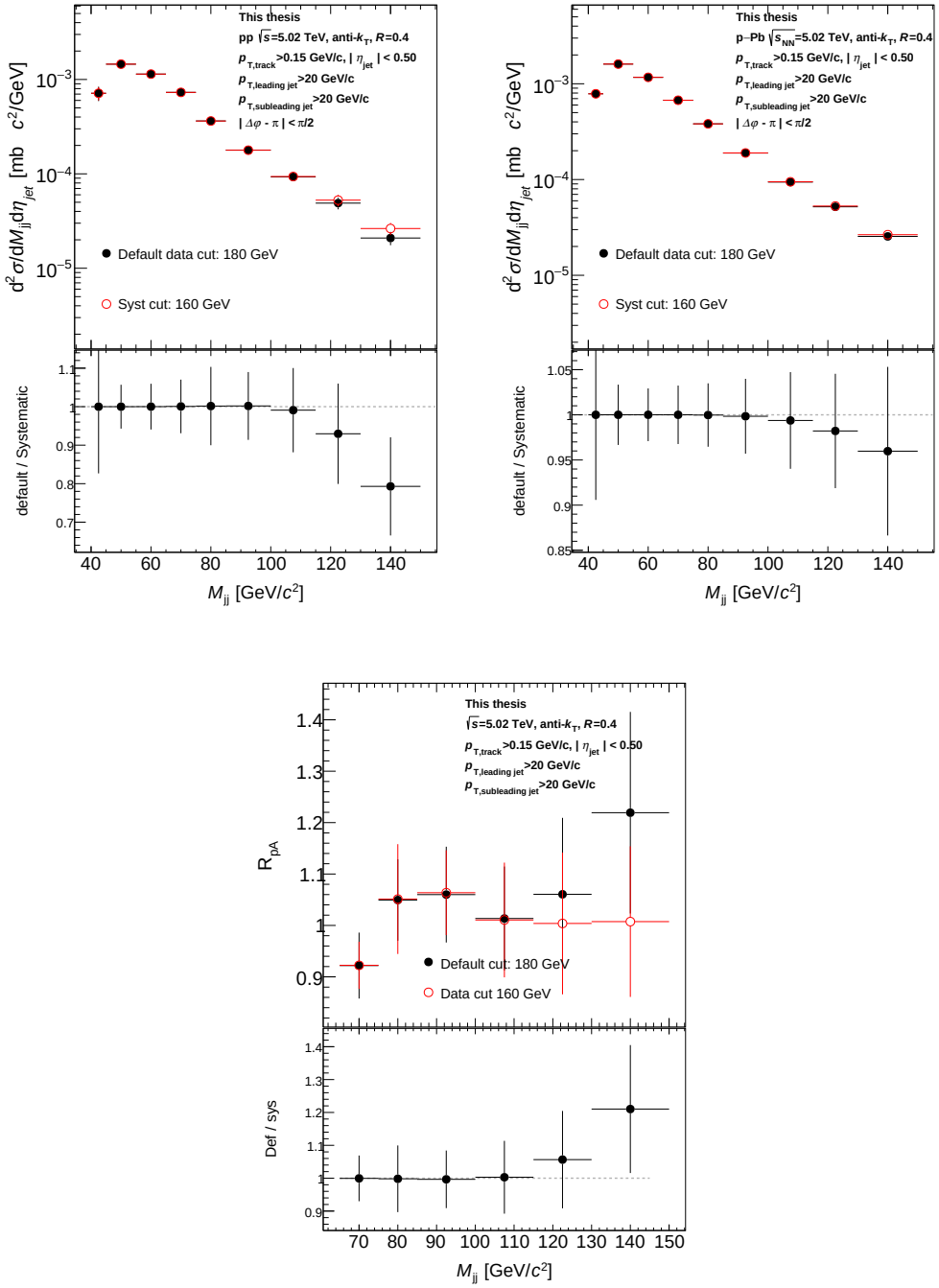


FIGURE 44 The truncation check of the unfolding procedure. The data is cut from above 180 GeV/c^2 by default, and we test how cutting the data above 160 GeV/c^2 affects the ratio.

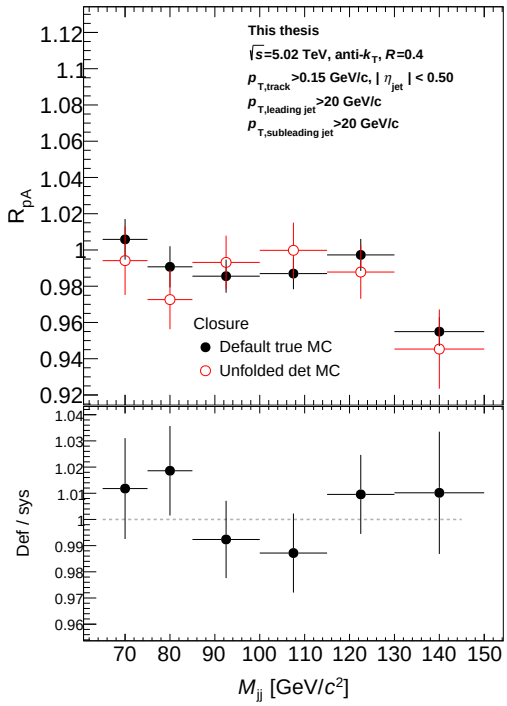


FIGURE 45 R_{pA} has been calculated for both the unfolded detector level MC and the truth level MC. A statistically separate dataset not used in the creation of the default unfolding matrix was used in the unfolding of the detector level MC.

The unfolding is performed with the altered binning and the results are compared with the default-binned unfolding results in Fig. 46. The final bin is the only one showing a notable difference between the binning, arising from the pp spectrum's final bin fluctuating with the binning change.

Prior choice The prior used in the Bayesian unfolding procedure can influence the unfolding results. The prior is therefore altered in order to gauge the systematic effect of the choice of the prior. We choose to weight the y-axis of the unfolding matrix with a once-unfolded data spectrum, using this as our guess on how the spectrum should appear, instead of the default procedure of using the MC truth level of the unfolding procedure.

The weights are taken from the once-unfolded data, separately for pp and p–Pb, and filled and scaled into histograms so that their integral matches the true spectrum in the $80 - 140$ GeV$/c^2$ mass area, in which the sample size is reasonable. The weight histogram is then divided by the truth level MC histogram, which is the default prior, to obtain a ratio between the default and the new prior. Finally, the ratio is used to scale each y-bin of the unfolding matrix to obtain a scaled default prior. The effect of the change of the prior is presented in Fig. 47. The final bin, again, is prone to fluctuation with the change of the prior. Otherwise the results differ by under 5%.

Unfolding algorithm The choice of the unfolding algorithm is also a source of systematic uncertainty. We take the Bayesian unfolding algorithm as the default method, and study the difference in the results when unfolding instead with the SVD algorithm. The comparison of the results is shown in Fig. 48. We find a good agreement of the unfolded results, with a slight increase in discrepancy in the final bin. The agreement is better in the p–Pb dataset, while for the pp dataset the differences between the two unfolding methods are slightly larger.

Number of iterations The number of iterations, *nIter*, has been selected for the Bayesian unfolding procedure to be 3. This number can be varied around itself by ± 1 to investigate the systematic effect of the parameter choice. This is shown in Fig. 49. The differences between the resulting R_{pA} are small, with the largest difference seen, again, in the final bin, of around 5 %.

The six sources of uncertainty resulting from or related to the unfolding procedure are depicted individually on a stacked histogram in Fig. 50. The sum of these histograms does not reflect the total unfolding systematical uncertainty. The major difference in the plots between passes is seen in the change of the main source of uncertainty in the 115–130 GeV$/c^2$ bin in the p–Pb histogram, and the final bin increasing in the pass2 pp and ratio histograms.

Before the change of binning, the total systematic uncertainty in the final pp dijet mass bin was over 0.25, but the systematic uncertainty in the ratio stayed similar. The last bin in the pp data is thus seen to be very prone to fluctuation, and this is visualized by the high systematic error of this last bin in the pp histogram in Fig. 50 (top right). In Fig. 51 are all the sources of systematic uncertainty related to unfolding similarly stacked for R_{pA}. These will be the final uncertainties used when combining the sources in the next section.

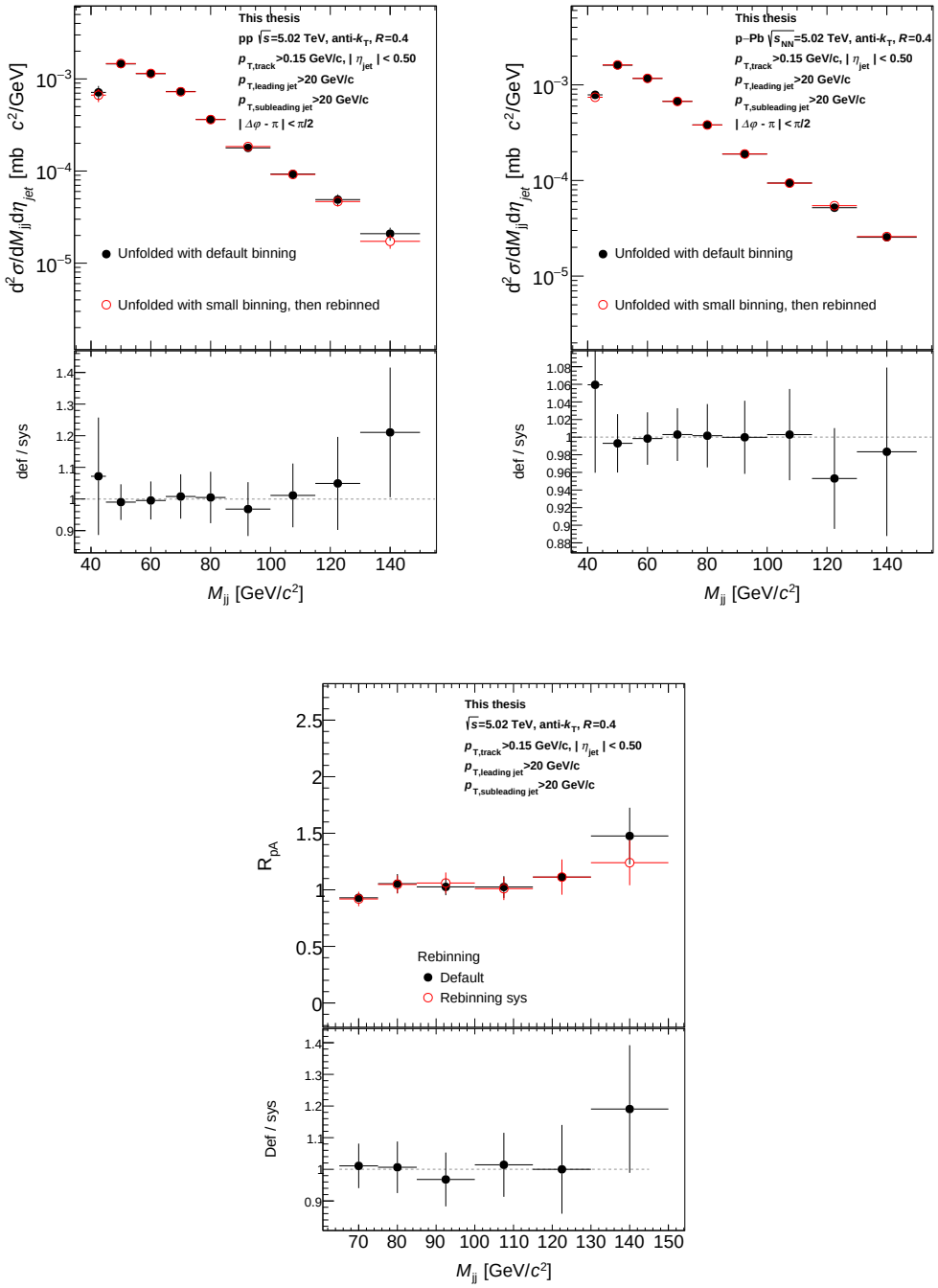


FIGURE 46 The effect of changing the binning in the unfolding, shown for pp (top left), p–Pb (top right) and the R_{pA} (bottom).

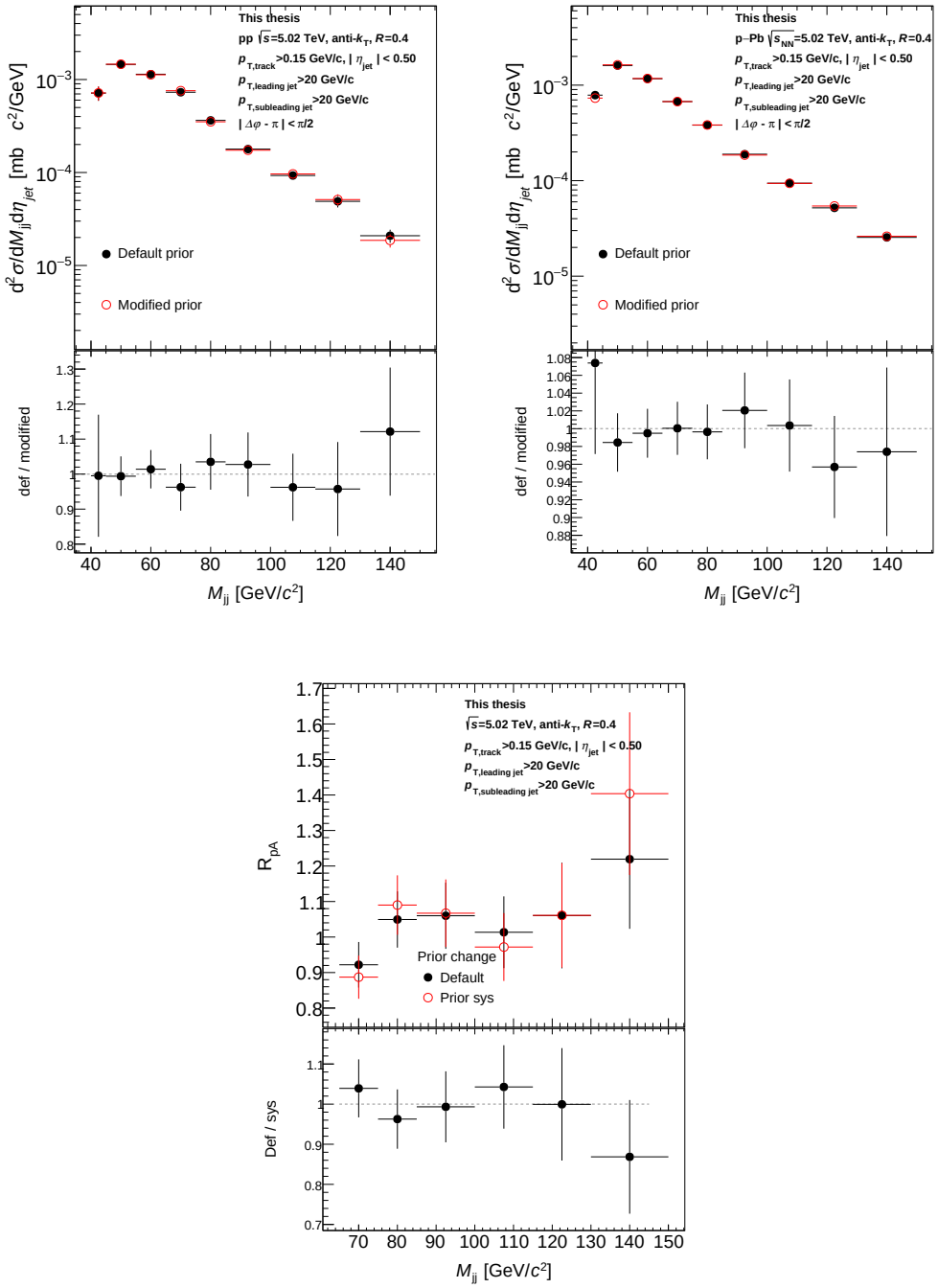


FIGURE 47 The effect of changing the prior of the unfolding, shown for pp (top left), p–Pb (top right) and the R_{pA} (bottom).

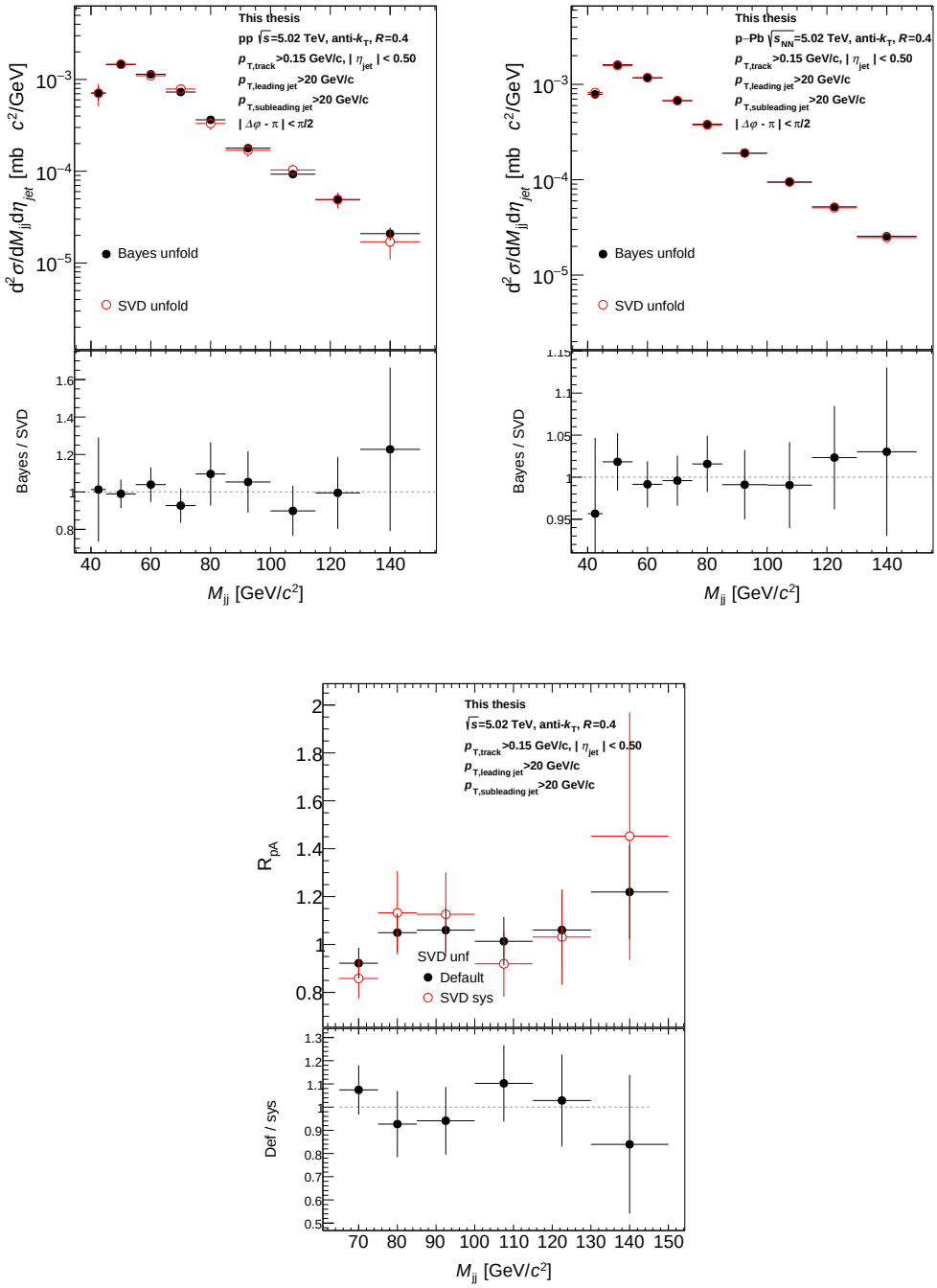


FIGURE 48 Comparing the Bayesian and SVD unfolding methods, shown for pp (top left), p–Pb (top right) and the R_{pA} (bottom).

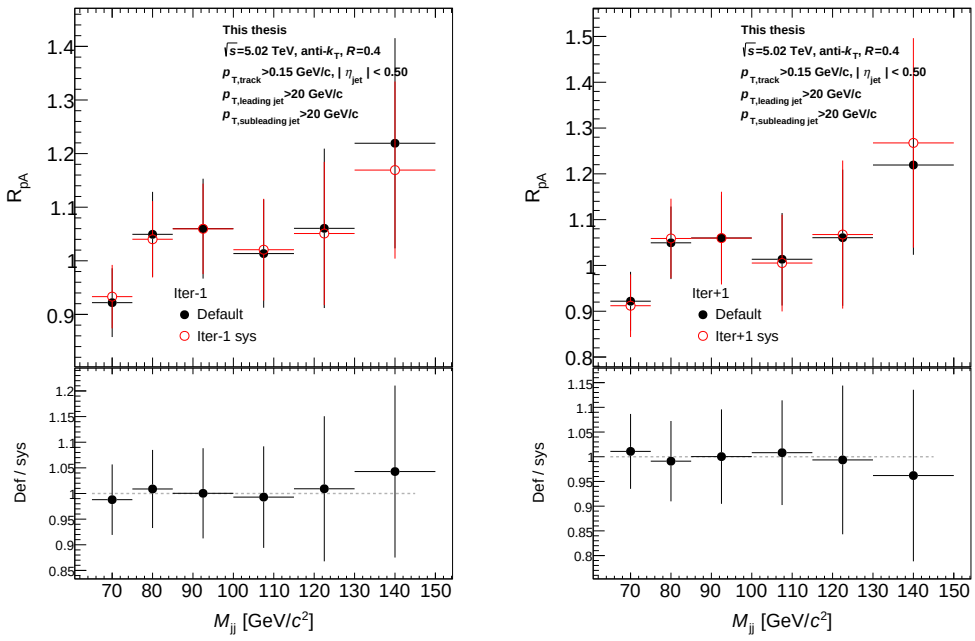


FIGURE 49 R_{pA} obtained by unfolding with the default number of iterations $nIter = 3$ compared to $nIter = 2$ (left) and $nIter = 4$ (right) iterations.

5.7.3 Total systematic uncertainty

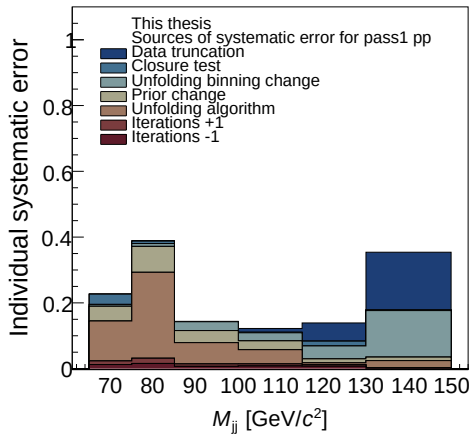
Following the reference [220], the total systematic uncertainties related to unfolding are combined as

$$\delta_{\text{unfolding}} = \sqrt{\frac{\sum_{i \in \text{unf syst}}^{N^{\text{tot unf}}} \delta_i^2}{N^{\text{tot unf}}}}, \tag{46}$$

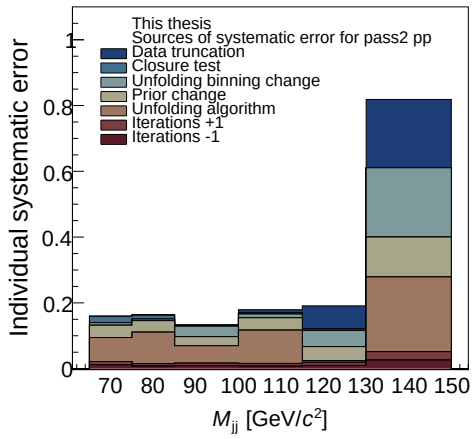
since each source of systematic uncertainty is comprised of independent measurements of the same underlying systematic uncertainty in the unfolding.

Both the total unfolding systematic and the tracking efficiency uncertainty are smoothed with the help of the ROOT `TH1::Smooth function`, which uses the 353QH smoothing algorithm twice. The algorithm is explained in detail in [221]. After smoothing the tracking and unfolding systematics, the total systematic uncertainty is calculated bin-by-bin by taking a square root of the sum of squares.

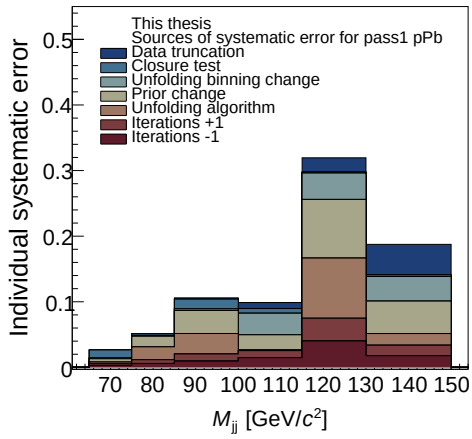
The total systematic uncertainties are presented in Fig. 52 for the updated pass2 results and the previous pass1 results for comparison. From the figure, we see that there is no significant change in the total systematics with the pass update aside from the final bin (130–150 GeV/c^2). In the final bin, the systematics of the p–Pb results are similar, but the pp results show a slight increase, with a greater increase in the effect from the unfolding (red dashed line) on the total systematics. In the systematic uncertainty of the ratio, the final bin stands out as the only major change – the systematic uncertainty arising from the combined unfolding systematics is nearly the same size as the contribution from the tracking systematic uncertainty, which increases the total systematic uncertainty from



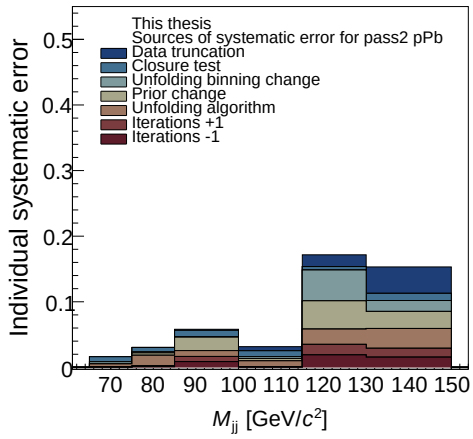
(a) pass1 pp



(b) pass2 pp



(c) pass1 p–Pb



(d) pass2 p–Pb

FIGURE 50 All of the unfolding-related systematic errors are presented simultaneously in a stacked histogram plot. Note that this is not the total unfolding systematic error, but the portions of each unfolding systematic error. The total unfolding systematic error is shown as a part of Fig. 52.

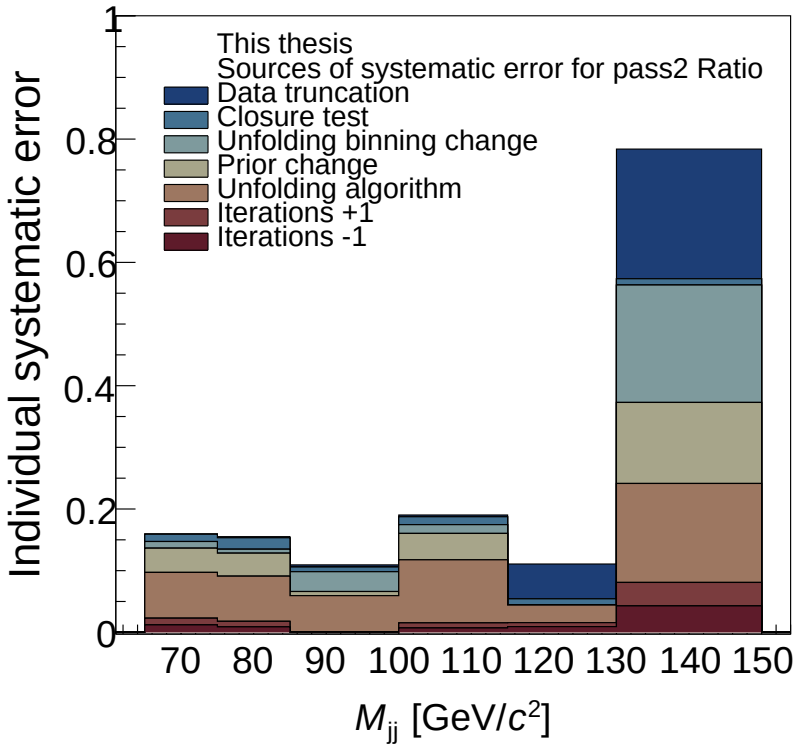


FIGURE 51 All of the unfolding-related systematic errors are presented simultaneously in a stacked histogram plot for R_{pA}.

around 0.16 to around 0.20.

5.8 Results

The dijet invariant mass cross section is calculated according to Eq. (41), and is presented in the top panel of Fig. 53. The pp spectrum has been scaled with the mass number of the lead ion, $A = 208$. In the bottom panel, the nuclear modification factor R_{pA}, defined in Eq. (44), is presented. The R_{pA} is found to agree with unity within uncertainties.

In order to compare the results with theory, we ran simulations for both pp and p–Pb collisions. The baseline pp simulations were a PYTHIA simulation, run using the CT14NLO [222] proton PDFs, and a simulation combining POWHEG and PYTHIA, using CT18ANLO [135] proton PDFs. In order to study cold nuclear matter effects resulting from nPDFs, these simulations were repeated with the proton PDF of the proton heading toward positive rapidity modified with nuclear modifications. For the PYTHIA simulation, the proton PDF was modified as described by EPPS16 [136], and for the POWHEG+PYTHIA simulation, it was modified according to EPPS21 [132].

The PYTHIA simulations were run in the $\hat{p}_T$ bins 15–30, 30–40, 40–65, 65–90, 90–120, 120–150, 150–200 and >200 GeV/c^2. In POWHEG, the rare hard events are

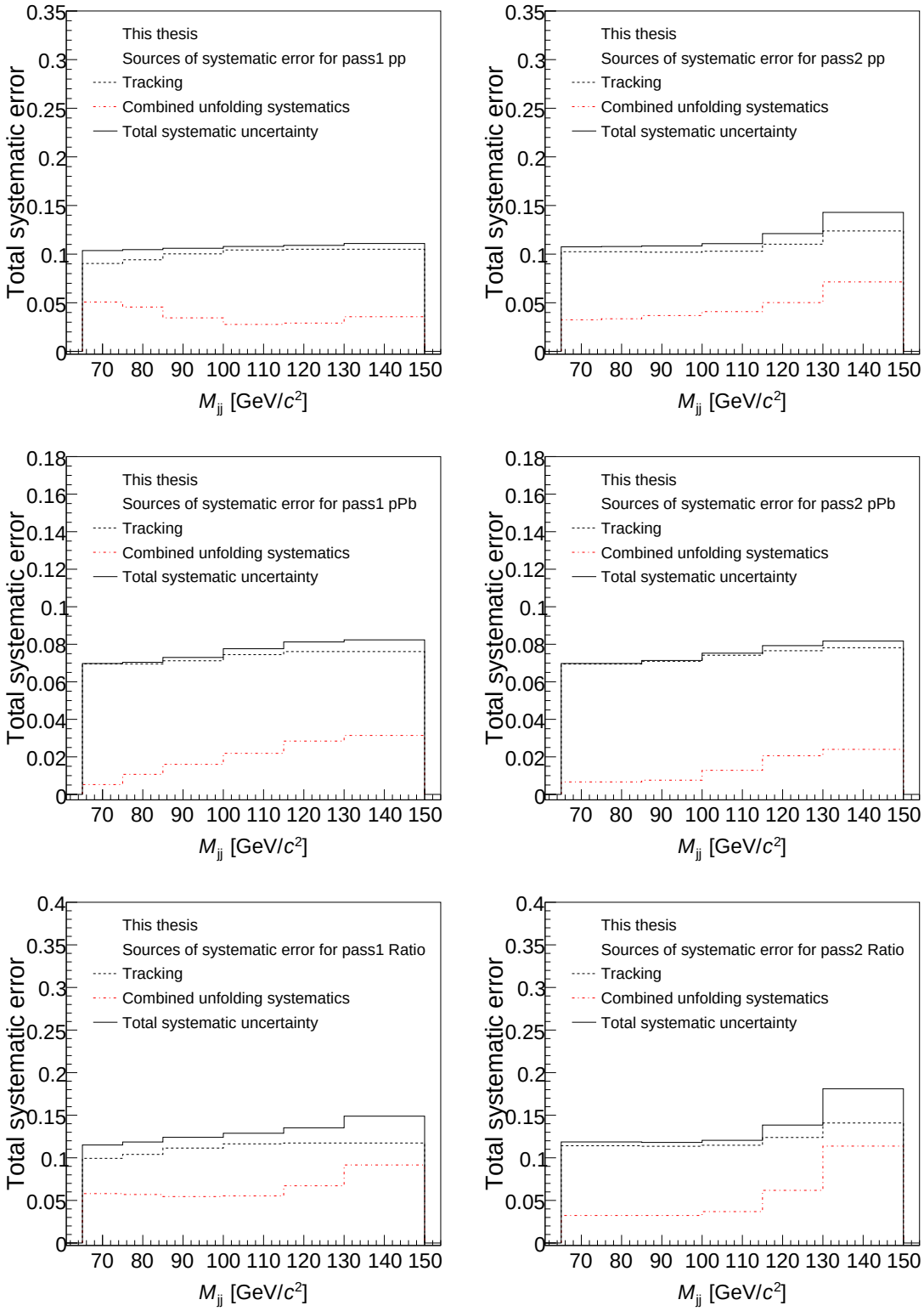


FIGURE 52 All systematic errors and the total systematic error after taking root of the sum of squares. Data from pass1 is on the left and pass2 on the right. To avoid fluctuations, the tracking and the combined unfolding systematic are smoothed out using ROOT's TH1::Smooth function.

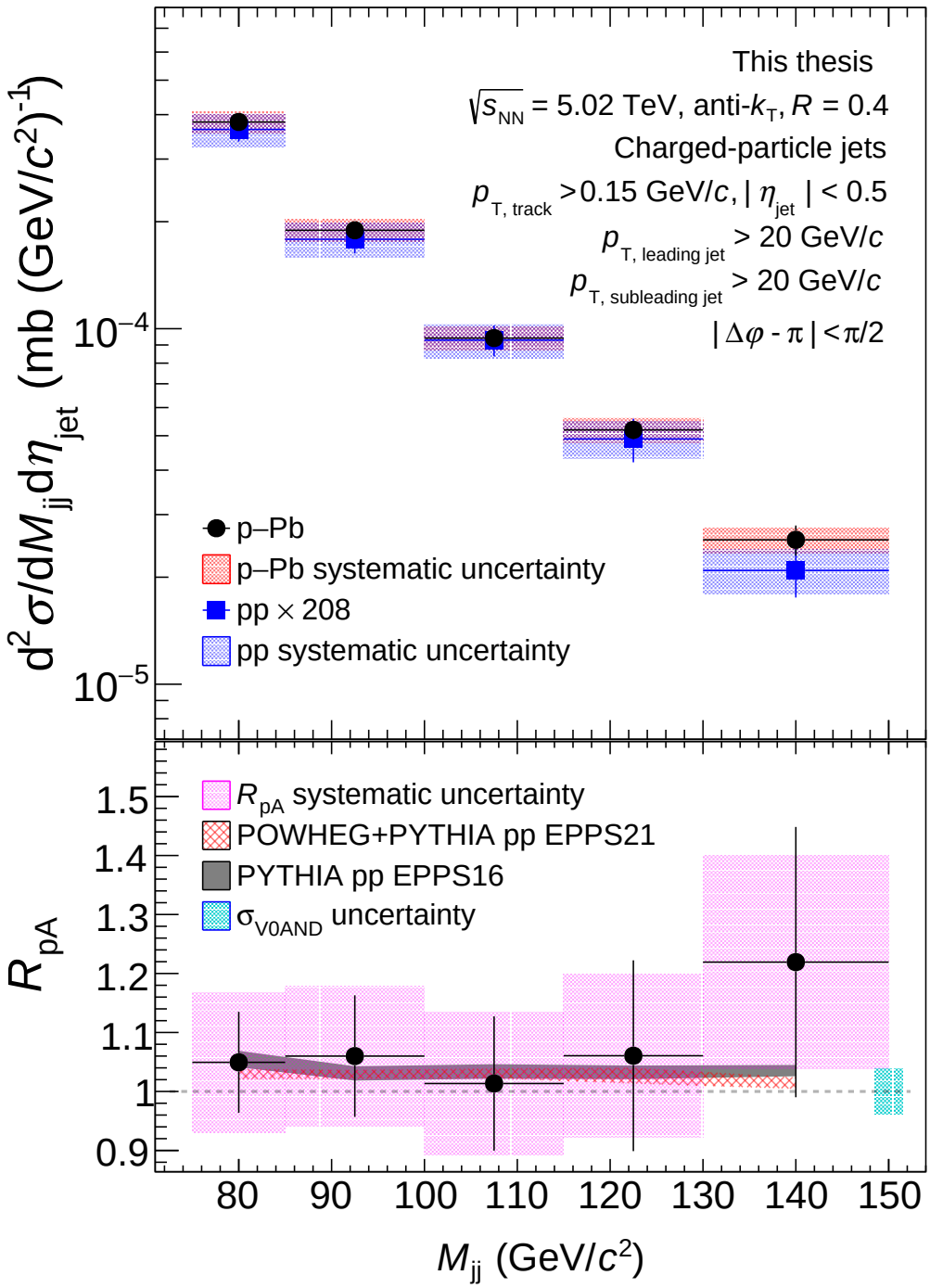


FIGURE 53 The dijet invariant mass for p-Pb and pp×208 at $\sqrt{s_{NN}}$ = 5.02 TeV. The result also includes model comparisons of the nuclear modification factor R_{pA} with PYTHIA and POWHEG+PYTHIA.

artificially enhanced during the event generation. Therefore, we account for it by scaling each generated event with the inverse of the enhancement. The rapidity shift of the center-of-mass system $\Delta y = -0.465$ described in Sec. 4.4 is taken into account by setting the collision energy for both beams asymmetrically for PYTHIA, or boosting the whole event by Δy for POWHEG+PYTHIA.

In order to obtain R_{pA} from the simulations, we divide the dijet mass cross section obtained from the simulations including nPDFs by the baseline simulations without the nPDFs. These results are shown together with the data in the bottom panel of Fig. 53, in gray and hashed red bands. Both of the simulation results show a slight enhancement of the dijet yield, as the R_{pA} is above unity. This, however, is a very slight difference in comparison with the statistical and systematic uncertainties.

The momentum fractions of the partons entering the hardest QCD interaction in simulated events were studied in Ref. [1]. It was found that the probed fractions x are close to the anti-shadowing region in nPDFs, and, thus, we attribute a slight enhancement in the R_{pA} to a hint of anti-shadowing. As the effect is subtle in comparison to the relatively large uncertainties in the data, the results are limited in their discriminating power. The simulations are consistent with the data within uncertainties. These results will provide a baseline for future measurements in Pb–Pb collisions.

6 JETS IN ALICE FORWARD CALORIMETER

6.1 FoCal test beams

A full-length prototype of the FoCal, which was described in detail in Sec. 2.4.1, has been built and tested in various test beam campaigns between 2021 and 2023 at the CERN PS and SPS beam lines [4]. Data were gathered in test beams using hadron beams of energies up to 50 GeV and electron beams of energies up to 300 GeV. One full 20-layer FoCal-E segment and a FoCal-H segment using the copper-scintillating-fiber spaghetti geometry were assembled for these test beams. In order to model the performance of the FoCal prototype, a MC simulation was developed to correspond to the setup used in the test beams. The software package GEANT4.10.7 [212] was used for the development of this simulation. Fig. 54 shows the simulated prototype, with a 100 GeV electron and pion initiated shower passing through the detector.

Between 2021 and 2023, the tested prototypes of FoCal increased in their size and complexity. Improvements in the systems and instrumentation and the development of the prototypes during this time is summarized in Fig. 55.

At the end of 2022, the prototype hardware resembled the envisioned FoCal

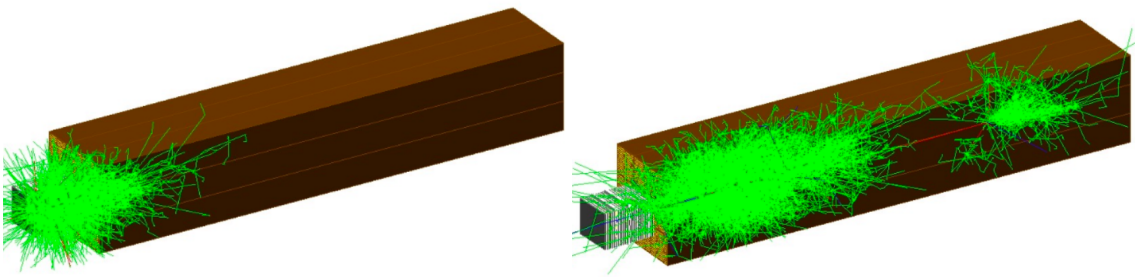


FIGURE 54 The simulated FoCal-E and FoCal-H prototypes through which a shower is created by a 100 GeV electron (left) and pion (right). The incident particle, not pictured, enters the detector from the left, where the FoCal-E lies in front of the FoCal-H. Figure from Ref. [4], reprinted under the license CC BY 4.0.

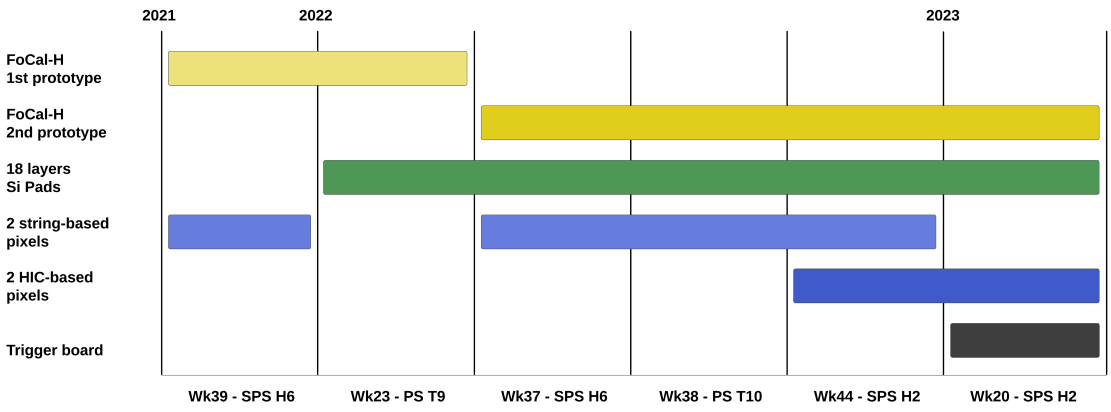


FIGURE 55 The timeline of the FoCal prototype development between test beam campaigns held in 2021 and 2023. Figure from Ref. [4], reprinted under the license CC BY 4.0.

detector. The full test beam setup used in the fall 2022 and 2023 test beams is shown in Fig. 56.

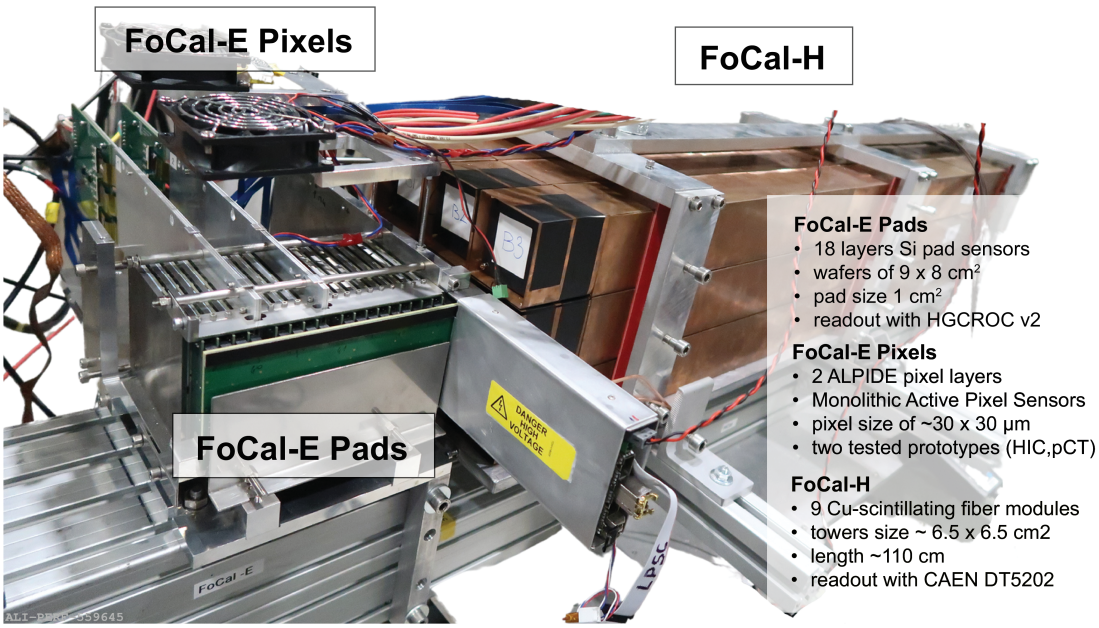


FIGURE 56 A picture of the setup of the FoCal prototype used in the late 2022 and 2023 test beam campaigns. The FoCal-E prototype is on the left, with 18 pad and two pixel layers. The FoCal-H prototype is on the right, corresponding to the spaghetti-type implementation. Figure from Ref. [4], reprinted under the license CC BY 4.0.

As the FoCal-E is longitudinally segmented, it is possible to measure the longitudinal shower profile of EM showers passing through the prototype, with a granularity $\approx 1X_0$. This has been done by measuring the signal charge for each pad layer, summing the signal around the leading cell in a $5 \times 5\,\mathrm{cm}^2$ square, and

calibrating the signal to charge. In Fig. 57, a summary of the measured longitudinal shower profiles is shown for the data from electron beams of 20–300 GeV. Generally, the simulation agrees well with the data.

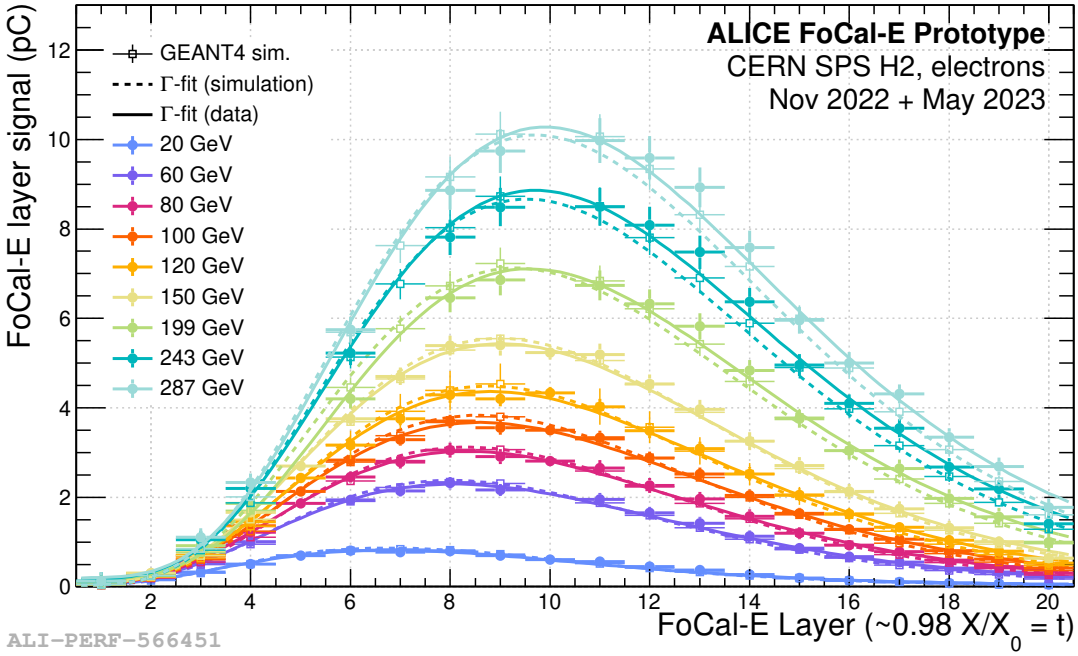


FIGURE 57 A summary of the longitudinal shower profiles measured and simulated with GEANT4, for electrons of energies 20–300 GeV. Both distributions are fitted with a Γ-distribution. Figure from Ref. [4], reprinted under the license CC BY 4.0.

Figure 58 shows the relative energy resolution of the FoCal-E pad layers, for both data sets defined above. Respective simulations have also been performed and their results are also shown. The energy resolution is expected to contain three contributions [223] as

$$\frac{\sigma_E}{E} = \frac{\sigma_{\text{stoch}}}{\sqrt{E/\text{GeV}}} \oplus \sigma_{\text{const}} \oplus \frac{\sigma_{\text{noise}}}{E/\text{GeV}}, \tag{47}$$

where σ_E and E are in units of GeV and σ_i are unit-less. The first term accounts for the stochastic shower fluctuations. The second term is a constant, describing the limitations of detector design and calibration. The third and final term describes the effect from the electronic noise. These terms and their numerical values are shown in the Fig. 58 for the fits to both data and simulations.

A required performance of the FoCal-E energy resolution was found in Ref. [44] to be 5%. This is achieved at energies at or above 200 GeV, in which the resolution of simulations and data are close to 3%, thus fulfilling the required performance. As energy is related to p_{T} and pseudorapidity η in the high-energy limit by $p_{\text{T}} \approx E/\cosh\eta$, and in the FoCal acceptance, $12 \lessapprox \cosh\eta \lessapprox 165$, we reach this energy at p_{T} values of 1.2 to 16 GeV.

In order to study the energy resolution of the FoCal-H, the calibrated ADC sum distribution was used for each beam energy. The distributions were fit with

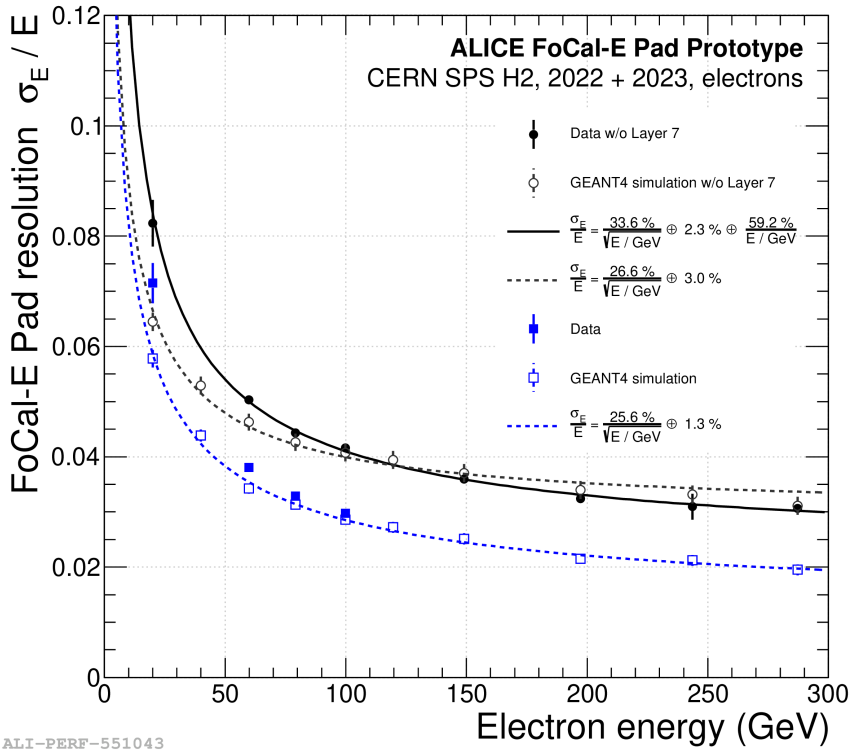


FIGURE 58 The relative energy resolution of the pad layers in FoCal-E. Data was measured in the May 2023 test beams, marked in blue, and data measured in November 2022 without pad layer 7 was combined with the May 2023 data to form the distribution marked in black. Simulation results are shown in open markers. Figure from Ref. [4], reprinted under the license CC BY 4.0.

a Gaussian and a Crystal ball fit function. These functions were used to compute the energy resolution σ_E/E, where E is the mean value of the final fit, and σ_E the width. The fitting procedure is described in detail in Ref. [4]. In Fig. 59 are shown the obtained resolutions for data and simulations, as a function of the hadron beam energy. Fit parameters of the energy resolution in Eq. (47) are listed in the caption of Fig. 59. At high energies, the stochastic term becomes less significant, and the constant term will dominate. The constant term is of the order of 10 %, and this value is consistent with the expected performance of FoCal-H at high energies [4].

In Fig. 60, the 10th pixel layer in FoCal-E is visualized as an electron beam at energy 287 GeV passes through. The shower event of a single electron passing through is shown on the left. A clear, pronounced core of the shower in the center is surrounded by a tail of less dense pixel hit occupancy. On the right, a two-electron shower event has been recorded, and the showers from each electron are clearly distinguishable from each other. This demonstrates the capability of the pixel layers to resolve the structure of particle showers on a sub-millimeter scale, vital to discriminate single-photon shower events from background π^0 decay photons.

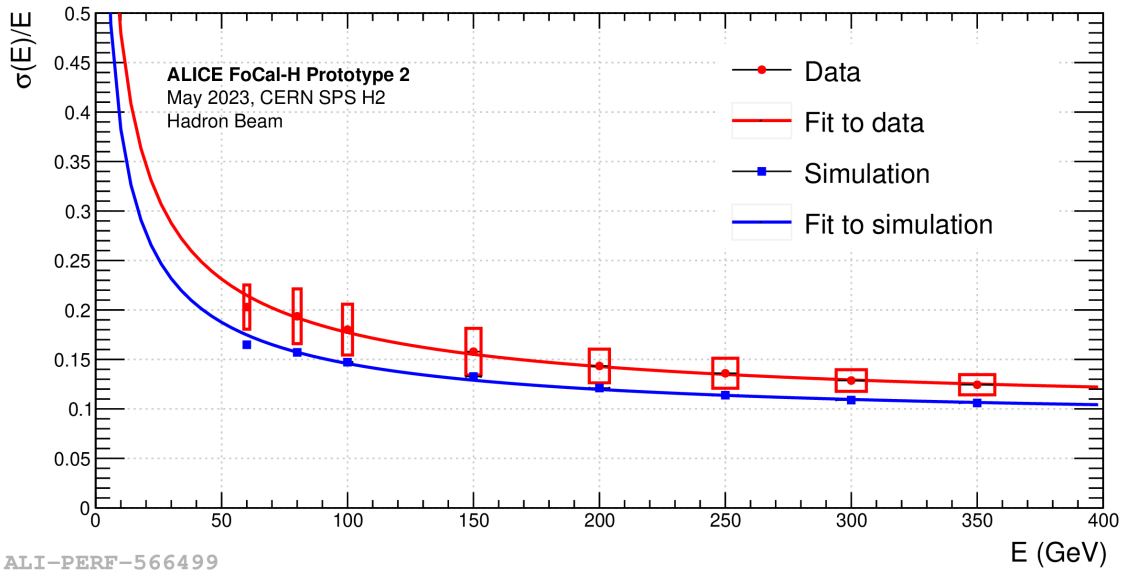


FIGURE 59 The energy resolution of the FoCal-H prototype in May 2023 as a function of energy of the hadron beam is shown in red. Simulation results are also shown, in blue. The numerical values of the fit are: $\sigma_{\text{stoch}} = 1.48$, $\sigma_{\text{noise}} = 0$, $\sigma_{\text{const}} = 0.10$ for data and $\sigma_{\text{stoch}} = 1.18$, $\sigma_{\text{noise}} = 0$, $\sigma_{\text{const}} = 0.09$ for the simulation. Figure from Ref. [4], reprinted under the license CC BY 4.0.

6.2 Full detector simulations

In order to perform simulation studies with the full detector system in Run 4, we simulate the entire ALICE detector system, including the full FoCal detector as described in Sec. 2.4.1. All Run 4 detectors, as well as the beam pipe and instrumentation along it are implemented in the GEANT [212] model used in these simulations.

While the true ALPIDE pixels used for the two pixel layers of FoCal-E have a pixel granularity of $30 \times 30\,\mu\text{m}^2$, to minimize simulation computing time, we simulate pixels with a granularity of $50 \times 50\,\mu\text{m}^2$. At the reconstruction level, these pixels are further summed into $0.5 \times 0.5\,\text{mm}^2$ cells. For these studies, we did not simulate charge diffusion or sharing charge between pixels. However, to simulate fluctuations in the deposited charge, the thickness of the sensitive silicon layer was given a thickness of $30\,\mu$m. Expected dead areas in the layers are included in the GEANT detector model.

While at the time of these studies, the working version of the FoCal-H geometry was the spaghetti design, also used in the test beam campaign, for the studies described in this section, the sandwich design [44] has been used in the detector-level simulation. This is due to the computational expensiveness of the spaghetti-type design. The sandwich design of FoCal-H in the simulation contains 34 layers of lead with a width of 3 cm each, interleaved with scintillator layers of width 0.2 cm. The simulated hits are summed longitudinally per tower, which are of transverse size $2.5 \times 2.5\,\text{cm}^2$. Studies [2] have shown that the selection between these two designs does not have a significant impact on the perfor-

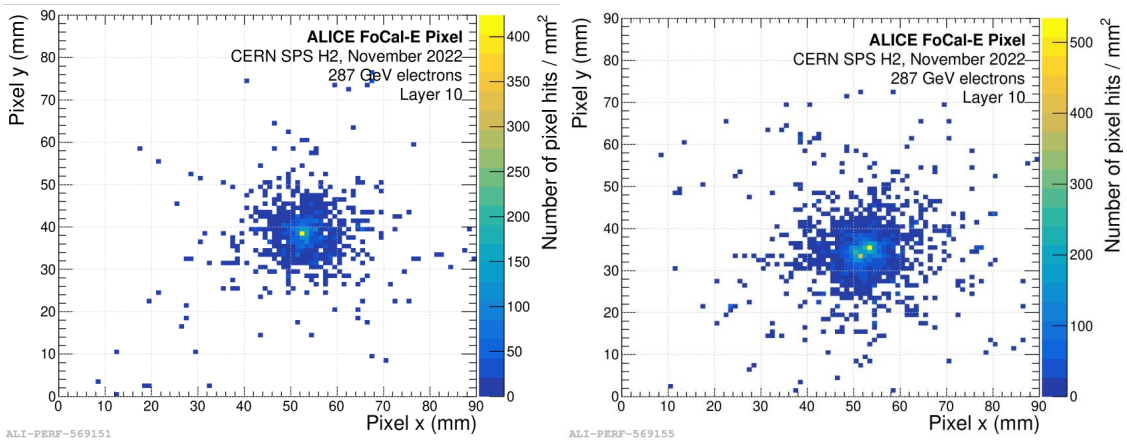


FIGURE 60 The event display of the pixel layer at layer 10 of the FoCal-E. In the left hand side plot, a 287 GeV electron shower has passed through the layer, while a two-electron shower has passed through in the right hand side plot. Figures from Ref. [4], reprinted under the license CC BY 4.0.

mance results, and any slight differences show an improvement in performance for the spaghetti design as compared to the sandwich design. Repeating the jet performance studies with the newest sheet-FoCal-H design is currently in development.

Signals are simulated with GEANT3 [211] based on shower development modeling and the energy deposition in the detector. We do not include a separate detector response simulation nor smearing, as their effects are expected to be negligible for high-energy photons and hadrons at very forward rapidities, as was found in the test beams [4]. The signals in the pixel layers are digitized such that any signal above the threshold energy of 4 keV is counted as a pixel hit. For the pad layers, the analog readout uses the energy deposition given by GEANT3 as the detector signal directly. Noise effects were found relevant only for low-energy showers, with a negligible impact on forward measurements where shower energies are large, above 50 GeV. While we do include a digitization step in the reconstruction chain of the simulation, there is no electronics response effects implemented in the simulation. Preliminary test beam results, however, show a good agreement with simulations when shower energies are above 100 GeV, which suggests that the implementation of the electronic response would not change the performance results significantly.

6.2.1 Reconstruction and clustering

A clustering algorithm is applied to the detector signals given by the simulation [2]. The detector signals are first grouped into six segments: each pixel layer forms a segment by itself, and pad segments are formed from signals in the layers 1–4, 6–9, 11–15 and 16–20. In these studies, the clustering has been applied per segment, although the procedure can also be performed for each layer. Equally spaced cells are constructed from deposited hits, which are summed from the spatial range of the cell size. The size is $1 \times 1\,\text{cm}^2$ for the pad segments where the cell

energy is the sum of the energy in all the layers of the segment, and $500 \times 500\,\mu\text{m}^2$ for the pixel segments.

First, cluster seeds are identified in each segment separately from a list of cells sorted by energy, from the first three segments. A minimum energy threshold is set for a seed, and a minimum distance between cluster seeds is also set. For each seed, all cells within a specified cluster radius are summed to form a cluster. A weight is assigned to each nearby cell of a seed, calculated by fitting a double exponential function to single-photon simulation data. The weight is dependent on the radial distance of cells from the seed and the energy of the seed. For the last three segments, pre-seeds are determined using the clusters identified in the first three segments, and they are used as the seeds. The parameters used in the clustering are tuned separately for each segment, in order to obtain a good efficiency and low fake rate arising from shower splitting [2].

Once clusters have been determined for each segment, they are combined into clusters for the full FoCal-E. The pad layer clusters are first matched and combined, after which the same is done for the pixel clusters. Finally, the pixel layer clusters are used in order to separate showers that are indistinguishable in the lower-resolution pad layers. This is done by matching the pixel and pad layer clusters geometrically, and if more than one pixel cluster is matched to one pad layer cluster, the pad layer cluster is then split into the number of corresponding pixel layer clusters, according to the energy partition of the pixel layer clusters. The final shower position is set as the average position of the two pixel layer clusters.

More detailed studies have been performed in Ref. [2] on the performance of this clustering algorithm. It is found that, when studying single-photon box simulations, in about 10% of all events there is no reconstructed cluster, and in about 15 % of all events there is more than one cluster. Ideally, the algorithm would lead to just one cluster per each event. However, factors like detector material between the interaction point and the FoCal detector, cluster splitting, and insensitive areas of the detector result in the differing amounts of clusters found in the events.

The clusterization parameters can also be optimized to improve performance in specific kinematic regions. It has been studied in Ref. [2] that by altering the parameters in order to split more clusters, the π^0 reconstruction efficiency can be improved at high p_{T}, high energy and small decay photon separation distance.

FoCal-H signals have not been clusterized, therefore tower signals are used in the simulation. There is ongoing work being done on implementing a clusterizer for the FoCal-H signals. This will be studied in future simulation studies, and will require a separate calibration procedure before it can be used to improve results for e.g. jet performance.

6.2.2 Calibration

It is necessary to calibrate the MC detector-level response of the calorimeter in order to accurately reconstruct the energy of incident particles and to account

for the different responses of the FoCal-E and FoCal-H. This is done using single charged pions and electrons, generated with the same energy using a box generator covering the entire FoCal acceptance. Electrons are expected to deposit a large portion of their energy in the FoCal-E, while charged pions deposit their energy primarily in the FoCal-H. The pions are also expected to exhibit large fluctuations in the position at which the shower is initiated, which can occur both in the FoCal-E and the FoCal-H.

To calibrate the energy, scaling factors are determined to form a linear combination of shower energies in both detectors,

$$E_{\text{tot}} = \alpha E_{\text{em}} + \beta E_{\text{had}}, \tag{48}$$

where E_{tot} is the calibrated total energy deposited in FoCal, E_{em} and E_{had} are the sums of the energies deposited in the FoCal-E and FoCal-H, respectively. The calibration parameters α and β are determined by studying the distributions of E_{tot} for electrons and charged pions with initial energy $E = 500$ GeV, and fitting a Crystal Ball function [224] to the distribution. The values of these parameters and the detailed calibration procedure are presented in Ref. [2].

6.3 Jet reconstruction in FoCal

At the truth, or particle, level, the jet reconstruction procedure takes into account all primary particles and decay daughters found within the FoCal detector acceptance. At the detector level, the reconstruction uses FoCal-E clusters assigned with a zero-mass, and tower signals from FoCal-H. We do not apply a cluster or cell energy threshold cut in this analysis.

We use the anti-k_{T} clustering algorithm with E-scheme recombination, as described in Sec. 4. We require the centroid of the jet to be within the FoCal acceptance, requiring $3.4 + R < \eta^{\text{jet}} < 5.5 - R$, where R is the jet radius. For the jet radius $R = 0.6$, this corresponds to a pseudorapidity acceptance of $4.0 < \eta^{\text{jet}} < 4.9$. In order to minimize effects from outlying MC tracks with a much higher p_{T} than is physically meaningful, we discard events in which there is a jet that has a p_{T} over three times the $\hat{p}_{\text{T}}$ of the event. To select particles for the analysis, we require that the particle is a physical primary, unless its decay daughters are physical primaries, in which case we choose to use the decay particles instead. A physical primary in ALICE is a particle with a mean proper lifetime τ larger than 1 cm/c, either produced directly in the interaction between colliding partners, or from decays of particles with a $\tau < 1$ cm/c, restricted to decay chains leading to the interaction [225].

In order to estimate the reconstruction efficiency of jets, we need to match the detector level reconstructed jet to the corresponding simulated particle level jet. We require the matched detector level and particle level jets to have a phase space separation of $\Delta R = \sqrt{\Delta\phi^2 + \Delta y^2} < 0.6$, where $\Delta\phi$ and Δy are the separation of the jet pair in azimuthal angle and rapidity, respectively. The requirement

$\Delta R < 0.6$ is chosen in order to ensure the jets overlap sufficiently, in order to reduce contributions from fake matches. If there are multiple matches found, we select the jet pair with the smallest value of ΔR.

The FoCal jet performance is studied using truth and detector level information obtained from simulations. For this analysis, we ran a PYTHIA simulation of pp events as projected to be measured by the ALICE detector in Run 4. The propagation of the generated particles through GEANT provides the detector level information. At the generator level, we simulate pp collisions using PYTHIA8 at an energy of $\sqrt{s} = 14$ TeV. In order to obtain decent statistics of jets in the simulation, we run the pp simulation in $\hat{p}_{\mathrm{T}}$ bins. The binning used is 5-10, 10-20, 20-30, 30-40, 40-60, 60-100, 100-200, 200-1000 GeV/c^2. These bins are then merged by scaling with the cross section of each bin and number of events. Simulations of box-generated single pions are also used for the jet analysis. In these simulations, single particle events have been generated for a specific particle species, in this case single pions, with an energy of 500 GeV.

In Fig. 61, we present the $p_{\mathrm{T}}^{\mathrm{jet}}$-differential inclusive jet cross section generated in these simulations, for anti-k_{T} jets using jet radius $R = 0.6$, within the FoCal acceptance. The particle level jet spectrum is shown in solid lines, and the detector level jet spectrum in dashed lines, in two different ranges of rapidity. As can be seen, the more forward jets, in red, present a lower cross section as compared to the less forward jets, in black. Also, comparing the detector level and particle level jets shows that, while their shapes are very similar, the detector level distributions are lower. This visualizes that the detector level jets need to be corrected for this difference. A power-law fit to the spectra yields a power of $n \approx 6$, which corresponds closely to the value obtained from the inclusive jet spectra at mid-rapidity in ALICE [226].

When studying jets at a very forward rapidity, jet reconstruction performance is significantly impacted by the finite area of the jets and the transverse dimensions of the EM and hadronic showers in FoCal. The phase space in (η, ϕ) becomes very compact at high η compared to mid-rapidity. In order to quantify the effect, we consider a plane transverse to the particle beam at a distance D from the interaction vertex. The radial distance r from the beamline is

$$r \approx \frac{2D}{e^{\eta}}, \tag{49}$$

where we use the small angle approximation. As η increases, the value of r decreases exponentially. Inserting the FoCal parameters, $D = 700$ cm and $3.2 < \eta < 5.8$, to the equation, we can study the values of r as a function of η on the front face of the FoCal detector. These values are presented in Table 4, along with spatial intervals corresponding to a jet size of $\Delta\eta \approx \Delta\phi \approx 0.4$.

As can be seen in the last two columns in Table 4, the physical size of the $R \approx 0.4$ jet gets increasingly smaller as we go towards more forward rapidities. Figure 62 shows the effect of moving more forward in rapidity on the phase space and spatial coordinates. In the figure, the simulated FoCal-H shower profile is shown for charged pions, averaged over 10000 showers. The simulated showers

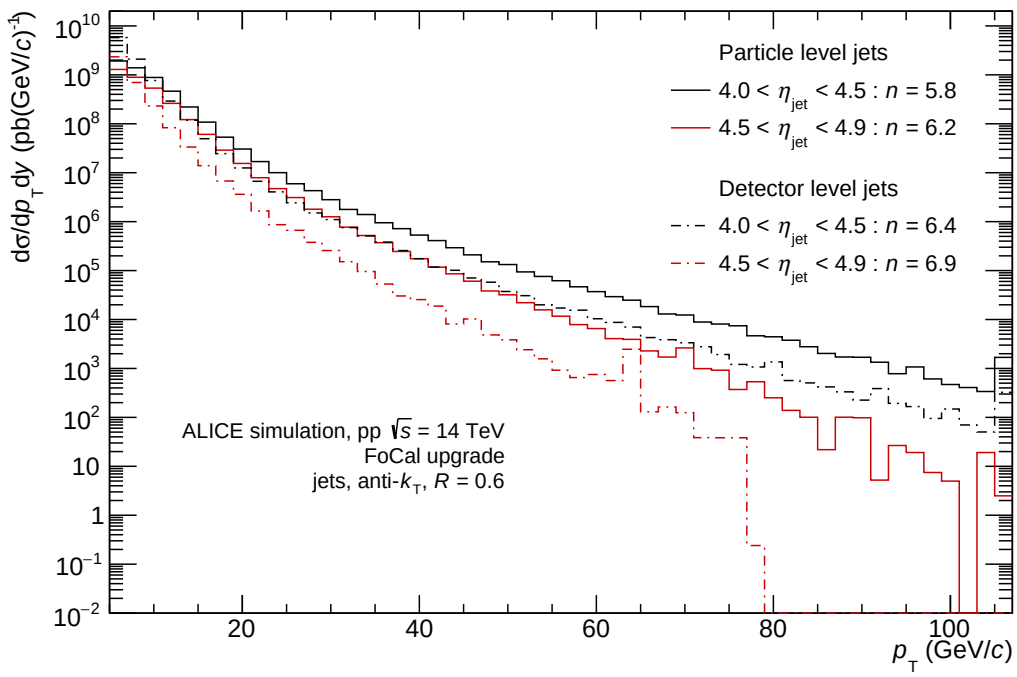


FIGURE 61 The p_{T}-differential inclusive jet cross section for generated pp collisions at $\sqrt{s}$=14 TeV for jets reconstructed with the anti-k_{T} algorithm. Jets are reconstructed with R=0.6 required to have their centroid within the FoCal acceptance. Particle-level (solid line) and detector-level jet (dashed line) spectra are presented in two intervals of η^{jet}, in black and red. The power determined by a power-law fit to each spectrum is given by the value n. Figure from Ref. [2], reprinted under the license CC BY 4.0.

TABLE 4 Radial distances r for the phase space intervals in a plane transverse to the beam as seen in the front face of the FoCal detector, at $D = 700$ cm from the interaction point. For selected values of η, the radial distance from the beamline is shown in cm according to Eq. (49). In the third column are the differences in r corresponding to a phase space increment of $\Delta\eta = 0.4$, and in the fourth column is the spatial distance for a phase space increment of $\Delta\phi = 0.4$

η	r (cm)	$r(\eta - 0.4) - r(\eta)$ (cm)	$r \cdot (\Delta\varphi = 0.4)$ (cm)
3.4	46.7		18.7
3.8	31.3	15.4	12.5
4.2	21.0	10.3	8.4
4.6	14.1	6.9	5.6
5.0	9.4	4.6	3.8
5.4	6.3	3.1	2.5
5.8	4.2	2.1	1.7

sample the full FoCal acceptance, but have been adjusted later to be situated at the same location in each of the figures. Thus, we generate an average hadronic shower profile in FoCal-H.

Figure 62 consists of three panels, in which the shower is centered around $(\eta, \phi) = (4.9, 0), (4.5, 0), (4.0, 0)$ in order from left to right. Each panel contains a red circle, corresponding to the nuclear interaction length of Cu, which contains 81% of the shower energy in (x, y) phase space. The black contour in each panel is drawn in (η, ϕ) phase space with a radius $R = 0.6$, centered around the same point as the shower in each panel. These black contours signify jet cones of jet radius $R = 0.6$. As is clearly visualized in the left hand panel, the contour for $\eta = 4.9$ captures only a fraction of the true jet energy, which is expected to create a large difference in the truth and detector level jets. This effect is also clear in the middle panel, corresponding to $\eta = 4.5$, while in the panel on the right, corresponding to $\eta = 4.0$, the effect is not as large.

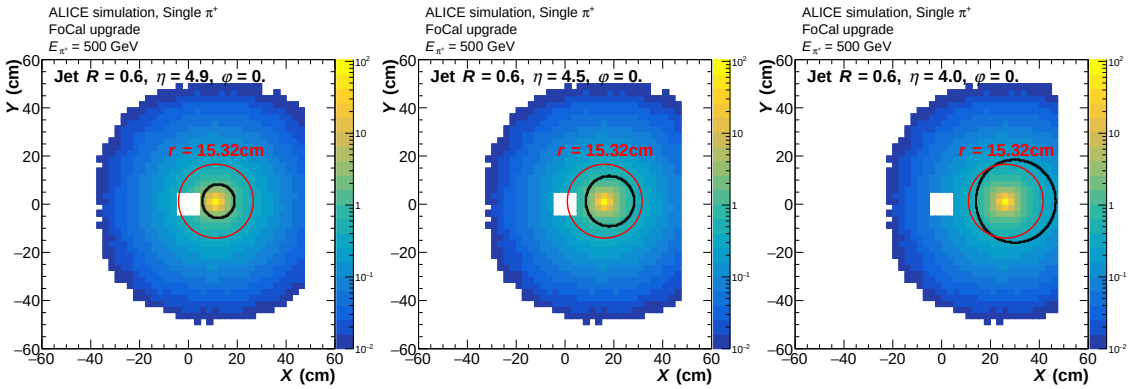


FIGURE 62 The averaged hadronic shower profile in FoCal-H, obtained from simulating 10k showers centered around $(\eta, \phi) = (4.9, 0), (4.5, 0), (4.0, 0)$ from left to right. The black contour represents a jet of radius $R = 0.6$. The red circle contains 81% of the total shower energy in (x, y) space. Figures from Ref. [2], reprinted under the license CC BY 4.0.

6.4 Jet performance results

In order study jet performance in simulations, we want to be able to quantify the relative energy or p_{T} difference of the matched detector level and particle level jets. This can be done by defining the following distributions,

$$\Delta E = \frac{E^{\mathrm{det}} - E^{\mathrm{part}}}{E^{\mathrm{part}}},$$
$$\Delta p_{\mathrm{T}} = \frac{p_{\mathrm{T}}^{\mathrm{det}} - p_{\mathrm{T}}^{\mathrm{part}}}{p_{\mathrm{T}}^{\mathrm{part}}}, \tag{50}$$

in which the energy E and transverse momentum p_{T} of the particle level jets, notated by "part", and the detector level jets, notated by "det", are compared. We define as the Jet Energy Scale (JES) the mean of the ΔE and Δp_{T} distributions, while the Jet Energy Resolution (JER) is defined as their standard deviation.

In order to determine the jet performance, we first focus on simulations of single energetic charged and neutral pions, since they are the simplest probes affected by the above discussion of phase space effects. We generate samples of charged and neutral pions separately, and apply the jet finding procedure described in the previous section on the combination of FoCal-E clusters and FoCal-H tower signals generated by the simulations.

In Fig. 63, we show the JES of ΔE as a function of the particle level jet energy for single pions in two different intervals of jet pseudorapidity η_{jet}. As all the values of the JES are negative, the results signify a deficit in detector level jet energy as compared to the particle level. This can be attributed to the phase space effects described in Sec. 6.3. As the jet reconstruction of the electromagnetic showers in FoCal-E induced by the neutral pions misses a fraction of the total energy, this causes the JES to be within 10–20 %. This is in line with the fact that the Molière radius of the FoCal-E is smaller than the spatial size of the $R = 0.6$ jets. As for the hadronic showers, largely within FoCal-H, the reconstruction of jets from charged pions exhibits a notably larger deficit in energy, with a JES of around 25–30 %. This is in line with the larger transverse size of showers in FoCal-H. At larger values of η, in red in Fig. 63, the effect is larger, which is consistent with the exponential nature of the spatial contraction of the phase space, as described in Eq. (49).

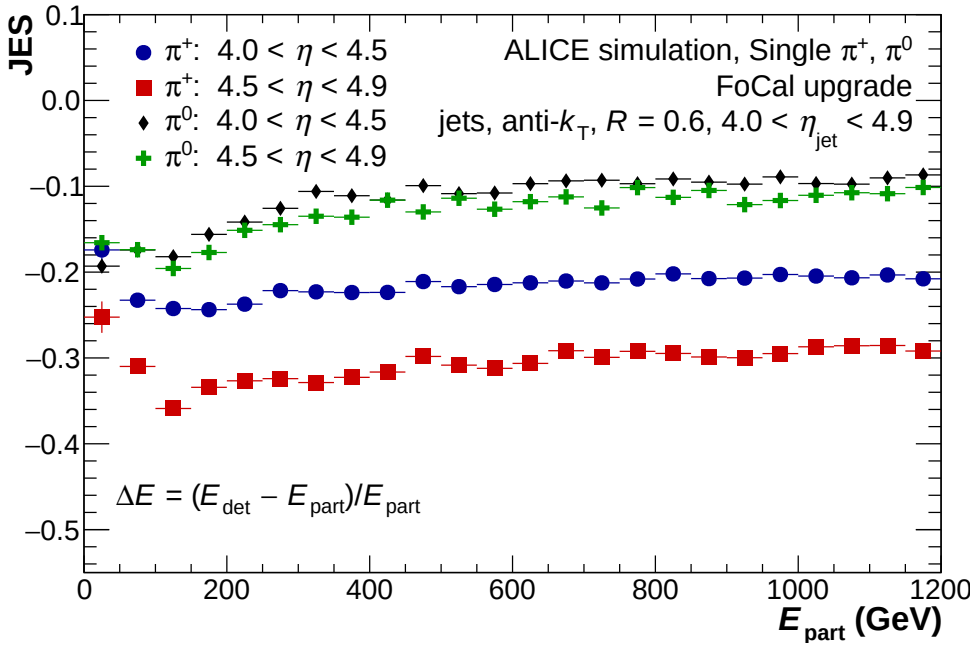


FIGURE 63 JES of ΔE for single charged and neutral pions as a function of particle level jet energy. Figure from Ref. [2], reprinted under the license CC BY 4.0.

Following the single pion study, we then study the results of jets arising from the simulated pp collisions at $\sqrt{s} = 14$ TeV. In Fig. 64, we show the response matrices in which we correlate matched particle and detector level jets as a function of jet energy E or transverse momentum p_{T}. The energy of the jets has been calculated with the assumption of zero mass of the particles initiating the shower. It is important to study both kinematic variables at this stage of the analysis, as there can be differences between jet response as a function of each kinematic variable. Looking at the response matrices, we note that the shapes of both matrices are similar, with a weight of both distributions lying above the diagonal, drawn in faint gray. This behavior corresponds to a lower value of the kinematic quantity at the detector level than at particle level, which is expected as described in the previous section.

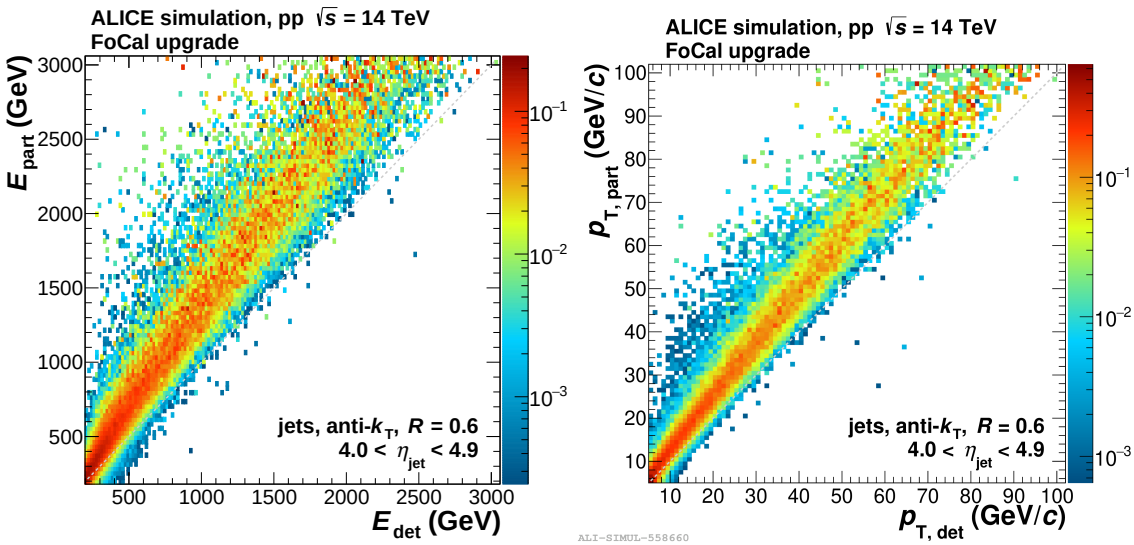


FIGURE 64 Response matrices as a function of energy (left) and transverse momentum (right) for the matched particle and detector level jets with $R = 0.6$, normalized to unity in slices of the corresponding y-axis: E^{part} and $p_{\mathrm{T}}^{\mathrm{part}}$. Figures from Ref. [2], reprinted under the license CC BY 4.0.

The distributions of ΔE and Δp_{T} are shown in Fig. 65, for three selected intervals of E_{part} and $p_{\mathrm{T,part}}$. A Gaussian function has been fit to each distribution. As the meaningfulness of the JES and JER quantities rely on a relatively Gaussian distribution, their characterization of the jet response is shown to be relatively reliable based on these Gaussian fits. However, there is a non-Gaussian tail appearing at the higher values of ΔE and Δp_{T}. This indicates that in order to make precise measurements of jets with FoCal, we need to correct for these effects by unfolding the full response using response matrices. This is work to be done in the future.

Presented in Fig. 66 are the JER and JES for the energy of jets, and in Fig. 67 for the jet transverse momentum. The results are shown for jets of radius $R = 0.6$ with their centroid lying within $4.0 < \eta_{\mathrm{jet}} < 4.9$. As in Fig. 63 for the single pions, the JES values are negative, signifying missing detector level jet energy or

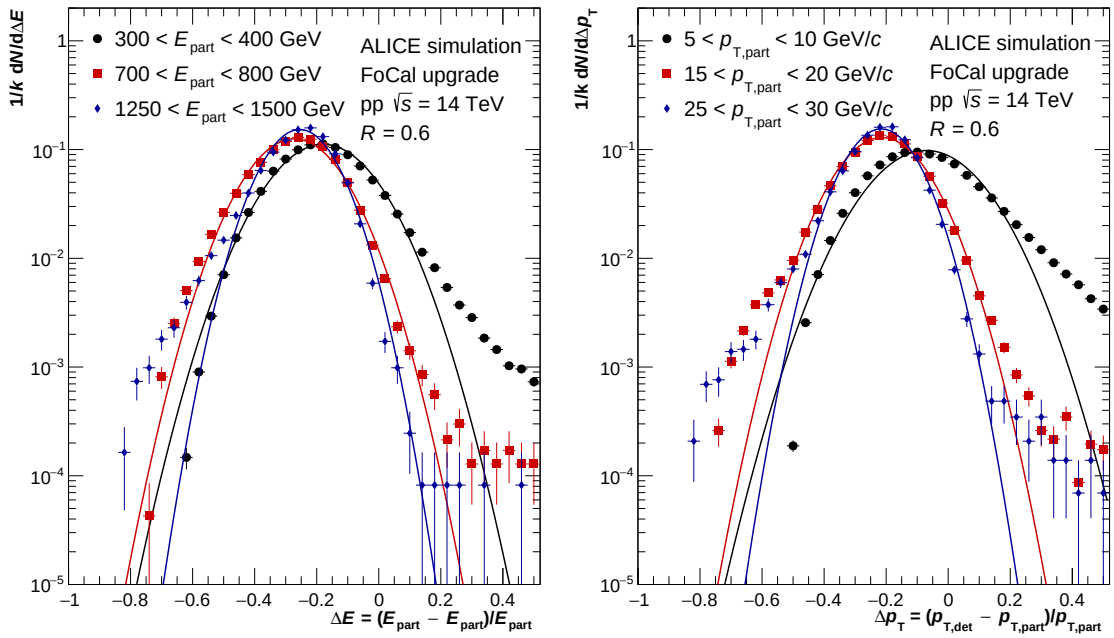


FIGURE 65 Distributions of ΔE (left) and Δp_{T} (right) are shown for jets with radius $R = 0.6$ with their centroid lying within $4.0 < \eta_{\mathrm{jet}} < 4.9$. The distributions are shown for selected intervals of E_{part} and $p_{\mathrm{T,part}}$ and a Gaussian fit is shown for each distribution. Figures from Ref. [2], reprinted under the license CC BY 4.0.

transverse momentum. At the particle level, particles at the edge of the jet cone contribute fully to the jet energy. At the detector level, however, a particle can create a large shower, which can lead to some detector-level energy of a track on the edge falling outside the reconstructed jet cone. Additionally, part of the observed detector-level missing energy may originate from the absence of a clustering algorithm in the FoCal-H calorimeter. If clustering were implemented, the total energy of a cluster with its reconstructed center point within the jet cone would be included in the jet. Without clustering, only the individual tower energies within the jet cone are summed and included. Also, the anti-k_{T} algorithm produces circular jets, which can exclude some energy that a clustering algorithm might otherwise capture.

Comparing the two JES distributions, a slightly more uniform JES is seen for the Δp_{T} case than for ΔE after the dip at $p_{\mathrm{T,part}} \approx 15$ GeV/c and $E_{\mathrm{part}} \approx 600$ GeV. The JER is below 15% for $p_{\mathrm{T,part}} > 10$ GeV/c and below 10% for $p_{\mathrm{T,part}} > 25$ GeV/c when calculated from the Gaussian fit. Similarly, the JER is below 15% for all values of E_{part} and below 10% when going toward higher energies of $E_{\mathrm{part}} >$ 1250 GeV. The difference between the JER values obtained from the two different methods reflects the contribution of the non-Gaussian tail.

Comparing the JER for two different intervals of jet pseudorapidity, shown

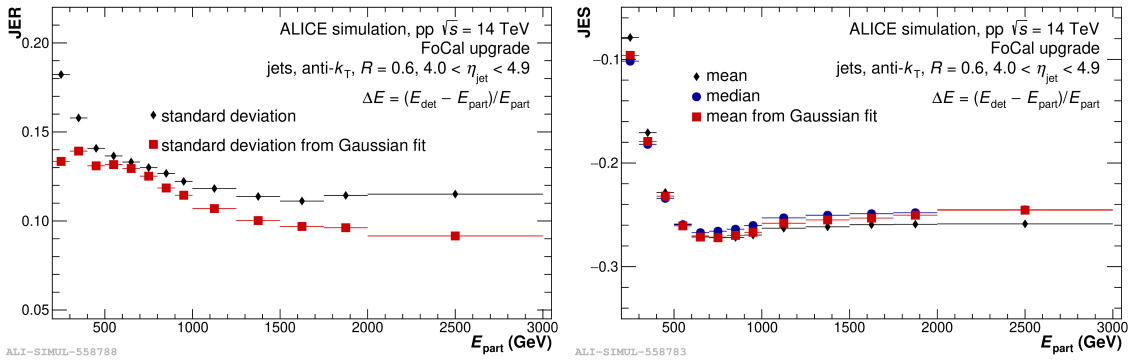


FIGURE 66 JER (left) and JES (right) for jets with radius $R = 0.6$, calculated using ΔE. The mean and standard deviation were calculated both using a numerical integral and by fitting a Gaussian function, as shown in Fig. 65. Figures from Ref. [2], reprinted under the license CC BY 4.0.

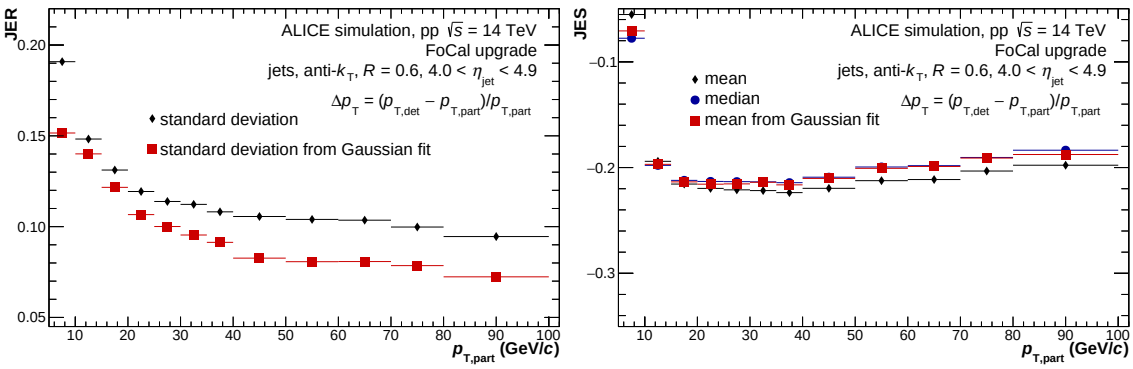


FIGURE 67 JER (left) and JES (right) for jets with radius $R = 0.6$, calculated using Δp_{T}. The mean and standard deviation were calculated both using a numerical integral and by fitting a Gaussian function, as shown in Fig. 65. Figures from Ref. [2], reprinted under the license CC BY 4.0.

in Fig. 68[1], we can compare the energy resolution of measurements of jets with their centroid within $4.0 < \eta_{\mathrm{jet}} < 4.5$ (red) and within $4.5 < \eta_{\mathrm{jet}} < 4.9$ (black). The most forward jets show an increase in the JER at higher energies, close to the total resolution of all jets, while the less-forward jets show a slight improvement in resolution after $E_{\mathrm{part}} \approx 600$ GeV, of the order of a few percent. This finding is supported by the discussion in the previous section about the phase space becoming more compact at larger rapidities, and thus the jet cone containing less of the total shower energy.

As was discussed, from Fig. 63 it is clear that the response of the FoCal detector is much better for single electromagnetic showers than it is for single hadronic showers. This can be understood as a consequence of the EM show-

[1] These results are based on earlier simulations and have minor differences compared to published results presented in Fig. 66. These results are included to illustrate the rapidity dependence.

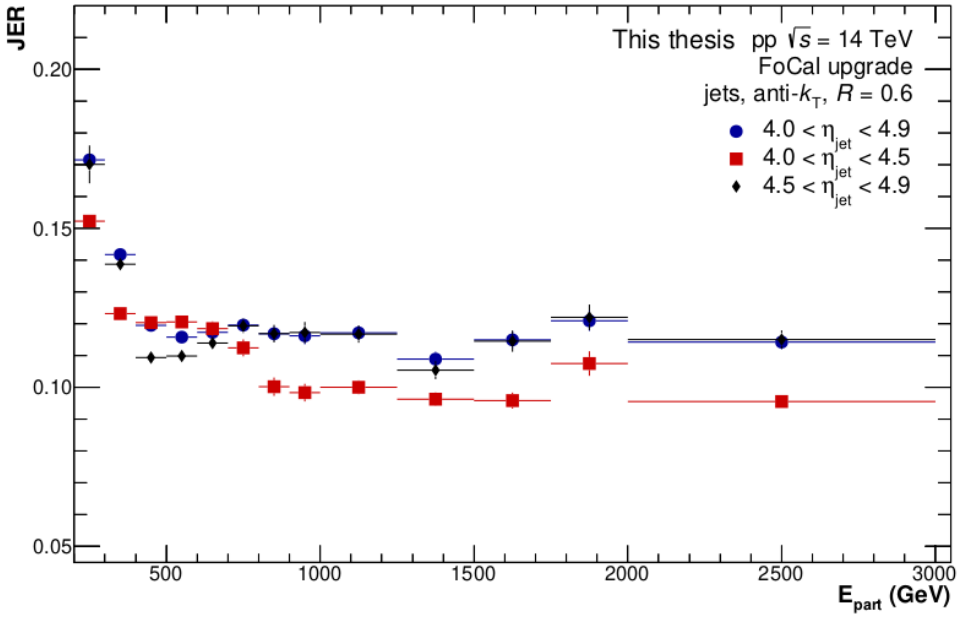


FIGURE 68 The JER calculated for jets with their centroid within $4.0 < \eta_{\text{jet}} < 4.5$ (red) and within $4.5 < \eta_{\text{jet}} < 4.9$ (black). Also shown is the total JER for all jets within the FoCal acceptance, in blue.

ers being more spatially compact when compared to hadronic showers, as well as through the geometric considerations discussed before. The jet performance would therefore be improved if only jets with a dominant neutral energy component were under study.

To quantify the improvement in jet performance for neutral jets, we can study the Neutral Energy Fraction (NEF) of jets at both particle and detector levels. From the event generator, at the particle level we can obtain information on the particle provided by the generator and whether it is neutral or charged. At the detector level, we can define the NEF as the ratio of neutral energy, measured in the FoCal-E, to the total energy, measured in both FoCal-E and FoCal-H combined.

In Fig. 69, the probability distributions of the NEF are shown for particle-level jets within the FoCal acceptance, for varying jet radii $R = 0.2, 0.4$ and 0.6, in six different bins of jet transverse momentum $p_{\text{T}}^{\text{jet}}$. In the highest bins of $p_{\text{T}}^{\text{jet}}$, jets with different radii behave very similarly, with the mean value of the NEF distribution peaking near $1/3$, as expected from isospin symmetry [227]. As jets become more collimated at higher p_{T}, the difference between the behavior of different jet cone sizes also diminishes. This can be due to a larger fraction of the jet energy being contained within a smaller cone radius due to the increased collimation. Heading towards lower bins of $p_{\text{T}}^{\text{jet}}$, the behavior of the distribution changes significantly for the lower-radius jets. Especially in the case of $R = 0.2$ jets, in the lowest $p_{\text{T}}^{\text{jet}}$ bins the distribution appears nearly flat. Based on these results, the selection of a larger value of R in jet reconstruction is required.

The peaks near the edges, at 0 and 1, seen in the case of $R = 0.2$ jets as well

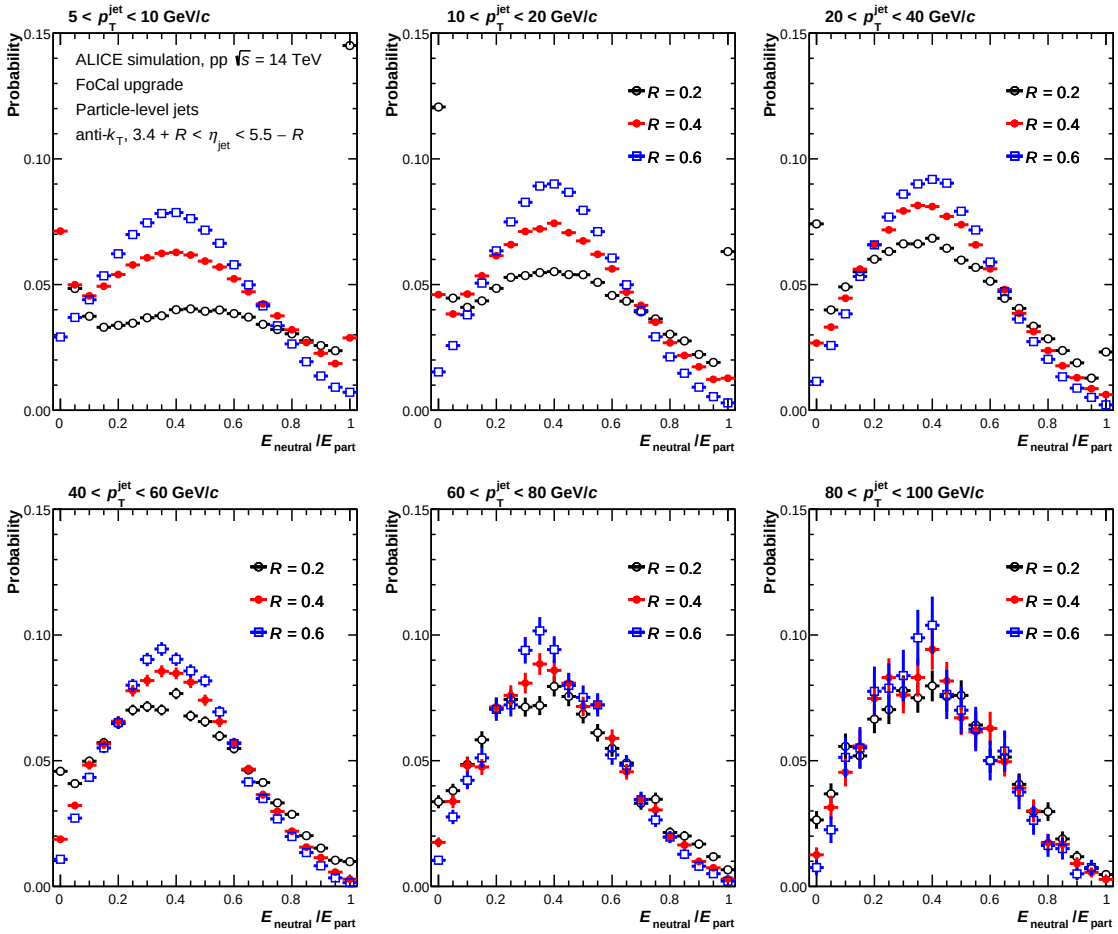


FIGURE 69 The probability distribution of the Neutral Energy Fraction (NEF) calculated at the particle level for Pythia-generated pp collisions at $\sqrt{s} = 14$ TeV. The results are shown for jets within the FoCal acceptance, for varying jet radii $R = 0.2, 0.4$ and 0.6, in six different bins of p_T^{jet}. Figures from Ref. [2], reprinted under the license CC BY 4.0.

as the $R = 0.4$ jets in the lowest bins of p_T^{jet} were studied in detail to find their origin. They were found to appear from jets that consist of only a small number of constituents, mainly from jets with 1–5 constituents. This was only the case for the jets with a smaller radius, as no such issue was found when studying the $R = 0.6$ jets. This is understood, as the larger size of the jet cone results in a, on average, larger amount of particles to coincide within the same jet cone.

We select $R = 0.6$ jets for further study, and plot three different p_T^{jet} bins in Fig. 70. Here we can see how, even in the lowest bin, the mean of the distribution peaks around $1/3$. While there is a slight shift of the distribution shape with varying p_T^{jet} bin, there is no significant dependence of the NEF probability distribution on the p_T^{jet}. Hence, we focus the following NEF studies in this section on $R = 0.6$ jets.

In Fig. 71, the correlation of the NEF in matched particle and detector level jets is presented, for selected bins in p_T^{jet}. As is visualized by the diagonal shape

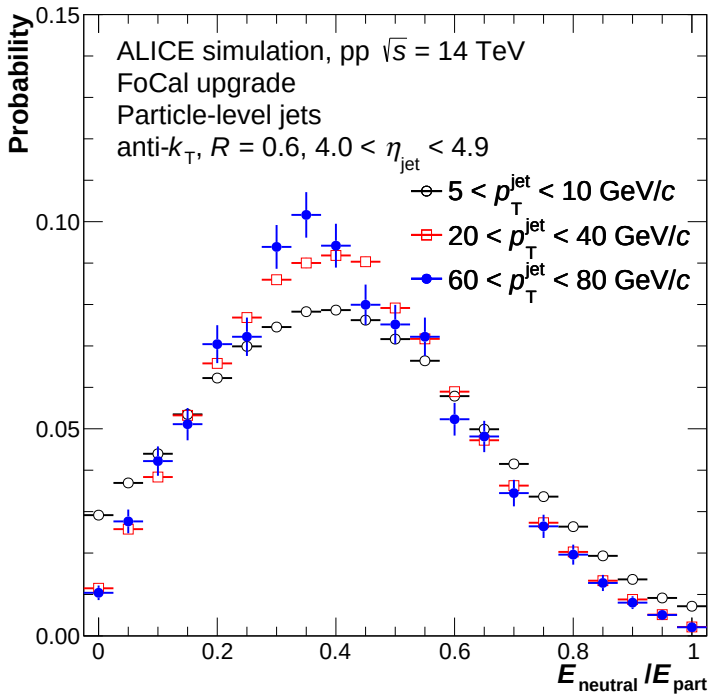


FIGURE 70 The NEF of particle level jets within the FoCal acceptance, with radius $R = 0.6$ from Pythia-generated pp collisions at $\sqrt{s} = 14$ TeV. Three bins of $p_{\mathrm{T}}^{\mathrm{jet}}$ are drawn to showcase the dependence of the jet NEF on jet transverse momentum. Figure from Ref. [2], reprinted under the license CC BY 4.0.

of the distribution, there is a high correlation of the NEF distributions between the particle and detector level jets.

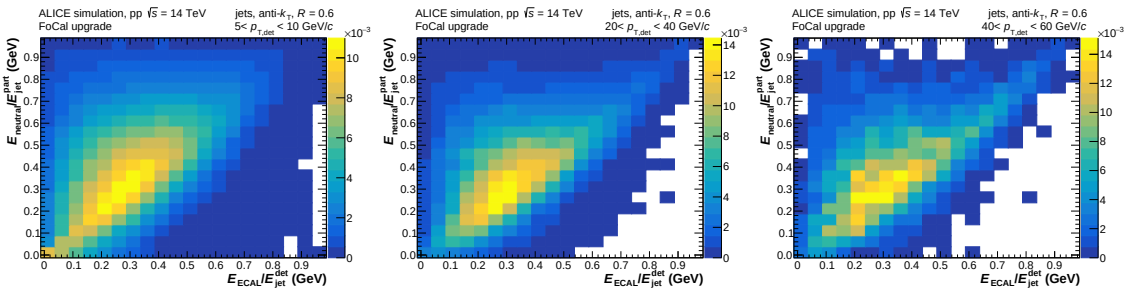


FIGURE 71 The correlation of the NEF of particle and detector level jets, for selected bins in $p_{\mathrm{T}}^{\mathrm{jet}}$. Figures from Ref. [2], reprinted under the license CC BY 4.0.

We then study whether the FoCal jet population can be biased towards having a larger NEF, i.e. a larger electromagnetic shower component, in order to improve the jet energy scale, JES. As, experimentally, we can only determine the detector level shower energy ratio, we study jets biased using the overlapping shower energy in FoCal-E and FoCal-H, denoted by $\mathrm{NEF}_{\mathrm{det}}$.

In Fig. 72, the JES and JER calculated from ΔE are shown. The black points

are the same as were shown in Fig. 66, but the blue and red distributions have been calculated for a biased jet population. The blue points represent jets with a suppressed EM shower fraction ($\mathrm{NEF}_{\mathrm{det}} < 1/3$), and the red squares represent jets with an enhanced EM shower fraction ($\mathrm{NEF}_{\mathrm{det}} > 2/3$) relative to the unbiased population. On the left hand side of the figure, we see that the JER is clearly insensitive to this biasing of the jet population. However, on the right, the JES is markedly improved, by around 10% in most regions of E_{part}, for the EM shower fraction enhanced jet population.

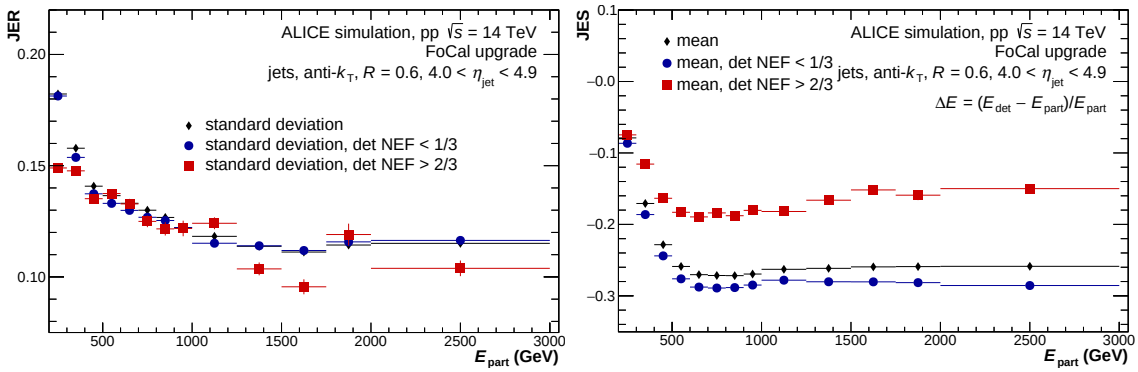


FIGURE 72 JER (left) and JES (right) of the full jet population (black) for jets with $R = 0.6$, and jet populations biased by their fraction of neutral energy as measured at the detector level, with cuts $\mathrm{NEF}_{\mathrm{det}} < 1/3$ (blue) and $\mathrm{NEF}_{\mathrm{det}} > 2/3$ (red). Figures from Ref. [2], reprinted under the license CC BY 4.0.

In Fig. 73, the JES of the jet population with enhanced EM shower fraction has been plotted for both detector level jets as well as particle level jets. In order to form an equal comparison of the NEF-bias for both particle and detector level jets, we select the highest and lowest 20 % of the jets when they are ordered in increasing levels of their NEF, defined as described above for the two jet populations. For particle level jets, the 20 % cut corresponds to jets with their $\mathrm{NEF}_{\mathrm{part}} > 0.63$, close to the 1/3 cut used in the previous study. The JES has been calculated from the mean of the Gaussian fit for this figure.

When comparing the highest 20 % $\mathrm{NEF}_{\mathrm{part}}$ JES in Fig 73 to the JES of the full jet population reported in Fig. 72, there is a clear improvement observed. Comparing JES of the highest and lowest 20 % NEF jets for both detector and particle level jets shows a significant difference in both cases when biasing on the EM shower fraction, with the greatest improvement seen for the detector level jets. The difference in JES of the detector level jets is heightened at the largest values of E_{part}, where the difference is over 10%. These studies suggest that a jet population that is biased by enhancing the electromagnetic shower fraction has an improved response relative to the unbiased jet population.

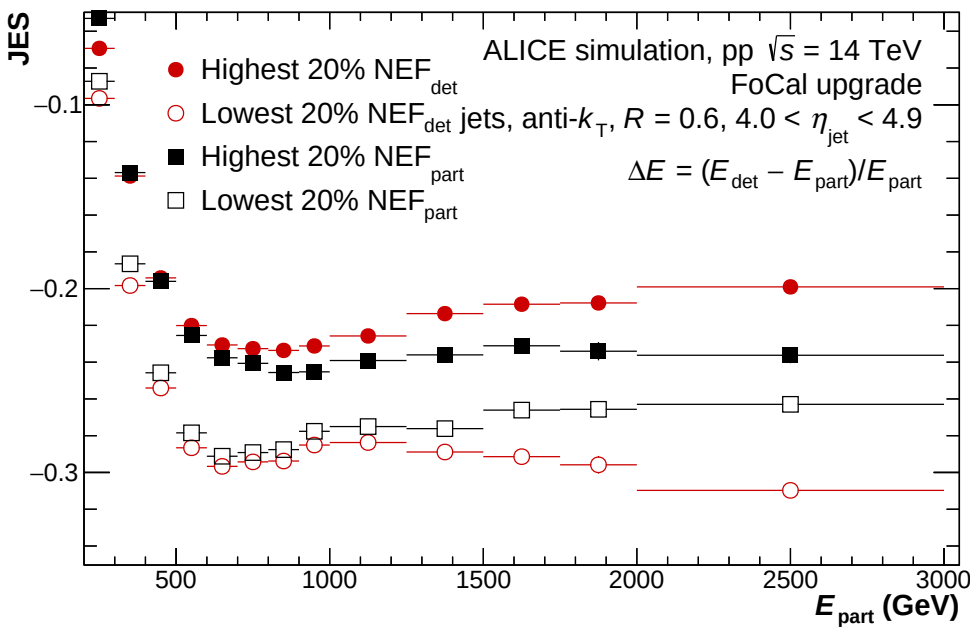


FIGURE 73 JES is presented for particle and detector level jets with $R = 0.6$, for biased jet populations. The lowest and highest 20 % of the jets, when they are ordered in increasing levels of their NEF, have been selected. Figure from Ref. [2], reprinted under the license CC BY 4.0.

7 CONCLUSIONS AND OUTLOOK

In this thesis, I have presented the analysis and results of two jet-related studies in ALICE. The measurement of the dijet invariant mass in pp and p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV with ALICE Run 2 data was reported. I have updated the analysis in Ref. [1] to incorporate the latest reconstruction pass of both datasets. The spectra of dijet invariant mass and the corresponding nuclear modification factor R_{pA} were determined in the mass range of $75 - 150\,\mathrm{GeV}/c^2$. Within measurement uncertainties, no nuclear modification was observed. However, the measured R_{pA} lies systematically slightly above unity, which is consistent with predictions from PYTHIA and POWHEG+PYTHIA simulations. These simulations suggest that the slight enhancement may arise from dijets probing the anti-shadowing region of the nuclear PDFs of lead. To confirm this effect with precision, larger dijet statistics are required. This would be possible by the significantly higher luminosities available in ALICE Run 3. A similar analysis with Run 3 data could thus provide a sensitive probe of the anti-shadowing region and cold nuclear matter effects. Since the pp and p–Pb measurements serve as a crucial baseline for Pb–Pb collisions, the next interesting study would be to extend the analysis to Pb–Pb data measured in Run 2 and Run 3.

Looking ahead, once Run 4 begins, ALICE will commence physics measurements with the upgraded FoCal detector. FoCal will facilitate measurements of a new kinematic regime in Bjorken-x, not covered by any existing experiment. Such measurements will complement those performed at other facilities, such as D-meson measurements in LHCb and, in the near future, DIS measurements at the Electron Ion Collider (EIC) under construction at BNL. With FoCal, measurements of prompt photons, dijets and various correlations are expected to provide new insight into gluon saturation phenomena in protons and nuclei by constraining the gluon density at small x. In addition, FoCal will extend the rapidity range of ALICE, enabling high-p_{T} neutral meson and jet measurements that enable further mapping of the QGP density as a function of rapidity.

The jet performance in FoCal was determined with simulation studies. The jet energy scale (JES), quantifying the fraction of truth-level jet energy missing from detector-level jets, demonstrated the need for unfolding procedures. For-

ward jet kinematics explain how the full hadronic shower can extend beyond the jet cone, which affects the JES and the jet energy resolution (JER). This effect is more pronounced for smaller jet radii, motivating the choice of $R = 0.6$ for the analysis. The JER for $R = 0.6$ jets was found to be better than 15% for $p_{\mathrm{T}} > 10$ GeV/c and better than 10% for $p_{\mathrm{T}} > 25$ GeV/c, consistent with the forward-jet performance reported by CMS [228]. Jets with an enhanced neutral energy fraction were found to have improved performance. Ongoing work on a dedicated clusterizer for the hadronic calorimeter is expected to further improve the general jet resolution.

Overall, the FoCal jet performance is sufficient for single-jet measurements, and by extension, also for forward dijet and photon-jet correlation measurements with FoCal. With Run 4 data, FoCal will provide a unique opportunity to explore nuclear PDFs and gluon saturation phenomena in the low-x regime, offering valuable input to the effort to understand the dynamics of QCD matter.

REFERENCES

[1] O. Saarimäki, *Dijet invariant mass spectrum in pp and p–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV*. PhD thesis, 2023. http://urn.fi/URN:ISBN:978-951-39-9662-8.

[2] **ALICE** collaboration, *Physics performance of the ALICE Forward Calorimeter upgrade*, 2023. https://cds.cern.ch/record/2869141.

[3] **ALICE** collaboration, L. Huhta, *Physics program and performance of the ALICE Forward Calorimeter upgrade (FoCal)*, *PoS* **DIS2024** (2025) 265 [arXiv:2409.08983 [hep-ex]].

[4] M. Aehle *et. al.*, *Performance of the electromagnetic and hadronic prototype segments of the ALICE Forward Calorimeter*, *JINST* **19** (2024) no. 07 P07006 [arXiv:2311.07413 [physics.ins-det]].

[5] G. Barr *et. al.*, *Particle Physics in the LHC Era*. Oxford University Press, 01, 2016.

[6] D. E. Morrissey and M. J. Ramsey-Musolf, *Electroweak baryogenesis*, *New J. Phys.* **14** (2012) 125003 [arXiv:1206.2942 [hep-ph]].

[7] L. Lee, C. Ohm, A. Soffer and T.-T. Yu, *Collider Searches for Long-Lived Particles Beyond the Standard Model*, *Prog. Part. Nucl. Phys.* **106** (2019) 210 [arXiv:1810.12602 [hep-ph]]. [Erratum: Prog.Part.Nucl.Phys. 122, 103912 (2022)].

[8] A. Sopczak, *Higgs Boson Searches at the LHC Beyond the Standard Model*, *MDPI Physics* **6** (2024) no. 3 1132 [arXiv:2409.14061 [hep-ex]].

[9] R. Franceschini, *Physics Beyond the Standard Model Associated with the Top Quark*, *Ann. Rev. Nucl. Part. Sci.* **73** (2023) 397 [arXiv:2301.04407 [hep-ph]].

[10] MissMJ, "Standard model of elementary particles." Wikimedia Commons, https://commons.wikimedia.org/wiki/File: Standard_Model_of_Elementary_Particles.svg, 2019. Released under the Creative Commons Attribution-Share Alike 3.0 Unported license.

[11] M. Gell-Mann, *The eightfold way: A theory of strong interaction symmetry*, tech. rep., California Inst. of Tech., Pasadena. Synchrotron Lab., 03, 1961.

[12] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking. Version 2*, pp. 22–101. 2, 1964.

[13] E. D. Bloom *et. al.*, *High-Energy Inelastic e p Scattering at 6-Degrees and 10-Degrees*, *Phys. Rev. Lett.* **23** (1969) 930.

[14] M. Breidenbach *et. al.*, *Observed behavior of highly inelastic electron-proton scattering*, *Phys. Rev. Lett.* **23** (1969) 935.

[15] G. 't Hooft, *A Planar Diagram Theory for Strong Interactions*, *Nucl. Phys. B* **72** (1974) 461.

[16] O. Garcia-Montero and S. Schlichting, *Effective theories for nuclei at high energies*, *Eur. Phys. J. A* **61** (2025) no. 3 54 [arXiv:2502.09721 [hep-ph]].

[17] **Particle Data Group** collaboration, S. Navas *et. al.*, *Review of particle physics*, *Phys. Rev. D* **110** (2024) no. 3 030001.

[18] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, *Phys. Rev. Lett.* **30** (1973) 1343.

[19] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346.

[20] L. C. P. Van Hove, *Theoretical prediction of a new state of matter, the 'quark-gluon plasma' (also called 'quark matter')*, tech. rep., CERN, 1987.

[21] CERN and BICEP2 Collaboration and NASA, "Illustration of the history of the universe." CERN Website, https://home.cern/news/series/lhc-physics-ten/recreating-big-bang-matter-earth, accessed 2025-09-08.

[22] A. Andronic, P. Braun-Munzinger and J. Stachel, *Hadron production in central nucleus-nucleus collisions at chemical freeze-out*, *Nucl. Phys. A* **772** (2006) 167 [arXiv:nucl-th/0511071].

[23] A. Andronic *et. al.*, *Decoding the phase structure of QCD via particle production at high energy*, *Nature* **561** (2018) no. 7723 321 [arXiv:1710.09425 [nucl-th]].

[24] J. Cleymans *et. al.*, *Comparison of chemical freeze-out criteria in heavy-ion collisions*, *Phys. Rev. C* **73** (2006) 034905 [arXiv:hep-ph/0511094].

[25] F. A. Flor, G. Olinger and R. Bellwied, *Flavour and Energy Dependence of Chemical Freeze-out Temperatures in Relativistic Heavy Ion Collisions from RHIC-BES to LHC Energies*, *Phys. Lett. B* **814** (2021) 136098 [arXiv:2009.14781 [nucl-ex]].

[26] S. Borsanyi *et. al.*, *QCD Crossover at Finite Chemical Potential from Lattice Simulations*, *Phys. Rev. Lett.* **125** (2020) no. 5 052001 [arXiv:2002.02821 [hep-lat]].

[27] **STAR** collaboration, G. Odyniec, *Beam Energy Scan Program at RHIC (BES I and BES II) – Probing QCD Phase Diagram with Heavy-Ion Collisions*, *PoS* **CORFU2018** (2019) 151.

[28] Y. Akiba *et. al.*, *The Hot QCD White Paper: Exploring the Phases of QCD at RHIC and the LHC*, arXiv:1502.02730 [nucl-ex].

[29] **ALICE** collaboration, *The ALICE experiment: a journey through QCD*, *Eur. Phys. J. C* **84** (2024) no. 8 813 [arXiv:2211.04384 [nucl-ex]].

[30] G. J. Odyniec in *4th International Workshop on Relativistic Aspects of Nuclear Physics*, pp. 189–209, 2, 1996.

[31] M. Harrison, T. Ludlam and S. Ozaki, *RHIC project overview*, *Nucl. Instrum. Meth. A* **499** (2003) 235.

[32] **ALICE** collaboration, *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002.

[33] **ALICE** collaboration, *Letter of Intent for A Large Ion Collider Experiment [ALICE]*, tech. rep., CERN, Geneva, 1993.

[34] G.-Y. Qin and X.-N. Wang, *Jet quenching in high-energy heavy-ion collisions*, *Int. J. Mod. Phys. E* **24** (2015) no. 11 1530014 [arXiv:1511.00790 [hep-ph]].

[35] M. Connors *et. al.*, *Jet measurements in heavy ion physics*, *Rev. Mod. Phys.* **90** (2018) 025005 [arXiv:1705.01974 [nucl-ex]].

[36] **CMS** collaboration, *Constraining nPDFs using dijet production in pPb collisions at 8.16 TeV with the CMS experiment*, tech. rep., CERN, Geneva, 2024.

[37] **ATLAS** collaboration, *Observation of a Centrality-Dependent Dijet Asymmetry in Lead-Lead Collisions at* $\sqrt{s_{\mathrm{NN}}} = 2.77$ *TeV with the ATLAS Detector at the LHC*, *Phys. Rev. Lett.* **105** (2010) 252303 [arXiv:1011.6182 [hep-ex]].

[38] **ATLAS** collaboration, *Measurements of the suppression and correlations of dijets in Pb+Pb collisions at sNN=5.02 TeV*, *Phys. Rev. C* **107** (2023) no. 5 054908 [arXiv:2205.00682 [nucl-ex]].

[39] **CMS** collaboration, *Jet momentum dependence of jet quenching in PbPb collisions at* $\sqrt{s_{NN}} = 2.76$ *TeV*, *Phys. Lett. B* **712** (2012) 176 [arXiv:1202.5022 [nucl-ex]].

[40] **ALICE** collaboration, *Transverse momentum spectra and nuclear modification factors of charged particles in pp, p-Pb and Pb-Pb collisions at the LHC*, *JHEP* **11** (2018) 013 [arXiv:1802.09145 [nucl-ex]].

[41] **PHENIX** collaboration, *Cold-nuclear-matter effects on heavy-quark production in d+Au collisions at* $\sqrt{s_{NN}} = 200$ *GeV*, *Phys. Rev. Lett.* **109** (2012) no. 24 242301 [arXiv:1208.1293 [nucl-ex]].

[42] **PHENIX** collaboration, *Heavy-flavor electron-muon correlations in $p+p$ and d+Au collisions at $\sqrt{s_{NN}}$ = 200 GeV*, *Phys. Rev. C* **89** (2014) no. 3 034915 [arXiv:1311.1427 [nucl-ex]].

[43] **PHENIX** collaboration, *Nuclear Modification of ψ', ξ_c, and J/ψ Production in d+Au Collisions at $\sqrt{s_{NN}}$=200 GeV*, *Phys. Rev. Lett.* **111** (2013) no. 20 202301 [arXiv:1305.5516 [nucl-ex]].

[44] ALICE Collaboration, CERN, *Letter of Intent: A Forward Calorimeter (FoCal) in the ALICE experiment*, tech. rep., CERN, Geneva, 2020.

[45] O. Brüning *et. al.*, *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004.

[46] M. Schaumann *et. al.* in *9th International Particle Accelerator Conference*, 6, 2018.

[47] **ALICE** collaboration, *Evidence of nuclear geometry-driven anisotropic flow in OO and Ne−Ne collisions at $\sqrt{s_{\rm NN}} = 5.36$ TeV*, tech. rep., CERN, Geneva, 2025.

[48] J. Vollaire *et. al.*, *Linac4 design report*, vol. 6/2020 of *CERN Yellow Reports: Monographs*. CERN, Geneva, 9, 2020.

[49] K. Hanke, *Past and present operation of the CERN PS Booster*, *Int. J. Mod. Phys. A* **28** (2013) 1330019. https://cds.cern.ch/record/1709399.

[50] S. Gilardoni and D. Manglunki, eds., *Fifty years of the CERN Proton Synchrotron: Volume 1*. CERN Yellow Reports: Monographs. 2011.

[51] J. Coupard *et. al.*, *LHC Injectors Upgrade, Technical Design Report: v.2: Ions*, tech. rep., CERN, Apr., 2016.

[52] N. Doble *et. al.*, *The Super Proton Synchrotron (SPS): A Tale of Two Lives*, *Adv. Ser. Direct. High Energy Phys.* **27** (2017) 135.

[53] **ATLAS** collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.

[54] **CMS** collaboration, *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004.

[55] **LHCb** collaboration, *The LHCb Detector at the LHC*, *JINST* **3** (2008) S08005.

[56] E. Lopienska, "The CERN accelerator complex, layout in 2022." CERN Document Server, 2022. General Photo. https://cds.cern.ch/record/2800984.

[57] **ALICE** collaboration, *ALICE Inner Tracking System (ITS)*. Technical design report. ALICE. CERN, Geneva, 1999.

[58] J. Alme *et. al.*, *The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **622** (Oct., 2010) 316–367. http://dx.doi.org/10.1016/j.nima.2010.04.042.

[59] **ALICE** collaboration, *ALICE Transition-Radiation Detector: Technical Design Report*, tech. rep., 2001.

[60] **ALICE** collaboration, G. Dellacasa *et. al.*, *ALICE Technical Design Report of the Time-of-Flight system (TOF)*, tech. rep., 2, 2000.

[61] **ALICE** collaboration, S. Beole *et. al.*, *ALICE technical design report: Detector for high momentum PID*, .

[62] **ALICE** collaboration, G. Dellacasa *et. al.*, *ALICE technical design report of the photon spectrometer (PHOS)*, .

[63] **ALICE** collaboration, *ALICE Electromagnetic Calorimeter Technical Design Report*, tech. rep., CERN, Sept., 2008.

[64] **ALICE** collaboration, *ALICE dimuon forward spectrometer*. Technical design report. ALICE. CERN, Geneva, 1999.

[65] **ALICE** collaboration, P. Cortese *et. al.*, *ALICE Technical Design Report on Forward Detectors: FMD, T0 and V0*, tech. rep., 9, 2004.

[66] **ALICE** collaboration, G. Dellacasa *et. al.*, *ALICE technical design report: Photon multiplicity detector (PMD)*, .

[67] **ALICE** collaboration, M. Gallio *et. al.*, *ALICE Zero-Degree Calorimeter (ZDC)*. Technical design report. ALICE. CERN, Geneva, 1999.

[68] **ALICE** collaboration, *ALICE Time Projection Chamber*. Technical design report. ALICE. CERN, Geneva, 2000.

[69] **ALICE** collaboration, C. Garabatos, *The ALICE TPC*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **535** (2004) no. 1 197. https://www.sciencedirect.com/science/article/pii/S0168900204016018. Proceedings of the 10th International Vienna Conference on Instrumentation.

[70] *Performance of the ALICE experiment at the CERN LHC*, *International Journal of Modern Physics A* **29** (2014) no. 24 1430044 [`arXiv:https://doi.org/10.1142/S0217751X14300440`]. https://doi.org/10.1142/S0217751X14300440.

[71] A. Maire, "Track reconstruction principle in ALICE for LHC run I and run II." General Photo, 2011.

[72] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **262** (1987) no. 2 444. https://www.sciencedirect.com/science/article/pii/0168900287908874.

[73] M. Arslandok *et. al.*, *Track Reconstruction in a High-Density Environment with ALICE*, *Particles* **5** (2022) no. 1 84 [arXiv:2203.10325 [physics.ins-det]].

[74] **ALICE** collaboration, B. Abelev *et. al.*, *Measurement of the inclusive differential jet cross section in pp collisions at* $\sqrt{s} = 2.76$ *TeV*, *Phys. Lett. B* **722** (2013) 262 [arXiv:1301.3475 [nucl-ex]].

[75] S. van der Meer, *Calibration of the Effective Beam Height in the ISR*, tech. rep., CERN, June, 1968.

[76] **ALICE** collaboration, *Measurement of inelastic, single- and double-diffraction cross sections in proton–proton collisions at the LHC with ALICE*, *Eur. Phys. J. C* **73** (2013) no. 6 2456 [arXiv:1208.4968 [hep-ex]].

[77] **ALICE** collaboration, *ALICE 2017 luminosity determination for pp collisions at* $\sqrt{s}$ *= 5 TeV*, . https://cds.cern.ch/record/2648933.

[78] **ALICE** collaboration, *Measurement of visible cross sections in proton-lead collisions at* $\sqrt{s_{\mathrm{NN}}}$ *= 5.02 TeV in van der Meer scans with the ALICE detector*, *JINST* **9** (2014) no. 11 P11003 [arXiv:1405.1849 [nucl-ex]].

[79] **ALICE** collaboration, *Upgrade of the ALICE Experiment: Letter Of Intent*, *Journal of Physics G: Nuclear and Particle Physics* **41** (jul, 2014) 087001. https://dx.doi.org/10.1088/0954-3899/41/8/087001.

[80] **ALICE** collaboration, *Addendum to the Technical Design Report for the Upgrade of the ALICE Time Projection Chamber*, tech. rep., 2015.

[81] **ALICE** collaboration, F. Reidt, *Upgrade of the ALICE ITS detector*, *Nucl. Instrum. Meth. A* **1032** (2022) 166632 [arXiv:2111.08301 [physics.ins-det]].

[82] **ALICE** collaboration, M. Slupecki, *Fast Interaction Trigger for ALICE upgrade*, *Nucl. Instrum. Meth. A* **1039** (2022) 167021.

[83] **ALICE** collaboration, D. Krupova, *Muon Forward Tracker in ALICE at the LHC*, *Nucl. Instrum. Meth. A* **1064** (2024) 169388.

[84] CERN, "Lhc long-term schedule (screenshot)." https://lhc-commissioning.web.cern.ch/schedule/LHC-long-term.htm, accessed 2025-09-08.

[85] P. Buncic, M. Krzewicki and P. Vande Vyvre, *Technical Design Report for the Upgrade of the Online-Offline Computing System*, tech. rep., 2015.

[86] V. Barroso *et. al.*, *The new ALICE data acquisition system (O^2/FLP) for LHC Run 3*, *EPJ Web Conf.* **295** (2024) 02029.

[87] **ALICE** collaboration, *Technical Design Report for the Upgrade of the ALICE Inner Tracking System*, *Journal of Physics G: Nuclear and Particle Physics* **41** (jul, 2014) 087002. https://dx.doi.org/10.1088/0954-3899/41/8/087002.

[88] **ALICE** collaboration, P. Martinengo, *The new Inner Tracking System of the ALICE experiment*, *Nuclear Physics A* **967** (2017) 900. https://www.sciencedirect.com/science/article/pii/S037594741730194X. The 26th International Conference on Ultra-relativistic Nucleus-Nucleus Collisions: Quark Matter 2017.

[89] *Technical Design Report of the ALICE Forward Calorimeter (FoCal)*, tech. rep., CERN, Geneva, 2024.

[90] **ALICE** collaboration, G. Aglieri Rinella, *The ALPIDE pixel sensor chip for the upgrade of the ALICE Inner Tracking System*, *Nucl. Instrum. Methods Phys. Res., A* **845** (2017) 583. https://cds.cern.ch/record/2239752.

[91] O. Bourrion *et. al.*, *Prototype electronics for the silicon pad layers of the future Forward Calorimeter (FoCal) of the ALICE experiment at the LHC*, *JINST* **18** (2023) no. 04 P04031 [`arXiv:2302.13912 [physics.ins-det]`].

[92] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization of Hard Processes in QCD*, *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1 [`arXiv:hep-ph/0409313`].

[93] "Proton-proton collision at 13 TeV registered by ALICE with 1200 bunches proton beams.." CERN Document Server, 2018. ALICE Event Display. https://cds.cern.ch/record/2315867.

[94] "2018 Heavy Ion Collision event (365573, 400666938) with three high transverse momentum jets.." CERN Document Server, 2018. ALICE Event Display. https://cds.cern.ch/record/2647094.

[95] "Pb ion collision from the last run of Run 2 recorded by ALICE." CERN Document Server, 2018. ALICE Event Display. https://cds.cern.ch/record/2649643.

[96] **ALICE** collaboration, C. Nattrass, *Measurements of Jets in ALICE*, *Acta Phys. Polon. B* **47** (2016) 1543.

[97] J. D. Bjorken, *Energy loss of energetic partons in quark-gluon plasma: possible extinction of high p_T jets in hadron-hadron collisions*, tech. rep., Aug., 1982. Preprint.

[98] **STAR** collaboration, *Direct observation of dijets in central Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV*, *Phys. Rev. Lett.* **97** (2006) 162301 [`arXiv:nucl-ex/0604018`].

[99] **PHENIX** collaboration, *Transverse momentum and centrality dependence of dihadron correlations in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV: Jet-quenching and the response of partonic matter*, *Phys. Rev. C* **77** (2008) 011901 [arXiv:0705.3238 [nucl-ex]].

[100] **ATLAS** collaboration, *Observation of a Centrality-Dependent Dijet Asymmetry in Lead-Lead Collisions at $\sqrt{s_{NN}} = 2.77$ TeV with the ATLAS Detector at the LHC*, *Phys. Rev. Lett.* **105** (2010) 252303 [arXiv:1011.6182 [hep-ex]].

[101] **CMS** collaboration, *Observation and studies of jet quenching in PbPb collisions at nucleon-nucleon center-of-mass energy = 2.76 TeV*, *Phys. Rev. C* **84** (2011) 024906 [arXiv:1102.1957 [nucl-ex]].

[102] **ALICE** collaboration, *Measurement of jet suppression in central Pb-Pb collisions at $\sqrt{s_{\rm NN}}$ = 2.76 TeV*, *Phys. Lett. B* **746** (2015) 1 [arXiv:1502.01689 [nucl-ex]].

[103] A. J. Larkoski *et. al.*, *Soft Drop*, *JHEP* **05** (2014) 146 [arXiv:1402.2657 [hep-ph]].

[104] **ATLAS** collaboration, *Measurement of substructure-dependent jet suppression in Pb+Pb collisions at 5.02 TeV with the ATLAS detector*, *Phys. Rev. C* **107** (2023) no. 5 054909 [arXiv:2211.11470 [nucl-ex]].

[105] **ALICE** collaboration, *Measurement of the groomed jet radius and momentum splitting fraction in pp and Pb−Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV*, *Phys. Rev. Lett.* **128** (2022) no. 10 102001 [arXiv:2107.12984 [nucl-ex]].

[106] J. Gallicchio and M. D. Schwartz, *Quark and Gluon Jet Substructure*, *JHEP* **04** (2013) 090 [arXiv:1211.7038 [hep-ph]].

[107] A. J. Larkoski, J. Thaler and W. J. Waalewijn, *Gaining (Mutual) Information about Quark/Gluon Discrimination*, *JHEP* **11** (2014) 129 [arXiv:1408.3122 [hep-ph]].

[108] A. J. Larkoski, G. P. Salam and J. Thaler, *Energy Correlation Functions for Jet Substructure*, *JHEP* **06** (2013) 108 [arXiv:1305.0007 [hep-ph]].

[109] f. t. A. C. Sheikh, Ashik Ikbal, *Measurements of heavy-flavor jets with ALICE at the LHC*, *Nucl. Part. Phys. Proc.* **309-311** (2020) 174 [arXiv:2009.02751 [nucl-ex]].

[110] J.-w. Qiu and I. Vitev, *Coherent QCD multiple scattering in proton-nucleus collisions*, *Phys. Lett. B* **632** (2006) 507 [arXiv:hep-ph/0405068].

[111] F. Gelis *et. al.*, *The Color Glass Condensate*, *Ann. Rev. Nucl. Part. Sci.* **60** (2010) 463 [arXiv:1002.0333 [hep-ph]].

[112] I. Vitev, *Non-Abelian energy loss in cold nuclear matter*, *Phys. Rev. C* **75** (2007) 064906 [arXiv:hep-ph/0703002].

[113] **ALICE** collaboration, *Measurement of inclusive charged-particle jet production in pp and p-Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV*, *JHEP* **05** (2024) 041 [arXiv:2307.10860 [nucl-ex]].

[114] J. W. Cronin *et. al.*, *Production of hadrons at large transverse momentum at 200, 300, and 400 GeV*, *Phys. Rev. D* **11** (Jun, 1975) 3105. https://link.aps.org/doi/10.1103/PhysRevD.11.3105.

[115] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029 [arXiv:hep-ph/0204244].

[116] J. S. Moreland, J. E. Bernhard and S. A. Bass, *Alternative ansatz to wounded nucleon and binary collision scaling in high-energy nuclear collisions*, *Phys. Rev. C* **92** (2015) no. 1 011901 [arXiv:1412.4708 [nucl-th]].

[117] B. Schenke, P. Tribedy and R. Venugopalan, *Fluctuating Glasma initial conditions and flow in heavy ion collisions*, *Phys. Rev. Lett.* **108** (2012) 252301 [arXiv:1202.6646 [nucl-th]].

[118] K. J. Eskola *et. al.*, *Scaling of transverse energies and multiplicities with atomic number and energy in ultrarelativistic nuclear collisions*, *Nucl. Phys. B* **570** (2000) 379 [arXiv:hep-ph/9909456].

[119] H. Niemi, K. J. Eskola and R. Paatelainen, *Event-by-event fluctuations in a perturbative QCD + saturation + hydrodynamics model: Determining QCD matter shear viscosity in ultrarelativistic heavy-ion collisions*, *Phys. Rev. C* **93** (2016) no. 2 024907 [arXiv:1505.02677 [hep-ph]].

[120] T. Sjöstrand in *23rd International Conference on High-Energy Physics*, 8, 1985.

[121] A. Dainese, C. Loizides and G. Paic, *Leading-particle suppression in high energy nucleus-nucleus collisions*, *Eur. Phys. J. C* **38** (2005) 461 [arXiv:hep-ph/0406201].

[122] K. J. Eskola *et. al.*, *The Fragility of high-p(T) hadron spectra as a hard probe*, *Nucl. Phys.* **A747** (2005) 511 [arXiv:hep-ph/0406319 [hep-ph]].

[123] T. Renk and K. Eskola, *Prospects of medium tomography using back-to-back hadron correlations*, *Phys. Rev.* **C75** (2007) 054910 [arXiv:hep-ph/0610059 [hep-ph]].

[124] **ATLAS** collaboration, *Search for New Physics in Dijet Mass and Angular Distributions in pp Collisions at $\sqrt{s} = 7$ TeV Measured with the ATLAS Detector*, *New J. Phys.* **13** (2011) 053044 [arXiv:1103.3864 [hep-ex]].

[125] **CMS** collaboration, *Search for Resonances in the Dijet Mass Spectrum from 7 TeV pp Collisions at CMS*, *Phys. Lett.* **B704** (2011) 123 [arXiv:1107.4771 [hep-ex]].

[126] **CMS** collaboration, *Search for Dijet Resonances in 7 TeV pp Collisions at CMS*, *Phys. Rev. Lett.* **105** (2010) 211801 [arXiv:1010.0203 [hep-ex]].

[127] **ATLAS** collaboration, *Search for dijet resonances in events with an isolated charged lepton using $\sqrt{s} = 13$ TeV proton-proton collision data collected by the ATLAS detector*, *JHEP* **06** (2020) 151 [arXiv:2002.11325 [hep-ex]].

[128] K. J. Eskola and H. Honkanen, *A Perturbative QCD analysis of charged particle distributions in hadronic and nuclear collisions*, *Nucl. Phys. A* **713** (2003) 167 [arXiv:hep-ph/0205048].

[129] J. Binnewies, B. A. Kniehl and G. Kramer, *Pion and kaon production in e+ e- and e p collisions at next-to-leading order*, *Phys. Rev. D* **52** (1995) 4947 [arXiv:hep-ph/9503464].

[130] B. A. Kniehl, G. Kramer and B. Potter, *Fragmentation functions for pions, kaons, and protons at next-to-leading order*, *Nucl. Phys. B* **582** (2000) 514 [arXiv:hep-ph/0010289].

[131] F. Arleo, *(Medium-modified) Fragmentation Functions*, *Eur. Phys. J. C* **61** (2009) 603 [arXiv:0810.1193 [hep-ph]].

[132] K. J. Eskola *et. al.*, *EPPS21: a global QCD analysis of nuclear PDFs*, *Eur. Phys. J. C* **82** (2022) no. 5 413 [arXiv:2112.12462 [hep-ph]].

[133] **H1, ZEUS** collaboration, H. Abramowicz *et. al.*, *Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data*, *Eur. Phys. J. C* **75** (2015) no. 12 580 [arXiv:1506.06042 [hep-ex]].

[134] **NNPDF** collaboration, R. D. Ball *et. al.*, *Parton distributions from high-precision collider data*, *Eur. Phys. J. C* **77** (2017) no. 10 663 [arXiv:1706.00428 [hep-ph]].

[135] T.-J. Hou *et. al.*, *New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC*, *Phys. Rev. D* **103** (2021) no. 1 014013 [arXiv:1912.10053 [hep-ph]].

[136] K. J. Eskola *et. al.*, *EPPS16: Nuclear parton distributions with LHC data*, *Eur. Phys. J. C* **77** (2017) no. 3 163 [arXiv:1612.05741 [hep-ph]].

[137] K. Kovarik *et. al.*, *nCTEQ15 - Global analysis of nuclear parton distributions with uncertainties in the CTEQ framework*, *Phys. Rev. D* **93** (2016) no. 8 085037 [arXiv:1509.00792 [hep-ph]].

[138] A. Khalek *et. al.*, *nNNPDF2.0: quark flavor separation in nuclei from LHC data*, *JHEP* **09** (2020) 183 [arXiv:2006.14629 [hep-ph]].

[139] A. Kusina *et. al.*, *Impact of LHC vector boson production in heavy ion collisions on strange PDFs*, *Eur. Phys. J. C* **80** (2020) no. 10 968 [arXiv:2007.09100 [hep-ph]].

[140] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438.

[141] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, *Nuclear Physics B* **126** (1977) no. 2 298. https://www.sciencedirect.com/science/article/pii/0550321377903844.

[142] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, *Sov. Phys. JETP* **46** (1977) 641.

[143] G. P. Salam, *An Introduction to leading and next-to-leading BFKL*, *Acta Phys. Polon. B* **30** (1999) 3679 [arXiv:hep-ph/9910492].

[144] A. H. Mueller, *A Simple derivation of the JIMWLK equation*, *Phys. Lett. B* **523** (2001) 243 [arXiv:hep-ph/0110169].

[145] T. Lappi and H. Mäntysaari, *Next-to-leading order Balitsky-Kovchegov equation with resummation*, *Phys. Rev. D* **93** (2016) no. 9 094004 [arXiv:1601.06598 [hep-ph]].

[146] A. Accardi *et. al.*, *Electron Ion Collider: The Next QCD Frontier: Understanding the glue that binds us all*, *Eur. Phys. J. A* **52** (2016) no. 9 268 [arXiv:1212.1701 [nucl-ex]].

[147] E. Iancu and R. Venugopalan, *The Color glass condensate and high-energy scattering in QCD*, pp. 249–3363. 3, 2003. arXiv:hep-ph/0303204.

[148] L. D. McLerran and R. Venugopalan, *Gluon distribution functions for very large nuclei at small transverse momentum*, *Phys. Rev. D* **49** (1994) 3352 [arXiv:hep-ph/9311205].

[149] F. Gelis, *Color Glass Condensate and Glasma*, *Int. J. Mod. Phys. A* **28** (2013) 1330001 [arXiv:1211.3327 [hep-ph]].

[150] **LHCb** collaboration, R. Aaij *et. al.*, *Study of prompt D^0 meson production in pPb collisions at $\sqrt{s_{\rm NN}} = 5$ TeV*, *JHEP* **10** (2017) 090 [arXiv:1707.02750 [hep-ex]].

[151] K. J. Eskola *et. al.*, *A QCD analysis of LHCb D-meson data in p+Pb collisions*, *JHEP* **05** (2020) 037 [arXiv:1906.02512 [hep-ph]].

[152] INCNLO, *INCNLO-direct photon and inclusive hadron production code - INCNLO version 1.4*, 2002. https://lapth.cnrs.fr/PHOX_FAMILY/readme_inc.html.

[153] F. Aversa *et. al.*, *QCD Corrections to Parton-Parton Scattering Processes*, *Nucl. Phys. B* **327** (1989) 105.

[154] P. Aurenche *et. al.*, *Prompt Photon Production at Large p(T) Scheme Invariant QCD Predictions and Comparison with Experiment*, *Nucl. Phys. B* **297** (1988) 661.

[155] P. Aurenche *et. al.*, *A Critical phenomenological study of inclusive photon production in hadronic collisions*, *Eur. Phys. J. C* **9** (1999) 107 [arXiv:hep-ph/9811382].

[156] P. Aurenche *et. al.*, *Large p(T) inclusive pi0 cross-sections and next-to-leading-order QCD predictions*, *Eur. Phys. J. C* **13** (2000) 347 [arXiv:hep-ph/9910252].

[157] R. A. Khalek *et. al.*, *nNNPDF3.0: Evidence for a modified partonic structure in heavy nuclei*, arXiv:2201.12363 [hep-ph].

[158] L. Wang *et. al.*, *Forward inclusive jet productions in pA collisions*, *Phys. Rev. D* **107** (2023) no. 1 016016 [arXiv:2211.08322 [hep-ph]].

[159] H.-y. Liu *et. al.*, *Single inclusive jet production in pA collisions at NLO in the small-x regime*, *JHEP* **07** (2022) 041 [arXiv:2204.03026 [hep-ph]].

[160] H. Mäntysaari and H. Paukkunen, *Saturation and forward jets in proton-lead collisions at the LHC*, *Phys. Rev. D* **100** (2019) no. 11 114029 [arXiv:1910.13116 [hep-ph]].

[161] A. van Hameren *et. al.*, *Forward di-jet production in p+Pb collisions in the small-x improved TMD factorization framework*, *JHEP* **12** (2016) 034 [arXiv:1607.03121 [hep-ph]]. [Erratum: JHEP 02, 158 (2019)].

[162] A. van Hameren, *KaTie : For parton-level event generation with k_T-dependent initial states*, *Comput. Phys. Commun.* **224** (2018) 371 [arXiv:1611.00680 [hep-ph]].

[163] Al-Mashad *et. al.*, *Dijet azimuthal correlations in p-p and p-Pb collisions at forward LHC calorimeters*, *JHEP* **12** (2022) 131 [arXiv:2210.06613 [hep-ph]].

[164] P. Kotko *et. al.*, *Improved TMD factorization for forward dijet production in dilute-dense hadronic collisions*, *JHEP* **09** (2015) 106 [arXiv:1503.03421 [hep-ph]].

[165] D. Gillberg, *Performance of the ATLAS Forward Calorimeters in First LHC Data*, tech. rep., CERN, Geneva, 2010.

[166] **ALICE** collaboration, *Physics of the ALICE Forward Calorimeter upgrade*, tech. rep., CERN, 2023. ALICE Public Note.

[167] **CMS** collaboration, *Studies of Dijet Transverse Momentum Balance and Pseudorapidity Distributions in pPb Collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV*, *Eur. Phys. J. C* **74** (2014) no. 7 2951 [arXiv:1401.4433 [nucl-ex]].

[168] G. C. Blazey *et. al.* in *Physics at Run II: QCD and Weak Boson Physics Workshop: Final General Meeting*, pp. 47–77, 5, 2000. arXiv:hep-ex/0005012.

[169] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, *Eur.Phys.J.* **C72** (2012) 1896 [arXiv:1111.6097 [hep-ph]].

[170] **CMS** collaboration, *Measurement of charged pion, kaon, and proton production in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. D* **96** (2017) no. 11 112003 [arXiv:1706.10194 [hep-ex]].

[171] **STAR** collaboration, *Systematic Measurements of Identified Particle Spectra in pp, d^+ Au and Au+Au Collisions from STAR*, *Phys. Rev. C* **79** (2009) 034909 [arXiv:0808.2041 [nucl-ex]].

[172] **STAR** collaboration, *Pion, kaon, and (anti)proton production in U+U collisions at sNN=193 GeV measured with the STAR detector*, *Phys. Rev. C* **107** (2023) no. 2 024901 [arXiv:2208.00653 [nucl-ex]].

[173] J. M. Butterworth *et. al.*, *KtJet: A C++ implementation of the K-perpendicular clustering algorithm*, *Comput. Phys. Commun.* **153** (2003) 85 [arXiv:hep-ph/0210022].

[174] M. Cacciari, G. P. Salam and G. Soyez, *The catchment area of jets*, *Journal of High Energy Physics* **2008** (Apr., 2008) 005–005. http://dx.doi.org/10.1088/1126-6708/2008/04/005.

[175] J. E. Huth *et. al.* in *1990 DPF Summer Study on High-energy Physics: Research Directions for the Decade (Snowmass 90)*, pp. 0134–136, 12, 1990.

[176] G. P. Salam and G. Soyez, *A practical seedless infrared-safe cone jet algorithm*, *Journal of High Energy Physics* **2007** (May, 2007) 086–086. http://dx.doi.org/10.1088/1126-6708/2007/05/086.

[177] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev. D* **48** (1993) 3160 [arXiv:hep-ph/9305266].

[178] M. Cacciari, G. P. Salam and G. Soyez, *The Anti-k(t) jet clustering algorithm*, *JHEP* **04** (2008) 063 [arXiv:0802.1189 [hep-ph]].

[179] Y. L. Dokshitzer *et. al.*, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001 [arXiv:hep-ph/9707323].

[180] M. Wobisch and T. Wengler in *Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting)*, pp. 270–279, 4, 1998. arXiv:hep-ph/9907280.

[181] G. D'Agostini, *A multidimensional unfolding method based on Bayes' Theorem*, tech. rep., DESY, Hamburg, 1994.

[182] V. Blobel, *Unfolding methods in high-energy physics experiments*, tech. rep., 1985.

[183] G. D'Agostini, *Improved iterative bayesian unfolding*, 2010. https://arxiv.org/abs/1010.0632.

[184] A. Hocker and V. Kartvelishvili, *SVD approach to data unfolding*, *Nucl. Instrum. Meth. A* **372** (1996) 469 [arXiv:hep-ph/9509307].

[185] K. Tackmann and A. Hoecker, *Svd-based unfolding: implementation and experience*, 2011. https://arxiv.org/abs/1112.2226.

[186] **ALICE** collaboration, *Measurement of Event Background Fluctuations for Charged Particle Jet Reconstruction in Pb-Pb collisions at* $\sqrt{s_{NN}} = 2.76$ *TeV*, *JHEP* **03** (2012) 053 [arXiv:1201.2423 [hep-ex]].

[187] **ALICE** collaboration, *Measurement of charged jet suppression in Pb-Pb collisions at* $\sqrt{s_{NN}}$ *= 2.76 TeV*, *JHEP* **03** (2014) 013 [arXiv:1311.0633 [nucl-ex]].

[188] **ALICE** collaboration, *Pseudorapidity density of charged particles in p + Pb collisions at* $\sqrt{s_{NN}} = 5.02$ *TeV*, *Phys. Rev. Lett.* **110** (2013) no. 3 032301 [arXiv:1210.3615 [nucl-ex]].

[189] **ALICE** collaboration, *Charged-particle multiplicity density at mid-rapidity in central Pb-Pb collisions at* $\sqrt{s_{NN}} = 2.76$ *TeV*, *Phys. Rev. Lett.* **105** (2010) 252301 [arXiv:1011.3916 [nucl-ex]].

[190] **CMS** collaboration, *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* **08** (2012) 130 [arXiv:1207.2392 [hep-ex]].

[191] **ALICE** collaboration, *Measurement of charged jet production cross sections and nuclear modification in p-Pb collisions at* $\sqrt{s_{\mathrm{NN}}} = 5.02$ *TeV*, *Phys. Lett. B* **749** (2015) 68 [arXiv:1503.00681 [nucl-ex]].

[192] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, *Physics Letters B* **659** (Jan., 2008) 119–126. http://dx.doi.org/10.1016/j.physletb.2007.09.077.

[193] G. Soyez *et. al.*, *Pileup subtraction for jet shapes*, *Physical Review Letters* **110** (Apr., 2013) . http://dx.doi.org/10.1103/PhysRevLett.110.162001.

[194] **CDF** collaboration, *Two-Particle Momentum Correlations in Jets Produced in* $p\bar{p}$ *Collisions at* $\sqrt{s}$ *= 1.96 TeV*, *Phys. Rev. D* **77** (2008) 092001 [arXiv:0802.3182 [hep-ex]].

[195] J. Bellm *et. al.*, *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J. C* **76** (2016) no. 4 196 [arXiv:1512.01178 [hep-ph]].

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

[197] T. Gleisberg *et. al.*, *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007 [arXiv:0811.4622 [hep-ph]].

[198] K. C. Zapp, *JEWEL 2.0.0: directions for use*, *Eur. Phys. J. C* **74** (2014) no. 2 2762 [arXiv:1311.0048 [hep-ph]].

[199] J. H. Putschke *et. al.*, *The JETSCAPE framework*, arXiv:1903.07706 [nucl-th].

[200] M. Gyulassy and X.-N. Wang, *HIJING 1.0: A Monte Carlo program for parton and particle production in high-energy hadronic and nuclear collisions*, *Comput. Phys. Commun.* **83** (1994) 307 [arXiv:nucl-th/9502021].

[201] T. Pierog *et. al.*, *EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider*, *Phys. Rev. C* **92** (2015) no. 3 034906 [arXiv:1306.0121 [hep-ph]].

[202] C. Bierlich *et. al.*, *The Angantyr model for Heavy-Ion Collisions in PYTHIA8*, *JHEP* **10** (2018) 134 [arXiv:1806.10820 [hep-ph]].

[203] J. Bellm and C. Bierlich, *PISTA: Posterior Ion STAcking*, arXiv:1807.01291 [hep-ph].

[204] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [arXiv:hep-ph/0409146].

[205] F. Buccioni *et. al.*, *OpenLoops 2*, *Eur. Phys. J. C* **79** (2019) no. 10 866 [arXiv:1907.13071 [hep-ph]].

[206] A. Buckley *et. al.*, *General-purpose event generators for LHC physics*, *Phys. Rept.* **504** (2011) 145 [arXiv:1101.2599 [hep-ph]].

[207] C. Bierlich *et. al.*, *A comprehensive guide to the physics and usage of PYTHIA 8.3*, 2022. https://arxiv.org/abs/2203.11601.

[208] A. Buckley *et. al.*, *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J. C* **75** (2015) 132 [arXiv:1412.7420 [hep-ph]].

[209] **ALICE** collaboration, *Centrality dependence of charged jet production in p-Pb collisions at* $\sqrt{s_{\mathrm{NN}}}$ *= 5.02 TeV*, arXiv:1603.03402 [nucl-ex].

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

[211] R. Brun *et. al.*, *GEANT: Detector Description and Simulation Tool*. CERN Program Library. CERN, Geneva, 1993. Long Writeup W5013.

[212] **GEANT4** collaboration, S. Agostinelli *et. al.*, *GEANT4 - A Simulation Toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.

[213] R. Esteve Bosch *et. al.*, *The ALTRO chip: A 16-channel A/D converter and digital processor for gas detectors*, *IEEE Trans. Nucl. Sci.* **50** (2003) 2460.

[214] **ALICE** collaboration, *AliPhysics*, 3, 2023. https://github.com/alisw/AliPhysics.

[215] **ALICE** collaboration, *Centrality dependence of particle production in p-Pb collisions at* $\sqrt{s_{\rm NN}}$*= 5.02 TeV*, *Phys. Rev. C* **91** (2015) no. 6 064905 [arXiv:1412.6828 [nucl-ex]].

[216] **CMS** collaboration, *CMS-TOTEM feasibility studies for single diffractive Z, W, Jpsi and central exclusive dijet production in pp collisions at 13 TeV*, tech. rep., CERN, Geneva, 2014.

[217] **TOTEM** collaboration, *The TOTEM experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08007.

[218] W. Florkowski, *Phenomenology of Ultra-Relativistic Heavy-Ion Collisions*. 3, 2010.

[219] **ALICE** collaboration, *Measurements of inclusive jet spectra in pp and central Pb-Pb collisions at* $\sqrt{s_{\rm NN}}$ *= 5.02 TeV*, *Phys. Rev. C* **101** (2020) no. 3 034911 [arXiv:1909.09718 [nucl-ex]].

[220] **ALICE** collaboration, *Measurement of inclusive and leading subjet fragmentation in pp and Pb–Pb collisions at* $\sqrt{s_{NN}}$ *= 5.02 TeV*, *Journal of High Energy Physics* **2023** (May, 2023) . http://dx.doi.org/10.1007/JHEP05(2023)245.

[221] CERN, *Proceedings of the 1974 CERN School of Computing*, (Geneva), CERN, 1974.

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

[223] M. Livan and R. Wigmans, *Calorimetry for Collider Physics, an Introduction*. UNITEXT for Physics. Springer, 2019.

[224] J. E. Gaiser, *Appendix-F Charmonium Spectroscopy from Radiative Decays of the J/Psi and Psi-Prime*. Ph.d. thesis, SLAC, 1982. Page 178.

[225] **ALICE** collaboration, *The ALICE definition of primary particles*, . https://cds.cern.ch/record/2270008.

[226] C.-Y. Wong and G. Wilk, *Tsallis fits to* p_T *spectra and multiple hard scattering in pp collisions at the LHC*, *Phys. Rev. D* **87** (2013) no. 11 114007 [arXiv:1305.2627 [hep-ph]].

[227] D. Wang, L. Wu, J. M. Yang and M. Zhang, *Photon-jet events as a probe of axionlike particles at the LHC*, *Phys. Rev. D* **104** (2021) no. 9 095016 [arXiv:2102.01532 [hep-ph]].

[228] **CMS** collaboration, *Jet energy scale and resolution measurement with Run 2 Legacy Data Collected by CMS at 13 TeV*, . https://cds.cern.ch/record/2792322.